The impact of EU secondary use of health data guidelines on multilateral research projects during the COVID-19 pandemic: A study of cross-border data sharing policies in EU collaborative research projects in context of a pandemic: an examination of policy and reality.

Arkesteijn, Marije

Leiden University

Master's Thesis

The impact of EU secondary use of health data guidelines on multilateral research projects during the COVID-19 pandemic

A study of cross-border data sharing policies in EU collaborative research projects in context of a pandemic: an examination of policy and reality.

Author:

Marije Arkesteijn

Student Number:

[redacted]

Supervisor:

Dr. Sarah Giest

Word count:

23759

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Public Administration: International and European Governance

Universiteit Leiden

Governance and Global Affairs

January 6, 2023

Abstract

The impact of EU secondary use of health data guidelines on multilateral research projects during the COVID-19 pandemic

By Marije Arkesteijn

During the COVID-19 pandemic, it was assumed that European cooperation, both individually and collectively, would produce better results than autonomous national self-interest. Especially the demand for increased cross-national cooperation to accelerate data exchange for multilateral COVID-19 research to inform public health policy-making was highly critical. However, sharing health data for secondary purposes such as research is difficult, as technical, political, and ethical issues were identified before the COVID-19 pandemic. This thesis focused on data management issues and barriers such as a lack of metadata standards and data interoperability. Facilitating cross-border secondary use of health data to inform public health decisions has been on the EU's agenda for some time, leading to the creation of the Joint Action Towards the European Health Data Space and the European Commission's recommendation on a European electronic health record exchange standard, among other things. The COVID-19 pandemic provided an excellent case study for determining whether these guidelines were adequate for guiding efficient data sharing in collaborative research. For instance, the EU made a significant investment in cooperative COVID-19 research projects with the goal of providing data to support public health policies. In this thesis, ReCoDID, ORCHESTRA, unCoVer, and SYNCHROS—four projects financed by the EU Horizon2020 program—are discussed in detail. The projects shed light on the challenges of sharing patient-level data from observational cohorts, particularly with regard to data management issues such as data interoperability. It was discovered that EU guidelines did enable the formation of research projects and that these projects were even aimed at improving data harmonisation and exchange in COVID-19 research. However, because there is still no EU-standardised agreement on the selection of data interoperability standards, this has become a difficult task. Specifically, none of the four projects examined was able to locate interoperability standards at the legal, policy, care process, information, application, or infrastructure levels.

Contents

List of Abbreviations

ALLEA: All European Academics

COVID-19: Coronavirus disease 2019

EASAC: European Academies Science Advisory

EBM: Evidence-based medicine

EC: European Commission

EIF: European Interoperability Framework

eEIF: eHealth European Interoperability Framework

EHDS: The European Health Data Space

EHR: Electronic health record

EU: European Union

FAIR: Findability, accessibility, interoperability, and reusability.

FEAM: Federation of European Academies of Medicine

GDPR: General Data Protection Regulation

HERA: Health Emergency Preparedness and Response Authority

ICT: Information Communication and Technology

IHR: International health regulations

MS: Member State

PHEIC: Public Health Emergency of International Concern

ReCoDID: Reconciliation of Cohort Data for Infectious Diseases

ReEIF: Refined eHealth European Interoperability Framework

SARS: Severe Acute Respiratory Syndrome

SYNCHROS: SYNergies for Cohorts in Health: integrating the Role of all Stakeholders

TEHDS: The Joint Action Towards the European Health Data Space

UN: United Nations

WHO: World Health Organisation

1. Introduction

During any health crisis, one is confronted with phrases like "you are only as healthy as your neighbour" or "a virus knows no borders." These words would emphasise the evident necessity for increased cross-border cooperation in the fight against public health emergencies like the recent Coronavirus disease 2019 (COVID-19) outbreak. In fact, the COVID-19 pandemic has shown the significance of international partnerships in addressing emerging public health challenges (The Lancet, 2022). Namely, countries rely on each other to strengthen the fight against the disease outbreak through collaborative research and information sharing, vaccine discovery and development, and travel policies (Jit, et al., 2021).

International collaboration on COVID-19 has been marked by a collision of scientific globalism, i.e., the pursuit of open and collaborative research to expand knowledge and address shared difficulties (Lee & Haupt, 2021) and scientific nationalism, i.e., the pursuit of local scientific achievement to advance national aims like economic competitiveness and foreign policy objectives (Sá & Sabzalieva, 2018). The COVID-19 pandemic prompted requests for increased cross-national cooperation to speed up data exchange from a "globalism" standpoint (Baker, et al., 2020; Chakraborty, et al., 2020; Gaisa, 2020). From the beginning, the international research community coordinated by exchanging laboratory and surveillance data, genome sequences, and details on clinical results (Liu, et al., 2020; Dagliati, Malovini, Tibollo, & Bellazzi, 2020). According to bibliometric research, more multilateral collaborations than the historical average were involved in the first months of 2020's publication of scientific articles (Lee & Haupt, 2021)

Despite several successful examples, there is still much more that might have been done in a number of crucial areas. For instance, insufficient international research collaboration has resulted in effort duplication. While there may be merit in trying different tactics, studies on the effectiveness of drugs like hydroxychloroquine have been done again under similar circumstances even after they were found to be unsuccessful (Pearson, 2021; KC, Ananthakrishnan, Painter, Butani, & Teerawattananon, 2021). Moreover, problems with the global sharing of experiences were identified by the Covid Health System Response Monitor developed by the European Observatory on Health Systems and Policies. That is, not enough has been learned from the experiences of other nations during COVID-19. When it came to measures for face protection, travel policies, social distancing, and other issues, decision-makers were frequently simply unaware of what was going on in neighbouring countries (European Observatory on Health Systems and Policies, 2021). Hence, the occurrence of the

current pandemic has once again highlighted the importance of global cooperation in research for both transmittable and non-transmissible diseases.

Specifically, the creation of novel digital health solutions has been sparked by the demand for timely, high-quality evidence to guide policy-makers' decision-making. There is widespread agreement on the importance of some programmes using COVID-19 patient data for research, timely sharing of research results, and accelerating research collaboration (for instance, Moorthy et al., 2020; COVID-19 Clinical Research Coalition, 2020). As an illustration, the examination of electronic health records (EHRs), which gather patient data from regular clinical care in real-time, has improved disease surveillance and generated data to support public-health choices (Satterfield et al., 2021; Sheikhtaheri et al., 2021). In addition to electronic health records, other secondary uses of health data are likely to be successful in informing public health decisions. Observational datasets, public epidemiological data, and clinical trial data are examples of this (Vukovic et al., 2022; Patil & Gopal, 2022; Iacob & Simonelli, 2020). As a result, calls for advanced cross-border data sharing on the ongoing COVID-19 pandemic have been made by researchers and international policymakers, who have also advocated to create platforms to make it easier to share patient-level data for those who have contracted the COVID-19 virus (Xu & Kraemer, 2020; Moorthy, et al., 2020).

Specifically, it has become more important than ever to acknowledge health data as a public good with procedures to promote fast data exchange and guiding strategies (Schwalbe, et al., 2020; Knottnerus, 2015; Hess & Ostrom, 2011). Despite these optimistic efforts, the COVID-19 pandemic, in which acquiring newly learned information became a life-or-death matter, made it more urgent than ever to overcome the barriers, holdups, and difficulties in health data exchange (LoTempio, et al., 2020). More specifically, it served as a catalyst for the academic and policy sectors to re-evaluate whether existing standard data governance methods are sufficient for utilising personal data for public health study and response in an environment that is now more uncertain and undergoing rapid change (ALLEA, EASAC, FEAM, 2021).

While cross-border data sharing for research is critical for global public health governance in general, the European Union (EU) is considered a reliable system for regional cooperation that demonstrates the conflict between short-term nationalistic incentives and long-term imperatives for collaboration to acquire public goods (Jit, et al., 2021). In the history of multilateral cooperation at the EU level, crises have resulted in both more and less European integration. For instance, despite the fact that the Schengen crisis in 2015 and the Euro crisis in 2010 had

many similarities, the outcomes were vastly different (Schimmelfennig, 2018). The Schengen crisis, in particular, has not resulted in greater EU integration, as it did during the Euro crisis.

The EU has the ability to act as a coordinated and supportive force during health emergency situations such as pandemics (Europa Nu, n.d.). The European Health Union was founded in October 2022 to formally implement this procedure (European Commission, n.d.). In doing so, the EU has put in place systems that enable closer coordination and the implementation of coordinated actions in the event of health emergencies (European Commission, n.d.). These initiatives merely supplement healthcare governance in EU countries: how each Member State (MS) manages its health policies is entirely up to them (Europa Nu, n.d.). Importantly, the majority of these key initiatives, such as the recently launched European Data Space, highlight how a data-driven and digitalised EU environment can aid in the process of these collaborations taking place (European Commission, 2022). Especially when it comes to making health data available in an open and secure cross-border flow across the EU to facilitate data access and use for citizens and public health stakeholders.

Remarkably, the investment in coordinated scientific research on the novel virus was one of the most visible EU responses to COVID-19. As of January 2020, the EC promoted new research projects under the Horizon2020 funding (European Commission, n.d.). The EU issued its first call in March 2020, with a budget of €48.5 million, followed by a second call in August 2020, with a budget of €129.5 million (European Commission, n.d.). These funds were designated for projects that would advance "knowledge for the clinical and public health response to the COVID-19 pandemic" as well as "innovative and rapid health-related approaches to respond to COVID-19 and deliver quick results for society for a higher level of health-system preparedness" (European Commission, n.d.). Accordingly, the EU's support and confidence in these funds appeared to be strong.

However, EU funding is not always enough to produce excellent results in the form of successful research projects that can inform public health decisions. While the European Health Union's key initiatives are mainly based on lessons learned from the recent COVID-19 pandemic, the EU has been working for some time to strengthen its multilateral networks in preparation for a data-enabled future, in particular the digital healthcare transformation (European Commission, n.d.). That is, long before the COVID-19 pandemic, the European Commission (EC) proposed policies and regulations to facilitate EU-wide implementation of consistent, interoperable, and collaborative health-data approaches to benefit both primary and secondary health data use (Boyd, et al., 2021). The latter, in particular, benefits the EU by

offering a uniform, reliable, and effective framework for the "use of health data for research, innovation, policy-making, and regulatory activities" (European Commission, 2022, para. 3).

These guidelines on the secondary use of data are just as important, if not more so, in health crises such as the COVID-19 pandemic, to aid previously discussed multilateral collaboration in research projects to inform public health decisions (Heymann, 2020; Kamradt-Scott, 2018). However, because they frequently consist of subdivisions of guidelines covering a wide range of digital public health governance-related topics, these regulations are not as clear as day (Boyd et al., 2021). Furthermore, these guidelines are updated on a regular basis, or new ones are created to supplement older regulations (Boyd et al., 2021). The majority of these guidelines are legal and technical in nature, such as the General Data Protection Regulation (European Council, 2018) or the EC Expert Group on FAIR Data Action Plan (European Commission, 2018). The technical guidelines for data management are especially important because, while the volume of data, the need for rapid development and transmission, and the importance of harmonious international research provide a completely unique environment for projects during the COVID-19 pandemic, problems with data access and sharing are not uncommon (Dron, et al., 2022).

In the past, cross-border collaborative research projects have suffered from data gaps and reduced research impact due to data gathering issues such as non-standardised data collection, variation in data terminologies, and long-term organisational, ethical, and political barriers to data sharing (van Panhuis, et al., 2014; Salas-Vega et al., 2015; Abboud, et al., 2021; ALLEA, EASAC, FEAM, 2021). This reveals a paradox, given that the EU has supported and facilitated secondary data uses through a number of regulations and recommendations. Given this contradiction, it is worth investigating whether these guidelines have resulted in similar issues in collaborative research projects during the COVID-19 pandemic.

To summarise, the COVID-19 pandemic has demonstrated the clear need for personal health data sharing across state lines to facilitate research for public health governance. Making this possible, particularly in multilateral systems such as the EU, could be beneficial in dealing with health emergency policy-making. Secondary data use in research projects, as well as the challenges that come with it, are not new in the EU; efforts to support and address both have been on the agenda for some time. The question is whether the guidelines were adequate when the pandemic began to support, first, the creation of research projects that the EC was eager to fund, and second, data management within these projects. All things considered, it creates an intriguing research topic to investigate further.

1.1. Research objective and questions

The purpose of this thesis is to examine EU efforts to improve cross-border health data sharing in the context of collaborative research during the COVID-19 pandemic. More specifically, this thesis aims to explain how the EU guidelines on secondary data use influenced multilateral COVID-19 research projects. It was decided to use a process-tracing theory testing method to guide this research objective. As a result of discussing the main concepts in the literature review, this thesis formulates and empirically tests its expectations. The first step is a thorough policy analysis to determine how accessible the EU policy environment of secondary data sharing is for collaborative research projects. Furthermore, regardless of whether these projects were simple or difficult to create, this thesis aims to identify potential challenges in this policy field, as well as whether these challenges hampered efficient data sharing within these projects. Document analyses, in which multiple official EU documents, project objectives, and evaluation reports are examined, are the primary method for locating evidence that could support the hypotheses. To correspond with the research objective described above, the thesis's main question is as follows:

How did the EU's guidelines on secondary health data use affect multilateral research collaborations during the COVID-19 pandemic?

The following sub-questions will be used to build the answer to the central question:

1) During the pandemic, which multilateral COVID-19 research projects were created to inform public health decisions?
2) Which cross-border data-sharing standards were present in these projects to enable efficient secondary data use in these multilateral projects?

These sub-questions are designed to aid in the research process. In other words, a within-case analysis is carried out within the context of the European Commission's research funding action to combat the COVID-19 pandemic, where the first sub-question seeks to provide the subunits of analysis, i.e., the identified projects, for detecting data sharing standards, which answers the second sub-question.

1.2. Academic and societal relevance

There appears to be a significant amount of academic research available on the topic of secondary use of health data and research for data-driven health governance in general. There have been many studies covering not only the benefits of cross-border data sharing in research to facilitate public health decisions, but also the potential ethical, political, and legal challenges

that may arise, particularly in a multilateral political system such as the EU (see, for example, Vukovic et al., 2022; Cuquet et al., 2017; Bruzzone & Debackere, 2021 and more). Thus, the theme is not a new phenomenon in this regard. However, because the COVID-19 pandemic is a new and recent occurrence, there haven't been many academic reviews that look at the secondary use of health data for data-driven health governance in this specific context. Therefore, this thesis has the potential to help close that gap in the academic literature.

Furthermore, given the importance of the secondary use of health data for an advanced EU digital health environment in general, but especially in health crises, the findings of this study could serve as a valuable message for all EU researchers, policymakers, and public health stakeholders. That is, the COVID-19 pandemic is viewed as a unique opportunity to assess whether the EU's previous efforts to facilitate cross-border data sharing for public health research have paid off. Since the COVID-19 pandemic was deemed a massive global issue, affecting citizens individually as well as a country's healthcare, economy, education, and so on (WHO, 2020), there has been a strong public interest in research related to potential answers to the best approaches in the fight against the disease outbreak. Therefore, any lessons or failures identified in this study could be applied to new strategies in the post-pandemic era or future pandemics.

1.3. Thesis outline

This thesis follows the following structure in order to answer the main research question and its sub-questions. The second chapter of this thesis begins with a review of the literature. This chapter conducts a thorough review of previously published academic research on the topics of multilateralism, data-driven policymaking, cross-border health data sharing, and secondary health data use. The expected causal mechanism is outlined at the end of this chapter, followed by the formation of three hypotheses. The third chapter presents the research design as an overview of the selected research methods, as well as their strengths and weaknesses. This chapter justifies the use of an explanatory deductive process-tracing approach, document analysis, and the operationalization of derived concepts from literature to measurable indicators for empirical testing.

The fourth chapter provides a detailed overview of the case, discussing the EU's efforts to improve cross-border health data sharing in the context of collaborative research during the COVID-19 pandemic. The fifth chapter contains both the results and analysis of the two sub-questions. The first sub-question addresses which multilateral COVID-19 research projects were established to inform public health decisions, potentially confirming the first part of one

of the three hypotheses. Then, four selected projects are examined to answer the second sub-question about which data-sharing standards are used, which could confirm the second part of one of the three hypotheses. In doing so, the analysis refers back to the theoretical arguments derived from the review of the literature. The sixth and final chapter is the conclusion, which provides an answer to the central research question. All research findings are concluded and summarised here, and potential avenues for future research are discussed.

2. Literature review

As stated in the introduction, the research question in this thesis is: *"How did the EU's guidelines on secondary data use affect multilateral research collaborations during the COVID-19 pandemic?"* The following chapter is intended to provide an overview of previous state-of-the-art research addressing the research question's main concepts. This data is gathered from a variety of sources, including books, papers, and scientific journal articles. A critical study of the most pertinent data that can be employed in the analysis to finally answer the research question is also included in the literature review.

The following concepts emerge during the examination of the research question: multilateralism, data-driven policymaking, cross-border health data sharing, and secondary health data use. This chapter begins with a political scholar's perspective on multilateralism in public health in the context of the EU. Following that, the literature review discusses the concept of data-driven policymaking and its significance in public health. Next, the concept of cross-border data sharing is described as necessary in order to enable data-driven policy-making in multilateral public health governance. The final concept, secondary use of health data, describes the type of data incorporated in these multilateral collaborations. The three proposed hypotheses are presented at the end of the literature review.

2.1. Multilateralism

Although the term "multilateralism" can refer to a variety of things, the focus of this thesis will be on multilateral relations, collaborations, and governance. Multilateral governance is defined as "the process of coordinating national policy in groups of three or more states," according to Robert Keohane (1990, p. 731). More specifically, multilateral governance refers to "international governance" or global governance of the "many" (Kahler, 1992). Furthermore, multilateralism was founded on the principle of "opposition [to] bilateral discriminatory arrangements that were believed to increase the leverage of the powerful over the weak and to

increase international conflict," according to Miles Kahler (1992, p. 681). Multilateral collaborations and agreements are examples of multilateral governance.

To narrow the scope of this study, it has been decided to concentrate solely on multilateral governance in the EU rather than multilateral governance at a global level. However, some multilateral collaborations may be initiated within the EU but involve non-member states as well. These collaborations or partnerships are not excluded from the study. Similarly, the term "European Integration" is used in this thesis to refer to all forms of collaboration between European nations, typically but not always including EU member states. More EU integration entails more shared political and legal institutions, legislation, and decision-making processes (Marks, Hooghe, & Blank, 1996). Therefore, the terms multilateral governance/collaboration, European cooperation, European governance, and European integration are used interchangeably in this thesis.

Specifically in times of crisis, multilateral governance is frequently chosen due to two specific advantages of multilateral cooperation: 1) the advantages of minimising the detrimental (or beneficial) spillover effects that the actions of states have, and 2) the advantages of providing global public goods (Derviş, 2020). Multilateral accords do, however, include a certain amount of pooling of authority, which could provoke resistance from some national governments (Jit, et al., 2021). However, the price of not participating in multilateral initiatives, such as losing potential entry into the global economy, can be expensive in a world increasingly ruled by multilateral agreements (Jit, et al., 2021).

Four causes of this challenge were identified by Frieden et al. (2012). These elements are related to the previously mentioned global public goods. First, establishing a common understanding of global public goods is difficult and requires collaboration (Frieden et al., 2012). Second, nations have domestic issues that limit their ability to deal with longer-term multilateral commitments to provide global public goods (Frieden et al., 2012). Third, governments have diverse goals. For instance, they may both want to avert a pandemic, but their security, economic, or industrial objectives may be very different (Frieden et al., 2012). Last, governments are doubtful of the capacity of the multilateral community to provide global public goods as a result of previous unsuccessful attempts (Frieden et al., 2012).

2.1.1. Multilateral Governance in public health

Multilateralism is a recurring theme in the field of public health. Specifically, the concept of international health collaboration is not a new phenomenon. International health rules have been

standardised since the first of fourteen International Sanitary Conferences, which took place centuries ago (Youde, 2012). However, the most notable beginning to a more unified global approach to public health came when there were a significant number of governance initiatives after the WHO's establishment in 1948 (McInnes, 2019). These initiatives were frequently a profusion of soft law declarations, best practice manuals, conventions, norms, and moral codes, among other things. Rather than being legally enforceable documents, most of the legislation and other governance mechanisms were either of a technical nature (e.g., vaccination distribution) or normative in nature (e.g., 1978 Declaration of Alma Ata on Primary Health Care) (McInnes, 2019). Global health crises, such as the Ebola outbreak in 2014, have highlighted the need for a more institutional and legal international approach to health governance (Fidler, 2015; Gostin & Friedman, 2015; Gill & Benatar, 2016). While global health governance as a policy field is important in general, this thesis focuses solely on the international politics of pandemics and health crises preparedness. That is, health emergencies, in particular, demonstrated the importance of formalising cross-border cooperation in order to avoid effort duplication and make decisions more quickly and efficiently (Kamradt-Scott, 2018). This was most visible in the field of medicine and treatment research, particularly in vaccine development (Kamradt-Scott, 2018). Not only to prevent medical errors but also to predict early outbreaks and track supply shortages, thereby informing public health policies (Ferguson, et al., 2006).

The governance frameworks surrounding disease outbreaks such as the influenza pandemic have a lengthy history. The global influenza monitoring network, which was started at the close of World War II, gradually widened its geographic scope while gathering biological data and information on possible precautions against a recurrence of the deadly 1918 Spanish flu pandemic (Hoyt, 2006). However, in the interim, the area of pandemic governance continues to be characterised by divergent viewpoints, objectives, and political motivations. Economic factors, security arguments, and biomedical knowledge that was later condensed and marketed as apolitical technical "proof" have all significantly influenced the institutional framework and stakeholders involved in reducing the likelihood of an influenza pandemic (Kamradt-Scott, 2018). Over the course of decades, these ambitions and opposing worldviews have changed and been pushed to adapt, punctuated by times of extreme awareness and severe disengagement.

In particular, two significant changes in the last 20 years have changed the politics of health emergency governance. First, a new narrative about global cooperation emerged, which led to a perception that global governance mechanisms needed to be given more attention (Fidler &

Drager, 2009; Lee, 2003). This implies that whenever such an event occurs, power should be distributed more globally and less locally (Cooper, Kirton, & Schreck, 2009; Fidler, 2003; Dodgson, Drager, & Lee, 2002). Second, the development of new organisations that are influencing public health governance and performance has reshaped the institutional architecture. That is, as new stakeholders and initiatives emerged, the environment for disease outbreaks and health emergency policies became more complex and unclear (Spicer & Walsh, 2011). This included an increase in the involvement of non-health actors in global health policies and governance, such as intergovernmental organisations like the World Bank and the UN Security Council (Youde, 2012).

Additionally, another trend in the once exclusively public sector of health governance is the rise of public-private partnerships (Williams & Rushton, 2011). According to Gill and Benatar, this is a negative development because it prioritises profit over need in the "market civilisation model of economic development" (Gill & Benatar, 2016, para. 1). Williams and Rushton (2011) present a more nuanced analysis that, while criticising some aspects, highlights other advantages, not the least of which is the increased funding for healthcare requirements. Furthermore, another shift in health emergency governance is the expansion of global health networks and their relevance. These transnational networks, according to some scholars, have likely played a significant role in the worldwide drop in morbidity and death that has occurred since 1990 (Shiffman, et al., 2016). Similarly, Jennifer Chan (2015) notes a trend of global interconnection in the promotion of HIV and AIDS awareness that might be repeated with other health crises. Specifically, the local and the global are closely related, and activists operating within states can leverage international agendas to further their goals, while global activities can have an impact locally (Chan, 2015).

Based on these two developments, McInnes (2019) identified a variety of political effects, including a greater focus on global health issues, particularly health crises; there has been a diffusion of power and authority; the development of a "market" for funding, which has implications for how the "power of the purse" functions; the nature of authority has become problematic due to rising expectations regarding the capacity of international institutions to intervene to prevent and address health crises; and competing norms have led to opposing agendas regarding the purpose of health (McInnes, 2019). In essence, the politics surrounding disease outbreaks such as the pandemic flu act as a microcosm of larger global health politics, complete with all the corresponding difficulties.

As authors Frenk and Moon (2012) argue, a sovereignty paradox emerges in multilateral public health governance. That is, today's world consists of sovereign nation-states and health is primarily a national concern. The factors affecting health, however, as well as the approaches to upholding this commitment, are becoming more worldwide (Jamison, Frenk, & Knaul, 1998). The health dangers brought on by globalisation cannot be controlled by one nation alone, nor can it produce an adequate solution to the majority of global concerns (Bettcher & Lee, 2002). Instead of states sacrificing their sovereignty, the solution to this paradox is for them to share it by organising an international united response through strong multilateral institutions (Frenk, Godal, Gómez-Dantés, & Store, 2022).

It is vital to create frameworks that will allow both states and international institutions to resolve the sovereignty paradox in order to benefit from advances within public health governance. More specifically, multilateralism has to produce a significant meta-innovation that will enable all innovations in health governance (Frenk, Godal, Gómez-Dantés, & Store, 2022). This necessity of changing the institutional framework of global governance has been heavily debated. The multilateral organisations' secretariats have received the majority of calls for change to date, and there are certainly many ways to enhance their effectiveness (Katz & Standley, 2019). Nevertheless, examining member states' actions, which frequently harm multilateral institutions inadvertently or knowingly, is also crucial (Kickbusch & Gleicher, 2012). Previous health crises, such as the Severe Acute Respiratory Syndrome (SARS) outbreak in 2002, highlight the urgent need to improve the international health regulations' ability to be effectively enforced through both incentives that encourage adherence to the rules and engagement and sanctions that deter non-compliance (Fidler, 2004). For sovereign states to embrace such enforcement authority, there must be a global public health convention or treaty that reshapes the principles and standards of global governance for health security (Duff et al., 2021).

A strong financial structure that shields the governing body of a global health convention from political sway and fluctuations is also necessary (Gostin & Mok, 2009; Fidler, 2010). Financial support must be reliable, sustainable, and sovereign. Instead of voluntary, unreliable, and designated gifts, such funding must mostly come from member states' regular, required contributions (Reddy, Mazhar, & Lencucha, 2018). In addition, multilateralism's meta-innovation cannot rely exclusively on regional and international organisations. These activities require networks of committed scientists and other organisations to support them (Chan, 2015).

Many developments and innovations can take place more easily with all of these factors contributing to a strengthened multilateral institutional public health environment. One of the innovations in multilateral public health governance that is discussed in the following chapter is the use of technological developments that can detect evidence or data to support policymaking within this environment.

2.2. Data-driven policymaking

Data-driven policymaking aims to profit from new data sources and to foster engagement with pertinent stakeholders and citizens (Janssen & Helbig, 2018; Ferroa, et al., 2013; Linders, 2012) by employing Information Communication and Technologies (ICTs) (Bertot & Choi, 2013; Janssen, et al., 2021). Data-driven policymaking expands on the idea of evidence-based policymaking. Three categories of evidence are seen as significant in the literature on evidence-based policymaking: “systematic ('scientific') research, programme management experience ('practise'), and political judgement” (Head, 2008, p. 1).

While acknowledging the value of these sorts of evidence, data-driven policymaking differs from evidence-based policymaking in that it places a greater emphasis on incorporating big and open data sources into decision-making as well as including individuals in the co-creation of policies. Moreover, data-driven policymaking strives to generate legitimacy in addition to producing more effective policies (Bijlsma, et al., 2011). Furthermore, due to citizens' growing mistrust of official data and statistics, the participation of citizens in the policy-making process is crucial (Davies, 2017).

This thesis focuses primarily on data-driven policymaking theory rather than the broader scope of evidence-based policymaking as it is more aligned with the concept and sub-concepts of ‘Cross-border sharing of health data’ and ‘Secondary use of health data’ discussed later in Chapter 2.3. and Chapter 2.4. However, the mention of evidence policymaking in this thesis can sometimes refer to a less extensive method of data-driven policymaking.

2.2.1. Data-driven policymaking in public health

The use of clinical data created by research to advance medical understanding and advise policy, particularly in global health politics, is not new (Pope, 2003). Furthermore, the desire for "evidence" to guide health crisis policymaking, such as pandemic influenza procedures, has increased since the introduction and standardisation of evidence-based medicine (EBM) practises in the mid-1990s (Kamradt-Scott, 2018). This emphasis on technical scientific data has been accompanied by the notion that such information is apolitical because it was obtained

objectively (or "scientifically"). EBM and evidence-based policy have now established a dominant position in pandemic influenza preparedness and planning, with policymakers routinely basing their explanations for particular policy decisions on the "evidence" (Kamradt-Scott, 2018). The significance now placed on having access to influenza vaccines and antiviral medications is one of the most notable instances of EBM methodologies informing and verifying particular policy responses (Kamradt-Scott, 2018). Despite the widespread use of EBM techniques and policymaking in public health governance, only a small number of government agencies have so far recognised that data-driven policymaking makes better use of sensor data and works with citizens to co-produce policies (van Veenstra & Kotterink, 2017).

Government policymaking in healthcare has evidently an impact on fostering population health. The data-driven methodology uses publicly available data to generate insights, disseminate knowledge, and evaluate the effects of policies and programmes on health (Karpati & Ellis, 2019). In order to inform feasible ways to solve a health problem, the data-driven technique creates information that indicates the scope and size of the issue. Governments use this data when communicating with stakeholders and the general public, and they create assessment plans to evaluate the execution and real effects of any enacted policy interventions (Karpati & Ellis, 2019). The model below is taken from the Palgrave Handbook of Global Health Data Methods for Policy and Practice (2019). It provides a comprehensive overview of data-driven policymaking in public health.

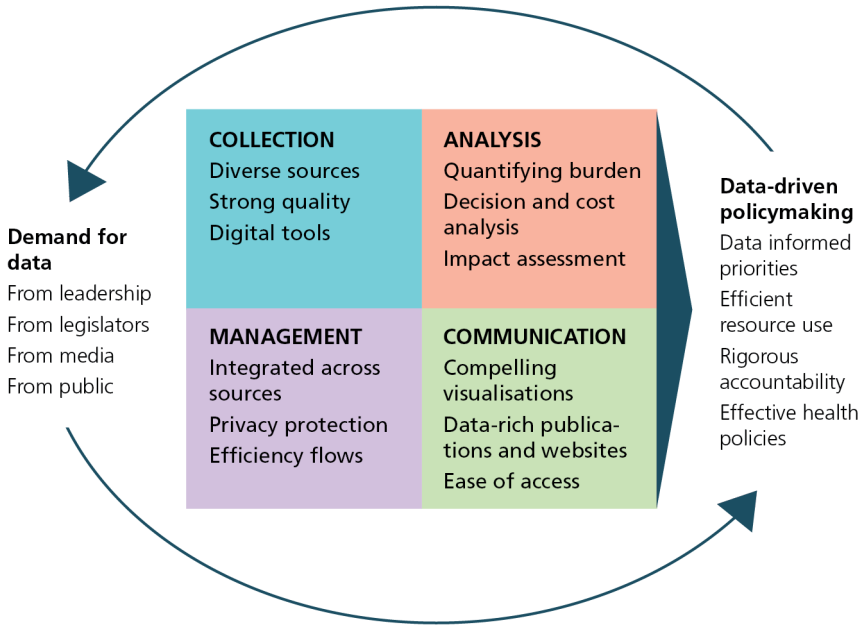


Figure 1. Framework for data-driven policymaking (Karpati & Ellis, 2019).

Data must be requested in order to be gathered, and in order for data to be requested, it must be made available. The diagram above portrays the process of data-driven policymaking, which begins and ends with a data-request and data-application architecture.

First, the health agency's systems for collecting data on disease surveillance, behavioural surveys, vital statistics, and other topics, as well as their operation, scope, cost, completeness, and timeliness, are referred to as *data collection.* Agencies draw information from a variety of internal and external sources, including "census, economic, and environmental data" (Karpati & Ellis, 2019, p. 88).

Furthermore, "the system of databases and other information technology solutions" required to store and incorporate the vast amount of data made accessible to public health organisations are known as *data management* (Karpati & Ellis, 2019, p. 88). Connecting and harmonising data from various sources, preserving privacy and discretion in personal data elements, and making data accessible to users inside and outside of government, including academics and citizens, are some examples.

Moreover, *data analysis* is the process of transforming raw data into relevant information via cleaning, coding, and analysis. Statistics and economic modelling are included in analyses, which might be qualitative or quantitative. Characteristics of health issues or risk behaviours, population health conditions, health system functionality, or programme efficiencies and impact are possible outputs of data analysis (Karpati & Ellis, 2019).

Lastly, data producers and users, such as "government decision-makers, lawmakers, the general public, the press, advocates and civil society representatives, and academics," are connected by *data communication* (Karpati & Ellis, 2019, p. 89). Sharing data with stakeholders has several benefits, including informing them of priorities, defending decisions about resource allocation and policy, and creating and communicating responsibility for achieving objectives and benchmarks (Karpati & Ellis, 2019).

Data sharing across-borders must be made available in an open, secure, and inclusive flow of data in order for data-driven policymaking in multilateral health governance to be possible.

2.3. Cross-border sharing of health data

Cross-border data sharing or cross-border data flows are defined as:

> The movement or transfer of information between servers across country borders. Data needs to be able to move freely so that no matter where you are, you have access to the

> information and services you need. Everyone from individuals to large corporations relies on transferring data. (BSA Software Alliance, 2017, p. 1)

Data transfer across-borders is essential for the services that support international trade, enhance health and safety, advance social good, and support emerging technology (BSA Software Alliance, 2017). Likewise, The World Economic Forum believes that it promotes collaboration and cross-pollination between individuals and businesses, as well as the exchange of knowledge and ideas in other sectors such as health and education (Baller, Dutta, & Lanvin, 2016).

When it comes to public health data, academics agree that health systems must drastically change their approach to using the data at their disposal if they are to effectively address the challenges of an ageing population, shifting illness patterns, complex healthcare needs, and limited resources (Oderkirk, Wenzl, & Slawomirsk, 2019). Sharing data, information, and knowledge across states has long been shown to be both feasible and advantageous for improving human health.

As the amount of public health data develops and becomes more interdependent, a diverse set of data is employed to inform research and guide international health governance. Schwalbe et al. (2020) classified data into four major archetypes: patient data, health systems data, routine public health data, and data from health research, thereby distinguishing data types and narrowing the scope of the concept of cross-border sharing of health data. The archetypes are built on the existing framework named: 'Healthcare data spectrum' (Feldman, Johnson, & Chawla, 2018). It should be emphasised that this classification describes the method by which these data are produced, not necessarily their usage or application.

The first archetype refers to "data generated at the patient level which include biomedical data (e.g., genomic data, proteomic data), electronic health records, and data generated by individuals themselves (e.g., data from wearables and social media)" (Schwalbe, et al., 2020, para. 6). Sharing patient data enables independent review of study findings to guarantee dependability, repeatability, and accountability. Sharing datasets can help academics and those creating health interventions find the right demographic data sources. By enhancing the evidence base on which clinical and regulatory choices are made, data sharing can significantly improve public health (Bull, Roberts, & Parker, 2015).

The archetype of health systems data includes "human resources data, service availability and utilisation data (e.g., number of hospital beds, insurance claims data), and performance metrics

(e.g., performance assessment results, patient satisfaction surveys)" (Schwalbe, et al., 2020, para. 7). This data is usually available in health management information systems datasets, even in resource-constrained environments. In order to accomplish shared public health objectives and outcomes, exchanging health systems data is likely to improve management and partnership between the public and private sectors. Insight into methods to increase the effectiveness and efficiency of health services as well as measurements of the effects of new health policies and treatments can be gained by sharing health systems data with the research community.

Public health data (routine and non-routine) is the third archetype. Routine sources stand for "health facility and community information systems" (Schwalbe, et al., 2020, para 8). On the other hand, "Household and other population-based surveys, censuses, civil registration (e.g., births and deaths) and vital statistics systems, disease surveillance systems, health facility surveys, and administrative data systems are examples of non-routine sources" (Schwalbe, et al., 2020, para 8). This information has previously been made public, either collectively or individually. Establishing global health standards, norms, and guidelines can be accomplished by exchanging both types of public health data. More precise estimates of morbidity and mortality, meaning estimates based on individual causes, can be produced using these data in particular.

The data mentioned in the above archetypes are mostly gathered for administrative needs or for managing and planning programmes. Likewise, observational and experimental study initiatives in public health provide a substantial amount of data. By producing information regarding the effectiveness and safety of treatments, the sharing of clinical trial data might hasten improvements in public health (Schwalbe, et al., 2020). There are countless examples of clinical trial data being utilised to undermine the efficacy of interventions or elevate patient care (Lo, 2015). The availability of such data encourages openness and replicability in scientific health studies.

Specifically in times of public health emergencies, cross-border data sharing seems essential (The Lancet, 2022; Kamradt-Scott, 2018). That is, new discoveries in research data are vital in international health management, especially at the beginning of an epidemic or disease outbreak (Heymann, 2020). This data consists of preliminary reports from the outbreak location and laboratories that assisted with the initial inquiry, as well as findings from other outbreaks involving similar species. Knowledge of the safety, effectiveness and manufacturing processes of prevention and treatment technologies benefits society at large (Heymann, 2020). Equally important, nevertheless, is what is unknown. As the outbreak spreads, it will be necessary to

gather more data to improve the risk assessment and guarantee that patients are treated as effectively as possible. More specifically, according to Heymann, this means the data of:

> Routes of transmission and transmissibility, the natural history of infection in humans, the populations at risk, the successful clinical practices that are being used to manage patients, the laboratory information needed to diagnose patients, and the genetic sequence information used to assess viral stability. (Heymann, 2020, para 3.)

Aside from these methods, which are important for data-driven policymaking in multilateral health governance in general and in health crises in particular, cross-border health data sharing can serve a variety of purposes or applications.

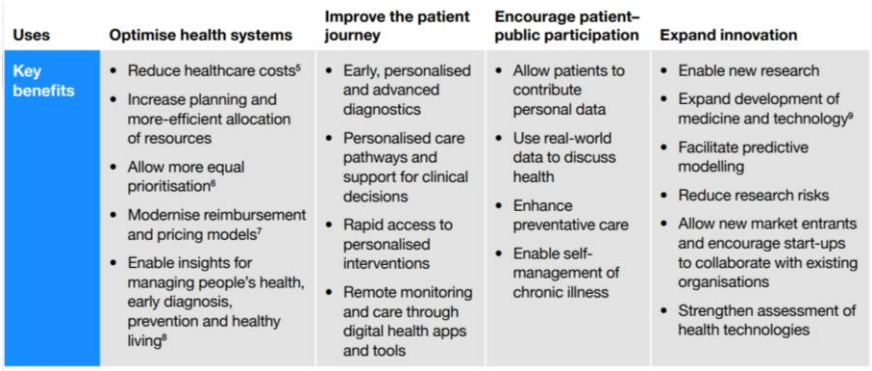
2.4. Secondary use of health data

In terms of health data usage or application, this thesis focuses on the secondary use of health data. The concept of secondary use implies the existence of primary use as well. In order to meet the needs of direct care, primary usage includes the collection and application of data in the context of individualised medical care in clinics and hospitals (Perera, Holbrook, Thabane, & Foster, 2011). On the other hand, according to researchers of the Open Data Institute (ODI), the secondary use of health data refers to:

> The use of aggregated health data from population-level sources, including electronic health records, wearable technologies, health-insurance claims data, health registry data (or burden of disease registries), clinical trials and other research, and drug consumption data to improve personal care planning, medicines development, safety monitoring, research and policymaking. (Boyd, Zimeta, Tennison, & Alassow, 2021).

The advancements quoted by Boyd et al. (2021) are made possible in large part by using primary health data for secondary purposes. Namely, secondary use of health data can boost the value of data already gathered from clinical settings, including data on healthcare occurrences and clinical trial outcomes, as well as data gathered from other sources, like records of illnesses and insurance claims, sensors, and wearable technology (Safran, et al., 2007). This information is typically referred to as "real-world data" (Boyd, Zimeta, Tennison, & Alassow, 2021). The data can then be aggregated, anonymized, and repurposed to enhance citizens' experiences, improve their health, develop more effective healthcare systems, and promote innovation (Boyd, Zimeta, Tennison, & Alassow, 2021).

Furthermore, other possibilities for secondary use of health data include ways to improve services, lessen health disparities through better resource allocation, or use it to improve individualised healthcare, such as by comparing treatments for patients with comparable features (Elkin, et al., 2010). By augmenting research data to evaluate whether novel treatments would be effective for a wider population, secondary usage of data can also support innovation (Boyd, Zimeta, Tennison, & Alassow, 2021). In other words, the entire healthcare system benefits from secondary uses of health data. A brief overview of the benefits of secondary usage of health data is provided by the Open Data Institute.



Uses	Optimise health systems	Improve the patient journey	Encourage patient–public participation	Expand innovation
Key benefits	• Reduce healthcare costs[6] • Increase planning and more-efficient allocation of resources • Allow more equal prioritisation[6] • Modernise reimbursement and pricing models[7] • Enable insights for managing people's health, early diagnosis, prevention and healthy living[8]	• Early, personalised and advanced diagnostics • Personalised care pathways and support for clinical decisions • Rapid access to personalised interventions • Remote monitoring and care through digital health apps and tools	• Allow patients to contribute personal data • Use real-world data to discuss health • Enhance preventative care • Enable self-management of chronic illness	• Enable new research • Expand development of medicine and technology[9] • Facilitate predictive modelling • Reduce research risks • Allow new market entrants and encourage start-ups to collaborate with existing organisations • Strengthen assessment of health technologies

Figure 2. Secondary use of health data and their benefits (Boyd, Zimeta, Tennison, & Alassow, 2021)

All European Academics (ALLEA), the European Academies Science Advisory Council (EASAC), and the Federation of European Academies of Medicine (FEAM) collaborated to launch a study on the 'International Sharing of Personal Health Data for Research' (2021). They base their paper on the idea that research data should be considered a public good, as other academics have stated (Knottnerus, 2015; Hess & Ostrom, 2011). This idea implies that pre-existing personal data should be accessible securely for predetermined research purposes. In addition, while it is generally agreed that research data should be made available to other researchers once it has been published for certain, well-defined goals, there are also many chances for data sharing throughout the actual research process (Richter, Meissner, Strangfeld, & Zink, 2016). Furthermore, the study confirms that the public sector's use of personal data sharing for health research has the potential to improve population health, the health of individual patients, and the social cohesiveness of European society as a whole (ALLEA, EASAC, FEAM, 2021).

Furthermore, by organising the scientific community in such a way that it is able to optimise the public interest of the rapidly growing body of data, developed data flow can improve the speed of diagnosis and promote competence and effectiveness in data examination, validation, and application (Martani, et al., 2019). In order to assure a suitably high study sample size,

reproduce findings, and reveal complex pathways, data on individual subjects are frequently pooled in medical research. For patients with uncommon conditions or rare diseases, sharing data gathered internationally may be particularly crucial (Courbier, Dimond, & Bros-Facer, 2019). Additionally, the research of ALLEA, EASAC and FEAM (2021) emphasises that if sizable amounts of data on personal health are shared, there are hitherto unheard-of possibilities for actual diagnostic and therapeutic decision-making. Furthermore, the report confirms that data sharing has become standard procedure in several academic fields, such as genetics and genomics, but it may have lagged behind in the field of public health until recently, which several academics have previously suggested (Walport & Brest, 2011; Carr & Littler, 2015; Salas-Vega, Haimann, & Mossia, 2015).

There are numerous opportunities to support public health decisions by research utilising secondary health data. One example is the secondary use of Electronic Health Records (EHRs), which is related to the first archetype "data at the patient level" identified by Schwalbe, et al. (2020). EHRs are personal files kept in an electronic system that is primarily intended to gather, save, and analyse patient data as well as make all patient data available securely (PAHO, n.d). Furthermore, an EHR system may provide tools to help guide clinical decisions or vital clinical data for patient care (PAHO, n.d.). EHRs can be used for secondary purposes in addition to their primary usage as data collection tools. More specifically, by increasing the effectiveness, quality, and affordability of clinical trials, EHRs have a significant potential to enhance clinical research (Fears, et al., 2014). In addition to clinical trials, health surveillance or comparative effectiveness quality measurement are other research-related activities that can be supported by optimised EHRs (Fears, et al., 2014). EHRs are especially important during a health emergency such as a pandemic because comprehensive and interoperable EHRs facilitate data sharing in both the national and international public health systems. In other words, these files could be used to keep track of and report on alleged and recorded cases, possible treatments, and unusual symptoms (ALLEA, EASAC, FEAM, 2021). Additionally, it enables a more rapid comprehension of how a pandemic develops within a certain demographic, allowing for earlier and more thorough containment and/or mitigation efforts (ALLEA, EASAC, FEAM, 2021).

While much academic literature has contributed to the benefits of secondary health data use, some academics have also discussed the challenges involved. For instance, although the secondary use of health data from EHRs to support research is fairly common, different sites participating in the same study may use different data collection tools (examples of these are clinical studies, observation cohorts, patient records, etc.). Significant variations in data

collection techniques can create sources of heterogeneity that can significantly skew integrated results (Granda & Blasczyk, 2016). In addition to a wide range of IT solutions for handling its data, there are many community-developed standard terminologies for research projects and for health data in general (Lenz, Beyer, & Kuhn, 2007). Similarly, heterogeneity in these terminologies can impede effective data sharing within research projects.

According to the Open Data Institute, data interoperability is one of the major issues confronting the EU health data ecosystem (Boyd, et al., 2021). Interoperability is defined as "the ability of different information technology systems and networks to communicate with one another; to exchange data accurately, effectively, and consistently; and to use that information" (PAHO, n.d.). Health data and health innovations in the EU must be in line with the levels of interoperability, be cross-border compatible, and come from a variety of different health-system divisions (Boyd et al., 2021). Additionally, this interoperability must permit nations to continue to make their own health-policy choices independently. For instance, while ensuring cross-border interoperability, local values, cultures, customs, and healthcare organisations are some variables that still must be respected (Boyd et al., 2021). Interoperability of health data is critical not only for tracking personal medical data for primary purposes, but also for developing standard interfaces, deciding on standard data sets, and establishing quality criteria for secondary purposes (Salas-Vega et al., 2015). Interoperability is especially difficult for academics and policymakers to coordinate and collaborate on because different countries have different interoperability regulations. Furthermore, disparities in clinical standards and languages impede the interoperability of EU data sets (Baeza-Yates, 2013). As a result, cross-border interoperability is regarded as essential in order to enable secondary data use.

In response to these challenges, guidelines in data management were created, such as the well-known FAIR data principles, which stand for Findable, Accessible, Interoperable, and Reusable (European Commission, 2018). These are the guidelines put forth by a group of researchers and organisations to facilitate the reuse of digital assets (European Commission, 2018). Although most data scientists are familiar with the FAIR data concepts, there has been little improvement in their application in the medical field (Mons, Schultes, Liu, & Jacobsen, 2020). Additionally, there is a glaring absence of reliable infrastructures for sharing data that are strong enough to speed up standardised protocols, track data management, and guarantee accountability (Salas-Vega, Haimann, & Mossia, 2015).

When it comes to database systems, for example, there is a distinction between centralised and federated infrastructures in the collection of health data. Each institution could move its data to

a centralised database in the centralised architecture, or to a federated database, which "define[s] the architecture and interconnect[s] databases to minimise central authority while supporting partial sharing and coordination among database systems" (McLeod & Heimbigner, 1985). Data for research purposes can be stored in centralised data repositories, but there are significant legal and moral issues that must be addressed (Turn, Shapiro, & Juncosa, 1976; Broekstra, Aris-Meijer, Maeckelberghe, Stolk, & Otten, 2020). Centralised data repositories in regulatory systems such as Europe would be made possible by a common language for granted approval that includes details about potential data uses and transnational data exchange (Duball, 2020). A centralised infrastructure, in particular, is required to create an open and harmonised data environment and better health research initiatives (Heijlen & Crompvoets, 2021). This may be possible for (observational) scientific research, but it is exceedingly difficult for EHRs from standard medical files (Chiasera, Toai, Bogoni, Armellin, & Jara, 2011). There are initiatives that use federated data analysis and machine learning in an effort to deal with the in-built regulatory complexity and privacy issues. While individual-level data is kept with its own settings, these techniques simply send aggregated data to the central node, where it is continually aggregated. However, in order to make the data discoverable but not automatically available, each institution should have a well-defined protocol (Vesteghem, et al., 2019).

Altogether, because a multilateral study with secondary uses of health data may have to deal with such a wide range of data sources and data collection systems, it appears necessary to harmonise common data and metadata standards, as well as standard methods to allow quick data ingestion and practise. To refer back to multilateralism in public health, a renewed multilateral framework that addresses these issues may sound promising, but when it comes to regulatory systems like the EU, a coordinated and centralised EU response to health data is not legally possible because EU Member States have chosen to retain their core public health sovereignty (Goossens, et al., 2022). As a result, in this thesis, the term secondary use of health data refers to the use of research data for public health purposes, specifically in health crises. It refers to data used for international health policymaking or observational studies. Secondary use of health also brings along its challenges.

2.5. Hypothesis

Certain expectations of the research outcome can be made based on theoretical and academic literature. As the research question — How did the EU's guidelines on secondary health data use affect multilateral research collaborations during the COVID-19 pandemic?— is divided into two sub-questions, so are the expectations in this study. To begin, this study anticipates

that the EU's guidelines on secondary health data use enabled (H1 or H3) or did not enable (H2) multilateral research collaborations during the COVID-19 pandemic. Second, the study anticipates that cross-border data sharing standards either enabled (H1) or prevented (H2 or H3) data sharing in these collaborations.

While multilateral cooperation during a global health crisis is not guaranteed, it is certainly desired. Data-driven or evidence-based policymaking has been at the heart of global governance during times of global health emergencies. In particular, cross-border data sharing of personal health data appears to be highly valued. As a result, personal health data is increasingly desired to be recognised as a public good. Health data as a public good is especially useful for secondary purposes such as policymaking, research, and clinical trials. Cross-border collaboration is required to ensure the free flow of research data to inform public health decisions. These cross-border research collaborations are only possible in the EU if the secondary use of health data policies facilitate and allow them. Both in the creation of these projects and in the data standards used within them.

The first possible answer to the central question is whether countries banded together to combat the COVID-19 outbreak and were willing to invest in cross-border collaborations, partly because the EU has demonstrated to be a strong multilateral system. If sufficient EU recommendations are made prior to the onset of the COVID-19 pandemic, it is expected that greater awareness of the importance of using secondary health data and its benefits in health crises, as well as complementary EU policies, will create an enabling environment that will allow for more cross-border research collaboration during the pandemic. Furthermore, EU guidelines aided data sharing standards to allow for efficient data sharing. As a result, the first hypothesis was developed:

H1: *During the COVID-19 pandemic, EU secondary health data use guidelines enabled greater multilateral research collaborations with adequate cross-border data-sharing standards.*

If it is discovered, on the other hand, that EU regulations and directives were only recently established, the question of whether these investments were made on time may arise. Or, because health policies are not prioritised at the EU level, countries choose a more nationalistic approach to health research collaborations, and EU guidelines were simply insufficient and legally binding enough, multilateral collaborations in COVID-19 research did not occur more frequently. Other well-defined data sharing challenges may also come into play, impeding

adequate cross-border data sharing. As a result, the alternative hypothesis presented below was developed:

H2: *During the COVID-19 pandemic, EU secondary health data use guidelines did not enable greater multilateral research collaborations with insufficient cross-border data-sharing standards.*

A final hypothesis is proposed that combines the first and second parts of H1 and H2. That is, another possible prediction is that the combination of EU investments and policies for a better digitalised European health system, as well as awareness of the importance of secondary health data sharing, sparked significant multilateral collaborations during the pandemic. The data sharing standards in these collaborations, however, were inadequate. As a result, the following hypothesis was developed:

H3: *During the COVID-19 pandemic, EU secondary health data use guidelines enabled greater multilateral research collaborations, but insufficient cross-border data-sharing standards resulted in ineffective data sharing within these partnerships*.

Because these three hypotheses are based on academic and theoretical literature, the next step is to put them to the test in a real-world scenario. However, before proceeding with the thesis analysis, a proper strategy for testing these research hypotheses in the most effective manner is required. As a result, the following chapter provides a solid research design that describes and justifies all methodological choices.

3. Research design

This chapter describes the methodological steps taken to investigate how the EU's efforts to improve secondary health data use influenced multilateral research collaborations during the COVID-19 pandemic. In order to find the answer to the main research question, an appropriate research design was chosen. That is to say, the following research design chapter describes the various methods and procedures employed in this thesis. This chapter is divided into five subchapters to provide a detailed overview of the sequentially chosen steps in this research.

The first subchapter discusses case selection and describes the research context that was chosen, i.e., the scope of analysis within secondary data use in COVID-19 research collaborations. The second subchapter, research methods, justifies the methodological approaches used in this thesis research. Following that, the third subchapter provides an overview of the chosen data collection and analysis method, i.e., qualitative document analysis. Thereafter, the concepts

discussed in the theory are operationalised to measurable indicators in the fourth subchapter. Before moving on to the case description, the fifth and final subchapter discusses the validity, reliability, and research limitations by outlining potential flaws of the research design.

3.1. Case selection

It would have been too broad for this thesis to investigate how all EU secondary health data guidelines affected research multilateral collaborations during COVID-19. A detailed conceptualization of both variables was required in order to test the hypothesis in a real-world case. As a result, the two themes are specified, and their selection is justified further below.

3.1.1. Cross-border interoperability

The term interoperability or interoperable data appears frequently in EU secondary health data use. According to the eHealth Network (2017), there are two definitions of interoperability to be identified: narrow and broad. The first focuses on an ICT system's ability to communicate with other ICT systems in order to make use of each other's capabilities or to give its human users composite capabilities. The latter indicates that ICT system interoperability is a means to a goal that enables "agencies, organisations, groups of users, municipalities, regions, or even nation states to interact more effectively and efficiently" (eHealth Network, 2017, p. 6). In this broader definition, interoperability's overarching goal is to enhance these organisational and healthcare exchanges in more detail. Therefore, broad interoperability addresses both the interoperability of healthcare-related organisations as well as the organisation of technology interoperability. This extensive definition served as the thesis' guiding principle.

Many definitions of the term occur within the data governance environment due to its importance and widespread use in secondary data use. When it comes to health data interoperability, there are typically six interoperability levels, which are summarised as follows:

1. On the **legal** level, compatible laws and regulations set restrictions on interoperability both within and beyond the borders of Member States (eHealth Network, 2017).
2. Formal agreements between organisations must be created at the **policy** level. That is, the collaboration's goal and value must be determined. Furthermore, trust and obligations between the organisations are codified at this level. Consequently, documents governing the organisation of cooperation serve as its foundation (eHealth Network, 2017).
3. Following the organisations' agreement to collaborate, certain **care processes** are examined and matched to produce joined care pathways and common workflows.

Furthermore, the workflow processes are tracked and managed at this level. Finally, the necessary information is specified by the shared workflow in order to provide integrated care (eHealth Network, 2017).

4. The **information** level characterises the operational description of the data model, the concepts and potential values (also known as data elements), and the connections between these data components and terminologies that specify their interoperability (eHealth Network, 2017).
5. Agreements are created at the **application** level regarding how the healthcare information systems will manage the import and export of medical data. This level of the technical specification describes how and with what communication principles information is conveyed. These standards must allow information systems to export and import. The incorporation and processing of shared information in user-friendly apps is another component at this level (eHealth Network, 2017).
6. The level of **IT infrastructure** is the last. The general networking and communication protocols and standards, as well as the storage, backup, and database engines, are offered here. To put it another way, it includes all of the common interoperability standards and protocols (eHealth Network, 2017).

As a result, any EU secondary data use guideline in the form of cross-border interoperability standards relating to any of these layers was given top priority for review.

3.1.1. Research data

Despite the fact that all sectors merit investigation, this thesis focuses on secondary data use guidelines within research data collaborations. The research collaborations are aimed at assisting EU policymakers as well as health ministers (broadly speaking, EU public health stakeholders). In other words, these research collaborations cross-borders and take place at the European level.

As explained in theory, allowing healthcare innovators to access patient-identifiable data such as EHR for secondary purposes, or those other than primary care, will advance knowledge and innovation. This can aid in a better understanding by policymakers of the effects of healthcare initiatives in practical contexts. As they can use information from many interventions to inform therapeutic decisions, it enables a broader analysis. Additionally, a more focused analysis since they may filter data to evaluate the effects of particular variables in the actual world, such as age or co-morbidity.

The needs of the above-mentioned stakeholders—policymakers and health ministers in the EU—should be laid out in order to construct a specific idea of the secondary use of health data that was employed throughout this thesis. The research findings could be used to detect disease outbreaks, monitor worrying trends, and develop early intervention alert systems, to name a few applications (ALLEA, EASAC, FEAM, 2021). Other requirements include "greater access to data to inform decision-making, interoperable data systems to share and link data, for example linking health data with environmental or industry data, data institutions and structures to support leadership and collaboration in the data economy" (Boyd et al., 2021, p. 11). Moreover, access to highly advanced health information can assist authorities and medical collaboration networks in managing public health by creating early warning systems, as has been observed during the COVID-19 pandemic (Boyd, et al., 2021). Cross-border research collaborations are frequently organised in projects or consortia to avoid effort duplication while also benefiting the EU multilateral system as a whole. Given the positive outcomes of the research, the EU must not only fund it but also provide appropriate guidelines. Therefore, this thesis focused on the research scope of multilateral projects or consortia for research data to inform public health decisions during COVID-19.

3.2. Research methods

The primary goal of this study was to explain how the EU's guidelines on secondary health data use affected multilateral research collaborations during the COVID-19 pandemic. This implies that the study began with explanatory goals in mind. An explanatory study aims to clarify the reasons behind why a particular phenomenon occurs or why a change in variable A results in a change in variable B (Yin, 1994). As a result, this thesis sought to identify a chain of causally important events that could explain the impact of EU guidelines on secondary health data use on multilateral research collaborations during the COVID-19 pandemic.

This study began with descriptive research to describe the general phenomenon of health data-sharing practices for secondary purposes and the importance of cross-border data flows in public health. These insights were then applied to a case study to provide explanations. This strategy is known as an explanatory case study and seeks to respond to "how" or "why" inquiries without the researcher having any control over the occurrence of events (Yin, 2014). Additionally, these case studies concentrate on phenomena in relation to current, actual-world circumstances (Yin, 2014). Thus, the effect of the EU's secondary health data use guidelines on multilateral research collaborations during COVID-19 seemed a suitable contemporary phenomenon that could not be influenced. Furthermore, this explanatory case study was created

with various design considerations, which resulted in the following structure. To begin, the design of the case study was based on the research approach of theory testing. The study aimed to explain the mechanism of the EU's secondary health data guidelines tied to a specific case and context, namely the multilateral research collaborations to inform public health decisions during COVID-19. Three hypotheses emerged to guide this theory-testing research process further. In other words, deduction was used to test these three hypotheses during the analysis.

In other words, the process tracing method was chosen to conduct this explanatory case study effectively. Process tracing is defined by Andrew Bennett and Jeffrey Checkel as "the use of evidence from within a case to make inferences about causal explanations of that case" (2015, p. 4). Finding this strongly connected chain of events is consistent with the thesis's stated purpose several times. Namely, finding a mechanism to explain how the EU's guidelines on secondary health data use affected multilateral research collaborations during the COVID-19 pandemic. Beach and Pedersen (2013) outline this process as follows: first, conceptualising the mechanisms, then operationalizing the mechanisms, and finally, gathering data from the case. This was accomplished by collecting data from previously completed scientific studies, as these studies had already identified potential causal explanation components and the ways in which they could be linked. Deduction was used to test these potential explanations in the form of hypotheses. Finally, because process tracing enables the use and combination of these components to explain specific cases and events, it was empirically tested in the context of the COVID-19 collaborative research.

3.2. Qualitative document analysis

During the case study, different units within the case were observed in this process tracing. These observations, like the majority of case studies (Gerring, 2007), were primarily qualitative. Process tracing requires extensive data collection, which can be accomplished using a variety of techniques (Toshkov, 2016). This thesis used quantitative rather than qualitative tools to detect a causal inference in this specific theme. This approach was chosen because Patton (1980) states that qualitative research is useful for studying the causes and effects of a problem. However, due to a lack of potential interviewees given the timeframe and scope of this thesis, conducting qualitative interviews was ruled out. As a result, a qualitative research tool known as document analysis has been chosen, in which documents are evaluated to provide meaning and reason around the research problem (Given, 2008).

A qualitative document analysis fits a typical process tracing research direction, which is to recover the institutional context as thoroughly as possible and retrace the sequence of events that led to the desired outcome (Toshkov, 2016). Furthermore, the following three goals are the foundation for the usage of document analysis: First, to provide the stakeholders' operating environment with some perspective. Second, it can be used to monitor progress and change. Thirdly, in order to support the data from other sources (Bowen, 2009). As a result, this method contributes to the stated goal of the research design, which is to gather empirical knowledge across multiple research collaborations in order to understand how EU guidelines on secondary health data sharing have influenced them.

A researcher can use a variety of methodologies when conducting a document analysis. The following steps from Altheide's (1996) 'Process of Document Analysis' were adapted for this thesis's qualitative document analysis process. This included (a) establishing document inclusion criteria, (b) collecting documents, and (c) articulating key areas of analysis. This method of qualitative document analysis offers a thorough and organised investigation of written document content. Furthermore, the method makes it easier to analyse written policies in a consistent and unbiased manner.

3.2.1. Document inclusion criteria

Setting document inclusion criteria comes first. Choosing which collaborations to include, the categories of papers to be examined, and the timing of document publication was all part of this process.

Criteria for collaborations

Certain criteria had to be met in order for documents from the collaborations to be selected for analysis. To begin, these collaborations are multilateral research partnerships to inform public health decisions during the COVID-19 pandemic. Furthermore, the partnership had to be multilateral; as known from theory, these partnerships must include three or more states. These partnerships should be public projects or consortia that were funded by the EU, but non-member states may also be included. So long as some of the countries involved were EU members. Finally, the project or consortium had to include personal health data sharing, such as data sharing from EHR.

Types of documents

In terms of the analysis, because primary sources are the most commonly used type of source in document analysis, the majority of the data came from them. Primary sources can be used to

get as close to the actual event as possible (Krause, 2020). Public records, in particular, are the official, continuous records of an organisation's operations. EU and organisational policies, plans, recommendations, and other papers of a similar nature were included in the document analysis for the first sub-question. In order to answer the second sub-question, it was necessary to analyse papers that would refer to collaborations "in action." These documents could be yearly reports or assessment, appraisal, or evaluation reports for a particular project, programme, or initiative.

Publication date

This thesis used the 'COVID-19 pandemic' as its timeline. It was critical to define this timeframe precisely in order to make it easier to include and exclude data. The time span in this thesis was chosen to correspond with the start and end of the COVID-19 pandemic. This is a broad term that depends on the context of the country to determine when it began and ended. Officially, the COVID-19 outbreak began at the end of 2019, when the virus was discovered in China in December (WHO, 2020). However, because this thesis focuses on EU guidelines and cross-border sharing from within the EU, the start of the COVID-19 pandemic in Europe, to which this thesis adheres, began in early 2020 (WHO, 2020). Concerning the end of the COVID-19 pandemic, it should be noted that the pandemic is still ongoing and no specific end date has been established. Because the purpose of this thesis was not to forecast, the end of the pandemic in this research timeline cannot be later than when this thesis was written. That is, the pandemic has not officially ended, but the scope of this research has. As a result, the research timeline in this thesis is January 2020 to December 2022. In other words, the documents gathered and analysed could not be published outside of this timeframe.

3.2.2. Collecting documents

Because the documents were all in the public domain, the data for this thesis was gathered through desk research. There was no need for actual field research to investigate the phenomenon of the EU's secondary health data use guidelines or the occurrence of multilateral research collaborations during COVID-19. All of the information needed to answer two sub-questions was available online. Between September 2022 and January 2023, desk research was conducted to collect cases of various multilateral research partnerships formed during the pandemic. These cases became the primary focus of analysis to determine whether cross-border data-sharing standards were sufficient to enable these multilateral collaborations.

This inclusion resulted in the following document overview. Part one:

Organisation/source	Number of documents from that source
European Commission reports	2
Cordis EU research results Horizon2020 ReCoDID project objectives	1
Cordis EU research results Horizon2020 I-MOVE-COVID-19 project objectives	1
Cordis EU research results Horizon2020 unCover project objectives	1
Cordis EU research results Horizon2020 RECOVER project objectives	1
Cordis EU research results Horizon2020 HERoS project objectives	1
Cordis EU research results Horizon2020 SYNCHROS project objectives	1
Cordis EU research results Horizon2020 ORCHESTRA project objectives	1
Cordis EU research results Horizon2020 EpiPose project objectives	1
Cordis EU research results Horizon2020 CORESMA project objectives	1
Cordis EU research results	1

Horizon2020 **EXSCALATE4CoV project objectives**	

Part two:

Organisation/source	Number of documents
unCoVer: **Objectives** **Protocol summary** **Process** **Legal and ethical guidelines** **Dissemination and exploitation plan**	5
ReCoDID: **Objectives** **Statistical guidance** **Evaluation 2x** **Data mapping**	5
ORCHESTRA: **Objectives** **Evaluation** **Privacy policy** **Work progress 2x** **Statistics**	6
SYNCHROS: **Objectives** **Strategy brief 2x**	3

While conducting document analysis, a numerical code is provided in addition to an APA in-text reference. Appendix A contains a comprehensive list of all the documents used to support the veracity of this thesis.

3.2.3. Articulating key areas of analysis

In order to analyse the documents consistently, some key themes had to be identified. It was desired to find the role of the independent variable (EU efforts for improved secondary use of health data) in these projects. Six interoperability layers were determined to be important and relevant for ensuring more efficient data sharing for research and policymaking purposes. These interoperability standards are divided into Legal and Regulatory, Policy, Care Process, Information, Applications, and IT Infrastructure layers. Not only was it critical to find the mention of one or more of these six interoperability levels, but it was also necessary to understand the exact meaning of these layers and the stakeholders involved in order to detect the interpretative meaning behind some text, as the same things could be implied using different types of terms.

The value of the qualitative document analysis is centred on this kind of qualitative analysis of content, meaning, and significance in context, which notably sets the methodology apart from a search for keywords. One of the interoperability levels, for example, is 'information,' which includes the operational description of the data model (e.g. data elements). Rather than searching for references to data elements in general, the goal was to determine whether the project or consortia intended for indicators that facilitate the information element interoperability within the partnerships. These are context-specific, but they could also include aspects like appropriate terminologies, formatting, and language usage.

3.3. Concepts and operationalisation

This thesis allowed for a holistic within-case analysis, with an emphasis on a general aspect (EU guidelines on secondary data use) of multiple cases (multilateral COVID-19 research collaborations). While an embedded case study divides the case into many units of analysis, a holistic case study only uses one unit of analysis for each case (DePoy & Gitlin, 2016). As a result, various multilateral partnerships formed in health data sharing for research to inform public health decisions during the pandemic are viewed as units of analysis within the overall core case, namely the EU's efforts to improve the secondary use of health data. In order to effectively answer the sub-questions, thus the research question, these variables were operationalized as indicators. The following table provides a clear overview of this operationalisation.

Variables/ Concepts	Definition	Indicators	Data Sources
Independent Variable			
EU secondary health data use guidelines	Efforts at the European level to improve data sharing standards supporting cross-border data access and sharing	Standards that align with the refined Framework for European Interoperability in Health. The framework includes six levels of interoperability: Legal and Regulatory, Policy, Care Process, Information, Applications, and IT Infrastructure.	Yearly reports or assessment, appraisal, or evaluation reports for a project, programme, or initiative
Dependent Variable			
Multilateral research collaborations during the Covid19 pandemic	Three or more member states involved in public health initiatives to share health data to facilitate research for policymaking.	Three or more partners are involved in public projects or consortia that were funded by the EU. Aimed at secondary data use for COVID-19 research supporting policymaking.	Organisational policies, plans, recommendations, and other papers of a similar nature

These two variables are linked to the thesis's sub-questions. The first sub-question is linked to the dependent variable and seeks to identify these projects, whereas the second sub-question is linked to the independent variable and seeks to detect interoperability standards in these projects. The hypotheses were confirmed or denied based on the answers to these sub-questions, thereby answering the research's main question.

3.4. Research strengths and limitations

Each researcher creates his or her research design in the manner that he or she believes best suits the research problem. To achieve legitimate and fair research, all researchers seek validity and reliability, two critical factors in any evaluation. Although full validity and reliability are not possible, they are ideals that are aimed to achieve (Neuman, 2014). However, there may be

limitations to the chosen methods that the researcher cannot control while conducting this study. All influences and aspects such as flaws or conditions that place "restrictions on the methodology and conclusions" are considered in these limitations (Dudovskiy, 2016).

3.4.1. Validity

Truthfulness is synonymous with validity. In qualitative research, there is more interest in achieving authenticity than achieving a single version of "Truth" (Neuman, 2014). Internal and external validity are the two components of research validity. Whereas internal validity refers to the absence of errors in the design of a proposed study that could lead to incorrect conclusions, external validity refers to the application of research findings in other contexts (Neuman, 2014). While both are important, it appears that maintaining high internal and external validity in the same research approach is difficult. That is, the greater the ability to control outside influences in a study (internal validity), the less generalizable the results (external validity) (Toshkov, 2016).

In general, case studies are not the best design for estimating an average causal effect. That is, if the theory is plausible or only provides a weak hypothesis, theory testing cannot be applied to a case (Toshkov, 2016). As a result, one can debate whether the chosen research design, specifically the explanatory goals and theory-testing approach, is strong enough for this specific research case. However, because the theory in chapter two already provides a strong connection between the two variables, the hypotheses formed afterwards provide a solid foundation for revealing the expected causal mechanism. In other words, the method used in this study ensured high internal validity. The theory-testing process tracing method enabled in-depth knowledge of the theory as well as details of the case's context.

On the other hand, the external validity of this study is not particularly high. That is, this study is linked to a specific environment, the COVID-19 pandemic. As described in the theory chapter, health emergencies such as pandemics necessitate cross-border data sharing to aid secondary data use, making the COVID-19 pandemic a unique paradigm for testing whether these variables demonstrate some causal link between the two. It cannot, however, be guaranteed that the same outcomes will occur in other research contexts, such as the next pandemic or disease outbreak. Nonetheless, since generalising causal effect was not the goal of this research, the lack of external validity does not diminish its scientific and social relevance because the EU health data interoperability guidelines and multilateral collaborations during COVID-19 are interesting and current research areas in and of themselves.

3.4.2. Reliability

Dependability and consistency are synonyms for reliability. In qualitative studies, a variety of techniques are used to consistently gather observations (Neuman, 2014). The goal is to be consistent in making observations, which is similar to the concept of stability and reliability. In other words, it is crucial that the outcomes are consistent if the entire study is measured again using the same methodology.

This research has relatively high reliability. The within-case analysis of the broader context of the EU efforts for improved data interoperability allowed for multiple units to be analysed, contributing to constant data comparison within the same case. Specifically, the same data collection tools were used to analyse different documents from several projects and consortia. Furthermore, the process tracing approach through document analysis contributed to the formal organisation of the data and established its authenticity.

However, time constraints limited the scope of the research. Because this study was written in a relatively short period of time, some data collection methods were automatically excluded from the research design. Conducting an expert interview, for example, to triangulate the research findings would have added reliability to the entire study. On the other hand, justifying why a specific interviewee was chosen for this research theme would have been difficult. That is, the two sub-questions sought to reveal what types of partnerships were formed in terms of COVID-19 research, as well as what role data interoperability played, the majority of which information is publicly available. As a result, the document analysis tool in process tracing became the most reliable approach, which is not necessarily considered a major limitation in this study.

3.4.3. Other limitations

Another limitation of this research is a lack of understanding of some of the side themes in this thesis. Because this thesis was written as part of the requirements for the Master of Science in Public Administration degree, the emphasis in this thesis was on the data governance and data-driven policy aspects of the overall theme. However, the connection with the health sector is another significant aspect of this thesis. As a result, some academic papers or documents analysed contained health-related jargon that was outside of the researcher's expertise. Similarly, the topic of secondary health data sharing is fraught with legal ramifications. While this is an important aspect, it would necessitate extensive knowledge of EU law. It happened that the studied texts and documents that were primarily concerned with legal or health issues had to be

excluded from the study. Nonetheless, both variables provided a comprehensive list of existing knowledge; the researcher simply had to be more selective.

This chapter described all of the steps chosen for the research design and justified the choices made. In short, existing knowledge from the theory chapter was used to form three hypotheses, which were then applied to an explanatory case study and deductively tested. This case study was characterised as a within-case study (EU's secondary health data use guidelines) with multiple units of analysis (multilateral research partnerships formed during COVID-19). The theory-testing process tracing method was carried out through document analysis that consisted of desk research of primary sources. A qualitative approach in document analysis was used to find evidence to confirm or deny the hypotheses. Finally, the research's potential limitations were examined.

4. Case description

Before the COVID-19 pandemic, infectious disease outbreaks caused a number of public health catastrophes that are of global importance since 2005 (Jit, et al., 2021). The SARS pandemic of 2002–2003 served as a learning opportunity for the WHO, which updated its international health regulations (IHR) in 2005 (Jit, et al., 2021). From this point forward, the WHO may declare a Public Health Emergency of International Concern (PHEIC) (Jit, et al., 2021). These health emergencies are defined as "an extraordinary event which is determined to constitute a public health risk to other States through the international spread of disease and to potentially require a coordinated international response" (WHO, 2019). On 30 January 2020, the WHO declared a PHEIC, and on 11 March 2020, it classified the COVID-19 outbreak as a pandemic (WHO, 2020). Since this announcement, several initiatives have been launched to help the necessary coordinated international response, including cooperative research and information sharing, the development of vaccines, and travel regulations (Jit, et al., 2021).

As a regional institutional system, effective multilateral cooperation appears to be essential for the EU to manage the pandemic's shocks (Guimón & Narula, 2020). Certain guiding principles, such as relationships built on trust and acceptance of shared responsibility, are necessary for multilateral projects. It is up to the EU to support these specific principles, especially through guidelines, initiatives and further responsive, agile coordinating methods. Analysing the policy environment addressing EU health governance in emergencies, according to the terms of the 1997 Amsterdam Treaty, the EU is required to consider human health when formulating and carrying out its policies (Treaty of Amsterdam, 1997). However, how each Member State

administers healthcare inside their borders is up to them (Europa Nu, n.d.). The EU policies are merely an addition to the member states. Nevertheless, the EU has the ability to act as a coordinated and supportive force during emergency situations such as pandemics (Europa Nu, n.d.). The European Health Union was founded by the EU in October 2022 to formally implement this procedure (European Commission, n.d.). In doing so, the EU has put in place systems that enable closer coordination and the implementation of coordinated actions in the event of health emergencies (European Commission, n.d.).

The Serious Cross-Border Threats to Health Regulation (European Commission, 2022), the Commission's new European Health Emergency Preparedness and Response Authority (HERA) (European Commission, 2021), the recently launched European Health Data Space (EHDS) (European Commission, 2022), the pharmaceutical strategy (European Commission, 2020), and Europe's Beating Cancer Plan (European Commission, n.d.) are among these initiatives. The EC's (n.d.) initiatives aim for the following:

> .. building a strong European Health Union in which all EU countries prepare and respond to health crises together, medical supplies are available, affordable, and innovative, and countries collaborate to improve disease prevention, treatment, and aftercare, such as cancer. The European Health Union will better protect our citizens' health, equip the EU and its Member States to better prevent and respond to future pandemics, and strengthen the resilience of Europe's health systems. (para 2.)

This European Health Union specifically highlights the requirement for prompt and well-developed evidence to guide policymakers during health crises, with support of the development of novel digital health technology. The collection, access, and examination of significant amounts of data are specifically necessary for the timely transmission of important information. For instance, the examination of EHRs has improved disease surveillance and generated data to guide public-health decisions. Since pandemics present an opportunity to rethink governance systems and draw lessons from the past (Gersovitz, 2014), it appears that the majority of the European Health Union’s initiatives are based on the lessons acquired from the COVID-19 pandemic. While this is a positive step, multilateral collaborations were just as important during the pandemic, and one might wonder what EU guidelines were available at the time to guide this.

To give an example, the COVID-19 pandemic has encouraged numerous cross-national cooperation in empirical research. In fact, when it came to the COVID-19 outbreak, the EU was

determined to invest in coordinated scientific research to support public health decisions. The EC promoted new transnational research projects under Horizon2020 funding as of January 2020. (European Commission, n.d.). In March 2020, the commission set aside €48.5 million for research projects aimed at "advancing knowledge for the clinical and public health response to the COVID-19 pandemic" (European Commission, n.d.). The European Commission issued a second call for new research projects to receive Horizon2020 funding in August 2020. The EU sought "innovative and rapid health-related approaches to respond to COVID-19 and deliver quick results for society for a higher level of health-system preparedness" this time (European Commission, n.d.). The EU has set aside €129.5 million in funding for these multilateral projects.

Although it sounds promising, there may have been uneven results during the COVID-19 pandemic when using observational datasets, public epidemiology data, and clinical trial data to drive public-health strategies (Dron, et al., 2022). That is, there are identified data gaps that could have resulted from data capture issues such as non-standardised data gathering, inconsistency in data terms, and persistent structural, ethical, and political constraints to data sharing, reducing the potential scientific impact of these data (Dron, et al., 2022).

These projects can only take place in a secure environment and balanced participation of countries, which is ideally provided by the EU. As a result, because these projects have the potential to shed light on the complexities of forming multilateral collaborations while sharing patient-level data from observational cohorts, the EU secondary health data guidelines that underpin them must be outlined. Before doing so, it is important to understand the potential challenges that may arise within collaborative projects, as some of the guidelines may have been based on the recognition of these secondary use of health data barriers in the EU.

4.1. Challenges of secondary use of health data in the EU

Many studies have been conducted recently on the barriers and challenges to the use of secondary health data, both internationally and in Europe. For instance, Van Panhuis et al. (2014) provided a thorough description of the obstacles to sharing health data. These impediments range from technological to motivational, from economic to political, and from legal to ethical (van Panhuis, et al., 2014). Salas-Vega et al. (2015) subdivided these barriers further into key challenges such as confidentiality and data security, information access, data reliability, interoperability, and management and governance (Salas-Vega et al., 2015).

Furthermore, the Joint Action Towards the European Health Data Space (TEHDS) (2022) determine the obstacles to and facilitators of cross-border sharing of health data for secondary use, including both personal and non-personal health data. The outcomes of the literature review were then categorised, and the following four distinct themes were formed by identifying essential patterns:

- **"Data**: including data management, data quality, data interoperability, data monitoring and analysis" (TEHDAS, 2022)
- **"Infrastructure**: including the governance structure of the health data system and access to data" (TEHDAS, 2022)
- **"Legal**: including semantics, legal frameworks, and national interpretations of General Data Protection Regulation (GDPR)" (TEHDAS, 2022)
- **"Trust and transparency**: including political, social and organisational factors and citizens' engagement" (TEHDAS, 2022)

Because this thesis focuses on EU public health stakeholders, i.e. policymakers and researchers, it is critical to examine the barriers or challenges from their perspective. Considering their requirements once again, different studies find that access to data for research purposes is restricted by largely legal and data management concerns brought on by inconsistent interpretations and implementation (Iacob & Simonelli, 2020; Vukovic et al., 2022; Boyd et al., 2021; Salas-Vega et al., 2015; TEHDAS, 2022; Schwalbe et al., 2020 and more)

The legal issues that arise from the collection, analysis, and use of data include questions about consent, privacy, copyright and more (van Panhuis, et al., 2014). Different interpretations of the GDPR, particularly in the EU, and ignorance about how to enable health data for research purposes while adhering to Europe's strict data privacy laws have been the most pressing issues for public health stakeholders (ALLEA et al., 2021; Boyd et al., 2021). Discussing the legal barriers to the secondary use of health data is an important, current, and interesting topic, as it would provide the dilemma of balancing the privacy of European citizens with cross-border digital healthcare innovation. However, this topic would have a too legal perspective, which is neither desirable nor necessary in a thesis for a master's programme in public administration.

As a result, it has been decided to concentrate on the problems associated with data management, particularly those brought on by "incomplete or lost data, restrictive or conflicting data formats, a lack of metadata and standards, a lack of dataset interoperability (e.g., structure or language), and a lack of suitable analytical approaches" (Schwalbe et al., 2020, para 10). As explained in

the research design chapter, this thesis will pay particular attention to the cross-border interoperability in the secondary use of health data. Now that the main challenge in this thesis analysis has been identified, the sections that follow summarise the EU guidelines for secure and advanced secondary data use.

4.2. Data-driven policymaking in the EU

When it comes to data-driven policymaking in the EU, it appears that there is an increasing number of proposed or even implemented strategies for a data-enabled future in numerous sectors. The EU has a name for this approach: the European data strategy, which aims to establish the EU as a pioneer in a data-driven world (European Commission, n.d.). Data will be able to freely move around the EU and between industries with the creation of a single market, which will be advantageous to businesses, researchers, and public authorities (European Commission, n.d.).

A variety of policies have been established by the EU that acknowledge the value of data in fostering an open, innovative society and economy. This set of regulations, policy supports, investments, and strategic directions aims to establish the proper laws and legal instruments that permit data sharing for bettering educational, economic and health outcomes for all EU members and its citizens. The scope of the policies and rules governing the harmonisation of European data is too great to cover in this thesis, but the most significant examples show how crucial this task is to governments and the management of the health system. These examples include the GDPR (European Council, 2018), the aforementioned European Strategy for Data (European Commission, n.d.) with the Data Governance Act (European Commission, n.d.), the White Paper on Artificial Intelligence (European Commission, 2020), the Final Report and Action Plan from the Commission Expert Group on FAIR Data (European Commission, 2018), and Europe's Digital Decade (European Commission, n.d.).

4.3. Digital transformation of healthcare

In more specific regard to the digital transformation of healthcare, the EU's institutions, regulations, and programmes aim to modernise many facets of the industry, boost cross-border and intra-member state interoperability, and support collaborative networks that produce renewed approaches to healthcare problems. The EU's key policies include the Communication on Enabling Digital Transformation of Health and Care (European Commission, 2018), the 1+ Million Genomes Initiative (European Commission, 2018), European Reference Networks (European Commission, 2017), Electronic Exchange of Social Security Information (European

Commission, n.d.), eHealth Network (European Commission, n.d.), and One Health Action Plan Against AMR (European Commission, 2017).

Cross-border health data flows and multilateral networks are inextricably linked, as international cooperation on the exchange of health data advances research, and innovation, and improves health and healthcare on a global level (Oderkirk, Wenzl, & Slawomirsk, 2019). That is, the need of pooling data and share information across nations has increased due to evolving disease patterns, scientific advancements, and an understanding of the complexity of the disease. The necessary technical framework has been produced via digital technology (Oderkirk, Wenzl, & Slawomirsk, 2019). The EU policies mentioned above aim to facilitate these multilateral networks carrying out projects or consortia sharing research in order to facilitate public decisions (European Commission, 2022).

4.4. Relevant EU-wide policy for secondary use of health data

Although the secondary use of data is more and more recognised as an opportunity in the fields of data strategy, the digital economy, and health in the EU, it has only recently been put into practice. As previously shown, the EU has attempted to advance both a European vision for a data-enabled future (chapter 4.2.) and a digital transformation of healthcare (chapter 4.3.) with the aid of policies and regulations. However, the two types of policies frequently function apart from one another (Boyd, et al., 2021). Strategies often make reference to other sectors' policies in that category that are related to Europe's vision for a data-enabled future. Moreover, the programmes and policies that fall under the concept of digital transformation of healthcare work more independently. This is evident in the complexity, lack of standard data models, and fragmentation of the various experimental and research programs being conducted as part of Europe's healthcare digital transformation initiatives (Boyd, et al., 2021).

These two currently disjointed policy fields might be connected by the third developing group of policy activity of the European Health Union (European Commission, 2022) that was referred to at the beginning of the case description. The projects give a chance for a renewed cooperative strategy that makes use of the secondary use of health data. In other words, establishing the capacity for secondary purposes can be the justification for policy achievement because it is a common factor to fulfil the broader policy goals mentioned in the papers (European Commission, 2022). It seems more promising that these more recent initiatives, in contrast to past healthcare modernization programmes, can locate other, related, ongoing work that encourages the rapid adoption of data-driven and technologically developed healthcare systems.

It seems that the European Commission has a great opportunity to build a well-functioning environment that permits the secondary use of health data. This calls for the implementation of standardised, cooperative, and interoperable health data initiatives by supporting member states. The European Strategy for Data by the Commission intends to establish a consciously ethical approach, with a strong focus on enhancing data access and facilitating data sharing for social good (European Commission, n.d.). Policies acknowledge the trend toward managing "big data" as a source of inspiration for innovation. In particular, there is confidence in the EHDS's (European Commission, 2022) ability to eliminate the fragmentation that currently exists and build an innovative open health data infrastructure for Europe. The goal of the EHDS is to advance secure patient data interchange, even when patients cross-borders, and citizen ownership over their personal health information (European Commission, 2022). Additionally, it aims to promote research on medicines, treatments, and technologies and the use of health data for research, legislation, and regulation while maintaining data privacy laws and a reliable governance system (European Commission, 2022). Finally, it intends to define the accountability and safety of artificial intelligence in healthcare and assist other digital health services (European Commission, 2022).

The first stage in building this network is the TEHDS. The ideal multi-stakeholder approach for this network will promote participation from all sectors, including pharmaceutical firms, public health authorities, medical professionals, operators of health surveillance, patient advocacy organisations, and researchers (TEHDAS, 2022). This network could promote the usage of open standards and shared data formats, as well as identify important uses and work priorities (TEHDAS, 2022). Additionally, it might consent to creating common systems and common datasets that could serve as the foundation for industry development (TEHDAS, 2022).

TEHDS describes its goal as follows: "The overarching goal of TEHDAS is to support Member States and the Commission in developing and promoting concepts for exchanging health data for secondary purposes such as research, policymaking, education and innovation across Europe." (TEHDAS, 2022). The eight work packages that makeup TEHDAS examine prospects and provide solutions for data governance framework and functions. They also consider choices for infrastructures for data quality control and data sharing, sustainability, and ethical models. The purpose of WP 5, "Sharing Data for Health," is to make suggestions to EU Member States on how to draft national legislation to permit the interchange of health data across borders (TEHDAS, 2022). From the viewpoint of data users, one mission within this context seeks to identify and develop the evidentiary base for the secondary use of health data

(for researchers and policymakers). Clinical trials, research cohorts, and biobank data are all included in the "secondary use of electronic health data," despite the fact that it is primarily focused on EHR exchange (TEHDAS, 2022).

Altogether, it appears that the EU is constructing a solid framework in which cross-border secondary use of health data is becoming increasingly possible. However, it appears that the European Health Union, which had not yet been established prior to the COVID-19 pandemic, is the most promising EU effort to facilitate secondary health data used for research. The section that follows will outline two EU guidelines aimed specifically at interoperability standards that were 'active' prior to the start of the pandemic.

4.4.1. The European Interoperability Framework

There are specific guidelines that address data management within secondary data use policies. For example, the European Interoperability Framework (EIF) is a set of guidelines that outline how governments, companies, and individuals can connect with one another both inside the EU and across the boundaries of Member States (eHealth Network, 2017). The EC (2010) adopted the following definition of an interoperability framework:

> "An interoperability framework is an agreed approach to interoperability for organisations that wish to work together towards the joint delivery of public services. Within its scope of applicability, it specifies a set of common elements such as vocabulary, concepts, principles, policies, guidelines, recommendations, standards, specifications and practices." (European Commission, 2010)

This framework has been modified specifically for use in the health sector: Refined Framework for European Interoperability in Health (ReEIF). This defines a common set of standards, profiles, and practises pertinent to the delivery of healthcare services electronically. When implementing any form of electronic health practices, Member States will be encouraged to use the framework. Unlike the original framework, which addressed solely four interoperability layers: technical, organisational, legal, and semantic, the redefined framework contains six layers and looks like this:

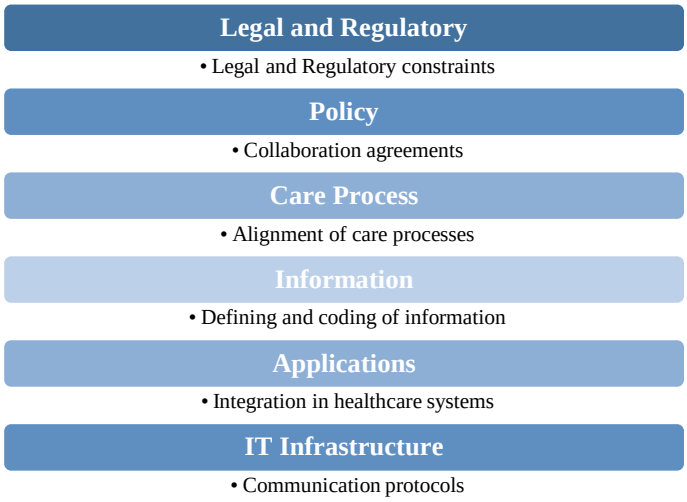


Figure 3. ReEIF model (eHealth Network, 2017).

The two tables that follow offer more details on the many facets of interoperability. The first illustrates the alignments required at various interoperability levels:

Organisation A	← →	Organisation B
Legal and Regulatory	Compatible legislation and regulations	**Legal and Regulatory**
Policy	Collaboration agreements	**Policy**
Care Process	Alignment of care processes and workflows	**Care Process**
Information	Data model, terminologies, formatting	**Information**
Applications	Integration in healthcare applications	**Applications**
IT Infrastructure	Communication- and network protocols	**IT Infrastructure**

Figure 4. Alignments interoperability levels (eHealth Network, 2017)

Another illustration of the stakeholders that may be involved in the various levels of interoperability is as follows:

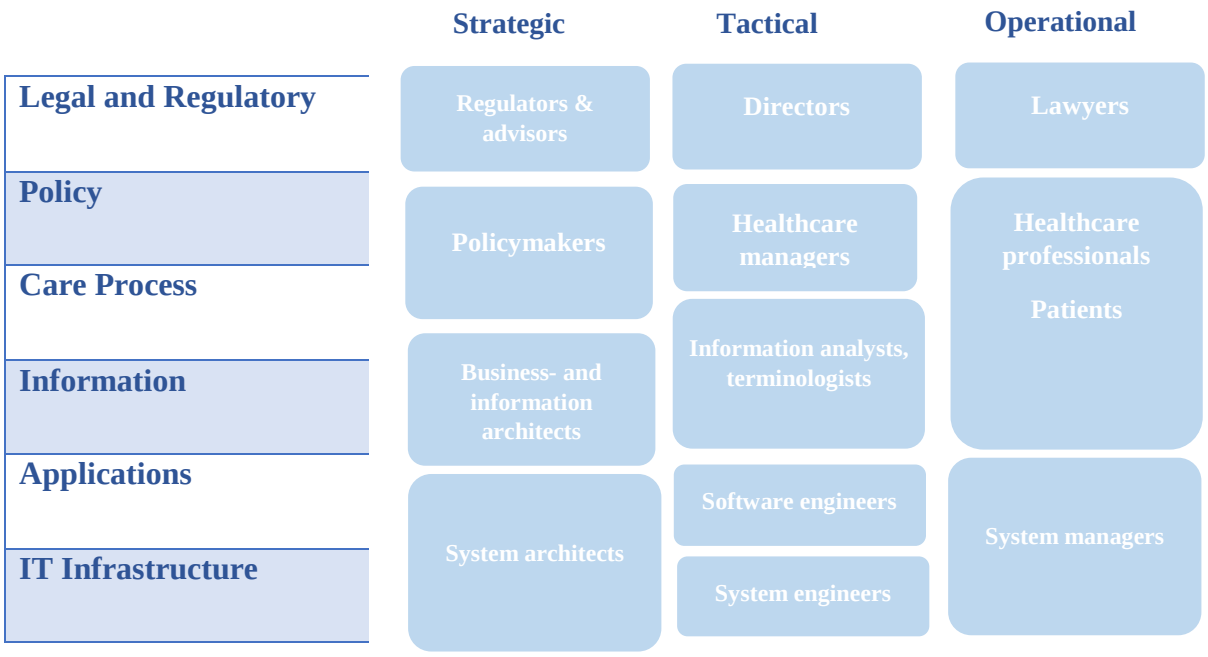


Figure 5. ReEIF model – stakeholders (eHealth Network, 2017)

4.4.2. Cross-border interoperability of EHRs

More specifically, the interoperability of EHRs enables the use and exchange of the gathered data between Member States of the EU that are neighbouring and those that are not. In terms of legal interoperability, six Member States had implemented legal requirements creating a foundation for international exchange as of 2014 (Milieu and Time.lex, 2014). Moreover, just a small minority of Member States put into practice particular technical guidelines and requirements (Milieu and Time.lex, 2014). Therefore, the majority of nations lacked legislative regulations governing the various interoperability layers. Neither the national nor the EU frameworks included a mandatory legal mandate for the establishment of EHR systems.

Following a 2017 EC open consultation, support for interoperability with harmonised standards and consensus on the need for eventual EU legislation on these concerns were given a high priority (European Commission, 2018). More specifically, to increase interoperability, the EU council requested the Member States and the EU Commission to emphasise the necessity for standard data structures, coding systems, and terminology (Council of the European Union, 2017). The EC's Recommendation, which will be examined in the following part, is a significant step toward achieving interoperability on the various layers described in the previous chapter.

To improve the movement of EHRs across borders, the EC has issued a recommendation on a European EHR exchange standard (European Commission, 2019). This recommendation builds on the previously mentioned initiatives and projects that seek to allow an increased cross-border flow of health data. This exchange format is assisting EU countries in their efforts to facilitate citizens, medical professionals, and public administrators to safely access and share health data from anywhere in the EU (European Commission, 2019).

In addition to promoting data best practices and policies to guarantee the privacy and integrity of health data, the recommendation sets detailed technical requirements for electronic health record access and interoperability (European Commission, 2019). Technical requirements are suggested as a place to start for upcoming growth, and it is advised that a governance approach involving all pertinent stakeholders be used (European Commission, 2019). Furthermore, the EC recommends that MSs adopt the resources offered by the European e-Health Services and adopt the necessary legislative and policy measures to encourage the deployment of interoperable EHR systems (European Commission, 2019).

Briefly stated, the framework consists of (a) the guidelines that should govern the access and exchange of EHRs across international boundaries; (b) a collection of standard technical requirements in specific health information domains (i.e., the foundation for the exchange format); and (c) a procedure to advance the further elaboration of the format (European Commission, 2019). Additionally, some interoperability criteria for representing and transmitting health data are included in the proposal of the baseline for the European EHR Exchange Format (European Commission, 2019). Future developments of the Commission's Exchange Format will be made through a collaborative coordination approach that takes into consideration the most recent advancements in technology and methodology (European Commission, 2019). It's crucial to emphasise that these recommendations still have no legal effect on how Member States may implement them at the national level. As a result, the EC recommendation is merely a soft law approach.

5. Analysis

According to the research design, the research question and hypotheses are divided into two parts to provide a logical structure for the answer to the research question. The first part, or sub-question, seeks to discover what multilateral projects were formed or were active in secondary data use in research for policymaking purposes during the pandemic. The second part, or sub-question, seeks to determine whether cross-border interoperability standards were sufficient to

enable efficient data sharing in these multilateral projects. The first question must be answered before selecting documents for analysis, or in other words, before answering the second question. As a result, the first section of this chapter may be regarded as the research findings that confirm or disprove the first part of the hypotheses. Furthermore, these findings will provide the ultimate research context for the document analysis, after which the outcome will then again provide the findings that could either confirm or disprove the second part of the hypotheses. If the text contains a document number indicated as #X, the document can be found in appendix A's list.

5.1. Research findings part I

The first sub-question seeks to identify potential multilateral projects formed or active in health data sharing for secondary purposes during the pandemic. The first part of the hypotheses, whether EU efforts resulted in more or fewer multilateral collaborations during COVID-19, could be confirmed or disproved in this way. These collaborations are represented as a project or consortia involving a multilateral relationship, as stated in the operationalisation and document inclusion criteria section. These projects also have the following features: they are EU-funded public projects or consortia (which may include non-EU Member States), they focus on secondary use of health data to facilitate COVID-19 virus research for policymaking, and they do so through, for example, EHR sharing.

As of January 2020, the EC promoted new research projects under the Horizon2020 funding (European Commission, n.d.). The commission set aside €48.2 million in March 2020 for 18 research projects. The 18 shortlisted projects were divided into four research themes: "preparedness and response," "diagnostics," "treatment," and "vaccines" (European Commission, n.d.). Only the first category was relevant to this thesis because the emphasis is on secondary data sharing for research with the goal of advanced public health policymaking.

Following a second call from the EC, 23 new research projects were shortlisted to receive Horizon2020 funding in August 2020. These 23 projects received €128.2 million in funding, which was divided into different themes: "rapid repurposing of manufacturing for vital medical supplies and equipment," "medical technologies, digital tools, and artificial intelligence analytics to improve surveillance and care at high Technology Readiness Levels," "behavioural, social, and economic impacts of outbreak responses," "Pan-European COVID-19 cohorts," and "collaboration of existing EU and international collaborators" (European Commission, n.d.). Only the projects in the last two themes were deemed relevant to this thesis.

There appear to have been numerous initiatives of collaboration in terms of secondary data use during the time period that this thesis' scope adheres to. Therefore, the following list is not an exhaustive list of collaborative projects that have been initiated and released. However, this list is as selective as possible. Organisational policies, plans, recommendations, and other related documents were analysed using the aforementioned indicators of dependent variables (chapter 3.3.) and inclusion criteria (chapter 3.2.1.). For this thesis, the following typical projects or consortia were chosen.

Name	Project or Consortium objective	Partners (EU or Non-EU)	EHR sharing
I-MOVE-COVID-19	"To obtain epidemiological, clinical and virological information on coronavirus and infected patients through the I-MOVE surveillance network spanning 11 countries." (European Commission, n.d.) (#1)	25 partners: AL, DE, ES(5), FR(5), IE, LT, NL(2), PT(2), RO, SE, UK(5)	Unclear
RECOVER	"To gather comprehensive data from clinical and epidemiological studies to strengthen Europe's clinical research preparedness for future emerging infectious diseases" (European Commission, n.d.) (#1)	11 partners: BE(2), CN, DE, FR(2), IT, NL(3), UK	Yes
unCover	"Unravelling Data for Rapid Evidence-Based Response to COVID-19" (European Commission, n.d.) (#2)	29 partners: BA, BE(2), BR, CO, EL, ES(5), HR(2), IE(2), IT(2), KR, LU, NO, PT(3), RO(2), SK, TR, UK, US	Yes
HERoS	"To improve the effectiveness and efficiency of the response to coronavirus outbreak by providing guidelines for improved crisis governance."	11 partners: FR, FI(2), IT, NL(2), PL(3), UK, US	Unclear

	(European Commission, n.d.) (#1)		
SYNCHROS	"Coordination and support for synchronising cohorts and population surveys in Europe and worldwide" (CORDIS | European Commission, n.d.-b) (#8)	11 partners: BE (2), CH, FR (2), NO, PL, ES (4)	Unclear
ReCoDID	"The consortium brings together a multidisciplinary team from four continents to fast track the research response to viruses and other pathogens by facilitating data and sample sharing between infectious disease cohort studies." (CORDIS | European Commission, n.d.) (#4)	15 partners: CA (4), CO, DE (2), DK, FR, NIC, NL (2), SN, USA (2),	Unclear
ORCHESTRA	"Connecting European Cohorts to Increase Common and Effective Response to SARSCoV-2 Pandemic" (European Commission, n.d.) (#2)	26 partners: AR, BE, CG, DE(6), ES(3), FR(3), GA, IT(6), LU, NL, RO, SK	Unclear
EpiPose	"To understand epidemiological characteristics COVID-19, social dynamics of the outbreak, public health preparedness and response, and assess economic impact" (European Commission, n.d.) (#1)	6 partners: BE(2), CH, IT, NL, UK	Unclear
CORESMA	"To help devise evidence-based response strategies by combining clinical, epidemiologic and immunological data from field studies and implementation research." (European Commission, n.d.) (#1)	7 partners: CH, CI, CN, DE(2), NL, NP	Unclear

EXSCALATE 4CoV	"To exploit powerful computing resources to identify molecules capable of targeting coronavirus and develop an effective tool to counter future pandemics" (European Commission, n.d.) (#1)	18 partners: BE, CH(2), DE(2), ES, IT(10), PL, SE	Unclear

5.2. Analysis part I

The EC and its Member States were unified and determined to invest in organised scientific research as defence against the novel disease at the dawn of 2020. These collaborations represent a step forward in multilateral governance, as Member States have been working together to reap the benefits of sharing health data in order to provide reliable research to inform policymaking for all EU citizens affected by the pandemic. Derviş (2020) cites cross-border collaboration to benefit from the spill-over effects as one of the reasons for preferring multilateral governance during a crisis. Looking back on the challenges identified by Frieden et al., all of the countries that collaborated had health data considered a public good and all had the same goal, which was to mitigate the effects of the COVID-19 pandemic. Clearly, the countries that contributed had faith in the multilateral community (Frieden et al., 2012); otherwise, they would not have entrusted their health data to these projects. Given the diversity of the countries that contributed, it is reasonable to assume that none of them had difficulty pooling authority (Jit, et al., 2021). Countries that did not participate may have domestic issues that make it difficult for them to deal with longer-term multilateral commitments to provide health data as a public good which is one of the challenges Frieden et al (2012) identified.

Furthermore, the increase in multilateral projects for COVID-19 research could be linked to the two developments in health governance politics that the theory refers to. To begin with, the new narrative about global health governance ensured that countries were more willing to distribute their initiatives on a global rather than a local scale, which was also stated by authors Cooper, Kirton, & Schreck (2009), Fidler (2003), and Dodgson, Drager, & Lee, (2002). Second, as the institutional architecture of health governance has changed, the actors involved have included not only public health ministers but also a large number of scientists and public health organisations. The importance of transnational networks in health governance was likewise identified by Shiffman et al. (2016). It is safe to assume that this united response has been brought about by the EU's strong multilateral system, thereby resolving the sovereignty paradox

mentioned by authors Frenk and Moon (2012). The guidelines that support the secondary use of health data collaborations in the EU multilateral system and facilitate these collaborations are outlined in the case description. Even though the EU guidelines are simply soft law, this could refer back to some sort of enforcement authority that was deemed essential for innovations in a multilateral institutional framework by Duff et al., 2021.

The financial structure that has supported these multilateral collaborations is most likely the most important. These projects were not only motivated but also shielded from political influence, thanks to European funding: Horizon2020. This can be traced back to what Gostin and Mok (2009) and Fidler (2010) meant when they stated that in order to be effective in health governance, a strong multilateral system required reliable and sustainable financial support. Furthermore, the collaborations were heavily reliant on dedicated scientists and other network organisations that made the research possible. This is also consistent with what Chan (2015) stated, that multilateral innovation cannot be achieved solely through regional and international organisations

It is clear that data-driven or evidence-driven policy-making has been chosen as a critical method of the EU's approach to fighting the pandemic. The projects intend to benefit from new health data sources and foster engagement with relevant stakeholders and citizens, to support policymaking when it comes to fighting the pandemic with ICTs, such as electronic health record sharing (Janssen & Helbig, 2018; Ferroa, et al., 2013; Linders, 2012). Alternatively put, the EC funded project initiatives with an emphasis on generating information to aid public health decisions. The majority of these projects used primary data for secondary purposes, i.e. research, as explained in theory by Safran, et al., 2007. There was an obvious investment in primary health data sharing to inform public health decisions, which corresponds to the definition of secondary use of health data outlined in the theory chapter (Boyd, et al., 2021). Furthermore, the chosen approach clearly deviates from evidence-based policymaking because the emphasis is on incorporating big and open data sources, such as health data, to inform decision-making and policy co-creation. Thus, the EU funded these projects attempting to find the most effective method of incorporating data into policymaking as authors Bijlsma et al (2011) stated in the theory chapter.

It not only demonstrated the importance of multilateral governance in these research cross-border partnerships, but it also demonstrated how the EU's digital environment has facilitated these partnerships. That is, facilitating cross-border data sharing is also considered relevant in these projects, as transnational sharing data in these projects aims to connect different cohorts

and data collections in order to support time-efficient and cost-effective policy responses to COVID-19. Similarly, authors Oderkirk, Wenzl, & Slawomirsk (2019) identified these benefits in cross-border data sharing. That is, the free flow of data allows researchers to compare and analyse different data sets while avoiding effort duplication.

Above all, it encourages the exchange of knowledge and expertise, resulting in the most informed decision-making, which has been deemed important in collaborations according to Baller, Dutta & Lanvin (2016). When examining the four archetypes of cross-border health data sharing identified by Schwalbe et al (2020), most projects make use of data generated at the patient level as well as data generated from observational and experimental research initiatives. Particularly important in these projects were data on the safety, effectiveness, and manufacturing processes of prevention and treatment technologies that benefit EU society as a whole, as well as data on risk assessment and ensuring that patients are treated as effectively as possible, which Heymann (2020) stated to be an essential approach in pandemics. As previously stated, the majority of the health data in this project is real-world data used for research to support public health decisions, implying that these projects make use of secondary health data. The majority of projects use not only pre-existing personal data but also research data made available to other researchers for data analysis, validation, and application by the scientific community.

The EU has encouraged the launch of 41 new COVID-19 pandemic research and innovation projects. This could imply that the COVID-19 pandemic has seen an increase in multilateral research projects or consortia as a result of EU efforts. Ten projects remained after being shortlisted based on the criteria in this thesis. This could still be considered a relatively large number of projects emerging in such a short period of time. This was due in part to EU funding, but it was also due to the digitised EU environment and the recognition of public health data as a public good. As a result, the following assumption can be made: the COVID-19 pandemic resulted in increased multilateral research projects to facilitate research for policymaking. Meaning that the first part of hypotheses one (H1) and three (H3) can be confirmed: During the COVID-19 pandemic, EU secondary health data use guidelines enabled greater multilateral research collaborations.

5.3. Research findings part II

The second sub-question seeks to determine which cross-border data-sharing standards were present in these projects to enable efficient secondary data use, and the answer will thus confirm

or refute the second part of the hypotheses: whether these guidelines led to sufficient or insufficient cross-border data-sharing standards within these partnerships. To make the search for interoperability standards easier, the projects chosen focus on data harmonisation and sharing within research cohorts, which can only be accomplished with the right data-sharing standards. As a result, detecting interoperability standards will now proceed through the projects by analysing what they struggled with when harmonising data or what they needed to improve to allow data harmonisation because it was now insufficient. Four of the projects mentioned above were chosen, with a focus on data harmonisation and exchange. The following multilateral EU-funded research initiatives address this: ReCoDID, ORCHESTRA, unCoVer, and SYNCHROS. To begin, a brief overview of the projects is provided to explain their goals.

ReCoDID stands for Reconciliation of Cohort Data for Infectious Diseases and aims to create a comprehensive, maintainable data collection infrastructure for infectious disease (ID) cohorts that enables the utilisation of clinical-epidemiological (CE) and high dimensional laboratory (HDL) data for the identification, analysis, and prevention of contamination with identified and unidentified pathogens (ReCoDiD, n.d.). Additionally, this project aims to eliminate common obstacles to data and human sample sharing and develop novel approaches for shared control, connected data and biorepositories, and collective examination (ReCoDiD, n.d.). (#4&15)

The ORCHESTRA Cohort ("Connecting European Cohorts to increase common and effective SARS-CoV-2 Response") (ORCHESTRA, n.d.) seeks to develop a pan-European cohort based on current and new extensive cohorts in EU and non-EU states of COVID-19 citizens infections of different generations and circumstances to evaluate risk factors, causes, and lasting effects (ORCHESTRA, n.d.). Additionally, this initiative seeks to create evidence-based guidelines for the efficient prevention, security, and medical care of COVID-19 patients, with an emphasis on the vulnerable community as well as to evaluate how lifestyle, socioeconomic variables, the environment, and containment methods affect the spread of the disease (ORCHESTRA, n.d.). The project also aims to offer a paradigm for data collection to improve reactivity to potential pandemic outbreaks in the future. (#9&16)

The unCoVer project, Unravelling Data for Rapid Evidence-Based Response to COVID-19 aims to keep track of, find and make it easier for people to access and use Real World Data relevant to COVID-19 (unCover, n.d.). Additionally, it seeks to discover data gaps, disadvantaged groups, and proactive synergies with COVID-19-related existing and new clinical databases (unCover, n.d.). It also intends to create a network for the utilisation of various data sources that can streamline ethical and legal considerations using innovative

software tools (unCover, n.d.). The final objective of the project is to combine knowledge on the application of innovative computational, epidemiological, and biostatistical methodologies to handle dissimilar and complicated data (unCover, n.d.). (#3&13)

SYNCHROS is the acronym for SYNergies for Cohorts in Health: integrating the Role of all Stakeholders (Synchros, n.d.). In order to create a future culture of unified demographic, patient, and clinical trial cohorts in Europe, this organisation was founded(Synchros, n.d.). Additionally, it seeks to map the cohort scene in the EU and major global projects (SYNCHROS Repository) (Synchros, n.d.). To facilitate the synchronisation of previous and future data collection, the study also aims to determine the optimal ways for merging cohort data (Synchros, n.d.). The initiative also seeks to find solutions to the practical, moral, and legal issues associated with combining data from a patient, clinical trial, and population cohorts (Synchros, n.d.). (#8&17)

5.4. Analysis part II

The interoperability standards of these projects are evaluated to determine whether there were sufficient or insufficient data-sharing standards present. It's interesting to note that the EC sought to support initiatives that would help to address the issue of potential heterogeneity sources in integrated results that can arise from significant disparities in data-gathering methods between sites. That is to say, there was some understanding of the value of thorough harmonisation, which enables data to be comparable and, ultimately, trustworthy and legitimate for integrated research analysis. Thus, the EU is well aware of the difficulties associated with secondary data use in terms of proper data management. To achieve sufficient statistical power to verify hypotheses and control for relevant biases, data collection and harmonisation across diverse cohorts could be helpful. However, this is a difficult mission because not all sites intend to perform data collection and analysis from the start. In other words, Granda and Blasczyk (2016) stated that different data collection techniques create sources of variability that can significantly skew integrated results. This second section of the analysis will look at whether these EU efforts were sufficient by detecting whether the six interoperability standards played a role in these projects. The first and second interoperability levels (legal and policy) and the fifth and sixth levels (application and infrastructure) are analysed together.

5.4.1. Legal and Policy interoperability

As stated in the case description, all EU laws and regulations pertaining to data interoperability are considered soft law. The legal interoperability level, as described in the research design chapter, is intended to guide organisations operating under various legal frameworks, policies, and strategies in order for them to collaborate (eHealth Network, 2017). Because all four

projects have partners from 35 different countries (European Commission, n.d.; European Commission, n.d.) (#1&2), it is possible that partners within the same project follow different legal interoperability guidelines, resulting in uneven data exchange/collection/analysis methods. Unfortunately, no consistent legal interoperability frameworks were discovered by one of the projects, as member states are still responsible for governing their national health policies, as stated in the theory chapter by Goossens et al. (2022) and in accordance with the case description of the current EU guidance. Even if EU legally binding interoperability standards exist, all four projects include non-EU members that may be unaware of these standards, which still results in heterogeneity.

Despite the fact that the majority of legal and policy interoperability standards between sites are non-existent or poorly documented, three of four projects mention the FAIR principles. Despite the theoretical debates that have occurred since its adoption (European Commission, 2018), Mons et al. (2020) have identified these principles as important. The application of the FAIR principle in the projects SYNCHROS, unCoVer, and ORCHESTRA could be considered a small step towards legal interoperability success (Bickenbach, 2020; Pealvo et al., 2021; Rinaldi et al., 2022) (#20, #18, #21). Furthermore, all four projects adhere to the GDPR in terms of medical data privacy (Bickenbach, 2020; Kabir, 2021; Orchestra-cohort, n.d.; Maxwell, 2020) (#20, #24, #25, #23). However, as the case description suggested, regulations or policies for data protection in the EU may obstruct proper data exchange (Iacob & Simonelli, 2020; Vukovic et al., 2022; Boyd et al., 2021; Salas-Vega et al., 2015) For example, the SYNCHROS project has discovered that some ambiguity in this regulation limits its application (Bickenbach, 2020) (#20). However, it was stated at the outset that only data management would be considered and that the legal aspects of data protection would not be examined in this thesis.

5.4.2. Care process interoperability

When it comes to the level of interoperability of the care process, the theory mentions the issue of partners being unaware of certain work tools or methods, which leads to inconsistent research results or confusion (Granda & Blasczyk, 2016). The ReEIF refers to this level as the alignment of care processes and workflows (eHealth Network, 2017). Because the projects' goals are to harmonise the received data, the inconsistency of data capture processes made this much more difficult. Data quality, missing data, and bias sources are all given special consideration in the ReCoDiD and unCoVer projects (Debray, 2022; Peñalvo, Blázquez, Chausa, Menasalvas, & Gómez, 2021) (#30) (#19). Due to the high degree of heterogeneity in the methods or tools used across sites, the results of all four projects are not directly comparable, and the projects are

unable to use the results as legitimate (Sialm & Sanchez-Niubo, 2020; de Lamballerie, 2022; Debray, 2022; Peñalvo, et al., 2021.) (#26)(#29)(#30)(#18).

To give an example, it is surprising that, with the exception of one, the goal of using EHRs does not appear to be a prominent aspect in any of the projects when looking at data collection tools within these projects. In other words, aside from the projects Recover (which was not chosen for further analysis) and unCoVer, it was not mentioned in any of the project objectives. Furthermore, only in the unCoVer project are EHRs used as the primary method of data collection and analysis (unCoVer, n.d.)(#13). Given the EU's prior investment in facilitating cross-border EHR sharing for both primary and secondary purposes, such as the EC recommendation (European Commission, 2019) and the THEDAS's primary focus on EHRs (TEHDAS, 2022), as well as the potential of EHRs to improve clinical research (Fears et al., 2014), it is surprising that these efforts did not contribute to projects with this specific focus. The topic of standardisation of care processes is rather technical in medical terms and beyond the scope of this thesis's expertise, making it difficult to detect detailed examples in the project documents. However, given that all four projects have mentioned the importance of open and harmonised sharing of work processes, it appears that a general care process standardisation has yet to be agreed upon.

5.4.3. Information interoperability

All four projects are probably the most concerned with achieving information-level interoperability. The ReEIF refers to this level as interoperability of defining and coding of information (eHealth Network, 2017). It has clearly been stated as one of the most difficult aspects of data management in these projects. In the SYNCHROS project, the approaches pertaining to cohort data comparability have been the main objective (Synchros, n.d.) (#17). According to SYNCHROS' cohort mapping process, cohort harmonisation methods and extensive information about cohort samples and data were not reported consistently (Sialm & Sanchez-Niubo, 2020)(#26). In fact, ORCHESTRA was directly confronted with this challenge: they matched 2,500 SARS-CoV-2 variables to globally standardised terminology (Rinaldi, et al., 2022)(#21). However, not every data component has an international code. As a result, innovative ideas have been presented to the relevant standard organisations in an effort to strengthen future data interchange internationally and develop a framework that will make data in preparation plans more comparable (Rinaldi, et al., 2022)(#21).

Similarly, the unCoVer project mentions semantic interoperability as a goal, which refers to the harmonisation of the accurate format and meaning of exchanged data and fits in the broader

level of information interoperability (Peñalvo, Blázquez, Chausa, Menasalvas, & Gómez, 2021)(#19). According to the unCoVer strategy plan, achieving full semantic interoperability would pave the way for innovative tools to be usable (Peñalvo, Blázquez, Chausa, Menasalvas, & Gómez, 2021) (#19). The ReCoDiD project is confronted with this level of information interoperability in a different way as they attempt to create linkages specifically between CE and HDL data within and across cohorts, which has previously led to inconsistency in research analysis (Maxwell, 2020) (#23). Given that all four projects have identified a lack of information interoperability, such as a lack of common language to define concepts, it is clear that general information interoperability is not yet agreed upon at all.

5.4.4. Application and Infrastructure interoperability

In terms of application and IT infrastructure interoperability, it appears from document analysis that these two levels are linked and thus must be analysed together. The application level, according to the ReEIF, addresses how healthcare information systems will manage the import and export of medical data, whereas the IT infrastructure level addresses general networking and communication protocols and standards, as well as storage, backup, and database engines (eHealth Network, 2017).

The ReCoDID project was mandated to create a harmonisation channel for COVID-19 cohorts in Europe, in collaboration with the European COVID-19 Data Portal hosted by the European Molecular Biology Laboratory (CORDIS | European Commission, n.d.)(#4). ORCHESTRA and ReCoDID work together to create pan-European cohort data sets for COVID-19 by combining information from both old and new large-scale cohort data sets across Europe (CORDIS | European Commission, n.d.) (#32). As a result, there are significant synergies between the two initiatives. Nonetheless, whereas ReCoDID's primary goal is to build a keen cohort data repository, and substantial attention was paid to legal and technical obstacles along the way, ORCHESTRA's primary goal is to implement consistent new data gathering between various cohort data to produce appropriate quality data advancing COVID-19 prevention and treatment (Tacconelli, et al., 2022) (#28). COVID-19 patients, risk groups, and healthcare professionals are only a few of the over 1,300,000 people who have been included in the ORCHESTRA cohort thus far (ORCHESTRA, 2022) (#33). That is to say, the projects take different approaches to federated vs. centralised data infrastructures. ORCHESTRA predominantly uses a federated infrastructure while ReCoDID adopts a centralised architecture (Tacconelli, et al., 2022) (#28). UnCoVer chose a federated architecture for the coordinated and support action project, and their primary data came from EHRs (Peñalvo, et al., 2021) (#18).

More than 22,000 COVID-19 patients who were hospitalised were included in this data (Peñalvo, et al., 2021) (#18). Additionally, these EHRs contained registries with almost 2,000,000 COVID-19-infected people, as well as national surveillance and screening data (Peñalvo, et al., 2021) (#18).

These different infrastructures led to different types of data applications. For instance, many cohorts are unable to exchange participant-level data directly because ORCHESTRA includes cohorts that existed prior to the project's inception (Orchestra, n.d.) (#27). A federated data analytics framework and machine learning network based on OPAL-DataSHIELD (Datashield, n.d.) has been created to allow the use of flexible and extensive analytics methodologies while avoiding direct data sharing (Tacconelli, et al., 2022) (#28). Even while the Opal-DataSHIELD now does not support the full range of analytic techniques and machine learning tools, it appears to be the best choice for datasets that cannot be exchanged in a straight data flow (Tacconelli, et al., 2022) (#28).

Opal-DataSHIELD is also used in the unCoVer project, as an open-source server programme that supports interoperability, and is used to characterise, harmonise, and incorporate heterogeneous data together into a multi-user data repository (Peñalvo et al., 2021) (#18). This federated infrastructure offers effective and safe methods for controlling extremely private patient data from EHRs, data that was not gathered for research purposes and necessitates a high-security solution as well as careful supervision of data protection compliance (Peñalvo et al., 2021) (#18). It should be emphasised that unCoVer had considerable ethical challenges when establishing the federated infrastructure during the early project start-up phase, which caused delays in the intended assessments that could have helped to mitigate pandemic effects (Peñalvo et al., 2021) (#18).

As for ReCoDID, the chosen infrastructure is a centralised data collection pipeline presented at the European COVID-19 data portal (via the European Molecular Biology Laboratory) due to the potential for combining significant OMICS (a branch of biology that ends in "omics," like genomics, transcriptomics, proteomics, or metabolomics) data with clinical-epidemiological data, as well as for future large-scale studies spanning various datasets and on the discrepancy between high-dimensional and well-characterised clinical-epidemiological data (CORDIS | European Commission, n.d.) (32). However, the European Molecular Biology Laboratory is considering some federated data analysis options, including national hubs of the European Genome Archive infrastructure (EGA) (CORDIS | European Commission, n.d.)(#32).

Overall, it appeared that all four projects took different approaches, as there is currently no standardised or common EU infrastructure. According to the theory, the best way to provide better access to health data for research is to have centralised databases (Heijlen & Crompvoets, 2021). However, because there are numerous privacy and ethical concerns here (Broekstra, Aris-Meijer, Maeckelberghe, Stolk, & Otten, 2020), centralised infrastructures should include informed consent for all personal health data. One of the best options so far is to use the OPAL DATASHIELD used by unCoVeR and ORCHESTRA. However, in order to make the data collection and analysis environment more homogeneous, the EU must provide a common infrastructure that allows for secure but open data applications in research.

When it comes to the data archetypes identified by Schwalbe et al., (2021), all projects use mostly patient data form observational which means a process from the first (patient data) to the fourth (research data) archetype, which has resulted in different interoperability challenges. To return to Karpati and Ellis' (2019) framework for data-driven policymaking, in order to provide research for appropriate data-driven policymaking, interoperability standards must be met at every stage of the process. The analysis in this thesis demonstrated that data-sharing standards tied to interoperability were insufficient to make data-driven policymaking easier in all four steps, data collection, data management, data analysis, and data communication. Choosing projects or consortia with the goal of data harmonisation in research to support health governance provided a clear picture of what these projects struggled with. They discovered that disparate data-sharing standards slowed the process of conducting effective research. To summarise, there were not enough standardised methods or systems to detect at all interoperability levels. A deep and sensitive analysis was hampered by the absence of coordination and harmonisation among the COVID-19 patient cohorts, which were developed with speed. In other words, a lack of clear rules for data sharing, understanding of legal requirements, or access to technical tools for standardising and sharing data frequently impedes scientific progress. Furthermore, there are numerous community-developed data standards available at the moment, making it an overlapping and confusing choice. Altogether, the second part of hypotheses two (H2) and three (H3) can be confirmed: During the COVID-19 pandemic, insufficient cross-border data sharing standards resulted in ineffective data sharing within these partnerships.

Given that the analysis part I confirmed hypotheses one (H1) and three (H3), it appears that the following hypothesis is definite: During the COVID-19 pandemic, EU secondary health data

use guidelines enabled greater multilateral research collaborations, but insufficient cross-border data sharing standards resulted in ineffective data sharing within these partnerships.

6. Conclusion

The purpose of this thesis was to investigate the EU guidelines on secondary use of health data in the context of COVID-19 research collaborations. More specifically, this thesis sought to explain how these guidelines affected multilateral research projects. The secondary use of health data in research to inform public health decisions is not a new phenomenon, nor are the EU guidelines that accompany them. However, because sharing personal health data for purposes other than primary care is a sensitive topic with significant political, ethical, technical, and legal challenges, these guidelines remain only recommendations and are not legally binding. For instance, the EU has sought to support cross-border interoperability of health data for primary and secondary purposes, which means that it can help citizens get personalised care wherever they are, but it can also help with international research or clinical trials to benefit public health in the EU as a whole. The creation of the eHealth Network is one example of this. These guidelines are highly crucial during health crises, such as the recent COVID-19 pandemic, to encourage timely and efficient data sharing for patient treatment, as well as research to inform public health decisions and health surveillance.

The COVID-19 pandemic in particular has brought to light the urgent need for international scientific cooperation in both the public and private sector sectors to research diagnostics, vaccines, and treatments to respond to health emergencies. This needed the free exchange of samples and data as well as quick access to it. As a result, cross-border collaborative research during COVID-19 provided a unique research field in which to investigate whether previously proposed EU guidelines facilitated and guided them. By doing so, it could be determined whether the EU's efforts were sufficient, or whether necessity truly is the mother of invention, implying that these guidelines require another round of revision. This thesis took a two-step approach, first determining whether these guidelines were appropriate for the formation of multilateral research projects, and then determining whether these guidelines assisted in the establishment of adequate standards for data sharing within these projects. Following both examinations, this thesis demonstrated how EU secondary health data use guidelines enabled greater multilateral research collaborations while failing to result in adequate cross-border data-sharing standards. The path to this answer is described in the following section.

6.1. Summary of findings

While describing the research context in which these collaborative COVID-19 projects were developed, it became clear that the EU has provided a significant amount of initiatives to support a data-driven policy environment in general and specific digitisation of healthcare in particular. However, when examining EU-wide policy for the secondary use of health data, it appeared that the most promising initiatives to support and guide the secondary use of health data were only introduced during the creation of the European Health Union, which occurred during and after the COVID-19 pandemic. This was the first indication that the environment for secondary data use was not as advanced and accessible as it should be to support international COVID-19 research projects to aid public health decisions. After identifying a lack of cross-border data interoperability as one of the most significant challenges in using secondary data in research projects, this thesis provided two specific guidelines initiated by the EU prior to the COVID-19 pandemic: the European Commission's recommendation on standardised EHR formats and the refined framework for European interoperability in healthcare. Although they are considered soft law, these two guidelines were most likely found to be important while developing the COVID-19 research projects.

This thesis found that the EU was eager to invest in research projects that would contribute to the sharing of disease knowledge in order to effectively inform public health decisions. This resulted in the funding of 41 new COVID-19 pandemic research and innovation projects through Horizon2020. This was believed to be a positive step forward in multilateral public health governance, as 35 different countries collaborated on 41 different projects (the research projects also included non-member states). More specifically, the research initiatives supported by this EU funding aim for multilateral collaborations rather than national self-interest. More importantly, this thesis revealed an intriguing number of projects that were specifically focused on data harmonisation and data sharing in order to create a balanced and coordinated international research environment to support public health governance during the pandemic. As a result, this thesis concluded that these multilateral research projects were facilitated not only because of the EU funds but also because of the digitalised environment and increased awareness of the importance of cross-border data sharing for secondary purposes delivered by EU efforts.

However, while investigating whether these projects were supported by suitable cross-border data-sharing standards, this thesis discovered that the projects experienced poor synchronisation and coordination, making thorough and delicate harmonization challenging. This thesis

demonstrated through document analysis that none of the six interoperability levels was deemed adequate in the research projects. That is, a lack of clear criteria for data exchange, understanding of legal and policy requirements, and technical expertise to homogenise and distribute data hampered the initiatives, frequently impeding the advancement of research. Additionally, it was challenging to choose between the several community-developed data standards as the range of choices is too complicated and unclear. In short, it has not yet been possible to create meta-level interoperability between these standards. As a result, according to this thesis, heterogeneous cross-border data-sharing standards (i.e., insufficient interoperability standards) hampered data sharing within multilateral research partnerships.

6.2. Academic relevance and future prospect

Given the ongoing COVID-19 pandemic, the goal of this thesis was to contribute to the academic literature on the secondary use of health data for data-driven health governance, but with a focus on the specific contemporary context in which the world now finds itself. The COVID-19 pandemic provided a once-in-a-lifetime case study opportunity to determine whether EU efforts to facilitate secondary use of health data in crisis situations were adequate. As a result of discovering that this was not the case, this thesis was able to contribute to closing a current gap in the academic literature. More importantly, the fact that EU guidelines were insufficient could serve as a learning experience for anyone dealing with EU public health governance or research, as the lessons learned are beneficial not only for future pandemics or health emergencies but also for public health in general. Despite their importance, research initiatives to guide public health governance are fraught with uncertainty for their survival in the post-pandemic era, as current challenges may evolve into deeper-rooted issues or be viewed as intractable. Especially since opting for collaborative efforts in research, but in any field, is not always guaranteed, addressing cross-border problems, such as health issues, is destined for failure without good guidelines and a strong multilateral system.

A collaborative atmosphere is created through large collaborative projects like RECODID, ORCHESTRA, and unCoVer, which promote scientific method development and discovery. Nevertheless, this does not totally eliminate the requirement for local assistance with data management and technological issues. Additional efforts are required from every Member State to create common standards for health data. Extraordinary dangers to public health, like the COVID-19 pandemic, should drive a review of some interoperability standards so that researchers can produce timely and meaningful findings. Ideally, Member States should offer uniform instructions on how to implement such rules in the event of a pandemic. The European

Health Data Space presents promising work to support the use of secondary health data in the EU and to address interoperability issues. That is, the EHDS would offer the crucial legal foundation for accessing and using secondary health data for research to support policymaking. Data governance, data quality, and interoperability are all critical in the data space. Establishing official digital health entities in each EU country, for example, would address and monitor interoperability policies and agreements while also providing standardised infrastructure to support transnational data exchange. The shared control of these infrastructures should be addressed by appropriate regulations. Furthermore, for electronic health data derived from EHRs, a self-certification structure is desired to achieve total interoperability in patient data. Thus, the policy framework within which these secondary uses of health data occur should be revisited; balancing policies between safety and openness has yet to be achieved. This could be accomplished with additional financial support and the involvement of the appropriate number of data and health experts.

6.3. Research limitations

Although the study was designed to produce the most reliable study results possible, it did have some limitations. The first part of the hypothesis, whether the EU guidelines enabled or did not enable more research projects using secondary health data, cannot be fully supported by evidence. That is, it is impossible to rule out the possibility that other factors contributed to the emergence of international research collaborations in the field of COVID-19. Nonetheless, many of these newly formed projects aimed at harmonising health data in order to create an open and international research environment to support public health policy. That is why it appears that previous EU efforts to promote and facilitate this has resulted in, first, increased awareness and, second, a more accessible policy field. Furthermore, the contribution from the Horizon2020 fund was beneficial, as financial support is also desirable for a strong multilateral system. The second part of the hypothesis, determining whether data sharing standards were capable of facilitating proper data exchange within these projects, was based on stronger evidence, as it was decided to measure this by detecting the mention of interoperability standards. Other data-sharing standards may have been missing or played a role in impeding efficient and fast scientific research, but these were outside the scope of this study. In any case, it is now clear that the EU guidelines have had little impact on interoperability standards and may require a new evaluation and adjustment. Making these guidelines more legally binding would be preferable, but this would raise a slew of legal issues, particularly concerning privacy. This leads to a recommendation for further research, namely how more legally binding

guidelines could perhaps ensure that the secondary use of health data can be made easier. Furthermore, the generalisation of this thesis's findings was not the goal, and because it is tied to a specific context of the COVID-19 pandemic, additional research outside of the COVID-19 pandemic or when new guidelines are developed is strongly advised.

Bibliography

Abboud, L., Cosgrove, S., Kesisoglou, I., Richards, R., Soares, F., Pinto, C., . . . Bowers, S. (2021). *Summary of results: case studies on barriers to cross-border sharing of health data for secondary use.* TEHDAS.

ALLEA, EASAC, FEAM. (2021, April). *International Sharing of Personal Health Data for Research*. Retrieved from https://easac.eu/fileadmin/PDF_s/reports_statements/Health_Data/International_Health_Data_Transfer_2021_web.pdf

Altheide, D. L. (1996). Process of Qualitative Document Analysis. *Qualitative Media Analysis*, 235-255.

Baeza-Yates, R. (2013). *Big Data or Right Data?* Barcelona: Labs Barcelona & Web Research Group, DTIC, Univ. Pompeu Fabra.

Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Chilton, J., Coraor, N., . . . Weaver, S. (2020). No more business as usual: agile and effective responses to emerging pathogen threats require open data and open analytics. *PLoS Pathog*.

Baller, S., Dutta, S., & Lanvin, B. (2016). *The Global Information Technology Report 2016.* Geneva: World Economic Forum.

Basrur, R., & Kliem, F. (2021). Covid-19 and international cooperation: IR paradigms at odds. *SN Social Sciences*, https://doi.org/10.1007/s43545-020-00006-4.

Bayati, M., Noroozi, R., Ghanbari-Jahromi, M., & Sadat Jalali, F. (2022). Inequality in the distribution of Covid-19 vaccine: a systematic review. *International Journal for Equity in Health*, https://doi.org/10.1186/s12939-022-01729-x.

Beach, D., & Pedersen, R. B. (2013). *Process-Tracing Methods: Foundations and Guidelines.* Ann Arbor: University of Michigan Press.

Bennett, A., & Checkel, J. (2015). *Process tracing : from metaphor to analytic tool.* Cambridge University Press.

Bertot, J. C., & Choi, H. (2013). Big data and e-government: issues, policies, and recommendations. *Proceedings of the 14th Annual International Conference on Digital Government Research*, 1-10.

Bettcher, D., & Lee, K. (2002). Globalisation and public health. *Journal of Epidemiology & Community Health*, 8-17.

Bickenbach, J. (2020). *Strategy Brief on practical, legal and ethical in the optimisation of cohort data in Europe D3.3.* SYNCHROS.

Bijlsma, R., Bots, P., Wolters, H., & Hoekstra, A. Y. (2011). An empirical analysis of stakeholders' influence on policy development: the role of uncertainty handling. *Ecology and society,*, 1-16.

Bowen, G. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 27-40.

Boyd, M., Zimeta, M., Tennison, J., & Alassow, M. (2021). *Secondary use of health data in Europe.* London: Open Data Institute.

Broekstra, R., Aris-Meijer, J., Maeckelberghe, E., Stolk, R., & Otten, S. (2020). Trust in Centralized Large-Scale Data Repository: A Qualitative Analysis. *Journal of Empirical Research on Human Research Ethics*, 365-378.

Bruzzone, G., & Debackere, K. (2021). As Open as Possible, as Closed as Needed: Challenges of the EU Strategy for Data. *les Nouvelles - Journal of the Licensing Executives Society, Volume LVI No. 1*.

BSA Software Alliance. (2017). *CROSS-BORDER DATA FLOWS.* Washington: BSA the Software Alliance.

Bull, S., Roberts, N., & Parker, M. (2015). Views of Ethical Best Practices in Sharing Individual-Level Data From Medical and Public Health Research: A Systematic Scoping Review. *Journal of Empirical Research on Human Research Ethics*, 225-238.

Carr, D., & Littler, K. (2015). Sharing Research Data to Improve Public Health. *Journal of Empirical Research on Human Research Ethics*, 314-316.

Chakraborty, C., Sharma, A., Sharma, G., Bhattacharya, M., Saha, R., & Lee, S. (2020). Extensive partnership, collaboration, and teamwork is required to stop the COVID-19 outbreak. *Arch Med Res, 51*, 728-730.

Chan, J. (2015). *Politics in the Corridor of Dying: AIDS Activism and Global Health Governance.* Baltimore, MD: Johns Hopkins University Press.

Chiasera, A., Toai, T., Bogoni, L., Armellin, G., & Jara, J. (2011). Federated EHR: How to improve data quality maintaining privacy. *011 IST-Africa Conference Proceedings*, 1-8.

Cooper, A. F., Kirton, J. J., & Schreck, T. (2009). *Governing Global Health.* Aldershot, UK: Ashgate.

CORDIS | European Commission. (n.d.). *RECODID*. Retrieved from https://cordis.europa.eu/project/id/825746

CORDIS | European Commission. (n.d.-b). *SYNCHROS*. Retrieved from https://cordis.europa.eu/project/id/825884

Council of the European Union. (2017). *Council Conclusions on Health in the Digital Society; Making.* 2017/C 440/05.

Courbier, S., Dimond, R., & Bros-Facer, V. (2019). Share and protect our health data: an evidence based approach to rare disease patients' perspectives on data sharing and data protection - quantitative survey and recommendations. *Orphanet J Rare Dis*, 175.

COVID-19 Clinical Research Coalition. (2020). Global coalition to accelerate COVID-19 clinical research in resource-limited settings. *The Lancet*, https://doi.org/10.1016/S0140-6736(20)30798-4.

Cuquet, M., Vega-Gorgojo, G., Lammerant, H., Finn, R., & ul Hassan, U. (2017). *Societal impacts of big data: challenges and opportunities in Europe.* Cornell University.

Dagliati, A., Malovini, A., Tibollo, V., & Bellazzi, R. (2020). Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Brief Bioinform*, 812-822.

Datashield. (n.d.). *A software solution for secure bioscience collaboration*. Retrieved from Data Shield secure bioscience collaboration: https://www.datashield.org/

Davies, W. (2017, January 19). *How statistics lost their power – and why we should fear what comes next*. Retrieved from the Guardian: https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy

de Lamballerie, X. (2022). *Optimised mapping of seroprevalence data according to time targeted populations, and categorised assays.* ReCoDiD.

Debray, T. (2022). *Reconciling measurements of individual cohort participants across heterogeneous data sets.* ReCoDiD.

DePoy, E., & Gitlin, L. N. (2016). *Introduction to Research Understanding and Applying Multiple Strategies.* Maryland Heights, Missouri: Mosby.

Derviş, K. (2020, November 17). *Multilateralism: What policy options to strengthen international cooperation?* Retrieved from Brookings: https://www.brookings.edu/research/multilateralism-what-policy-options-to-strengthen-international-cooperation/#footnote-2

Derviş, K., & Strauss, S. (2020, March 6). *What COVID-19 Means for International Cooperation*. Retrieved from Project Syndicate: https://www.project-syndicate.org/commentary/global-cooperation-can-prevent-next-pandemic-by-kemal-dervis-and-sebasti-n-strauss-2020-03?barrier=accesspaylog

Dodgson, R., Drager, N., & Lee, K. (2002). *Global Health Governance: A Conceptual Review*. Retrieved from Discussion Paper for the WHO, London School of Hygiene and Tropical Medicine.: https://www.researchgate.net/profile/Kelley_Lee/publication/242472817_Global_Health_Governance_A_CONCEPTUAL_REVIEW/links/0046352b0e94294750000000.pdf.

Dron, L., Kalatharan, V., Gupta, A., Haggstrom, J., Zariffa, N., Morris, A. D., . . . Park, J. (2022). Data capture and sharing in the COVID-19 pandemic: a cause for concern . *the Lancet*, 748-756.

Duball, J. (2020, April 28). *Centralized vs. decentralized: EU's contact tracing privacy conundrum*. Retrieved from Iapp: https://iapp.org/news/a/centralized-vs-decentralized-eus-contact-tracing-privacy-conundrum/

Dudovskiy, J. (2016). *The Ultimate Guide to Writing a Dissertation in Business Studies: A Step-by-Step Assistance.* UK: Scientific Research Publishing.

Duff, J., Liu, A., & Saavedra, J. (2021). A global public health convention for the 21st century. *Lancet*, 428-433.

eHealth Network. (2017). *Refined eHealth European Interoperability Framework.* Brussels: eHealth Network.

Elkin, P. L., Trusko, B. E., Koppel, R., Speroff, T., Mohrer, D., & Sakji, S. (2010). Secondary Use of Clinical Data. *Studies in Health Technology and Informatics. Seamless Care – Safe Care*, 14–29.

Europa Nu. (n.d.). *Volksgezondheidsbeleid*. Retrieved from Europa Nu: https://www.europa-nu.nl/id/vg9pl2emdcyl/volksgezondheidsbeleid

European Commission. (2010, December 16). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions 'Towards interoperability for European public services'.* Retrieved from European Interoperability Framework (EIF) for European public services: https://www.earonline.nl/images/earpub/9/96/EIF2.0.pdf

European Commission. (2017). *A European One Health Action Plan Against AMR.* https://health.ec.europa.eu/system/files/2020-01/amr_2017_action-plan_0.pdf: European Commission.

European Commission. (2017). *European reference networks*. Retrieved from European Commission: https://health.ec.europa.eu/european-reference-networks/overview_en

European Commission. (2018). *Communication on enabling the digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society*. Retrieved from European Commission: https://digital-strategy.ec.europa.eu/en/library/communication-enabling-digital-transformation-health-and-care-digital-single-market-empowering

European Commission. (2018). *Consultation: Transformation Health and Care in the Digital Single Market.* Luxembourg: Publication Office of the European Union.

European Commission. (2018). *European '1+ Million Genomes' Initiative*. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes

European Commission. (2018). *Turning FAIR into reality.* Brussels: European Commission.

European Commission. (2019). *Annex to the Commission Recommendation on a European Electronic Health Record Exchange Format.* Brussels: COM, 800 final.

European Commission. (2019). *Commission Recommendation (EU) 2019/243 of February 6, 2019 on a European Electronic Health Record Exchange Format.* Brussels: COM, 800 final.

European Commission. (2019). *Exchange of electronic health records across the EU.* Retrieved from Shaping Europe's digital future: https://digital-strategy.ec.europa.eu/en/policies/electronic-health-records

European Commission. (2020). *A pharmaceutical strategy for Europe*. Retrieved from European Commission: https://health.ec.europa.eu/medicinal-products/pharmaceutical-strategy-europe_en

European Commission. (2020). *White Paper on Artificial Intelligence.* Brussels: European Commission.

European Commission. (2021, September 16). *European Health Emergency preparedness and Response Authority (HERA): Getting ready for future health emergencies*. Retrieved from European Commission: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_4672

European Commission. (2022, October 24). *European Health Union: building a stronger EU health response*. Retrieved from European Commission: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6363

European Commission. (2022, May 3). *Questions and answers - EU Health: European Health Data Space (EHDS)*. Retrieved from European Commission: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_2712

European Commission. (n.d.). *A cancer plan for Europe*. Retrieved from European Commission: https://ec.europa.eu/info/strategy/priorities-2019-2024/promoting-our-european-way-life/european-health-union/cancer-plan-europe_en

European Commission. (n.d.). *eHealth Network*. Retrieved from European Commission: https://health.ec.europa.eu/ehealth-digital-health-and-care/eu-cooperation/ehealth-network_en

European Commission. (n.d.). *Electronic Exchange of Social Security Information (EESSI)*. Retrieved from European Commission:

https://ec.europa.eu/social/main.jsp?catId=1544&langId=en#:~:text=EESSI%20is%20an%20IT%20system,the%20EU%20rules%20on%20social

European Commission. (n.d.). *European Data Governance Act*. Retrieved from European Commission: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

European Commission. (n.d.). *European data strategy*. Retrieved from European Commission: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

European Commission. (n.d.). *European Health Union*. Retrieved from European Comission: https://ec.europa.eu/info/strategy/priorities-2019-2024/promoting-our-european-way-life/european-health-union_en

European Commission. (n.d.). *Europe's Digital Decade*. Retrieved from European Commission: https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade

European Commission. (n.d.). *New research projects on Coronavirus (August 2020)*. Retrieved from European comission: https://research-and-innovation.ec.europa.eu/research-area/health/coronavirus/coronavirus-projects_en

European Commission. (n.d.). *New research projects on Coronavirus (March 2020)*. Retrieved from European Commission: https://research-and-innovation.ec.europa.eu/research-area/health/coronavirus/coronavirus-projects_en

European Council. (2018). *The general data protection regulation*. Retrieved from European Council: https://www.consilium.europa.eu/en/policies/data-protection/data-protection-regulation/

European Observatory on Health Systems and Policies. (2021). *COVID-19 Health System Response Monitor (HSRM)*. Retrieved from European Observatory on Health Systems and Policies: https://eurohealthobservatory.who.int/monitors/hsrm/overview

Fears, R., Brand, H., Frackowiak, R., Pastoret, P., Souhami, R., & Thompson, B. (2014). Data protection regulation and the promotion of health research: getting the balance right. *QJM*, 3-5.

Feldman, K., Johnson, R. A., & Chawla, N. V. (2018). The State of Data in Healthcare: Path Towards Standardization. *Journal of Healthcare Informatics Research*, 248–271.

Ferguson, N., Cummings, D. A., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke , D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 448–452.

Ferroa, E., Loukisb, E. N., Charalabidisc, Y., & Osella, M. (2013). Policy making 2.0: From theory to practice. *Government Information Quarterly*, 359–368.

Fidler, D. P. (2003). SARS: Political Pathology of the First Post-Westphalian Plague. *Journal of Law, Medicine and Ethics*, 485–505.

Fidler, D. P. (2004). *SARS, Governance and the Globalization of Disease.* New York: Palgrave Macmillan.

Fidler, D. P. (2010). The Challenges of Global Health Governance. *Council on Foreign Relations*.

Fidler, D. P. (2015). The Ebola Outbreak and the Future of Global Health Security. *the Lancet*, 385: 9980.

Fidler, D. P., & Drager, N. (2009). Global Health and Foreign Policy: Strategic Opportunities and Challenges. *Background paper for the Secretary General's Report on Global Health and Foreign Policy*.

Frenk, J., & Moon, S. (2012). Governance challenges in global health. *New England Journal of Medicine*, 936-942.

Frenk, J., Godal, T., Gómez-Dantés, O., & Store, J. G. (2022). A reinvigorated multilateralism in health: lessons and innovations from the COVID-19 pandemic. *the Lancet*.

Frieden, J., Pettis, M., Rodrik, D., & Zedillo, E. (2012). *After the Fall: The Future of Global Cooperation.* Geneva: Geneva Reports on the World Economy 14.

Gaisa, M. (2020). COVID-19: a call for action, collaboration, reason, and unity. *Eur Urol Focus*, *6*, 1026-1027.

Gerring, J. (2007). *Case Study Research: Principles and Practices.* Cambridge University Press.

Gersovitz, M. (2014). Infectious disease externalities. *Encyclopedia of Health Economics*, 35-39.

Gill, S., & Benatar, S. R. (2016). Global Health Governance and Global Power. *International Journal of Health Services*, 346–365.

Given, L. M. (2008). *The sage encyclopedia of qualitative research methods.* Thousand Oaks, CA: Sage.

Goossens, H., Derde, L., Horby, P., & Bonten, M. (2022). The European clinical research response to optimise treatment of patients with COVID-19: lessons learned, future perspective, and recommendations. *Lancet Infect Dis.*, 153-158.

Gostin, L. O., & Friedman, E. A. (2015). A Retrospective and Prospective Analysis of the West African Ebola Virus Disease Epidemic: Robust National Health Systems at the Foundation and an Empowered WHO at the Apex. *the Lancet 385 (9980)*, 1902–1909.

Gostin, L. O., & Mok, E. A. (2009). *Grand challenges in global health governance.* Washington: The O'Neill Institute for National and Global Health Law.

Granda, P., & Blasczyk, E. (2016). *Data Harmonization Guidelines for Best Practice in Cross-Cultural Surveys.* Ann Arbor: Institute for Social Research, University of Michigan.

Guimón, J., & Narula, R. (2020). Ending the COVID-19 pandemic requires more international collaboration. *Research-Technology Management*, 38-41.

Head, B. W. (2008). Three Lenses of Evidence-Based Policy. *The Australian Journal of Public Administration*, 1-11.

Heijlen, R., & Crompvoets, J. (2021). Open health data: Mapping the ecosystem. *Open health data: Mapping the ecosystem. DIGITAL HEALTH*, https://doi.org/10.1177/20552076211050167.

Hess, C., & Ostrom, E. (2011). *Understanding Knowledge as a Commons. From Theory to Practice.* MIT press.

Heymann, D. L. (2020). Data sharing and outbreaks: best practice exemplified. *the Lancet*, 469-470.

Hoyt, K. (2006). Vaccine Innovation: Lessons from World War II. *Journal of Public Health Policy*, 38–57.

Iacob, N., & Simonelli, F. (2020). Towards a European Health Data Ecosystem. *European Journal of Risk Regulation*, 884-893.

Jamison, D., Frenk, J., & Knaul, F. (1998). International collective action in health: objectives, functions, and rationale. *the Lancet*, 514-517.

Janssen, M., & Helbig, N. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly*, 99-105.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2021). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 258-268.

Jit, M., Ananthakrishnan, A., McKee, M., Wouters, O. J., Beutels, P., & Teerawattananonch, Y. (2021). Multi-country collaboration in responding to global infectious disease threats: lessons for Europe from the COVID-19 pandemic. *The Lancet Regional Health - Europe*.

Kabir, Z. (2021). *Legal and ethical guidelines: a scoping exercise.* unCoVer.

Kahler, M. (1992). Multilateralism with Small and Large Numbers. *International Organization*, 681-708.

Kamradt-Scott, A. (2018). The Politics of Pandemic Influenza Preparedness. In C. McInnes, K. Lee, J. Youde, & (eds), *The Oxford Handbook of Global Health Politics* (pp. 530–550). New York: Oxford University Press.

Karpati, A., & Ellis, J. (2019). Measure, Inform, Build: Enabling Data-Driven Government Policymaking. In S. Macfarlane, C. AbouZahr, & (eds), *The Palgrave Handbook of Global Health Data Methods for Policy and Practice* (pp. 85–102). London: Palgrave Macmillan.

Katz, R., & Standley, C. (2019). Regional approaches for enhancing global health security. *BMC Public Health*, 473.

KC, S., Ananthakrishnan, A., Painter, C., Butani, D., & Teerawattananon, Y. (2021). Research collaboration is needed to inform quarantine policies for health-care workers. *The Lancet*, 2334.

Keohane, R. O. (1990). Multilateralism: An Agenda for Research. *International Journal*, 731-764.

Kickbusch, I., & Gleicher, D. (2012). *Governance for health in the 21st century*. World Health Organization.

Knottnerus, J. A. (2015). Research data as a global public good. *Journal of Clinical Epidemiology*, 270-271.

Krause, S. D. (2020, May 18). *Primary Research Versus Secondary Research*. Retrieved from Humanities, LibreText: https://human.libretexts.org/Bookshelves/Composition/Book%3A_The_Process_of_Research_Writing_(Krause)/02%3A_Thinking_Critically_About_Research/2.03%3A_Primary_Research_Versus_Secondary_Research

Lee, J., & Haupt, J. (2021). Scientific globalism during a global crisis: research collaboration and open access publications on COVID-19. *High Educ, 81*, 949-966.

Lee, K. (2003). *Globalization and Health: An Introduction.* Basingstoke, UK: Macmillan.

Lenz, R., Beyer, M., & Kuhn, K. A. (2007). Semantic integration in healthcare networks. *International Journal of Medical Informatics*, 201-207.

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, 446-454.

Liu, N., Chee, M. L., Niu, C., Pek, P., Siddiqui, F., Ansah, J., . . . Eng Hock Ong, M. (2020). Coronavirus disease 2019 (COVID-19): an evidence map of medical literature. *BMC Med Res Methodol, 20*, 177.

Lo, B. (2015, February 24). *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. Retrieved from JAMA network: https://jamanetwork.com/journals/jama/fullarticle/2091787

LoTempio, J., Spencer, D., Yarvitz, R., Vilain, A. D., Vilain, E., & Délot, E. (2020). We Can Do Better: Lessons Learned on Data Sharing in COVID-19 Pandemic Can Inform Future Outbreak Preparedness and Response. *Science and Diplomacy Vol. 9, No. 2,* https://www.sciencediplomacy.org/article/2020/we-can-do-betterlessons-learned-data-sharing-in-covid-19-pandemic-can-inform-future.

Marks, G., Hooghe, L., & Blank, K. (1996). European Integration from the 1980s: State centric v. Multi-level Governance. *Journal of Common Market Studies*, 341-378.

Martani, A., Geneviève, L., Pauli-Magnus, C., McLennan, S., & Elger, B. S. (2019). Regulating the Secondary Use of Data for Research: Arguments Against Genetic Exceptionalism. *Front. Genet.*, https://doi.org/10.3389/fgene.2019.01254.

Maxwell, L. (2020). *Report on the 1st round online survey and interviews related to perceived benefits and risks of sharing among cohort investigators.* ReCoDid.

McInnes, C. (2019). Global Health Governance. In C. McInnes, K. Lee, & J. Youde, *The Oxford Handbook of Global Health Politics.* Oxford Handbooks.

McLeod, & Heimbigner. (1985). A Federated Architecture for information management. *ACM Transactions on Information Systems, Volume 3, Issue 3*, 253–278.

Milieu and Time.lex. (2014). *Overview of the national laws on electronic health records in the EU Member States and their interaction with the provision of cross-border eHealth services Report.* Brussels: Milieu Ltd, Time.lex.

Mons, B., Schultes, E., Liu, F., & Jacobsen, A. (2020). The FAIR Principles: First Generation Implementation Choices and Challenges. *Data Intelligence*, 1-9.

Moorthy, V., Restrepo, A., Preziosi, M.-P., & Swaminathan, S. (2020). Data sharing for novel coronavirus (COVID-19). *Bull World Health Organ*, doi: 10.2471/BLT.20.251561.

Nature. (2021, April 6). *The world must learn from COVID before diving into a pandemic treaty*. Retrieved from Nature: https://www.nature.com/articles/d41586-021-00866-7

Neuman, L. W. (2014). *Social Research Methods: Qualitative and Quantitative Approaches.* Essex: Pearson Education Limited.

Oderkirk, J., Wenzl, M., & Slawomirsk, L. (2019). 6. Data without borders: Boosting knowledge and innovation. In O. H. Studies, *Health in the 21st Century Putting Data to Work for Stronger Health Systems.* Paris: OECD Publishing.

ORCHESTRA. (2022). *Statistics*. Retrieved from Orchestra: https://orchestra-cohort.eu/statistics/

Orchestra. (n.d.). *Periodic Reporting for period 1 - ORCHESTRA (Connecting European Cohorts to Increase Common and Effective Response to SARS-CoV-2 Pandemic: ORCHESTRA)*. Retrieved from Cordis: https://cordis.europa.eu/project/id/101016167/reporting

ORCHESTRA. (n.d.). *Work-packages*. Retrieved from ORCHESTRA: https://orchestra-cohort.eu/work-packages/

Orchestra-cohort. (n.d.). *Privacy Policy*. Retrieved from Orchestra Cohort: https://orchestra-cohort.eu/privacy-policy/

PAHO. (n.d.). *Electronic Health Records and Interoperability: Understanding two key concepts for a better public health response*. Retrieved from Pan American Health Organization: https://www3.paho.org/ish/images/docs/Factsheets-Digital_Health-EHR-Interoperability-English-Final-March_31.pdf?ua=1

Patil, R., & Gopal , I. (2022). The Importance of Data Sharing in Managing Public Health Crises. *Global Pandemic and Human Security*, 393–412.

Patton, M. (1980). *Qualitative research and evaluation methods.* Sage Publications Inc.

Pearson, H. (2021, May 12). How COVID broke the evidence pipeline. *Nature*, 182-185. Retrieved from Nature: https://www.nature.com/articles/d41586-021-01246-x

Peñalvo JL, M. E. (2021). Unravelling data for rapid evidence-based response to COVID-19: a summary of the unCoVer protocol. *BMJ Open 2021*, doi: 10.1136/bmjopen-2021-055630.

Peñalvo, J., Blázquez, A., Chausa, P., Menasalvas, E., & Gómez, E. (2021). *COVID-19 RWD processing: Report on the data processing activities related to data harvest, curation, normalization and validation of the COVID-19 data, as identified in WP1*. unCoVer.

Peñalvo, J., Blázquez, A., Chausa, P., Menasalvas, E., & Gómez, E. (2021). *Deliverable No. 3.1 COVID-19 RWD processing: Report on the data processing activities related to data harvest, curation, normalization and validation of the COVID-19 data, as identified in WP1*. unCoVer.

Perera, G., Holbrook, A., Thabane, L., & Foster, G. (2011). Views on health information sharing and privacy from primary care practices using electronic medical records. *International Journal of Medical Informatics*, 94-101.

Pope, C. (2003). Resisting Evidence: The Study of Evidence-Based Medicine as a Contemporary Social Movement. *Health*, 267–282.

ReCoDiD. (n.d.). *Objectives*. Retrieved from ReCoDiD: https://recodid.eu/projects/objectives/

Reddy, S., Mazhar, S., & Lencucha, R. (2018). The financial sustainability of the World Health Organization and the political economy of global health governance: a review of funding proposals. *Global Health*, 119.

Richter, A., Meissner, Y., Strangfeld, A., & Zink, A. (2016). Primary and secondary patient data in contrast: the use of observational studies like RABBIT. *Clin. Exp. Rheumatol*, 79–86.

Rinaldi, E., Stellmach, C., Moses Raj Rajkumar, N., Caroccia, N., Dellacasa, C., Giannella, M., . . . Thun , S. (2022). Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic. *npj digital magazine*.

Sá, C., & Sabzalieva, E. (2018). Chapter 9: Scientific nationalism in a globalizing world. In B. Cantwell, H. Coates, & R. King, *Handbook on the Politics of Higher Education* (pp. 149-166). Edward Elgar Publishing.

Safran, C., Bloomrosen, M., Hammond, W., Labkoff, S., Markel-Fox, S., Tang, P., & Detmer, D. (2007). Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc*, 1-9.

Salas-Vega, S., Haimann, A., & Mossia, E. (2015). Big Data and Health Care: Challenges and Opportunities for Coordinated Policy Development in the EU. *Health Systems & Reform*, 285-300.

Samadashvili, S. (2020). A brave, post–COVID-19 Europe. *European View*, 122-130.

Satterfield, B. A., Dikilitas, O., & Kullo, I. J. (2021). Leveraging the Electronic Health Record to Address the COVID-19 Pandemic. *Mayo Clinic*, 1592-1608.

Schimmelfennig. (2018). European integration (theory) in times of crisis. A comparison of the euro and Schengen crises. *Journal of European Public Policy*, pp. 969-989.

Schwalbe, N., Wahl, B., Song, J., & Lehtimaki, S. (2020). Data Sharing and Global Public Health: Defining What We Mean by Data. *Frontiers Digital Health*, doi: 10.3389/fdgth.2020.612339.

Sheikhtaheri, A., Tabatabaee Jabali, S., Bitaraf, E., TehraniYazdi, A., & Kabir, A. (2022). A near real-time electronic health record-based COVID-19 surveillance system: An experience from a developing country. *Health Inf Manag*, doi: 10.1177/18333583221104213.

Shiffman, J., Quissell, K., Schmitz, H. P., Pelletier, D. L., Smith, S. L., Berlan, D., . . . Walt, G. (2016). A Framework on the Emergence and Effectiveness of Global Health Networks. *Health Policy and Planning*, 3-16.

Sialm, A., & Sanchez-Niubo, A. (2020). *Strategy brief on harmonisation and integration methods, and analytic approaches to maximise the value of cohort data.* SYNCHROS.

Spicer, N., & Walsh, A. (2011). 10 Best Resources on … the Current Effects of Global Health Initiatives on Country Health Systems. *Health Policy and Planning* .

Synchros. (n.d.). *Synchros Project*. Retrieved from Synchros: https://synchros.eu/synchros-project/

Tacconelli, E., Gorska, A., Carrara, E., Davis, R. J., Bonten, M., Friedrich, A. W., & Glasner, C. (2022). Challenges of data sharing in European Covid-19 projects: A learning opportunity for advancing pandemic preparednes. *The Lancet Regional Health - Europe*, https://doi.org/10.1016/j.lanepe.2022.100467.

TEHDAS. (2022). *Report on secondary use of health data through European case studies.* TEHDAS Consortium Partners.

The Lancet. (2022). Lessons from COVID-19 are shaping the response to monkeypox outbreak. *The Lancet Regional Health − Europe*, https://doi.org/10.1016/j.lanepe.2022.100463.

Toshkov, D. (2016). *Research Design in Political Sience.* Palgrave.

Treaty of Amsterdam. (1997, November 10). *Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts.* Retrieved from EUR-lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:11997D/TXT

Turn, R., Shapiro, N. Z., & Juncosa, M. L. (1976). Privacy and Security in Centralized vs. Decentralized Databank Systems. *Policy Sciences*, 17-29.

unCover. (n.d.). *About us*. Retrieved from unCoVer: https://uncover-eu.net/mission/

van Panhuis, W. G., Proma, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., . . . Burke, D. S. (2014). A systematic review of barriers to data sharing in public health. *BMC Public Health*, 1144.

van Veenstra, A. F., & Kotterink, B. (2017). Data-Driven Policy Making: The Policy Lab Approach. *Electronic Participation*, 100–111.

Vesteghem, C., Brøndum, R., Sønderkær, M., Sommer, M., Schmitz, A., Støve Bødker, J., . . . Bøgsted, M. (2019). Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. *Briefings in Bioinformatics*, 936–945.

Vukovic, J., Ivankovic, D., Habl, C., & Dimnjakovic, J. (2022). Enablers and barriers to the secondary use of health data in Europe: general data protection regulation perspective. *Archives of Public Health*, 115.

Walport, M., & Brest, P. (2011). Sharing research data to improve public health. *the Lancet*, 537–539.

WHO. (2019, December 19). *Emergencies: International health regulations and emergency committees*. Retrieved from World Health Organization: https://www.who.int/news-room/questions-and-answers/item/emergencies-international-health-regulations-and-emergency-committees

WHO. (2020, April 27). *Archived: WHO Timeline - COVID-19*. Retrieved from WHO: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19

WHO. (2020, October 13). *Impact of COVID-19 on people's livelihoods, their health and our food systems*. Retrieved from World Health Organisation: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems#:~:text=The%20economic%20and%20social%20disruption,the%20end%20of%20the%20year.

Williams, O. D., & Rushton, S. (2011). *Private Actors in Global Health Governance.* Basingstoke, UK: Palgrave.

Xu, B., & Kraemer, M. U. (2020). Open access epidemiological data from the COVID-19 outbreak. *Lancet Infect Diseases*, doi: 10.1016/S1473-3099(20)30119-5.

Yin, R. K. (1994). *Case study research: design and methods.* University of Michigan: Sage Publications.

Yin, R. K. (2014). *Case Study Research: Design and Methods.* California: SAGE, Thousand Oaks.

Youde, J. (2012). *Global Health Governance.* Cambridge, UK: Polity.

Appendix A. List of documents

Doc number	Name	Link
1	European Commission New research projects on Coronavirus (March 2020)	Report downloaded at https://research-and-innovation.ec.europa.eu/research-area/health/coronavirus/coronavirus-projects_en
2	European Commission New research projects on Coronavirus (August 2020)	Report downloaded at https://research-and-innovation.ec.europa.eu/research-area/health/coronavirus/coronavirus-projects_en
3	CORDIS EU research results unCoVer	https://cordis.europa.eu/project/id/101016216
4	CORDIS EU research results ReCoDID	https://cordis.europa.eu/project/id/825746
5	CORDIS EU research results I-MOVE-COVID-19	https://cordis.europa.eu/project/id/101003673
6	CORDIS EU research results RECOVER	https://cordis.europa.eu/project/id/101003589
7	CORDIS EU research results HERoS	https://cordis.europa.eu/project/id/101003606
8	CORDIS EU research results SYNCHROS	https://cordis.europa.eu/project/id/825884
9	CORDIS EU research results ORCHESTRA	https://cordis.europa.eu/project/id/101016167
10	CORDIS EU research results EpiPose	https://cordis.europa.eu/project/id/101003688
11	CORDIS EU research results CORESMA	https://cordis.europa.eu/project/id/101003480
12	CORDIS EU research results EXSCALATE4CoV	https://cordis.europa.eu/project/id/101003551
13	unCoVer project objectives	https://uncover-eu.net/mission/

15	ReCoDID project objectives	https://recodid.eu/projects/objectives/
16	ORCHESTRA project objectives	https://orchestra-cohort.eu/work-packages/
17	SYNCHROS project objectives	https://synchros.eu/synchros-project/
18	A summary of the unCoVer protocol on behalf of the unCoVer network	https://bmjopen.bmj.com/content/11/11/e055630.citation-tools
19	unCoVer Report on the data processing activities related to data harvest, curation, normalization and validation of the COVID-19 data, as identified in WP1	Report downloaded at: https://cordis.europa.eu/project/id/101016216/results
20	SYNCHROS Strategy Brief on practical, legal and ethical in the optimisation of cohort data in Europe D3.3	Report downloaded at: https://cordis.europa.eu/project/id/825884/results
21	ORCHESTRA Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic	https://www.nature.com/articles/s41746-022-00620-x#Sec7
22	ReCoDID Statistical guidance and approaches for dealing with heterogeneity, missing data and measurement error in pooled cohort data sets	Report downloaded at: https://cordis.europa.eu/project/id/825746/results
23	ReCoDiD Report on the 1st round online survey and	Report downloaded at: https://cordis.europa.eu/project/id/825746/results

	interviews related to perceived benefits and risks of sharing among cohort investigators	
24	UnCoVer Legal and ethical guidelines: a scoping exercise	Report downloaded at: https://cordis.europa.eu/project/id/101016216/results
25	ORCHESTRA Privacy policy Storage of personal data	https://orchestra-cohort.eu/privacy-policy/
26	SYNCHROS Strategy brief on harmonisation and integration methods, and analytic approaches to maximise the value of cohort data	Report downloaded at: https://cordis.europa.eu/project/id/825884/results
27	ORCHESTRA Work performed from the beginning of the project to the end of the period covered by the report and main results achieved so far	Text can be found at 'reporting': https://cordis.europa.eu/project/id/101016167/reporting
28	ORCHESTRA Challenges of data sharing in European Covid-19 projects: A learning opportunity for advancing pandemic preparedness and response	https://www.thelancet.com/journals/lanepe/article/PIIS2666-7762(22)00163-6/fulltext#back-bib0038
29	ReCoDiD Optimised mapping of seroprevalence data according to time,	Downloaded at: https://cordis.europa.eu/project/id/825746/results

	targeted populations, and categorised assays	
30	ReCoDiD Reconciling measurements of individual cohort participants across heterogeneous data sets	Report downloaded at: https://cordis.europa.eu/project/id/825746/results
31	unCoVer Dissemination and Exploitation Plan	Report downloaded at: https://cordis.europa.eu/project/id/101016216/results
32	ReCoDiD	Text can be found under: "Work performed from the beginning of the project to the end of the period covered by the report and main results achieved so far" And "Progress beyond the state of the art and expected potential impact (including the socio-economic impact and the wider societal implications of the project so far)" https://cordis.europa.eu/project/id/825746/reporting
33	ORCHESTRA Statistics	https://orchestra-cohort.eu/statistics/