Good science at heart

Lessons on ethics from big data research in sudden cardiac arrest

Bak, M.A.R.

Publication date
2021

1 | Introduction and outline

Everything's got a moral, if only you can find it.
(Lewis Carroll, *Alice's Adventures in Wonderland*. 1865; p. 71)

This would not be a book on medical research ethics if it did not mention the Second World War. So let me begin with the beginning. Not with the medical atrocities performed by Josef Mengele and others, nor with the Nuremberg Trials that resulted in the similarly named Code about ethical research with humans, but with something else: with information. Gregory Bateson defined information as "a difference that makes a difference".[1] Separate datapoints exist through a lack of uniformity, and when data are endowed with meaning, they can make all the difference in the world.

In the 1930s, my grandfather Dirk Bak (1891 - 1951) worked as manager at the regional newspaper De Rotterdammer. The paper was discontinued in October 1941 because it did not want to cooperate with the "new order".[i] During the war Dirk helped to supply underground resistance newspapers with paper, before he went on to work as director for the printing company Libertas in Utrecht.[ii] While the clandestine printing paper helped inform the resistance against the Nazis, on the other side paper records were used to the detriment of the Jewish community. Compared to the rest of Western Europe, the relatively largest number of Jewish people were deported from the Netherlands, partly due to our country's well-organised census. This comprehensive system for population registration was created before the occupation, for administrative purposes and social research, but used by the Nazis for genocide.[2] Recordkeeping produced data which were turned into information for mass murder; committed by the people who had access to this information. 'I have nothing to hide' is thus a sentence uttered only by the ignorant: we do not always know who has our personal data, why, and how they will use the information encoded in our data.[3]

A few kilometres from where I am writing this thesis, Anne Frank wrote in her famous diary: "in spite of everything I still believe that people are really *good at heart*."[iii] It is common knowledge that someone informed against her family causing their deportation, but less well-known is the fact that Anne's diary contained two hidden pages which had been pasted over with brown paper – these pages concealed literary contemplations about sexual matters. Her self-editing may have been a preparation

[i] Nieuwe Leidsche Courant, Saturday 3 March 1951. 'D. Bak te Utrecht overleden'.
[ii] At the time of writing in 2021, the company is still operational under the name Libertas-Pascal and is responsible for printing this thesis in an environmentally sustainable manner.
[iii] Anne Frank, *Het Achterhuis [The Secret Annex]*, Saturday 15 July 1944.

for a new version of the diary that she hoped would be published as an eyewitness account after the war.[iv] Self-censuring is not uncommon for writers, and may go much further: we know that Jewish writer Franz Kafka included a provision in his will to have all his manuscripts burned after his death but, luckily for us, the executor of the will did not honour this request. Anne Frank's wish was not upheld either. In 2018, researchers could uncover the text on the two pages for the first time, with the help of new photo-imaging technology that worked without touching the paper.

Technology thus creates new possibilities for interacting with existing sources of information. Currently, big data analytics are accelerating the potential to do good (or bad). What are the ethical aspects associated with such large-scale data collection? Should we respect the wishes and privacy concerns of the dead? What are good uses of data? In the field of medical research, these questions are highly relevant as state-of-the-art research is being conducted with health data of vulnerable patients. In this thesis I study what 'good' means for data-based research in sudden cardiac arrest.

1.1 The ethics of big health data research

Along with developments in computing power and data storage prices, increasingly large numbers of data can be stored and analysed.[4] The use of 'Big Data' in research may help to improve the quality of healthcare, to gain understanding of the underlying factors of health and disease, and to lower costs.[5] The standard definition of big data utilises a number of terms starting with a 'v' to explain what is so *big* about these data.[6, 7] If we focus on to the most common ones, big data is characterised by:

1. *Volume*: the sheer amount of available data, which has increased significantly in the past years and is estimated to double in the next five years.[v]
2. *Velocity*: the high speed at which data are collected, processed and analysed.
3. *Variety*: the large number of data sources and heterogeneity of datasets.
4. *Veracity*: the accuracy of the information encoded in the data, which is generally lower when data are collected from many different sources.
5. *Variability*: the consistency of data over time, e.g. to what extent the outcomes change when new data are added to a dataset.
6. *Value*: the worth of the data, which depends on whether they are analysed in an appropriate and scientifically valid manner.

[iv] Nina Siegal, May 15 2018, The New York Times. 'Researchers Uncover Two Hidden Pages in Anne Frank's Diary'. [https://www.nytimes.com/2018/05/15/books/anne-frank-diary-new-pages.html]

[v] Statista, June 2021. 'Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025'. [https://www.statista.com/statistics/871513/worldwide-data-created/]

In medicine, secondary analysis of big data is used among others to identify new risk factors and inform the design of clinical trials, which can be done at much lower cost than if the researchers would have collected the data themselves.[8] Valuable data may include clinical trial data, genetic sequencing data, electronic health records, administrative hospital or pharmacy data, information in cause-of-death registries, data generated through social media platforms or smartphone applications and wearable devices, and so on. Modern information technologies are used to link and analyse these data, with a growing reliance on algorithms and since recently also on advanced artificial intelligence (AI) models.[9] An example of a practical application is the identification of skin cancer in images taken with smartphones, which neural networks can do just as well as expert oncologists.[10] The majority of healthcare organisations are now using, or planning to use, big data and AI-based methods, and the economic value of the global market of big data analytics in healthcare was estimated at 24 billion USD in the year 2020.[vi] Despite their potential for improving care, however, big data analytics also create ethical concerns due to the sensitivity of health data.

Even before the Economist proclaimed in 2017 that "the world's most valuable resource is no longer oil, but data", others had referred to big data as the new nuclear energy: "we should treat personal electronic data with the same care and respect as weapons-grade plutonium – it is dangerous, long-lasting and once it has leaked there's no getting it back".[vii] Personal data is data which can be traced back to identifiable persons without unreasonable effort, and it is vital to properly store and handle such data. Data breaches are common in the online environment, including for health data. For example, in 2014, a study of thousands of health-related websites showed that in about 90% of website visits, health information was shared with third parties like advertisers and data brokers.[11] An intriguing insight into online data privacy is given by a documentary ('I love Alaska')[viii] about a major data breach by the US-based internet provider AOL that posted the searches of their clients online, coded by user number – the document had been meant for academic research, but could be used by reporters to identify individuals in the dataset. In the documentary, we get to know user #711391 by their searches, which reveal the incredibly personal nature of these data. For instance, on March 2, 2006 she searched for the following phrases: *people are not always how they seem over the internet*; *houston texas is one hot place to live*; *gay churches in houston texas*; *can liver problems cause you to lose your hair*; *pimple that gets white head on it and never goes away*. Within a few lines, we know about this internet user's sexual orientation, religion, possible health problems, and place of residence. Combining all her searches, a complete profile can be created.

[vi] Visiongain Ltd., June 2021. 'Big Data Analytics in the Healthcare Market 2021-2031'.

[vii] Cory Doctorow, 15 January 2008, the Guardian. 'Personal data is as hot as nuclear waste'. [https://www.theguardian.com/technology/2008/jan/15/data.security]

[viii] Lemert Engelberts & Sander Plug, 2008. 'I love Alaska'. Available on YouTube.

The ethical, legal and social implications (ELSI) of big data have received increasing attention in the past years, including in the medical field.[12] This includes questions around informed consent, privacy and data protection, ownership, and discrimination. Ethical aspects can be viewed using a deontological or consequentialist perspective, i.e. defining the good in terms of moral duties or consequences of actions, respectively. The ethics of medical research are also frequently analysed using the principlist approach of Beauchamp and Childress who proposed that four principles lie at the core of medical ethics: beneficence, non-maleficence, autonomy, and justice.[13] Autonomy, which refers to the right to make informed decisions about oneself, has become increasingly important in modern day care and forms the foundation for the moral and legal right to privacy (the latter is further discussed in Section 1.2). Traditional ethical guidelines like the Helsinki Declaration and the Nuremberg Code were initially drafted for experimental research with human subjects rather than for observational studies and they are not legally binding. However, aspects of these codes such as the need for informed consent, are found in data protection legislation.[14] In addition, guidelines for data-driven health research have been created in recent year.[15,16] These documents generally agree that because of the privacy risks associated with personal health data processing, participants in observational studies should have the ability to control the sharing and uses of their data, including about whether they do *not* want to know certain information, such as genetic findings that may potentially impact their health.[17] However, the rights of data subjects may conflict with the needs of science so a careful weighing of values is needed.[18,19]

Some comments about scope and terminology are in order. While big data analytics can be used in healthcare organisations in various ways, including for improving staffing management, I will focus on the use of (big) data for observational research aimed at precision health in sudden cardiac arrest. The word big is bracketed because the issues discussed in this thesis also apply to datasets not matching all the traits of big data: the difference between small and big data is one of degree, not of type. For readability the terms 'big data research', 'data-driven research', 'observational research' and 'health data research' are used interchangeably throughout this thesis. For the same reason, I refer to people whose data are used for research as 'subjects' or 'participants', even though strictly speaking the word participant may be reserved for people who have a more active role in the research. Lastly, I mostly speak of *health* data rather than biomedical data. The translational potential of data-driven studies in medicine had initially been indicated with the terminology of 'personalised medicine', which was later rephrased as 'precision medicine' so not to overpromise the empowerment of patients, and finally to 'precision health'.[20] The move from medicine to health is needed to show the broader applications for public health and because big data research also includes data on people's behaviour and social environment.

1.2 Data protection legislation in Europe

Early mentions of a right to informational privacy (otherwise known as informational self-determination) date back to the late 19th century when US-based scholars Warren and Brandeis advocated for privacy rights in relation to new technologies such as photography and audio recordings.[21] The right to privacy was first established in international law through the Universal Declaration of Human Rights (UDHR) in 1948 and shortly thereafter in 1950 the right to private life was affirmed in the European Convention of Human Rights.[ix] In Europe, where the case study that grounds this thesis is situated, the right to personal data protection came to be recognised as a fundamental right from the 1970s onwards. This right to data protection is distinct from the right to private life by being more modern and active – although the values that these rights protect, namely autonomy and human dignity, are similar. Article 8 of the European Union's Charter of Fundamental Rights, which was proclaimed in the year 2000 to promote the coherence of the EU, details the key principles of data protection. These had already been enshrined in the EU Data Protection Directive 95/46/EC which was in force between 1995 and 2018, until the new General Data Protection Directive (GDPR) was implemented after a two-year transition period.

The GDPR is meant to bring a higher level of harmonisation of data protection practices, as it is directly applicable in Member States through national implementation laws. Moreover, the GDPR has an extended territorial scope since it applies to all EU citizens even when the data processing takes place outside of the EU. In Chapter 2 of this thesis, an overview is provided of the relevant differences between the former Directive and the new regulation. The GDPR recognises health-related, genetic and biometric data as a 'special category of data' which may not be processed in principle, unless exceptions apply. In general, the legal framework aims to find a good balance between protecting the privacy of data subjects and ensuring that much-needed research can still be conducted. The initial draft of the GDPR was especially strict about the requirement for informed consent, which led to worries among researchers who feared that medical research would be harmed.[22,23]

As a result of political comprises during the negotiations on the final version of the law (between the European Commission, European Parliament, and the Council of the European Union), the GDPR now leaves room in several places for Member States to create their own rules. Importantly, it allows countries to decide about exceptions from informed consent for observational research (Article 9.2.j). Thus,

[ix] The UDHR is a non-binding declaration but forms the basis of international human rights law. The only legally binding international instrument in the field of data protection, is the lesser known Convention for the Protection of Individuals with Regard to the Automatic Processing of Personal Data (Convention 108) which is a 1981 treaty by the Council of Euro[pe which is also open for signature to countries outside of Europe.

many of the specifics of safeguarding data subjects' rights in research, such as these 'research exemptions' from consent, are delegated to national law. Moreover, as we shall see in later chapters, the scope of the GDPR is not all-encompassing as it only applies to identifiable natural persons, i.e. the law does not cover anonymised data and deceased data subjects. Guidance on the use of data from incapacitated adults and the options for proxy consent is also lacking, while this would be valuable especially in the critical care setting.[24] In practice, the application of the GDPR thus remains a responsibility of national governments and requires analysis of national data protection laws and relevant health law (that usually regulates the issue of doctor-patient confidentiality), as well as European human rights law. For instance, in the Netherlands data protection for health research is regulated by the GDPR Implementation Act (*Uitvoeringswet Algemene Verordening Gegevensbescherming, UAVG*), as well as by the Medical Treatment Act *(WGBO)* and a code of conduct related to the use of data in health research *(Gedragscode Gezondheidsonderzoek)*. These laws are further discussed in Chapters 7 and 8 which reflect specifically on the Dutch situation.

1.3 The case of sudden cardiac arrest research

Sudden cardiac arrest (SCA) is deadly within minutes when left untreated and accounts for approximately one-fifth of all deaths in Europe, with average survival rates of only 8-14% even when cardiopulmonary resuscitation is attempted.[25] SCA occurs when the heart acutely stops beating due to electrical malfunctions causing lack of, or irregular, contractions of the heart. In colloquial language, SCA ('cardiac arrest') is often confused with myocardial infarction ('heart attack') which is caused by blockage of the coronary arteries – to use the analogy of house renovation, the former problem calls for an electrician whereas the latter requires a plumber. The most common cardiac arrythmia resulting in SCA is ventricular fibrillation (VF) which is an atypical contraction of the heart chambers.[26] The onset of SCA is associated with various risk factors, which can be either inherited (e.g., sex, ethnicity or specific genetic variants) or acquired (e.g. age, weight, stress, comorbidities, use of certain medication, or environmental factors such as pollution and socioeconomic status).[27-30]

Treatment of SCA consists of cardiopulmonary resuscitation (CPR) which aims to preserve blood circulation by giving chest compressions, often together with artificial ventilation, until spontaneous circulation returns. In addition, the use of automated external defibrillators (AEDs) is recommended to restore the heart rhythm, as the use of onsite AEDs leads to twice the amount of people surviving with their neurological capacities intact.[31,32] People who are known to be at increased risk of SCA may receive an implantable cardioverter-defibrillator (ICD) which works in a similar fashion by giving shocks that help to 'reset' the heart to a normal rhythm.[33] Survival rates vary widely cross Europe and are influenced by factors related to the

SCA event, e.g. whether it was witnessed, if there was a shockable rhythm such as VF, and the time that passed before arrival of emergency medical services (EMS).[25] In addition, the differences in the organisation of healthcare systems may contribute to this variation. For instance, higher survival rates are found in regions and countries where EMS dispatch centres employ 'first responders' such as firefighters, police officers, and trained citizens to perform CPR on SCA victims.[34] Figure 1-1 shows the *chain of survival* depicting the phases of care for an SCA victim, where early recognition and calling the alarm number are initially important, followed by early CPR and defibrillation, and finally by high quality post-resuscitation care in the hospital.[35]

Differences in incidence and survival rates may reflect true differences but can also result from variation in reporting of SCA across different regions. Moreover, observational research in SCA is complicated by the variety of data sources needed for data collection that reflect the links in the chain of survival. Initial data collection is done through the dispatch centre, the EMS attending to the victim (and in some cases from the first-responders and AED devices), and the treating hospital.[36] Data collection is challenging because of the velocity of data about the resuscitation attempt, e.g. ambulance ECGs need to be collected quickly as they may be overwritten when the device is used on an new patient. Some SCA cohort studies collect additional data, such as the Amsterdam Resuscitation Studies (ARREST) researchers who obtain information about the patient's medication use and medical history, as well as DNA samples of certain patient groups.[37] These data are then also linked with statistics registries containing socio-economic data. Personal data are stored in a coded manner and surviving patients are asked for consent. While the ARREST registry collects highly detailed data from patients in several regions in the Netherlands, obtaining further understanding of the complex causal factors of SCA requires large-scale databases that necessarily span multiple countries. For this reason, the ESCAPE-NET project was created: the European Sudden Cardiac Arrest network towards Prevention, Education and New Effective treatment.

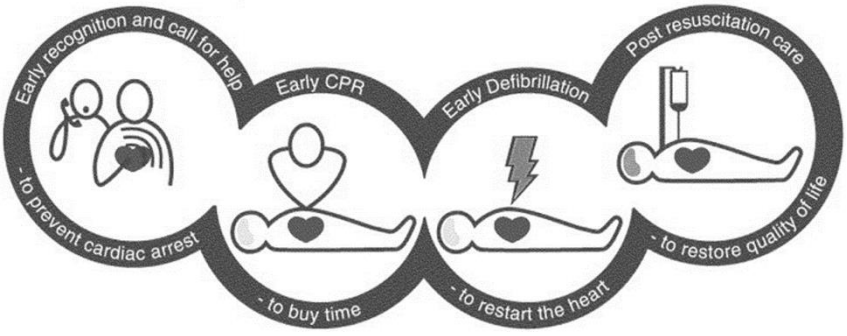

Figure 1-1. The chain of survival after sudden cardiac arrest. *Source: Nolan et al., 2006.*[35]

ECAPE-NET is an international consortium funded through a Horizon 2020 grant by the European Commission, which combines and harmonises SCA cohorts from various countries into a 'big' database of pseudonymised data.[27,38] Figure 1-2 shows the countries contributing to ESCAPE-NET and regions where collection of deep-phenotyped data takes place. Through this database, the project aims to increase understanding of SCA risk factors and improve prevention and treatment strategies. However, the processing of data from these vulnerable patients brings about a number of ethical concerns, especially at this large European scale. The veracity of data could be questioned due to the heterogeneity of sources contributing to ESCAPE-NET, and the international differences in ethical policies may complicate the conduct of joint research while respecting patients' privacy rights. These issues are exacerbated by the medical emergency setting where patients cannot give their prior informed consent for data processing. A large majority of people does not survive the SCA at all, while a smaller group survives with diminished mental capabilities. How should researchers then handle the personal data of SCA patients? That is the question driving the research presented in this thesis, which is elaborated on hereafter.

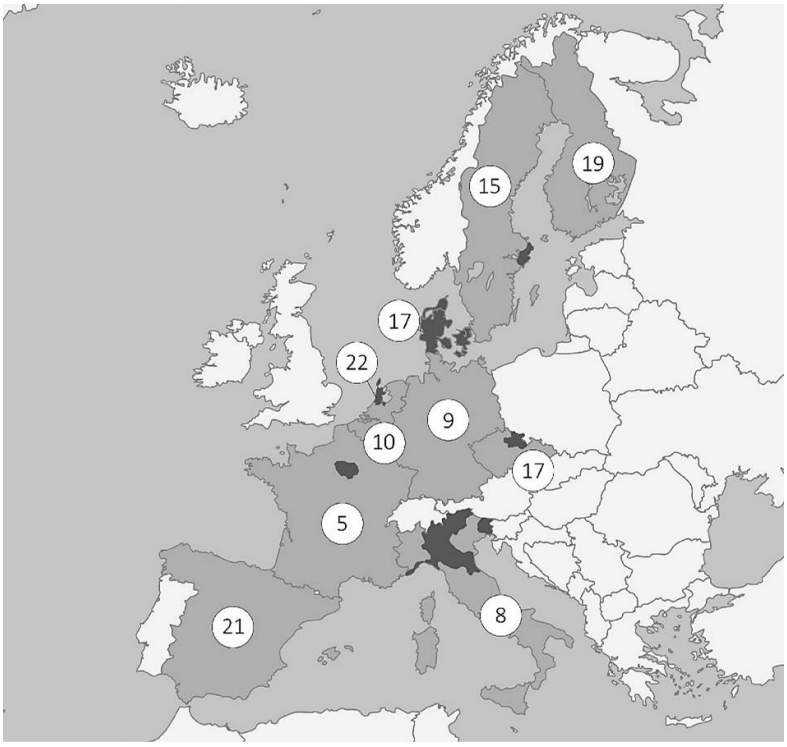

Figure 1-2. Differences in survival rates after SCA across Europe. Countries in middle grey indicate study sites of ESCAPE-NET; the catchment areas in these countries are indicated in dark grey. Numbers indicate survival rates (%) at hospital discharge of patients in whom CPR was attempted. *Source: Empana et al., 2018.*[27]

1.4 Aims and methods

The studies presented in this thesis are situated in the ESCAPE-NET project where the 'ethics' work package serves to identify and address relevant ethical aspects. The primary aim of this thesis is therefore to explore the ethical challenges – as well as potential solutions to these – of large-scale observational SCA research. The secondary aim is to provide guidance for other big health data initiatives, especially regarding two areas that have received less attention in the literature and in guidance documents: the use of deceased persons' data for research purposes, and the potential role of data-driven research in care. My main research question can be formulated in the following manner: *What are good ways of combining, analysing and translating large (international) collections of SCA research data?* The good in this sentence, and in the thesis's title, can be interpreted in two ways. Good science may be interpreted as valid and reliable science, but methodological scientific requirements are outside my scope and expertise.[x] I am concerned instead with the interpretation of the good as the ethically justifiable. The sub-questions addressed in this thesis are the following:

1. What are the potential *benefits and harms* associated with data-driven SCA research? How are these different for large-scale international databases?
2. What are good and transnationally justifiable practices for collecting, storing, and utilising data from SCA patients including those *who did not survive*?
3. How can *(personalized) treatment and prevention* strategies on the basis of SCA data research be created in a responsible manner?

ESCAPE-NET is used as case study to answer these questions as well as to fulfil the secondary aim of providing guidance for health data research in general.[xi] Acute and critical illness provides a valuable case for the study of medical research ethics because it functions as a *critical case*, meaning one that "can make a point quite dramatically" (p. 236).[39] Firstly, as described earlier, SCA patients are particularly vulnerable because they are acutely unconscious and cannot give prior consent (or at all if they die) for study procedures and data collection. Secondly, SCA research has greater life-saving impact, given low survival rates and limited options for improvement (e.g. people are currently unlikely to survive an unwitnessed cardiac arrest); this also means that clinical decision-making about treatment and prevention comes with greater responsibility (e.g. the decision whether to implant an ICD makes the difference between life or death). The emergency setting thus magnifies the ethical issues

[x] Good science, interpreted in this way, may be a tautology because only studies that lead to actual knowledge, would deserve to be called part of the scientific enterprise.

[xi] Within the ESCAPE-NET consortium, each participating research group functioned as a case-within-case for the case study of ESCAPE-NET that forms the basis for this thesis.

of health data research, making them easier to study while findings from the case of ESCAPE-NET can still be logically generalised to other types of illness. Logical or analytical generalisation refers to generalising theoretical propositions, as opposed to statistical generalization.[39,40] For instance, if this thesis were to find that deceased patients' data should always be anonymised for research, then it is logical to assume that the same requirement should apply to living persons' data.

1.4.1 Theoretical framework

Case studies are germane to *ethics parallel research*, that is, research where bioethicists are embedded within medical research consortia and tasked with providing guidance on and studying the ELSI of a specific project, such as ESCAPE-NET. This parallel and embedded study provides an opportunity for ethicists to constructively guide the normative implications of innovations in biomedicine, while these are being developed.[41-43] SCA research and its products can influence the wellbeing of participants and society at large, so the mitigation of risks requires involving patients and the public, as well as other relevant stakeholders. Engagement of stakeholders is required both from a democratic perspective and to benefit from the practical knowledge and experiences of these stakeholders. For instance, interview studies with SCA survivors have shown how the near-death experience disrupted people's lives not only in terms of physical, emotional and social challenges, but also had an existential impact; leading to suggestions for addressing the unmet needs of survivors in terms of support for coping with this turning point in life.[44,45] In ethics, the recognition of the patient perspective and a desire for evidence-based guidance have led to increasing integration of *empirical* research into bioethics, and contemporary ethics parallel research projects typically include empirical methods.[43,46] This 'empirical turn' is relatively recent: according to Robert Zussman who conducted extensive field research in intensive care units in the US, the dominant paradigm for medical ethics in the early 1990s was still characterised by a neglect of context, in terms of ignoring actual practices of care and the social background of patients and health care providers.[47]

What is empirical ethics? Empirical ethics research is descriptive as well as normative, and its ultimate aim is to improve the context-sensitivity of bioethics.[48] In this view, ethics guidance should be finetuned to the particular context (as the nature of principles and scope of their application depend on contextual factors), which can be done through empirical study of that context. Moreover, real-world practices are themselves sources of morality from which relevant ethical principles derive, rather than having these externally imposed. Thus, there is an intrinsic link between ethical judgements and the social sciences. Musschenga describes two meta-ethical paradigms that view context in this way: *epistemic contextualism* and the Rawlsian method of *reflective equilibrium* (which is a form of coherentism).[48] While contextualists aim to

reconstruct the values and norms internal to a certain practice, coherentists use reflective equilibrium to reveal whether moral judgements are coherent with theoretical principles and contextual facts. Discussion of these meta-ethical paradigms is beyond my scope; it suffices to say that the studies in thesis are grounded in contextualism, as I think this paradigm better captures the social nature of ethics than its alternative, coherentism, which arguably provides insufficient basis for moral justification.[49,50]

The …*at heart* from the title of this thesis is a phrase used to say what someone or something is really like, so my main research question may be re-formulated in the language of empirical ethics as: 'What is ethically justifiable data research about patients whose heart acutely stops beating, *really like*?'. Pols describes how empirical ethics can be used to analyse the intra-normativity of practices in health care, with a particular focus on the relations between people and with technologies.[51] This relational account can be illustrated with the citation from Lewis Carroll's famous story at the opening of this chapter. During a game of croquet, the Duchess tells Alice that everything's got a moral, if only you can find it: morality is in the object (emphasis on got) as well as in the observer (emphasis on you).[52] This leads to this thesis's background assumption that in SCA data research, neither the data themselves nor the researchers' uses of the data are morally neutral. Moreover, the quote illustrates that my findings reflect both the perspectives of interviewed stakeholders as well as of myself as interpreter.[53] Lastly, when the word 'if' in the Duchess's statement is emphasised, the two parts of the sentence reveal a dialectical process in which it is up to the observer to tease out the moral. This dialectical method stems from Hegel and Gadamer and constitutes another important aspect of the empirical ethics practised in this thesis. According to Widdershoven, the role of the bioethicist is to be in dialogue with relevant stakeholders and help them explicate their moral views and address ethical challenges.[54] This is also the role of the ethicists in ESCAPE-NET: I do not speak for SCA researchers, but hope that my findings can help these researchers in further developing their moral views and ethical practices.

1.4.2 Qualitative research methods

Various studies in this thesis report empirical findings (Chapters 3, 4, 5 and 11) and the more theoretical chapters also remain grounded in practice as a result of being embedded in an SCA research consortium. The empirical studies are of qualitative nature and include interviews as well as an invitational 'round table' conference. The value of qualitative research for ethics lies in the fact that it provides highly contextual knowledge and a holistic overview of an issue.[55] Qualitative research is especially useful when relatively little is known about a topic, which is the case for the ethics of data-driven SCA research.[56] For this thesis, I have conducted a total of 53 detailed semi-structured interviews that lasted between approximately one and two hours

each. Interviewed stakeholders included SCA survivors, their next-of-kin, researchers working with SCA registries and/or genetic biobanks, and ethical and legal experts. I also conducted ethnographic observations and document analysis related to data collection procedures and policies at various ESCAPE-NET partners, particularly at the Dutch ARREST group that I joined for the past years. In addition to the qualitative interview studies in this thesis, I reviewed the empirical literature about the post-mortem use of health research data, which included the perspectives of a total of approximately 5,400 individuals. Studies in this thesis were exempted from review by a research ethics committee, but respondent identifiers were still pseudonymised and consent for participation was asked.

The majority of interviews were conducted, analysed and discussed together with other researchers and supervised master's students, which increases the reliability of the conclusions drawn.[57] Of note is that I did the interviews with SCA researchers alone, because these involved visits to and observations of various ESCAPE-NET study sites. The internal validity of qualitative studies was safeguarded by sharing transcripts and findings with participants to enable corrections.[58] Coding of interview transcripts was done through an iterative process in which the coding tree was modified after each subsequent interview, and involved open coding followed by a mapping of these codes to general themes found through Chapter 2's review of the literature.[53] Scholarly works in bioethics and other relevant fields such as law, sociology and political philosophy, were used to analyse findings and confront these with theory and practice. Details about the methods of each specific study are provided in the corresponding chapters. A reference list with the literature consulted for this thesis is found after the general discussion (Chapter 12). Lastly, it should be remarked that the various chapters were initially written as stand-alone papers for scientific journals, so the transitions are less smooth than in a monograph.

1.5 Thesis outline

This thesis consists of three connected parts. Part I discusses the particular ethical, legal and social implications (ELSI) of observational sudden cardiac arrest research. In *Chapter 2*, the findings of a narrative literature review on this topic are presented. Specific themes found through this review are discussed with SCA survivors and their next-of-kin in *Chapter 3*. These themes are further debated in *Chapter 4,* for which I interviewed an international group of SCA researchers as well as experts in ethics and law of data-driven research, and organised a 'round table' conference.

Part II describes the ethical and legal aspects of research with data after the subject has died, an important issue for SCA research. *Chapter 5* details the findings of a systematic review of empirical studies on the post-mortem use of health data for research. In *Chapter 6* the underlying philosophical question on the existence of a

moral right to posthumous privacy for research participants, is analysed using the theory of Information Ethics. *Chapter 7* is a practice-oriented article about the rules for publication of case reports that recognisably describe deceased patients.

Part III presents three separate cases where SCA data research is applied in life-saving cardiac care. *Chapter 8* is an opinion piece about the harm due to the non-use of data contained in AEDs, as systematic collection of the data would benefit research but could also directly save lives. The AED is also the subject of *Chapter 9,* where I discuss the use of computer modeling to determine the fair placement of these devices throughout areas. *Chapter 10* is an opinion article that introduces the PROFID project, which aims to use AI for personalized placement of ICDs, and the chapter provides a first assessment of the ethical acceptability of this aim.

Chapter 11 functions as an epilogue to the case study of ESCAPE-NET guiding this thesis, which concludes with the suggestion of embedding trust-based governance in health data research. Finally, *Chapter 12* is a general discussion of the findings in the foregoing chapters and provides suggestions for practice and further study.