

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Unravelling data for rapid evidence-based response to COVID-19: A summary of the unCoVer protocol

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-055630
Article Type:	Protocol
Date Submitted by the Author:	18-Jul-2021
Complete List of Authors:	<p>Penalvo, Jose; Institute of Tropical Medicine, Unit of Non-Communicable Diseases</p> <p>Mertens, Elly; Institute of Tropical Medicine,</p> <p>Ademović, Enisa; University of Sarajevo, Department of Epidemiology and Biostatistics</p> <p>Akgun, Seval; Baskent University Faculty of Medicine, Public Health Department</p> <p>Baltazar, Ana Lúcia; Polytechnic Institute of Coimbra, Scientific-Pedagogical Unit of Dietetics and Nutrition</p> <p>Buonfrate, Dora; IRCCS Ospedale Sacro Cuore Don Calabria, Centre for Tropical Diseases</p> <p>Čoklo, Miran; Institut za antropologiju, Centre for Applied Bioanthropology</p> <p>Devleeschauwer, Brecht; Sciensano, Epidemiology and public health; Ghent University Faculty of Veterinary Medicine, Department of Veterinary Public Health and Food Safety</p> <p>Diaz Valencia, Paula Andrea; University of Antioquia, Epidemiology Group, National College of Public Health</p> <p>Fernandes, João C; Universidade Católica Portuguesa Escola Superior de Biotecnologia, Centro de Biotecnologia e Química Fina (CBQF)</p> <p>Gómez, Enrique Javier; Universidad Politécnica de Madrid, Center for Biomedical Technology; Universidad Politécnica de Madrid, Biomedical Engineering and Telemedicine Centre</p> <p>Hynds, Paul; Technological University Dublin, Environmental Sustainability & Health Institute</p> <p>Kabir, Zubair; University College Cork, School of Public Health</p> <p>Klein, Jörn; University of South-Eastern Norway,</p> <p>Kostoulas, P; University of Thessaly</p> <p>Llanos Jiménez, Lucía; Instituto de Investigacion Sanitaria de la Fundacion Jimenez Diaz, Clinical Research Unit</p> <p>Lotrean, Lucia Maria; Iuliu Hagieganu University of Medicine and Pharmacy Faculty of Medicine</p> <p>Majdan, Marek; Trnava University in Trnava,</p> <p>Menasalvas, Ernestina; Universidad Politecnica de Madrid Escuela Universitaria de Ingenieria Tecnica de Telecomunicacion, Biomedical Engineering and Telemedicine Centre</p> <p>Nguewa, paul; Universidad de Navarra</p> <p>Oh, In-Hwan; Kyung Hee University</p> <p>O'Sullivan, Georgie; University College Cork, School of Public Health</p> <p>Pereira, David M.; Universidade do Porto, 23 REQUIMTE/LAQV, Laboratório de Farmacognosia, Departamento de Química</p>

	<p>Reina Ortiz, Miguel; University of South Florida, Riva, Silvia; St Mary's University Twickenham, Department of Public Health</p> <p>Soriano, Gloria; Institute of Tropical Medicine, Department of Public Health</p> <p>Soriano, Joan; Universidad Autonoma de Madrid, Servicio de Neumología, Hospital Universitario de La Princesa</p> <p>Spilki, Fernando; FEEVALE University</p> <p>Tamang, Mary Elizabeth; Azienda ULSS 6 Euganea</p> <p>Trofor, Antigona; Universitatea de Medicina si Farmacie Gr T Popa Iasi, Pulmonary Diseases I</p> <p>Vaillant, Michel; Luxembourg Institute of Health, Competence Centre for Methodology and Statistics</p> <p>Van Ierssel, Sabrina; University Hospital Antwerp, Department of General Internal Medicine, Infectious diseases and Tropical Medicine</p> <p>Vukovic, Jakov; Croatian Institute of Public Health</p> <p>Castellano, Jose Maria; Hospital Universitario HM Montepíncipe, Cardiology Department</p>
Keywords:	COVID-19, STATISTICS & RESEARCH METHODS, Public health < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1
2
3
4
5
6 **Unravelling data for rapid evidence-based response to COVID-19: A summary of the unCoVer**
7
8 **protocol**
9

10
11 José L. Peñalvo^{1*}, Elly Mertens¹, Enisa Ademović², Seval Akgun³, Ana Lúcia Baltazar⁴, Dora Buonfrate⁵,
12
13 Miran Čoklo⁶, Brecht Devleeschauwer^{7,8}, Paula Andrea Diaz Valencia⁹, João C. Fernandes¹⁰, Enrique
14
15 Javier Gómez^{11,12,13}, Paul Hynds¹⁴, Zubair Kabir¹⁵, Jörn Klein¹⁶, Polychronis Kostoulas¹⁷, Lucía Llanos
16
17 Jiménez¹⁸, Lucia Maria Lotrean¹⁹, Marek Majdan²⁰, Ernestina Menasalvas¹², Paul Nguewa²¹, In-Hwan
18
19 Oh²², Georgie O'Sullivan¹⁵, David M. Pereira²³, Miguel Reina Ortiz²⁴, Silvia Riva²⁵, Gloria Soriano¹, Joan
20
21 B. Soriano^{26,27,28}, Fernando Spilki²⁹, Mary Elizabeth Tamang³⁰, Antigona Carmen Trofor^{31,32}, Michel
22
23 Vaillant³³, Sabrina Van Ierssel³⁴, Jakov Vukovic³⁵, and José M. Castellano^{36,37,38} **on behalf of the unCoVer**
24
25

26
27 **Network**
28
29
30
31

- 32
33 1 Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium.
34
35 2 Department of Epidemiology and Biostatistics, Faculty of Medicine, University of Sarajevo, Bosnia
36
37 and Herzegovina
38
39 3 Public Health Department, Baskent University School of Medicine, Turkey
40
41 4 Scientific-Pedagogical Unit of Dietetics and Nutrition, Coimbra Health School, Polytechnic
42
43 Institute of Coimbra, Portugal
44
45 5 Department of Infectious Tropical diseases and Microbiology, IRCCS Sacro Cuore Don Calabria
46
47 hospital, Verona, Italy
48
49 6 Centre for Applied Bioanthropology, Institute for Anthropological Research, Zagreb, Croatia
50
51 7 Department of Epidemiology and Public Health, Sciensano, Belgium
52
53 8 Department of Veterinary Public Health and Food Safety, Faculty of Veterinary Medicine, Ghent
54
55 University, Belgium
56
57
58
59
60

- 1
- 2
- 3 9 Epidemiology group, National College of Public Health, University of Antioquia, Medellin,
- 4
- 5 Colombia.
- 6
- 7
- 8 10 Centro de Biotecnologia e Química Fina (CBQF), Escola Superior de Biotecnologia da Universidade
- 9
- 10 Católica Portuguesa, Porto, Portugal
- 11
- 12 11 Center for Biomedical Technology, Universidad Politécnica de Madrid, Pozuelo de Alarcón, Spain.
- 13
- 14 12 Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación, Universidad Politécnica
- 15
- 16 de Madrid, Madrid, Spain.
- 17
- 18 13 Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina
- 19
- 20 (CIBER-BBN), Madrid, Spain.
- 21
- 22
- 23 14 Environmental Sustainability & Health Institute, Technological University Dublin, Ireland
- 24
- 25 15 School of Public Health, University College Cork, Ireland
- 26
- 27 16 University of South-Eastern Norway, Department of Nursing and Health Sciences, Norway
- 28
- 29 17 Faculty of Public Health, University of Thessaly, Greece
- 30
- 31 18 Clinical Research Unit, Instituto Investigación Sanitaria Fundación Jiménez Díaz (IIS-FJD, UAM),
- 32
- 33 Spain
- 34
- 35 19 Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania
- 36
- 37 20 Institute for Global Health and Epidemiology, Trnava University, Trnava, Slovakia
- 38
- 39 21 Instituto de Salud Tropical (ISTUN), Department of Microbiology and Parasitology, Navarra
- 40
- 41 Institute for Health Research (IdiSNA), University of Navarra, Pamplona, Spain.
- 42
- 43
- 44 22 Department of preventive Medicine, Kyung Hee University, Seoul, Republic of Korea
- 45
- 46 23 REQUIMTE/LAQV, Laboratório de Farmacognosia, Departamento de Química, Faculdade de
- 47
- 48 Farmácia, Universidade do Porto, Portugal
- 49
- 50
- 51 24 College of Public Health, University of South Florida, MI, USA
- 52
- 53 25 Department of Psychology and Pedagogic Science, St Mary's University Twickenham, London, UK
- 54
- 55 26 Servicio de Neumología, Hospital Universitario de La Princesa, Universidad Autónoma de Madrid,
- 56
- 57 Madrid, Spain
- 58
- 59
- 60

- 1
2
3 27 Centro de Investigación en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos
4 III (ISCIII), Madrid, Spain.
5
6
7 28 COVID-19 Clinical Management Team, WHO Health Emergency Programme, World Health
8 Organization HQ, Geneva, Switzerland
9
10
11
12 29 Feevale University, Novo Hamburgo, Brazil
13
14 30 Azienda ULSS6 Euganea, Padova, Italy
15
16 31 University of Medicine and Pharmacy Grigore T. Popa Iasi, Romania
17
18 32 Clinical Hospital of Pulmonary Diseases Iasi, Romania
19
20
21 33 Translational Medicine Operations Hub, Competence Center for Methodology and Statistics,
22 Luxembourg Institute of Health, Luxembourg
23
24
25 34 Department of General Internal Medicine, Infectious diseases and Tropical Medicine, Antwerp
26 University Hospital, Edegem, Belgium
27
28
29 35 Croatian Institute of Public Health, Zagreb, Croatia
30
31
32 36 Cardiology Department, Hospital Universitario HM Montepríncipe, HM Hospitales, Madrid, Spain.
33
34 37 Centro Nacional de Investigaciones Cardiovasculares (CNIC), Instituto de Salud Carlos III, Madrid,
35 Spain
36
37
38 38 Fundación de Investigación HM Hospitales, Madrid, Spain
39
40
41
42
43
44
45
46
47

48 *Corresponding Author: José L. Peñalvo, PhD, Department of Public Health, Institute of Tropical
49 Medicine, Antwerp (Belgium) Tel: +32(0)32476251 jpenalvo@itg.be | [@JosePenalvo](https://twitter.com/JosePenalvo)
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Introduction. unCoVer - Unravelling data for rapid evidence-based response to COVID-19 - is a HORIZON 2020-funded, network of 29 partners from 18 countries capable of collecting and utilizing real-world data (RWD) derived from the response and provision of care to COVID-19 patients by health systems across Europe and elsewhere. unCoVer aims to exploit the full potential of these information to rapidly address clinical and epidemiological research questions arising from the evolving pandemic.

Methods and analysis. From the onset of COVID-19 pandemic, partners are gathering RWD from electronic health records currently including information from over 22,000 COVID-19 hospitalized patients, and national surveillance and screening data, and registries with over 1,900,000 COVID-19 cases across Europe, with continuous updates. These heterogeneous datasets will be described, harmonised, and integrated into a multi-user data repository operated through Opal-DataSHIELD, an interoperable open-source server application. Federated data analyses, without sharing or disclosing any individual-level data, will be performed with the objective to reveal patients' baseline characteristics, biomarkers, determinants of COVID-19 prognosis, safety and effectiveness of treatments and potential strategies against COVID-19, as well as epidemiological patterns. These analyses will complement evidence from efficacy/safety clinical trials where vulnerable, more complex/heterogeneous populations, and those most at risk of severe COVID-19, are often excluded.

Ethics and dissemination. After strict ethical considerations, databases will be available through a federated data analysis platform allows processing available COVID-19 RWD without disclosing identification information to analysts and limiting output to data aggregates. Dissemination of unCoVer's activities will be related to the access and use of dissimilar RWD as well as the results generated by the pooled analyses. Dissemination will include training and educational activities, scientific publications, and conference communications.

Article Summary

- The unCoVer network includes a sizable number of partners that will exploit the full potential of real-world data derived from the provision of care to COVID-19 patients by health systems across Europe and elsewhere.
- With the availability of international harmonized RWD provided by the unCoVer network, a large amount of information on COVID-19 patients will be studied to better understand aspects of the pathophysiology, progression and treatment, and epidemiological patterns of this novel disease as well as to grasp the less understood, and potentially unearthed, risk factors associated with COVID-19 severity.
- The development and deployment of a federated data platform for combined analysis meet patients' data protection principles and comply with ethical standards, including GDPR and national data privacy legislation while allowing for advanced analytics.
- Continuous process evaluation will be carried out throughout the project life to identify limitations and barriers to the harmonised use of data and, simultaneously providing advice on improving data systems for rapid response to future public health crises.

INTRODUCTION

The outbreak of the coronavirus disease (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was declared a Public Health Emergency of International Concern by the World Health Organisation on 30 January 2020, and a pandemic, on 11 March 2020. Despite the deployment of public health measures such as restrictions of movements, gathering, and personal protection as well as massive vaccination campaigns, the coronavirus is still largely affecting mortality and morbidity worldwide, including persisting symptoms after the infection, what has now been termed post-COVID condition¹. Early epidemiological data of COVID-19 showed a higher risk of severe disease among older individuals, in particular those with chronic respiratory, cardiometabolic, and other chronic diseases including mental disorders and immunosuppressed individuals²⁻⁴. While the pathogenesis of certain chronic diseases predisposes to serious COVID-19 outcomes⁵, other factors such as common chronic medications, might also increase this risk due to interaction between SARS-CoV-2 infection, and the complex metabolic pathways^{6 7}. Also, common disease risk factors such as smoking and overweight/obesity, have been identified as key predictors of hospitalization and critical illness, even in young adults with no underlying conditions^{8 9}.

While the pandemic is evolving and countries are adapting their health systems to new phases of preventive measures, the research community is trying to fully elucidate the transmission and progression of the infection, as well as the most effective ways of treating and preventing new cases in preparation for any new waves, particularly due to new variants of SARS-CoV-2. The multi-dimensional and dynamic nature of the inter-related factors associated with individual responses to SARS-CoV-2 infection, and the diversity of long-term complications require a multidisciplinary research approach to unravel the natural history of this pandemic. Responding to the COVID-19 pandemic in real-time required a colossal effort from health systems worldwide, and across Europe where several countries have been severely affected. As a result, a wealth of data has been accumulated as part of

1
2
3 the health systems' efforts to fight COVID-19. These RWD reflect the impact of COVID-19 in patient's
4 health and characterize the protocols of health care in different health system settings. These close-
5 to-reality data allow for studies into patients' characteristics, determinants of disease prognosis, and
6 effectiveness of potential strategies against COVID-19 in real-world settings. They also complement
7 findings from ongoing efficacy/safety clinical trials where vulnerable/heterogeneous populations,
8 those most at risk of severe COVID-19, are often excluded. Harmonization of data from different
9 sources allows for comparison across health systems and improves patients' characterization using the
10 wider heterogeneity of the information. Still, to date, these RWD sources related to COVID-19 have
11 been exploited in a limited way and for specific questions, hence there is an untapped opportunity to
12 exploit the full potential of these data through identification, harmonisation, and big data analysis.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 **OBJECTIVES**

29
30
31
32 The unCoVer network aims to provide a research platform for the expert use of RWD, by bringing
33 together complementary data, medical and scientific expertise to address the still urgent questions
34 regarding determinants of COVID-19 prognosis to inform more effective medical and public health
35 strategies. Specifically, the network aims to facilitate access to otherwise scattered RWD sources, and
36 hereby provide opportunities for enhanced risk characterisation and robust risk prediction algorithms,
37 tackling the current pandemic, and eventually any future epidemics. This approach should lead to
38 control measures that will eventually relieve the pressure on the health systems, improve patient
39 prognosis especially among those more vulnerable (e.g., chronic patients, immunosuppressed
40 individuals, and population subgroups with limited access to health care, among others), and mitigate
41 the burden of COVID-19. The specific objectives of unCoVer are:
42
43
44
45
46
47
48
49
50
51
52
53

- 54
55
56
57 1. To bring together European and international expertise and data to synchronize collaborative
58 research efforts in addressing the ongoing COVID-19 pandemic in a common platform.
59
60

2. To continuously monitor, identify and facilitate the access and use of COVID-19-related RWD to fully exploit the potential of this routinely collected information, as a reflection of common medical practices.
3. To identify data gaps, and marginalized populations to proactively seek synergies with complementary existing and planned clinical databases related to COVID-19.
4. To provide a platform for the use of dissimilar data sources capable of streamlining ethical and legal aspects and anticipating the needs for data harmonization by innovative computational resources and integrated information for enhanced impact.
5. To bring together expertise on the use of advanced computational, epidemiological and biostatistical methods to handle heterogeneous, and multi-layered information to facilitate rapid queries and data outputs related to SARS-CoV-2 infection, underlying drivers of COVID-19 prognosis, the safety and effectiveness of treatments, and sequelae, as well as the impact of COVID-19 in health system resources.
6. To broadcast the use and results of the platform to attract new partners and to pursue complementarity with existing similar networks in Europe and internationally to save lives and optimize resources.

METHODS AND ANALYSIS

unCoVer is conceptualized as a functional network of partners, capable of harvesting and analysing RWD derived from the response and provision of care to patients by the health systems across Europe, and other countries such as Brazil and Colombia during the COVID-19 pandemic.

Setting

unCoVer comprises 29 partners from 25 institutions in the European Union and 4 non-EU partners representing 18 countries namely Belgium, Bosnia and Herzegovina, Brazil, Colombia, Croatia, Greece, Ireland, Italy, Luxembourg, Norway, Portugal, Romania, Slovakia, South Korea, Spain, Turkey, United

1
2
3 Kingdom, and the United States of America ([Figure 1](#)). Partners provide data mostly from front-line
4 hospitals but also national health agencies, registries, and investigator-lead observational studies, and
5 represent complementary scientific and medical fields, as well as expertise in research ethics, data
6 management and statistical modelling. unCoVer, thus, works as a functional network bridging clinical
7 expertise and data analytics, intending to exploit the full potential of the routine healthcare data
8 already collected from patients during the pandemic. The set-up of the network therefore relies on a
9 continuous iteration process between a) clinical partners, who will guide the development of research
10 questions needed to improve patient's care and inform public health strategies; and b) epidemiologists
11 and analytical experts, who will operationalize the research questions with advanced data processing,
12 analysis, and simulation tools capable of generating innovative solutions. The work of the network is
13 further supported by three external advisory boards ([Figure 2](#)) that provide expert counselling
14 concerning the relevance of the medical research and findings (External Advisory Board, EAB), data
15 protection (Data Protection and Ethics Advisory Board, DP-EAB), and stakeholder involvement (Societal
16 and Regulatory Advisory Board, SRAB).

36 **Databases**

37
38 unCoVer hosts longitudinal observational secondary-level data, which are largely collected for non-
39 research purposes, RWD, and refer to data generated during patient encounters with the health
40 system which have established information technology protocols and tools for retrieving and storing
41 information about the healthcare provided. To date, the unCoVer network incorporates 16 databases
42 of electronic medical records from 10 different countries, 6 national registries, 4 observational cohorts,
43 and 2 databases on population screening ([Table 1](#)). The data available to unCoVer has information on
44 hospitalization of COVID-19 patients with at least two-time points of data collection, at admission and
45 discharge. In addition to demographics, and clinical/epidemiological data, other data types such as bio-
46 specimens, imaging data, social network-/contact-tracing related data, movement-related data and
47 mental health data are also available but with limited access. Clinical/epidemiological data include case
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 report forms (CRF) at patient's admission (i.e. date of symptom onset and/or admission, signs and
4 symptoms at admission, laboratory results, pre-admission medication, comorbidities & risk factors),
5 during hospitalization (i.e. signs and symptoms, laboratory results, supportive treatment, admission to
6 ICU) and at discharge (i.e. date of discharge, outcome, as well as medication and complications)
7
8
9
10
11
12 [\(Supplementary Table 1\)](#).

13 14 15 16 **STRUCTURE AND ACTIVITIES**

17
18 Three blocks of activities, grouping work packages (WPs), were designed to build a functional network,
19 and aim to facilitate the flow of information for rapid assessment of research questions:
20
21
22

23
24
25 **Block 1: Definition, design, and data harvesting.** This first block forms the architectural foundation
26 and the core of the unCoVer network. In order to provide a comprehensive repository of available data,
27 WP1 'Data Identification' collects and catalogues all data in a standardised way, including a common
28 codebook that specifies the harmonized variables, with standardised variable names and data format
29 and labels, in preparation for data harmonization processes *i.e.* the key for the development of a
30 unified pool of data for analyses. In parallel, and acknowledging the sensitive nature of health data,
31 and personal information compliance with ethical and legal aspects, a checklist for assessing the risks
32 involved in data processing is implemented by WP2 'Ethics and Data Protection', with due
33 consideration to legal and regulatory issues concerning data protection, privacy and information
34 security. The checklist includes questions on the nature of the data (e.g. clinical data, hospital records
35 or publicly available data, personal data, data collected in vulnerable groups, availability of follow-up
36 data), informed consent (e.g. explicit consent or assent obtained), data protection (e.g. data protection
37 officer identified and data protection impact assessment completed), ethical approval (obtained or
38 pending), data privacy protection (e.g. anonymisation vs pseudonymisation, data minimisation), and
39 data transfer and use (data transfer agreement needed, name of the data controller, data processor,
40 joint data controller if applicable, and whether international transfer outside EU). Each data provider
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 within the network populates this checklist, which is reviewed on a case-by-case basis by the unCoVer
4 research ethics expert team together with the independent DP-EAB, and accordingly informed
5 decisions on risk mitigation are taken. In short, WP2 ensures that GDPR or equivalent guidelines are
6 adhered to during the data processing activity. Once a dataset is categorised as low risk, it will be
7 available to proceed with subsequent steps ([Supplemental Table 2](#)).
8
9
10
11
12
13
14
15

16 For the secure multi-party computation of unCoVer data, WP3 'Data Harmonization', developed an
17 infrastructure based on Opal 4.1 (OBiBa suite, Maelstrom Research, Montreal, Canada) to facilitate
18 interoperability of the data, including data management, harmonization and dissemination in a
19 secured environment ¹⁰. The Opal server application provides the necessary key features for data
20 encryption and decryption managed through Public Key Infrastructure as well as participant identifiers
21 management and user authentication/authorization for access via a rights and roles management with
22 username/password. Steps to achieve data harmonisation and secured data sharing and use are: 1)
23 Set-up the Opal server for each data provider and import relevant datasets, 2) Configure a harmonized
24 data description in each Opal server, and 3) Run distributed queries on harmonized datasets through
25 the DataSHIELD application that enables individual-level data analyses across multiple Opal servers
26 without sharing and disclosing any individual-level data ([Figure 3](#)). Thus, by using computational power
27 and standardising dissimilar information, while complying with ethical and legal requirements, a data
28 repository of anonymized and harmonized COVID-19 RWD will be made available for secured data
29 analyses.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 **Block 2: Analytical development, data use, and demonstration.** The broad range of medical, public
51 health, and research expertise available within the unCoVer network is at the heart of the WP4 'Data
52 analysis and Outcomes' dedicated to unCoVer findings from the data acquired to maximize their use
53 in informing COVID-19 response. Activities of this block aim to facilitate and streamline rapid response
54 to identified research gaps using the unCoVer infrastructure developed in Block 1. With the availability
55
56
57
58
59
60

1
2
3 of cross-national harmonized RWD within the unCoVer network, a large amount of hospitalised and
4 discharged COVID-19 patients will be studied in depth, together with complementary epidemiological
5 data, to understand the pathophysiology, progression, treatment, (long-term) complications, and (less
6 frequent) risk factors for early prevention of this novel disease as well as to grasp the cross-national
7 heterogeneity of the COVID-19 burden. To this end, an analytical toolbox, including both traditional
8 statistical methods and machine learning techniques and a Bayesian estimation framework, will be
9 developed for identifying relationships between early clinical and diagnostic profiles and the future
10 course of the infection, and for a detailed clinical and epidemiological characterization of COVID-19
11 patients, being able to generate patients' risk classification and risk prediction. The application of this
12 toolbox in real patient data, available within the unCoVer network, will then allow uncovering real-
13 world insights that would support policies and protocols for optimization of health resources of the
14 hospital and critical care. Therefore, the activities of this block will be related to the iterative trial use
15 of the Opal server and the toolbox by end-users (*i.e.* data analysts) as well as the lessons learned and
16 potentially preparing actions for the sustainability of the unCoVer platform, including the repository
17 and toolbox.

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39 **Block 3: Project management, communication, and exploitation.** The outputs of the network
40 including the scientific and technological knowledge and outcomes, provided by the previous two
41 blocks, are streamlined through scientific publications, training and educational activities, organisation
42 and participation in events, among others, steered by the WP5 'Communication, Exploitation &
43 Dissemination'. This last block is also dedicated to maintaining the functional network both internally
44 and externally, and with special attention to the management of intellectual property by utilizing best
45 practices in project coordination, as outlined in WP6 'Coordination'. Concerning this, the unCoVer
46 organization structure works under a Consortium Agreement signed by all partners and includes the
47 following key bodies within the consortium and management structure ([Figure 2](#)). A steering
48 committee formed by the principal investigators of the 29 partner institutions at the decision-making
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 level. A General Assembly (GA), involving WP leaders manages the network, and coordinates WP and
4
5 tasks leaders, assisted also by a Management Support Team (MST) to reinforce partners'
6
7 representation. An internal Exploitation and Dissemination Committee (EDC) also collaborates in the
8
9 overall management of the GA. Finally, the project coordinators communicate with the sponsor, and
10
11 facilitate crosstalk between the network and the external advisory boards.
12
13
14
15

16 **ETHICS AND DISSEMINATION**

17
18
19
20
21 Ethical aspects are of utmost importance in unCoVer. Upon the project start date, the roles, and
22
23 responsibilities of the independent DP-EAB were described, including the selection procedure of the
24
25 board members and its final composition, and mandate. A 'scoping exercise' was also conducted across
26
27 the network to ensure that all partners are aware of the common obligations in terms of data
28
29 processing activities using health or health-related data according to European and international
30
31 guidelines. Moreover, to be compliant with the General Data Protection Regulation (GDPR) and meet
32
33 the ethics requirements, the unCoVer network will follow the data processing steps represented in
34
35 [Figure 4](#), in the following sequence:
36
37
38
39
40

- 41 1. The unCoVer master checklist of data processing activities in network partners' is circulated within
42
43 the network to be completed by the data providers, and data providers are required to provide the
44
45 supporting documentation of each indicator of this list, such as informed consents, ethical
46
47 approvals and Data Protection Impact Assessment (DPIA). This information is processed by the
48
49 research ethics team, responsible for categorizing the datasets into three different categories: low,
50
51 medium, or high data privacy risk.
52
53
- 54 2. Datasets categorized as "low-risk" will be available to proceed with the harmonization process
55
56 and, therefore, Opal-DataSHIELD servers will be installed. Within Opal, the patient identifiers will
57
58 be separated from the patient study data by employing two databases: (1) the ID database that
59
60

1
2
3 stores the patient identifiers accessible by the data provider only, and (2) the study database that
4 stores pseudo-anonymised patient's data to be used for data analyses accessible, through code
5 only, by data analysts. The "medium-" and "high-risk" datasets will be subject to further review
6 and requirements before harmonization processes.
7
8
9

- 10
11
12 3. Finally, the installation of the servers will allow the consortium to analyse the available RWD
13 through an anonymisation layer to answer the pre-identified research questions. The system also
14 facilitates the definition of analytical projects and the specific databases and/or variables that will
15 be used for a specific project. As a rule, all output of data analytics will be restricted to the
16 presentation of data aggregates or to line listing deprived of personal identifiers so that the identity
17 of the study patient cannot be deduced (no backward identification).
18
19
20
21
22
23
24
25
26
27

28 To maximise the unCoVer network's output, dissemination, and exploitation strategies, as planned by
29 the EDC and advised the SRAB, i.e., a non-executive consulting substructure composed of several key
30 stakeholders from the regulatory, governance, civil society level, and patient's public initiatives, will be
31 segmented according to the network block activities, the potential users and the most adequate
32 channels of dissemination and interaction with potential users. The website (uncover-eu.net/), social
33 media accounts (Twitter @uncoverEU, LinkedIn, YouTube), and the project newsletter will be the
34 channels to reach all partners and stakeholders of unCoVer, both devoted to providing regular updates
35 on project activities and announcing upcoming milestones and events. The website will serve as a
36 repository of the project goals and activities and deliverables in an easy-to-understand language, as
37 well as publications, lectures and expert documents hosted for access by the partners or stakeholders.
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 **Patient and Public Involvement**

53
54 The unCoVer network has been designed to facilitate interactions and enhanced outreach to COVID-
55 19 stakeholders included in external advisory boards, as well as a prominent work package on
56 dissemination activities, that include but are not limited to:
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- *Scientific community.* To contribute to the body of knowledge in the field, two types of publications in peer-reviewed journals under Open Access schemes are foreseen: the unCoVer-network publications, *i.e.* for implementation and/or application of the unCoVer platform as a whole, and the unCoVer-partners publication, *i.e.* for specific collaborations between two or more unCoVer partners. In addition to both types of scientific publications, jointly organised workshops, virtual trainings, and virtual conferences, will be instrumental channels for the dissemination of the unCoVer activities and results to the scientific community. In such a yearly organised workshop, the application of the unCoVer repository and toolbox will be presented and expert feedback will be sought for further improvement. These activities of dissemination to the scientific community will result in overall awareness and international recognition of the unCoVer network, simultaneously strengthening the visibility and competitiveness of the institutions involved as centres of excellence.
 - *European platforms and data-driven initiatives.* Cooperation with other European projects dedicated to COVID-19 data sharing such Orchestra (orchestra-cohort.eu), Synchros (synchros.eu), Dragon (imi.europa.eu/projects-results/project-factsheets/dragon) RecodID (recodid.eu), and EC-COVID-19 Data Platform (covid19dataportal.org), as well as large networks such as the European Burden of Disease Network (burden-eu.net), and initiatives on data sharing infrastructures such as the Population Health Information Research Infrastructure (phiri.eu), will be established for the co-organisation of dissemination events along with seeking alignment and synergies to avoid duplication of efforts.
 - *Policy makers.* The accumulated prior experience and contact networks in the regulatory, policymaking framework of several members of the unCoVer network will be used to ensure that the work and output created can reach regulatory entities and policymakers, thus contributing to the impacts of the project in the decision-making process. This appears relevant as the UnCoVer network is willing to merge different sources of medical data with social, economic, mental, and

1
2
3 geographical data with the potential to identify highly tailored profiles of risk for community
4 prevention programmes and educational goals in different countries

- 5
6
7 - *Engagement with key stakeholders.* Results from the project will be further disseminated through
8 involvement with societal, regulatory, and administrative partners, as ensured by the external
9 advisory boards to the network, with the goal to warrant the project's impact according to the
10 stakeholder's needs and expectations.
11
12 - *General public.* To maximize awareness of unCoVer among the general public, the project activities
13 and milestones will be broadcasted via social media.
14
15 - *Patient associations and clinicians.* Given the foreseen impact of unCoVer's output, patient
16 associations and clinicians are identified as end-users that benefit from the data-oriented results,
17 and subsequently, these can be translated to them via seminars, lecture, and infographics as made
18 available on the project website, among others.
19
20
21
22
23
24
25
26
27
28
29
30
31

32 CONCLUSION

33 unCoVer brings a new global perspective of disease management in which protection of health requires
34 not only looking at regional or national aspects but also the development of data networks; A
35 worldwide outbreak needs to be tackled by a worldwide research network. The impact of the COVID-
36 19 pandemic in low-resource settings, and other vulnerable populations is a clear reminder of the
37 importance of equity and solidarity, also in medical research. As we move forward, it is vital that we
38 explore the drivers of the pandemic, learn from the global response, and become more prepared for
39 the similar future situations through global networks aiming at transcontinental data integration and
40 analysis. The impact of the COVID-19 pandemic in low-resource settings, and other vulnerable
41 populations is a clear reminder of the importance of equity and solidarity, also in medical research.
42 unCoVer wants to underpin some of the major global issues arising from the pandemic, including the
43 geography of health, the role of social and lifestyle determinants, and other bases of vulnerability as
44 prognostic factors of severity, and effectiveness of treatments and guidelines recommendations,
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

leveraging the use of big data within the remit of GDPR or its equivalent, for maximal benefit to population's health.

For peer review only

AUTHORS CONTRIBUTION

José L Peñalvo, Elly Mertens, Gloria Soriano, Zubair Kabir, Ernestina Menasalva, José M Castellano, David M Pereira, Silvia Riva, and Joan B Soriano wrote the manuscript. All authors have critically reviewed the manuscript, and approved the final version.

FUNDING STATEMENT

This project is funded by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 101016216.

DATA AVAILABILITY

unCoVer data will be available via a tiered-access web-based interface and complying with data protection requirements.

COMPETING INTERESTS

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

The unCoVer network: Institute of Tropical Medicine and Antwerp University Hospital, Belgium (José L. Peñalvo, Elly Mertens, Marianne Van der Sande, Patrick Soentjens, Diana Sagastume, James Cottam, Gloria Soriano, Sabrina Van Ierssel, and Hanne Van Tiggelen); Fundación Investigación HM Hospitales, Spain (José M. Castellano, Paula Villares, José Barberán, Mercedes Villareal, Justo Menéndez, Nerea Ruiz del Árbol, and Alberto Estirado); Universidad Politécnica de Madrid, Spain (Ernestina Menasalvas, Enrique Javier Gómez, Alberto Blázquez Herranz, David Fernandez Lobón, and Paloma Chausa); Universidad de Navarra, Spain (Paul Nguewa); Universidade do Porto, Portugal (David M. Pereira, and Morteza Hosseini); Technological University Dublin, Ireland (Paul Hynds, John Kelleher, and Elizabeth Hunter); University College Cork, Ireland (Zubair Kabir, Ella Arensman, Brendan Palmer, and Georgie

1
2
3 O'Sullivan); Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca, Romania (Milena Man,
4 Lucia Maria Lotrean, Mihaela Lupse, and Mira Florea); Universitatea De Medicina Si Farmacie Grigore
5 T Popa Din Iasi, Romania (Antigona Carmen Trofor, Andrei Tudor Cernomaz, Radu Adrian Crisan-Dabija,
6 and Cristina Grigorescu); Luxembourg Institute of Health (Michel Vaillant, and Guy Fagherazzi);
7 Universidade Católica Portuguesa, Portugal (João C Fernandes, and João Silva); Trnava University,
8 Slovakia (Marek Majdan, Daria Rabarova, Adriana Krsakova, Jaroslava Brnova, Janka Prnova, Jaroslav
9 Slany and Dominika Plancikova); Instituto Politécnico de Coimbra, Portugal (Ana Lucía Baltazar);
10 Hospital Universitario de La Princesa, Spain (Joan B Soriano, Julio Ancochea, Nisa Boukichou
11 Abdelkader, Adrián Peláez, Elena Ávalos, and Gorane Iturricastillo); Instituto Investigación Sanitaria
12 Fundación Jiménez Díaz, Spain (Lucía Llanos, Miguel Górgolas, Olga Sánchez-Pernaute, Arnoldo Santos
13 Oviedo, Sergio Luis Lima, Antonio Herrero, and Pablo Minguez), Panepistimio Thessalias, Greece
14 (Polychronis Kostoulas, Olympia Lioupi, Eleftherios Meletis, Konstantinos Pateras, and Costas Tsiamis);
15 Universitetet I Sorost-Norge, Norway (Jörn Klein, and Mustafa Asfari); Istituto Don Calabria, Italy (Dora
16 Buonfrate, Tamara Ursini, and Nicoletta De Santis); Sciensano, Belgium (Brecht Devleesschauwer,
17 Petronille Bogaert, Koen Blot, Miriam Saso, and Mathil Vandromme); Croatian Institute of Public
18 Health, Croatia (Jakov Vukovic, Ivan Pristas, Tamara Poljicanin, Jelena Dimnjakovic, Marko Brkic, and
19 Marija Svajda); Institut Za Antropologiju, Croatia (Miran Čoklo, Saša Missoni, Jelena Šarac, Natalija
20 Novokmet, Luka Bočkor, Ivan Dolanc, Antonija Jonjić, and Iva Šunić); Baskent Universitesi Vakfi, Turkey
21 (Seval Akgun, Tugba Gürgen Erdogan, Süleyman Çetinküner, Cenk Belibağlı, Kübra Demir, Mustafa
22 Görür, Turgut Bulut, and K.R. Nayar); St Mary's University Twickenham, United Kingdom (Silvia Riva);
23 Azienda ULSS6 Euganea, Italy (Mary Elizabeth Tamang, Carlo Giordani, and Petra Golin); Korea
24 University, South Korea (In-Hwan OH, and Seok Jun Yoon); University of South Florida, United States
25 (Miguel Reina Ortiz); Universidad de Antioquía, Colombia (Paula Andrea Diaz Valencia, Lina Ruíz, Juan
26 Pablo Pérez Bedoya, Oscar Ignacio Mendoza, Camilo Hincapie, Boris Rodriguez, and Noël Barengo);
27 Feevale University, Associacao Pro Ensino Superior Em Novo Hamburgo, Brazil (Fernando Spilki, Juliane
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

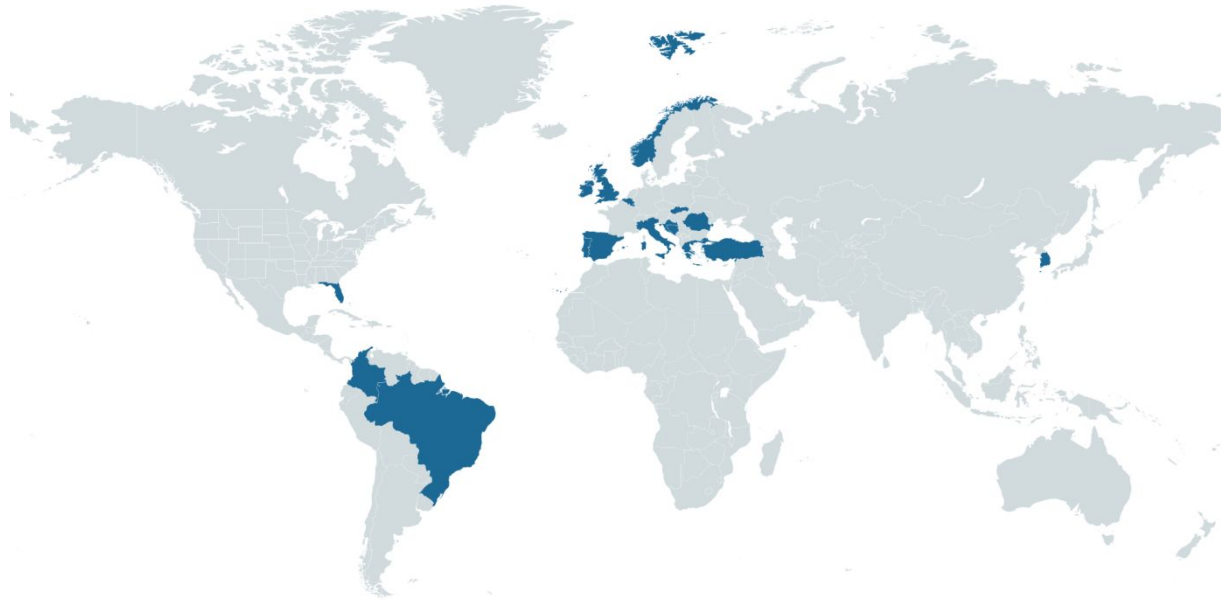
1
2
3 Deise Fleck, and Matheus Nunes Weber); Univerzitet U Sarajevu, Bosnia and Herzegovina (Enisa
4 Ademović, Lejla Burnazović-Ristić, Semra Čavaljuga, Džan Ahmed Jesenković, and Lejla Džananović).
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

REFERENCES

1. World Health Organization (WHO). Expanding our understanding of post COVID-19 condition: report of a WHO webinar, 9 February 2021., 2021.
2. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* 2020;395(10229):1054-62. doi: 10.1016/S0140-6736(20)30566-3
3. Shi Y, Yu X, Zhao H, et al. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. *Critical Care* 2020;24(1):108. doi: 10.1186/s13054-020-2833-7
4. World Health Organization (WHO). COVID-19 Clinical management: living guidance, 2021.
5. Jin Y, Ji W, Yang H, et al. Endothelial activation and dysfunction in COVID-19: from basic mechanisms to potential therapeutic approaches. *Signal Transduction and Targeted Therapy* 2020;5(1):293. doi: 10.1038/s41392-020-00454-7
6. Wan Y, Shang J, Graham R, et al. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology* 2020;94(7):e00127-20. doi: 10.1128/jvi.00127-20
7. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med* 2020;8(4):e21. doi: 10.1016/S2213-2600(20)30116-8 [published Online First: 2020/03/15]
8. Kassir R. Risk of COVID-19 for patients with obesity. *Obesity Reviews*;21(6):e13034. doi: 10.1111/obr.13034
9. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020;360:m1966. doi: 10.1101/2020.04.08.20057794
10. Doiron D, Marcon Y, Fortier I, et al. Software Application Profile: Opal and Mica: open-source software solutions for epidemiological data management, harmonization and dissemination. *Int J Epidemiol* 2017;46(5):1372-78. doi: 10.1093/ije/dyx180 [published Online First: 2017/10/13]

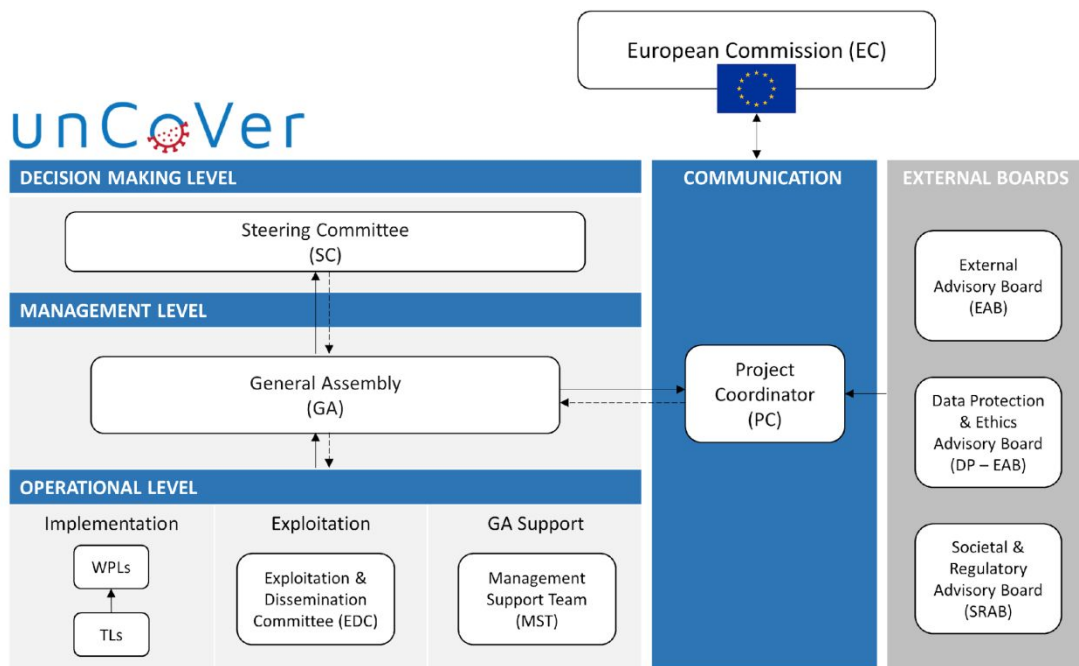
Figure 1. Geographical distribution of the unCoVer network (Belgium, Bosnia and Herzegovina, Brazil, Colombia, Croatia, Greece, Ireland, Italy, Luxembourg, Norway, Portugal, Romania, Slovakia, South Korea, Spain, Turkey, United Kingdom, and the United States of America)



For review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2. Management structure of the unCoVer network



Review only

Figure 3. Secure multi-party computation of unCoVer data based on Opal/DataShield infrastructure

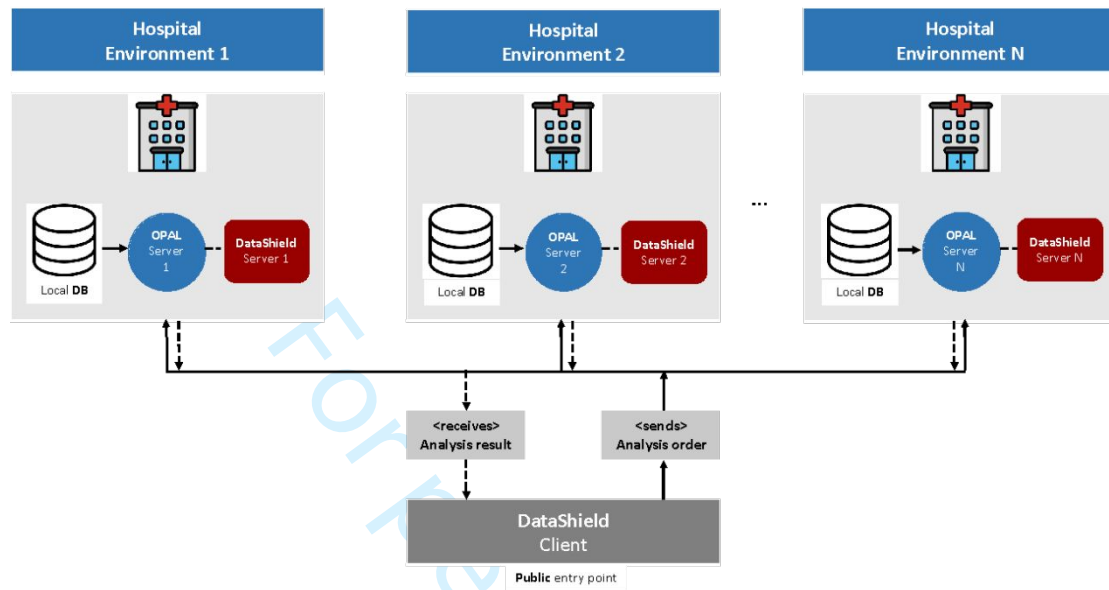


Figure 4. Data management process of the COVID-19 related data available within the unCoVer network

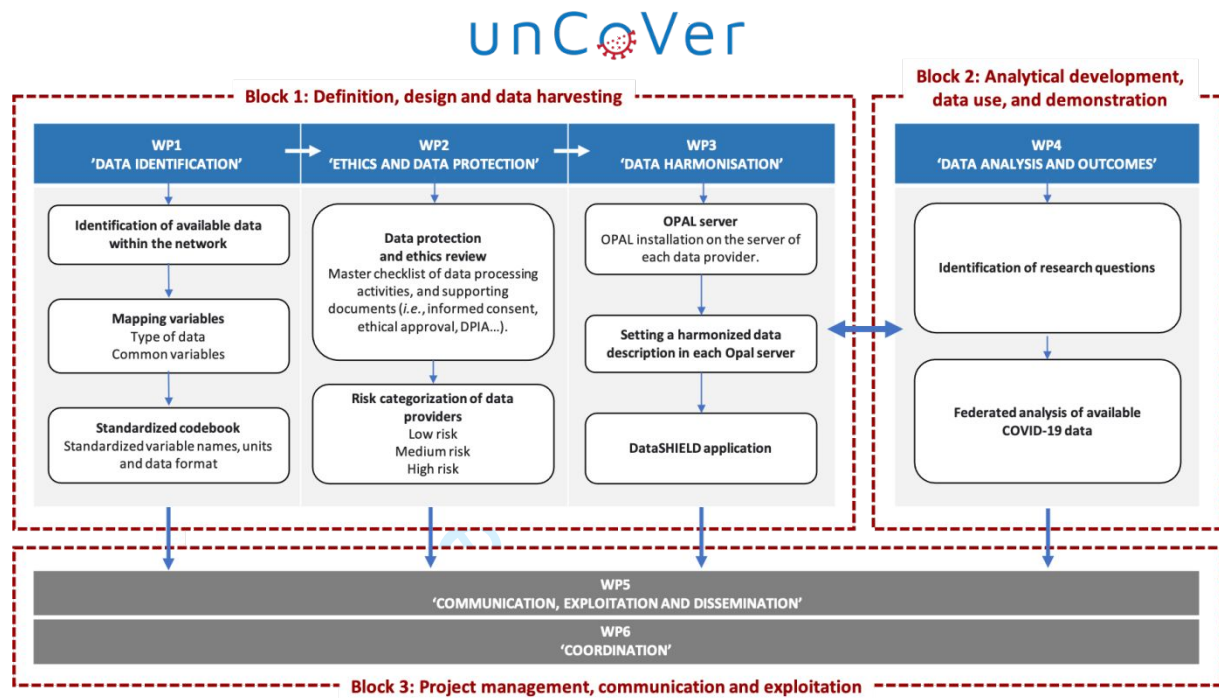


Table 1. Overview of the COVID-19 real-world data among the unCoVer network

Country	Institution ¹	Num. centres ²	Start date	End date	Num. patients ²	Study Population ³	Num. time points	Variable type																
								Demographics	Clinical/Epidemiology	Human OMICs	Pathogen OMICs	Bio-specimens	Imaging	Social network	Movement-related	Mental health	Economic	Diet	Screening					
Patient data from Electronic Medical Records																								
BA	UNSA	10	17/03/2020	20/05/2021	2,000	1, 2, 3	2	X	X					X										
BE	ITM-UZA	1	22/04/2020	Ongoing	187	1	2	X	X					X										
BR	ASPEUR	40	26/03/2020	Ongoing	2,000	1, 3	15	X	X		X											X		
ES	FIHM	12	01/03/2020	15/02/2021	4,480	1	2	X	X					X										
ES	HULPr	1	01/01/2020	29/09/2020	2,217	1	2	X	X															
ES	IIS-FJD	1	07/03/2020	31/05/2020	1,861	1	2	X	X					X										
ES	UNAV	2	01/04/2020	01/01/2021	100	1	1	X	X															
HR	INANTRO	2	01/08/2020	Ongoing	200	1	1	X	X		X			X										
IT	IRCCS - DB1	1	03/03/2020	Ongoing	200	1	1	X	X					X										
IT	IRCCS - DB4	1	01/03/2020	09/05/2020	355	4	1	X	X															
IT	ULSS6	4	01/03/2020	Ongoing	1,000	1	3	X	X					X										
RO	UMF Cluj	1			100	1	4	X	X					X										
RO	UMF IASI	1	01/03/2020	Ongoing	150	1	2	X	X					X										
SK	TU	1	01/10/2020	Ongoing	800	1	1	X	X															
TR	BU - DB1	6	01/03/2020	Ongoing	7,000	1	4	X	X															
Public Health Surveillance data and Registers																								
SK	TU	240	01/03/2020	Ongoing	776,000	2	1	X	X													X		
BE	Sciensano	98	14/03/2020	Ongoing	50,000	1	2	X	X					X										
CO	UdeA	1,314	14/03/2020	Ongoing	3,997,021	2	1	X	X															
HR	CIPH	NA	25/02/2020	Ongoing	220,000	2	1	X	X															
IE	TUDublin	48	29/02/2020	30/11/2020	74,000	2	1	X	X						X									
NO	USN	40	31/03/2020	Ongoing	2,313	1	10	X	X												X			
PT	UPORTO	NA	03/03/2020	Ongoing	830,000	2	1	X	X															

bmjopen-2021-055630 on 18 November 2021. Downloaded from <http://bmjopen.bmj.com/> on April 18, 2024 by guest. Protected by copyright.

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Country	Institution ¹	Num. centres ²	Start date	End date	Num. patients ²	Study Population ³	Num. time points	Variable type													
								Demographics	Clinical/Epidemiology	Human OMICs	Pathogen OMICs	Bio-specimens	Imaging	Social network	Movement-related	Mental health	Economic	Diet	Screening		
Observational Research data																					
IT	IRCCS - DB3	1	22/04/2020	25/05/2020	1,515	5	1	X	X												
LU	LIH	1	15/04/2020	01/04/2021	900	2	5	X	X			X									
PT	IPC	1	01/03/2020	01/07/2020	550	5	1	X												X	
UK	SMUC	1	01/07/2020	01/01/2021	200	2	1	X						X							
Screening data																					
IT	IRCCS - DB2	1	31/03/2020	12/05/2020	1,635	6	1	X	X												X
TR	BU – DB2	1	01/03/2020	Ongoing	18,000	5,7	1	X													X

¹Institution’s acronyms: ITM-UZA, Institute of Tropical Medicine and Antwerp University Hospital; FIHM, Fundación Investigación HM Hospitales; UNAV, Universidad de Navarra; UPORTO, Universidade Do Porto; TUDublin, Technological University Dublin; UMF Cluj, Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca; UMF IASI, Universitatea De Medicina Si Farmacie Grigore T Popa Din Iași; LIH, Luxembourg Institute of Health; TU, Trnavska Univerzita V Trnave; IPC, Instituto Politécnico de Coimbra; HULPr, Hospital Universitario de La Princesa; IIS-FJD, Instituto Investigación Sanitaria Fundación Jiménez Díaz; USN, Universitetet I Sorost-Norge; IRCCS, Istituto Don Calabria; CIPH, Croatian Institute of Public Health; INANTRO, Institut Za Antropologiju; BU, Baskent Universitesi Vakfi; SMUC, St Mary’s University Twickenham; ULSS6, Azienda ULSS6 Euganea; UdeA, Universidad de Antioquia; ASPEUR, Associacao Pro Ensino Superior Em Novo Hamburgo; UNSA, Univerzitet U Sarajevu. ² Number of centres providing information, and number of patients/individuals in the datasets could change due to the continuous update of the information; NA, Not available at the moment. ³ Study population refers to, 1) COVID-19 hospitalized patients, 2) COVID-19 cases, 3) COVID-19 patients attending primary care services, 4) COVID-19 patients attending emergency services, 5) general population, 6) health care workers, and 7) tourists.

Supplemental Table 1. Clinical and epidemiological data available in unCoVer network databases containing hospitalized patient data.

Country	Institution ¹	Patient admission CFR						Daily CFR during hospitalization						Patient outcome CRF			
		Demographics	Onset and/or Admission date	Signs and symptoms	Laboratory results	Pre-admission medication	Co-morbidities & risk factors	At the ward			In ICU			Medication treatment	Complications	Discharge date	Outcome status
								Signs and symptoms	Laboratory results	Supportive Treatment	ICU admission date	Signs and symptoms	Laboratory results				
Patient data from Electronic Medical Records																	
BA	UNSA	X	X	X	X	X	X			X	X	X	X	X		X	X
BE	ITM-UZA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
BR	ASPEUR	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ES	FIHM	X	X	X ²	X ²	X ²	X ²	X	X ²	X	X	X	X ²	X	X	X	X
ES	HULPr	X	X	X ²		X	X		X ²	X ³	X	X	X	X	X ⁴	X	X
ES	IIS-FJD	X	X	X	X	X	X		X	X ⁴	X	X	X	X	X	X	X
ES	UNAV	X	X		X												
HR	INANTRO	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
IT	IRCCS - DB1	X	X	X	X	X	X	X	X	X ³	X	X	X	X	X	X	X
IT	IRCCS - DB4	X	X	X	X		X ³										
IT	ULSS6	X	X	X ²	X	X	X	X	X	X	X	X	X	X	X	X	X
RO	UMFCluj	X	X	X			X	X	X	X	X	X	X	X	X ⁴	X	X
RO	UMFIASI	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
SK	TU	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
TR	BU – DB1	X	X	X			X						X				
TR	BU – DB2	X	X	X		X	X	X	X	X	X	X	X		X	X	X
Patient data from Public Health Surveillance and National Registers																	
SK	TU	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
BE	Sciensano	X	X	X	X		X		X		X	X	X	X	X	X	X
CO	UdeA	X													X	X	
HR	CIPH	X	X				X								X	X	
IE	TUDublin	X	X				X										
NO	USN	X	X	X	X		X						X	X	X	X	X

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

PT	UPORTO	X	X				X ²			X ³	X						X	
Patient data from Observational Research																		
IT	IRCCS - DB2	X		X			X ³											
IT	IRCCS - DB3	X		X			X ³											
LU	LIH	X	X	X	X	X	X	X	X	X	X ³	X	X	X	X	X	X	X

¹Acronyms: ITM-UZA, Institute of Tropical Medicine and Antwerp University Hospital; FIHM, Fundación Investigación HM Hospitales; UNAV, Universidad de Navarra; UPORTO, Universidade Do Porto; TUDublin, Technological University Dublin; UMF Cluj, Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca; UMF IASI, Universitatea De Medicina Si Farmacie Grigore T Popa Din Ia i; LIH, Luxembourg Institute of Health; TU, Trnavska Univerzita V Trnave; IPC, Instituto Politécnico de Coimbra; HULPr, Hospital Universitario de La Princesa; IIS-FJD, Instituto Investigación Sanitaria Fundación Jiménez Díaz; USN, Universitetet I Sorost-Norge; IRCCS, Istituto Don Calabria; CIPH, Croatian Institute of Public Health; INANTRO, Institut Za Antropologiju; BU, Baskent Universitesi Vakfi; SMUC, St Mary’s University Twickenham; ULSS6, Azienda ULSS6 Euganea; UdeA, Universidad de Antioquia; ASPEUR, Associacao Pro Ensino Superior Em Novo Hamburgo; UNSA, Univerzitet U Sarajevu.²Data needs to be extracted from text variables; ³Coded as yes/no without any further specifications; ⁴Length of stay in days instead of exact date of admission and/or discharge from hospital and/or ICU

Peer review only

Supplemental Table 2. Data protection Risk Assessment checklist

<p>> Was personal data/sensitive collected? Yes: What permissions and safeguarding are in place for the individual (evaluated through question below)? No: What data was collected and what are the risks?</p>		
<p>> Was Explicit Consent or Assent received (if necessary)? Yes: Method of requesting individual consent or assuming assent must be provided to unCoVer No: Should it have been collected?</p>		
<p>> Is Ethical Approval necessary? Yes: Ethical Approval must be provided to unCoVer No: If Ethical Approval is not necessary, personal data must not be identifiable (GDPR does not apply) and the data should be 'public' to some extent</p>		
<p>> Has the partner received Ethical Approval? Yes: Does unCoVer need any additional information following Ethical Approval? No: Hold until Ethical Approval is provided and assess</p>		
<p>> Anonymisation / Pseudonymization: Is the process used to anonymise or pseudonymise the data to be shared with uncover known? Who undertook this process? Yes: Transparent and appropriate methods used were clear and acceptable. Parties with data access are identified. Still ensure no personal data is passed onto unCoVer No: if these processes were undertaken, is there follow-up required?</p>		
<p>> Is there a letter from the Data Protection Officer (DPO) confirming the standardisation of Ethical Approval? Yes: Letter has been signed and returned by local DPO No: Follow-up with partner is needed</p>		
Low risk	Medium risk (a combination of some of the below)	High risk (a combination of some of the below)
Ethical approval received or not applicable	Ethical approval received or pending	Ethical approval not received
Explicit consent & assent received or not applicable	Explicit consent & assent received or waiting confirmation	Unclear details of consent / assent
Anonymisation and/or pseudonymization transparent & appropriate	Waiting confirmation of anonymization/ pseudonymization process(es)	Anonymization or pseudonymization incomplete. Or process unexplained
Data transfer agreement in place if necessary	Data transfer pending or in place if necessary	Non-EU partners without appropriate Data transfer agreements
DPO letter received	DPO letter pending/received	DPO letter pending

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Abstract, page 3
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported			Pag 5, 6
Objectives	3	State specific objectives, including any prespecified hypotheses			Pag 6, 7
Methods					
Study Design	4	Present key elements of study design early in the paper			Pag 7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection			Pag 7, 8, and Figure 1
Participants	6	(a) <i>Cohort study</i> - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.	Pag 7, 8, 9, and Table 1 and Supplemental Table 1

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18		<p><i>Case-control study</i> - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</p> <p><i>Cross-sectional study</i> - Give the eligibility criteria, and the sources and methods of selection of participants</p> <p><i>(b) Cohort study</i> - For matched studies, give matching criteria and number of exposed and unexposed</p> <p><i>Case-control study</i> - For matched studies, give matching criteria and the number of controls per case</p>		<p>RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.</p> <p>RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.</p>	
19 20 21 22 23 24 25	Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Table 1 and Supplemental Table 1
26 27 28 29 30 31 32 33	Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group		Table 1 and Supplemental Table 1
34 35	Bias	9	Describe any efforts to address potential sources of bias		Pag 9, 10, 11
36 37	Study size	10	Explain how the study size was arrived at		Pag 9, 10, 11
38 39 40 41 42	Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why		Pag 9, 10, 11, and Figure 3

43
44
45
46
47

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> - If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> - If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses		Pag 9, 10, 11, and Figure 3
20 21 22 23 24 25 26 27 28	Data access and cleaning methods		..	RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Figure 3
29 30 31 32 33 34 35	Linkage		..	RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Figure 3
Results					
36 37 38 39 40 41 42 43 44	Participants	13	(a) Report the numbers of individuals at each stage of the study (<i>e.g.</i> , numbers potentially eligible, examined for eligibility, confirmed eligible, included in the	RECORD 13.1: Describe in detail the selection of the persons included in the study (<i>i.e.</i> , study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in	NA (protocol paper not including results)

1		study, completing follow-up, and analysed)		the text and/or by means of the study flow diagram.	
2		(b) Give reasons for non-			
3		participation at each stage.			
4		(c) Consider use of a flow diagram			
5					
6	Descriptive data	14	(a) Give characteristics of study participants (<i>e.g.</i> , demographic, clinical, social) and information on exposures and potential confounders		NA (protocol paper not including results)
7			(b) Indicate the number of participants with missing data for each variable of interest		
8			(c) <i>Cohort study</i> - summarise follow-up time (<i>e.g.</i> , average and total amount)		
9					
10					
11					
12					
13					
14					
15					
16					
17					
18	Outcome data	15	<i>Cohort study</i> - Report numbers of outcome events or summary measures over time		NA (protocol paper not including results)
19			<i>Case-control study</i> - Report numbers in each exposure category, or summary measures of exposure		
20			<i>Cross-sectional study</i> - Report numbers of outcome events or summary measures		
21					
22					
23					
24					
25					
26					
27					
28					
29					
30	Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (<i>e.g.</i> , 95% confidence interval). Make clear which confounders were adjusted for and why they were included		NA (protocol paper not including results)
31			(b) Report category boundaries when continuous variables were categorized		
32			(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1 2 3 4	Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			NA (protocol paper not including results)
5	Discussion					
6 7 8	Key results	18	Summarise key results with reference to study objectives			NA (protocol paper not including results)
9 10 11 12 13 14 15 16 17	Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Pag 15
18 19 20 21 22 23 24	Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			Pag 15
25 26 27	Generalisability	21	Discuss the generalisability (external validity) of the study results			Pag 15
28	Other Information					
29 30 31 32 33 34	Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based			Pag 16
35 36 37 38 39 40	Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Pag 17

*Reference: Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

1 *Checklist is protected under Creative Commons Attribution ([CC BY](#)) license.
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

For peer review only

BMJ Open

Unravelling data for rapid evidence-based response to COVID-19: A summary of the unCoVer protocol

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-055630.R1
Article Type:	Protocol
Date Submitted by the Author:	20-Sep-2021
Complete List of Authors:	<p>Penalvo, Jose; Institute of Tropical Medicine, Unit of Non-Communicable Diseases Mertens, Elly; Institute of Tropical Medicine, Ademović, Enisa; University of Sarajevo, Department of Epidemiology and Biostatistics Akgun, Seval; Baskent University Faculty of Medicine, Public Health Department Baltazar, Ana Lúcia; Polytechnic Institute of Coimbra, Scientific-Pedagogical Unit of Dietetics and Nutrition Buonfrate, Dora; IRCCS Ospedale Sacro Cuore Don Calabria, Centre for Tropical Diseases Čoklo, Miran; Institut za antropologiju, Centre for Applied Bioanthropology Devleesschauwer, Brecht; Sciensano, Epidemiology and public health; Ghent University Faculty of Veterinary Medicine, Department of Veterinary Public Health and Food Safety Diaz Valencia, Paula Andrea; University of Antioquia, Epidemiology Group, National College of Public Health Fernandes, João C; Universidade Católica Portuguesa Escola Superior de Biotecnologia, Centro de Biotecnologia e Química Fina (CBQF) Gómez, Enrique Javier; Universidad Politécnica de Madrid, Center for Biomedical Technology; Universidad Politécnica de Madrid, Biomedical Engineering and Telemedicine Centre Hynds, Paul; Technological University Dublin, Environmental Sustainability & Health Institute Kabir, Zubair; University College Cork, School of Public Health Klein, Jörn; University of South-Eastern Norway, Kostoulas, P; University of Thessaly Llanos Jiménez, Lucía; Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz, Clinical Research Unit Lotrean, Lucia Maria; Iuliu Hagieganu University of Medicine and Pharmacy Faculty of Medicine Majdan, Marek; Trnava University in Trnava, Menasalvas, Ernestina; Universidad Politecnica de Madrid Escuela Universitaria de Ingeniería Técnica de Telecomunicación, Biomedical Engineering and Telemedicine Centre Nguewa, Paul; Universidad de Navarra Oh, In-Hwan; Kyung Hee University O'Sullivan, Georgie; University College Cork, School of Public Health Pereira, David M.; Universidade do Porto, 23 REQUIMTE/LAQV, Laboratório de Farmacognosia, Departamento de Química</p>

	<p>Reina Ortiz, Miguel; University of South Florida, Riva, Silvia; St Mary's University Twickenham, Department of Public Health</p> <p>Soriano, Gloria; Institute of Tropical Medicine, Department of Public Health</p> <p>Soriano, Joan; Universidad Autonoma de Madrid, Servicio de Neumología, Hospital Universitario de La Princesa</p> <p>Spilki, Fernando; FEEVALE University</p> <p>Tamang, Mary Elizabeth; Azienda ULSS 6 Euganea</p> <p>Trofor, Antigona; Universitatea de Medicina si Farmacie Gr T Popa Iasi, Pulmonary Diseases I</p> <p>Vaillant, Michel; Luxembourg Institute of Health, Competence Centre for Methodology and Statistics</p> <p>Van Ierssel, Sabrina; University Hospital Antwerp, Department of General Internal Medicine, Infectious diseases and Tropical Medicine</p> <p>Vukovic, Jakov; Croatian Institute of Public Health</p> <p>Castellano, Jose Maria; Hospital Universitario HM Montepíncipe, Cardiology Department</p>
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Ethics, Epidemiology, Infectious diseases, Public health
Keywords:	COVID-19, STATISTICS & RESEARCH METHODS, Public health < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1
2
3 1 **Unravelling data for rapid evidence-based response to COVID-19: A summary of the unCoVer**
4
5 2 **protocol**
6
7

8 3 José L. Peñalvo^{1*}, Elly Mertens¹, Enisa Ademović², Seval Akgun³, Ana Lúcia Baltazar⁴, Dora Buonfrate⁵,
9
10 4 Miran Čoklo⁶, Brecht Devleeschauwer^{7,8}, Paula Andrea Diaz Valencia⁹, João C. Fernandes¹⁰, Enrique
11
12 5 Javier Gómez^{11,12,13}, Paul Hynds¹⁴, Zubair Kabir¹⁵, Jörn Klein¹⁶, Polychronis Kostoulas¹⁷, Lucía Llanos
13
14 6 Jiménez¹⁸, Lucia Maria Lotrean¹⁹, Marek Majdan²⁰, Ernestina Menasalvas¹², Paul Nguewa²¹, In-Hwan
15
16 7 Oh²², Georgie O'Sullivan¹⁵, David M. Pereira²³, Miguel Reina Ortiz²⁴, Silvia Riva²⁵, Gloria Soriano¹, Joan
17
18 8 B. Soriano^{26,27,28}, Fernando Spilki²⁹, Mary Elizabeth Tamang³⁰, Antigona Carmen Trofor^{31,32}, Michel
19
20 9 Vaillant³³, Sabrina Van Ierssel³⁴, Jakov Vukovic³⁵, and José M. Castellano^{36,37,38} **on behalf of the unCoVer**
21
22 10 **network**
23
24
25
26
27
28

- 29 12 1 Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium.
30
31 13 2 Department of Epidemiology and Biostatistics, Faculty of Medicine, University of Sarajevo, Bosnia
32
33 14 and Herzegovina
34
35 15 3 Public Health Department, Baskent University School of Medicine, Turkey
36
37 16 4 Scientific-Pedagogical Unit of Dietetics and Nutrition, Coimbra Health School, Polytechnic
38
39 17 Institute of Coimbra, Portugal
40
41 18 5 Department of Infectious Tropical diseases and Microbiology, IRCCS Sacro Cuore Don Calabria
42
43 19 hospital, Verona, Italy
44
45 20 6 Centre for Applied Bioanthropology, Institute for Anthropological Research, Zagreb, Croatia
46
47 21 7 Department of Epidemiology and Public Health, Sciensano, Belgium
48
49 22 8 Department of Veterinary Public Health and Food Safety, Faculty of Veterinary Medicine, Ghent
50
51 23 University, Belgium
52
53 24 9 Epidemiology group, National College of Public Health, University of Antioquia, Medellin,
54
55 25 Colombia.
56
57
58
59
60

- 1
2
3 26 10 Centro de Biotecnologia e Química Fina (CBQF), Escola Superior de Biotecnologia da Universidade
4
5 27 Católica Portuguesa, Porto, Portugal
6
7 28 11 Center for Biomedical Technology, Universidad Politécnica de Madrid, Pozuelo de Alarcón, Spain.
8
9 29 12 Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación, Universidad Politécnica
10
11 30 de Madrid, Madrid, Spain.
12
13 31 13 Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina
14
15 32 (CIBER-BBN), Madrid, Spain.
16
17 33 14 Environmental Sustainability & Health Institute, Technological University Dublin, Ireland
18
19 34 15 School of Public Health, University College Cork, Ireland
20
21 35 16 University of South-Eastern Norway, Department of Nursing and Health Sciences, Norway
22
23 36 17 Faculty of Public Health, University of Thessaly, Greece
24
25 37 18 Clinical Research Unit, Instituto Investigación Sanitaria Fundación Jiménez Díaz (IIS-FJD, UAM),
26
27 38 Spain
28
29 39 19 Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania
30
31 40 20 Institute for Global Health and Epidemiology, Trnava University, Trnava, Slovakia
32
33 41 21 Instituto de Salud Tropical (ISTUN), Department of Microbiology and Parasitology, Navarra
34
35 42 Institute for Health Research (IdiSNA), University of Navarra, Pamplona, Spain.
36
37 43 22 Department of preventive Medicine, Kyung Hee University, Seoul, Republic of Korea
38
39 44 23 REQUIMTE/LAQV, Laboratório de Farmacognosia, Departamento de Química, Faculdade de
40
41 45 Farmácia, Universidade do Porto, Portugal
42
43 46 24 College of Public Health, University of South Florida, MI, USA
44
45 47 25 Department of Psychology and Pedagogic Science, St Mary's University Twickenham, London, UK
46
47 48 26 Servicio de Neumología, Hospital Universitario de La Princesa, Universidad Autónoma de Madrid,
48
49 49 Madrid, Spain
50
51 50 27 Centro de Investigación en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos
52
53 51 III (ISCIII), Madrid, Spain.
54
55
56
57
58
59
60

- 1
2
3 52 28 COVID-19 Clinical Management Team, WHO Health Emergency Programme, World Health
4
5 53 Organization HQ, Geneva, Switzerland
6
7 54 29 Feevale University, Novo Hamburgo, Brazil
8
9
10 55 30 Azienda ULSS6 Euganea, Padova, Italy
11
12 56 31 University of Medicine and Pharmacy Grigore T. Popa Iasi, Romania
13
14 57 32 Clinical Hospital of Pulmonary Diseases Iasi, Romania
15
16 58 33 Translational Medicine Operations Hub, Competence Center for Methodology and Statistics,
17
18 59 Luxembourg Institute of Health, Luxembourg
19
20
21 60 34 Department of General Internal Medicine, Infectious diseases and Tropical Medicine, Antwerp
22
23 61 University Hospital, Edegem, Belgium
24
25 62 35 Croatian Institute of Public Health, Zagreb, Croatia
26
27 63 36 Cardiology Department, Hospital Universitario HM Montepríncipe, HM Hospitales, Madrid, Spain.
28
29 64 37 Centro Nacional de Investigaciones Cardiovasculares (CNIC), Instituto de Salud Carlos III, Madrid,
30
31 65 Spain
32
33
34 66 38 Fundación de Investigación HM Hospitales, Madrid, Spain
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

70 *Corresponding Author: José L. Peñalvo, PhD, Department of Public Health, Institute of Tropical
71 Medicine, Antwerp (Belgium) Tel: +32(0)32476251 jpenalvo@itg.be | [@JosePenalvo](https://twitter.com/JosePenalvo)

1
2
3 76 **ABSTRACT**
4

5 77 **Introduction.** unCoVer - Unravelling data for rapid evidence-based response to COVID-19 - is a
6
7 78 HORIZON 2020-funded, network of 29 partners from 18 countries capable of collecting and utilizing
8
9 79 real-world data (RWD) derived from the response and provision of care to COVID-19 patients by health
10
11
12 80 systems across Europe and elsewhere. unCoVer aims to exploit the full potential of these information
13
14 81 to rapidly address clinical and epidemiological research questions arising from the evolving pandemic.

15
16 82 **Methods and analysis.** From the onset of COVID-19 pandemic, partners are gathering RWD from
17
18 83 electronic health records currently including information from over 22,000 COVID-19 hospitalized
19
20 84 patients, and national surveillance and screening data, and registries with over 1,900,000 COVID-19
21
22 85 cases across Europe, with continuous updates. These heterogeneous datasets will be described,
23
24 86 harmonised, and integrated into a multi-user data repository operated through Opal-DataSHIELD, an
25
26 87 interoperable open-source server application. Federated data analyses, without sharing or disclosing
27
28 88 any individual-level data, will be performed with the objective to reveal patients' baseline
29
30 89 characteristics, biomarkers, determinants of COVID-19 prognosis, safety and effectiveness of
31
32 90 treatments and potential strategies against COVID-19, as well as epidemiological patterns. These
33
34 91 analyses will complement evidence from efficacy/safety clinical trials where vulnerable, more
35
36 92 complex/heterogeneous populations, and those most at risk of severe COVID-19, are often excluded.

37
38 93 **Ethics and dissemination.** After strict ethical considerations, databases will be available through a
39
40 94 federated data analysis platform allows processing available COVID-19 RWD without disclosing
41
42 95 identification information to analysts and limiting output to data aggregates. Dissemination of unCoVer's
43
44 96 activities will be related to the access and use of dissimilar RWD as well as the results generated by the
45
46 97 pooled analyses. Dissemination will include training and educational activities, scientific publications,
47
48 98 and conference communications.
49

50
51
52
53
54 99
55
56
57
58
59
60

1
2
3 100 **Article Summary**
4
5 101
6
7 102 - The unCoVer network includes a sizable number of partners that will exploit the full potential
8
9 103 of real-world data derived from the provision of care to COVID-19 patients by health systems
10
11 104 across Europe and elsewhere.
12
13
14 105 - With the availability of international harmonized RWD provided by the unCoVer network, a
15
16 106 large amount of information on COVID-19 patients will be studied to better understand aspects
17
18 107 of the pathophysiology, progression and treatment, and epidemiological patterns of this novel
19
20 108 disease as well as to grasp the less understood, and potentially unearthed, risk factors
21
22 109 associated with COVID-19 severity.
23
24
25 110 - The development and deployment of a federated data platform for combined analysis meet
26
27 111 patients' data protection principles and comply with ethical standards, including GDPR and
28
29 112 national data privacy legislation while allowing for advanced analytics.
30
31
32 113 - Continuous process evaluation will be carried out throughout the project life to identify
33
34 114 limitations and barriers to the harmonised use of data and, simultaneously providing advice
35
36 115 on improving data systems for rapid response to future public health crises.
37
38
39 116
40
41 117
42
43 118
44
45 119
46
47 120
48
49
50
51
52
53
54
55
56
57
58
59
60

121 INTRODUCTION

122

123 The outbreak of the coronavirus disease (COVID-19), caused by severe acute respiratory syndrome
124 coronavirus 2 (SARS-CoV-2), was declared a Public Health Emergency of International Concern by the
125 World Health Organisation on 30 January 2020, and a pandemic, on 11 March 2020. Despite the
126 deployment of public health measures such as restrictions of movements, gathering, and personal
127 protection as well as massive vaccination campaigns, the coronavirus is still largely affecting mortality
128 and morbidity worldwide, including persisting symptoms after the infection, what has now been
129 termed post-COVID condition¹. Early epidemiological data of COVID-19 showed a higher risk of severe
130 disease among older individuals, in particular those with chronic respiratory, cardiometabolic, and
131 other chronic diseases including mental disorders and immunosuppressed individuals²⁻⁴. While the
132 pathogenesis of certain chronic diseases predisposes to serious COVID-19 outcomes⁵, other factors
133 such as common chronic medications, might also increase this risk due to interaction between SARS-
134 CoV-2 infection, and the complex metabolic pathways^{6 7}. Also, common disease risk factors such as
135 smoking and overweight/obesity, have been identified as key predictors of hospitalization and critical
136 illness, even in young adults with no underlying conditions^{8 9}.

137

138 While the pandemic is evolving and countries are adapting their health systems to new phases of
139 preventive measures, the research community is trying to fully elucidate the transmission and
140 progression of the infection, as well as the most effective ways of treating and preventing new cases
141 in preparation for any new waves, particularly due to new variants of SARS-CoV-2. The multi-
142 dimensional and dynamic nature of the inter-related factors associated with individual responses to
143 SARS-CoV-2 infection, and the diversity of long-term complications require a multidisciplinary research
144 approach to unravel the natural history of this pandemic. Responding to the COVID-19 pandemic in
145 real-time required a colossal effort from health systems worldwide, and across Europe where several
146 countries have been severely affected. As a result, a wealth of data has been accumulated as part of

1
2
3 147 the health systems' efforts to fight COVID-19. These RWD reflect the impact of COVID-19 in patient's
4
5 148 health and characterize the protocols of health care in different health system settings. These close-
6
7 149 to-reality data allow for studies into patients' characteristics, determinants of disease prognosis, and
8
9 150 effectiveness of potential strategies against COVID-19 in real-world settings. They also complement
10
11 151 findings from ongoing efficacy/safety clinical trials where vulnerable/heterogeneous populations,
12
13 152 those most at risk of severe COVID-19, are often excluded. Harmonization of data from different
14
15 153 sources allows for comparison across health systems and improves patients' characterization using the
16
17 154 wider heterogeneity of the information. Still, to date, these RWD sources related to COVID-19 have
18
19 155 been exploited in a limited way and for specific questions, hence there is an untapped opportunity to
20
21 156 exploit the full potential of these data through identification, harmonisation, and big data analysis.
22
23
24
25
26

157

158 **OBJECTIVES**

159

160 The unCoVer network aims to provide a research platform for the expert use of RWD, by bringing
161 together complementary data, medical and scientific expertise to address the still urgent questions
162 regarding determinants of COVID-19 prognosis to inform more effective medical and public health
163 strategies. Specifically, the network aims to facilitate access to otherwise scattered RWD sources, and
164 hereby provide opportunities for enhanced risk characterisation and robust risk prediction algorithms,
165 tackling the current pandemic, and eventually any future epidemics. This approach should lead to
166 control measures that will eventually relieve the pressure on the health systems, improve patient
167 prognosis especially among those more vulnerable (e.g., chronic patients, immunosuppressed
168 individuals, and population subgroups with limited access to health care, among others), and mitigate
169 the burden of COVID-19. The specific objectives of unCoVer are:

170

- 171 1. To bring together European and international expertise and data to synchronize collaborative
172 research efforts in addressing the ongoing COVID-19 pandemic in a common platform.

- 1
2
3 173 2. To continuously monitor, identify and facilitate the access and use of COVID-19-related RWD to
4
5 174 fully exploit the potential of this routinely collected information, as a reflection of common medical
6
7 175 practices.
8
9
10 176 3. To identify data gaps, and marginalized populations to proactively seek synergies with
11
12 177 complementary existing and planned clinical databases related to COVID-19.
13
14 178 4. To provide a platform for the use of dissimilar data sources capable of streamlining ethical and
15
16 179 legal aspects and anticipating the needs for data harmonization by innovative computational
17
18 180 resources and integrated information for enhanced impact.
19
20
21 181 5. To bring together expertise on the use of advanced computational, epidemiological and
22
23 182 biostatistical methods to handle heterogeneous, and multi-layered information to facilitate rapid
24
25 183 queries and data outputs related to SARS-CoV-2 infection, underlying drivers of COVID-19
26
27 184 prognosis, the safety and effectiveness of treatments, and sequelae, as well as the impact of
28
29 185 COVID-19 in health system resources.
30
31
32 186 6. To broadcast the use and results of the platform to attract new partners and to pursue
33
34 187 complementarity with existing similar networks in Europe and internationally to save lives and
35
36 188 optimize resources.
37
38
39
40

41 190 **METHODS AND ANALYSIS**

42
43 191 unCoVer is conceptualized as a functional network of partners, capable of harvesting and analysing
44
45 192 RWD derived from the response and provision of care to patients by the health systems across Europe,
46
47 193 and other countries such as Brazil and Colombia during the COVID-19 pandemic.
48
49

50 194 51 52 195 **Setting**

53
54 196 unCoVer comprises 29 partners from 25 institutions in the European Union and 4 non-EU partners
55
56 197 representing 18 countries namely Belgium, Bosnia and Herzegovina, Brazil, Colombia, Croatia, Greece,
57
58 198 Ireland, Italy, Luxembourg, Norway, Portugal, Romania, Slovakia, South Korea, Spain, Turkey, United
59
60

1
2
3 199 Kingdom, and the United States of America ([Figure 1](#)). Partners provide data mostly from front-line
4
5 200 hospitals but also national health agencies, registries, and investigator-lead observational studies, and
6
7 201 represent complementary scientific and medical fields, as well as expertise in research ethics, data
8
9 202 management and statistical modelling. unCoVer, thus, works as a functional network bridging clinical
10
11 203 expertise and data analytics, intending to exploit the full potential of the routine healthcare data
12
13
14 204 already collected from patients during the pandemic. The set-up of the network therefore relies on a
15
16 205 continuous iteration process between a) clinical partners, who will guide the development of research
17
18 206 questions needed to improve patient's care and inform public health strategies; and b) epidemiologists
19
20 207 and analytical experts, who will operationalize the research questions with advanced data processing,
21
22 208 analysis, and simulation tools capable of generating innovative solutions. The work of the network is
23
24 209 further supported by three external advisory boards ([Figure 2](#)) that provide expert counselling
25
26 210 concerning the relevance of the medical research and findings (External Advisory Board, EAB), data
27
28 211 protection (Data Protection and Ethics Advisory Board, DP-EAB), and stakeholder involvement (Societal
29
30 212 and Regulatory Advisory Board, SRAB).
31
32
33
34 213

36 214 **Databases**

38
39 215 unCoVer facilitates access to observational data for secondary analyses. These data are largely
40
41 216 collected for non-research purposes, RWD, and refer to data generated during patient encounters with
42
43 217 the health system which have established information technology protocols and tools for retrieving
44
45 218 and storing information about the healthcare provided. To date, the unCoVer network incorporates 16
46
47 219 databases of electronic medical records from 10 different countries, 6 national registries, 4
48
49 220 observational cohorts, and 2 databases on population screening ([Table 1](#)). The data available to
50
51 221 unCoVer has information on hospitalization of COVID-19 patients with at least two-time points of data
52
53 222 collection, at admission and discharge. In addition to demographics, and clinical/epidemiological data,
54
55 223 other data types such as bio-specimens, imaging data, social network-/contact-tracing related data,
56
57 224 movement-related data and mental health data are also available but with limited access.
58
59
60

1
2
3 225 Clinical/epidemiological data include case report forms (CRF) at patient's admission (i.e. date of
4
5 226 symptom onset and/or admission, signs and symptoms at admission, laboratory results, pre-admission
6
7 227 medication, comorbidities & risk factors), during hospitalization (i.e. signs and symptoms, laboratory
8
9 228 results, supportive treatment, admission to ICU) and at discharge (i.e. date of discharge, outcome, as
10
11
12 229 well as medication and complications) ([Supplementary Table 1](#)).

13
14 230

15 16 231 **STRUCTURE AND ACTIVITIES**

17
18 232 Three blocks of activities, grouping work packages (WPs), were designed to build a functional network
19
20 233 over a period of 36 months from November 2020 onwards, and aim to facilitate the flow of information
21
22 234 for rapid assessment of research questions.

23
24
25 235

26 27 236 **Block 1: Definition, design, and data harvesting.**

28
29 237 This first block forms the architectural foundation and the core of the unCoVer network. In order to
30
31 238 provide a comprehensive repository of available data, WP1 'Data Identification' collects and catalogues
32
33 239 all data in a standardised way, including a common codebook that specifies the harmonized variables,
34
35 240 with standardised variable names and data format and labels, and range of plausible values, in
36
37 241 preparation for data harmonization and validation processes *i.e.* the key for the development of a
38
39 242 unified pool of data for analyses. In parallel, and acknowledging the sensitive nature of health data,
40
41 243 and personal information compliance with ethical and legal aspects, a checklist for assessing the risks
42
43 244 involved in data processing is implemented by WP2 'Ethics and Data Protection', with due
44
45 245 consideration to legal and regulatory issues concerning data protection, privacy and information
46
47 246 security. The checklist includes questions on the nature of the data (e.g. clinical data, hospital records
48
49 247 or publicly available data, personal data, data collected in vulnerable groups, availability of follow-up
50
51 248 data), informed consent (e.g. explicit consent or assent obtained), data protection (e.g. data protection
52
53 249 officer identified and data protection impact assessment completed), ethical approval (obtained or
54
55 250 pending), data privacy protection (e.g. anonymisation vs pseudonymisation, data minimisation), and
56
57
58
59
60

1
2
3 251 data transfer and use (data transfer agreement needed, name of the data controller, data processor,
4
5 252 joint data controller if applicable, and whether international transfer outside EU). Each data provider
6
7 253 within the network populates this checklist, which is reviewed on a case-by-case basis by the unCoVer
8
9 254 research ethics expert team together with the independent DP-EAB, and accordingly informed
10
11 255 decisions on risk mitigation are taken. In short, WP2 ensures that GDPR or equivalent guidelines are
12
13 256 adhered to during the data processing activity. Once a dataset is categorised as low risk, it will be
14
15 257 available to proceed with subsequent steps ([Supplemental Table 2](#)).

16
17
18 258 For the secure multi-party computation of unCoVer data, WP3 'Data Harmonization', developed an
19
20 259 infrastructure based on Opal 4.1 (OBiBa suite, Maelstrom Research, Montreal, Canada) to facilitate
21
22 260 interoperability of the data, including data management, harmonization and dissemination in a
23
24 261 secured environment ¹⁰. The Opal server application provides the necessary key features for data
25
26 262 encryption and decryption managed through Public Key Infrastructure as well as participant identifiers
27
28 263 management and user authentication/authorization for access via a rights and roles management with
29
30 264 username/password. Steps to achieve data harmonisation and secured data sharing and use are: 1)
31
32 265 Set-up the Opal server for each data provider and import relevant datasets, 2) Configure a harmonized
33
34 266 data description in each Opal server, and 3) Run distributed queries on harmonized datasets through
35
36 267 the DataSHIELD application that enables individual-level data analyses across multiple Opal servers
37
38 268 without sharing and disclosing any individual-level data ([Figure 3](#)). Thus, by using computational power
39
40 269 and standardising dissimilar information, while complying with ethical and legal requirements, a data
41
42 270 repository of anonymized and harmonized COVID-19 RWD will be made available for secured data
43
44 271 analyses.

45
46
47
48
49
50 272

51 273 **Block 2: Analytical development, data use, and demonstration.**

52
53 274 The broad range of medical, public health, and research expertise available within the unCoVer
54
55 275 network is at the heart of the WP4 'Data analysis and Outcomes' dedicated to unCoVer findings from
56
57 276 the data acquired to maximize their use in informing COVID-19 response. Activities of this block aim to
58
59
60

1
2
3 277 facilitate and streamline rapid response to identified research gaps using the unCoVer infrastructure
4
5 278 developed in Block 1. With the availability of cross-national harmonized RWD within the unCoVer
6
7 279 network, a large amount of hospitalised and discharged COVID-19 patients will be studied in depth,
8
9 280 together with complementary epidemiological data, to understand the pathophysiology, progression,
10
11 281 treatment, (long-term) complications, and (less frequent) risk factors for early prevention of this novel
12
13 282 disease as well as to grasp the cross-national heterogeneity of the COVID-19 burden. To this end, an
14
15 283 analytical toolbox, including both traditional statistical methods and machine learning techniques and
16
17 284 a Bayesian estimation framework, will be developed for identifying relationships between early clinical
18
19 285 and diagnostic profiles and the future course of the infection, and for a detailed clinical and
20
21 286 epidemiological characterization of COVID-19 patients, being able to generate patients' risk
22
23 287 classification and risk prediction. The application of this toolbox in real patient data, available within
24
25 288 the unCoVer network, will then allow uncovering real-world insights that would support policies and
26
27 289 protocols for optimization of health resources of the hospital and critical care. Therefore, the activities
28
29 290 of this block will be related to the iterative trial use of the Opal server and the toolbox by end-users
30
31 291 (*i.e.* data analysts) as well as the lessons learned and potentially preparing actions for the sustainability
32
33 292 of the unCoVer platform, including the repository and toolbox.
34
35
36
37
38
39
40

41 294 **Block 3: Project management, communication, and exploitation.**

42
43 295 The outputs of the network including the scientific and technological knowledge and outcomes,
44
45 296 provided by the previous two blocks, are streamlined through scientific publications, training and
46
47 297 educational activities, organisation and participation in events, among others, steered by the WP5
48
49 298 'Communication, Exploitation & Dissemination'. This last block is also dedicated to maintaining the
50
51 299 functional network both internally and externally, and with special attention to the management of
52
53 300 intellectual property by utilizing best practices in project coordination, as outlined in WP6
54
55 301 'Coordination'. Concerning this, the unCoVer organization structure works under a Consortium
56
57 302 Agreement signed by all partners and includes the following key bodies within the consortium and
58
59
60

1
2
3 303 management structure ([Figure 2](#)). A steering committee formed by the principal investigators of the
4
5 304 29 partner institutions at the decision-making level. A General Assembly (GA), involving WP leaders
6
7 305 manages the network, and coordinates WP and tasks leaders, assisted also by a Management Support
8
9 306 Team (MST) to reinforce partners' representation. An internal Exploitation and Dissemination
10
11 307 Committee (EDC) also collaborates in the overall management of the GA. Finally, the project
12
13 308 coordinators communicate with the sponsor, and facilitate crosstalk between the network and the
14
15 309 external advisory boards.
16
17
18
19
20

21 311 **ETHICS AND DISSEMINATION**

22
23 312 The unCoVer study has been approved by the Institutional Review Board of the Institute of Tropical
24
25 313 Medicine in Antwerp (IRB/RR/ac/151, protocol number 1524/21). Ethical aspects are of utmost
26
27 314 importance in unCoVer. Upon the project start date, the roles, and responsibilities of the independent
28
29 315 DP-EAB were described, including the selection procedure of the board members and its final
30
31 316 composition, and mandate. A 'scoping exercise' was also conducted across the network to ensure that
32
33 317 all partners are aware of the common obligations in terms of data processing activities using health or
34
35 318 health-related data according to European and international guidelines. Moreover, to be compliant
36
37 319 with the General Data Protection Regulation (GDPR) and meet the ethics requirements, the unCoVer
38
39 320 network will follow the data processing steps represented in [Figure 4](#), in the following sequence:
40
41
42

43 321
44
45 322 1. The unCoVer master checklist of data processing activities in network partners' is circulated within
46
47 323 the network to be completed by the data providers, and data providers are required to provide the
48
49 324 supporting documentation of each indicator of this list, such as informed consents, ethical
50
51 325 approvals, and Data Protection Impact Assessment (DPIA). This information is processed by the
52
53 326 research ethics team, responsible for categorizing the datasets into three different categories: low,
54
55 327 medium, or high data privacy risk.
56
57
58
59
60

- 1
2
3 328 2. Datasets categorized as “low-risk” will be available to proceed with the harmonization process
4
5 329 and, therefore, Opal-DataSHIELD servers will be installed. Within Opal, the patient identifiers will
6
7 330 be separated from the patient study data by employing two databases: (1) the ID database that
8
9 331 stores the patient identifiers accessible by the data provider only, and (2) the study database that
10
11 332 stores pseudo-anonymised patient’s data to be used for data analyses accessible, through code
12
13 333 only, by data analysts. The “medium-” and “high-risk” datasets will be subject to further review
14
15 334 and requirements before harmonization processes.
16
17
18 335 3. Finally, the installation of the servers will allow the consortium to analyse the available RWD
19
20 336 through an anonymisation layer to answer the pre-identified research questions. The system also
21
22 337 facilitates the definition of analytical projects and the specific databases and/or variables that will
23
24 338 be used for a specific project. As a rule, all output of data analytics will be restricted to the
25
26 339 presentation of data aggregates or to line listing deprived of personal identifiers so that the identity
27
28 340 of the study patient cannot be deduced (no backward identification).
29
30
31
32 341

33
34 342 To maximise the unCoVer network’s output, dissemination, and exploitation strategies, as planned by
35
36 343 the EDC and advised the SRAB, i.e., a non-executive consulting substructure composed of several key
37
38 344 stakeholders from the regulatory, governance, civil society level, and patient’s public initiatives, will be
39
40 345 segmented according to the network block activities, the potential users and the most adequate
41
42 346 channels of dissemination and interaction with potential users. The website (uncover-eu.net/), social
43
44 347 media accounts (Twitter @uncoverEU, LinkedIn, YouTube), and the project newsletter will be the
45
46 348 channels to reach all partners and stakeholders of unCoVer, both devoted to providing regular updates
47
48 349 on project activities and announcing upcoming milestones and events. The website will serve as a
49
50 350 repository of the project goals and activities and deliverables in an easy-to-understand language, as
51
52 351 well as publications, lectures and expert documents hosted for access by the partners or stakeholders.
53
54
55
56
57
58

59 353 **Patient and Public Involvement**
60

1
2
3 354 The unCoVer network has been designed to facilitate interactions and enhanced outreach to COVID-
4
5 355 19 stakeholders included in external advisory boards, as well as a prominent work package on
6
7 356 dissemination activities, that include but are not limited to:

9
10 357

11
12 358 - *Scientific community.* To contribute to the body of knowledge in the field, two types of publications
13
14 359 in peer-reviewed journals under Open Access schemes are foreseen: the unCoVer-network
15
16 360 publications, *i.e.* for implementation and/or application of the unCoVer platform as a whole, and
17
18 361 the unCoVer-partners publication, *i.e.* for specific collaborations between two or more unCoVer
19
20 362 partners. In addition to both types of scientific publications, jointly organised workshops, virtual
21
22 363 trainings, and virtual conferences, will be instrumental channels for the dissemination of the
23
24 364 unCoVer activities and results to the scientific community. In such a yearly organised workshop,
25
26 365 the application of the unCoVer repository and toolbox will be presented and expert feedback will
27
28 366 be sought for further improvement. These activities of dissemination to the scientific community
29
30 367 will result in overall awareness and international recognition of the unCoVer network,
31
32 368 simultaneously strengthening the visibility and competitiveness of the institutions involved as
33
34 369 centres of excellence.

35
36
37
38
39 370 - *European platforms and data-driven initiatives.* Cooperation with other European projects
40
41 371 dedicated to COVID-19 data sharing such Orchestra (orchestra-cohort.eu), Synchronos (synchronos.eu),
42
43 372 Dragon (imi.europa.eu/projects-results/project-factsheets/dragon) RecodID (recodid.eu), and EC-
44
45 373 COVID-19 Data Platform (covid19dataportal.org), as well as large networks such as the European
46
47 374 Burden of Disease Network (burden-eu.net), and initiatives on data sharing infrastructures such as
48
49 375 the Population Health Information Research Infrastructure (phiri.eu), will be established for the
50
51 376 co-organisation of dissemination events along with seeking alignment and synergies to avoid
52
53 377 duplication of efforts.

54
55
56
57 378 - *Policy makers.* The accumulated prior experience and contact networks in the regulatory,
58
59 379 policymaking framework of several members of the unCoVer network will be used to ensure that
60

1
2
3 380 the work and output created can reach regulatory entities and policymakers, thus contributing to
4
5 381 the impacts of the project in the decision-making process. This appears relevant as the UnCoVer
6
7 382 network is willing to merge different sources of medical data with social, economic, mental, and
8
9 383 geographical data with the potential to identify highly tailored profiles of risk for community
10
11 384 prevention programmes and educational goals in different countries

12
13
14 385 - *Engagement with key stakeholders.* Results from the project will be further disseminated through
15
16 386 involvement with societal, regulatory, and administrative partners, as ensured by the external
17
18 387 advisory boards to the network, with the goal to warrant the project's impact according to the
19
20 388 stakeholder's needs and expectations.

21
22
23 389 - *General public.* To maximize awareness of unCoVer among the general public, the project activities
24
25 390 and milestones will be broadcasted via social media.

26
27 391 - *Patient associations and clinicians.* Given the foreseen impact of unCoVer's output, patient
28
29 392 associations and clinicians are identified as end-users that benefit from the data-oriented results,
30
31 393 and subsequently, these can be translated to them via seminars, lecture, and infographics as made
32
33 394 available on the project website, among others.

34
35
36 395

37 38 39 396 **DISCUSSION**

40
41 397 During the early phases of the pandemic, the unCoVer network grew organically from initial
42
43 398 partnerships accessing individual databases to answer pressing COVID-19 questions. These initial
44
45 399 collaborations soon recognized the need for accessing extended information to develop more robust
46
47 400 analytical models and outputs. To this end, the concept of unCoVer originates in the shared interest of
48
49 401 its partners to synchronize the research efforts needed for exploiting and valorising the underutilised
50
51 402 and inexpensive RWD in addressing the ongoing COVID-19 pandemic. Within a limited duration of two
52
53 403 years, efforts of the first year were mainly focussed on the organisation of the network and the set-up
54
55 404 of a collaborative federated data infrastructure, accounting for local, national and international ethical
56
57 405 and data protection guidelines and streamlining procedures for data processing activities, including
58
59
60

1
2
3 406 data identification, sharing, harmonisation, validation and analytics. Entering the second year, the
4
5 407 unCoVer network aims to demonstrate the usability of the developed infrastructure on the combined
6
7 408 use of RWD to address clinical and epidemiological research questions related to the COVID-19
8
9 409 pandemic, both revising questions answered with limited datasets and new questions arising from the
10
11 410 evolving pandemic. The unCoVer network, therefore, serves as a proof of concept for building a
12
13 411 federated data infrastructure facilitating data interoperability in a secured environment, while
14
15 412 complying with ethical and data protection guidelines. Moving forward, the established framework of
16
17 413 unCoVer provide valuable input for the use of complementary RWD with robust methodologies at hand
18
19 414 for the still ongoing current pandemic and future pandemics.
20
21
22

23 415

24
25 416

26
27 417

28
29 418

30
31 419

32
33 420
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 421 **AUTHORS CONTRIBUTION**
4

5 422 JLP coordinates the unCoVer network. GS manages overall project's activities. JLP, EM, EMEN, EJC, ZK,
6
7 423 JMC, and DMP lead project's work packages. JLP, EM, EA, SA, ALB, DB, MC, BD, PADV, PH, JK, LLJ, LML,
8
9 424 MM, PN, IHO, DMP, SR, JBS, FS, MET, ACT, MV, SVI, JV, and JMC characterized the data available to the
10
11 425 network. JLP, EM, EJG, EMEN, GS designed the data harmonization system and access to the federated
12
13 426 infrastructure. JLP, ZK, EMEN, GOS, SR and GS supervised compliance with legal and ethical
14
15 427 requirements of the data use. JLP, JCF, JK, PK, MM, EM, DMP, MRO, and GS are responsible for the
16
17 428 dissemination of the network's activities. JLP, EM, GS, ZK, EMEN, JMC, DMP, SR, and JBS drafted the
18
19 429 manuscript and designed the figures. All authors have critically reviewed and approved the final version
20
21 430 of the manuscript. The unCoVer network consist of all individual partner institutions and investigators
22
23 431 who are responsible for the identification, characterization, compliance with legal and ethical
24
25 432 requirements of the data use, as well as dissemination and coordination activities.
26
27
28
29

30 433

31
32 434 **FUNDING STATEMENT**
33

34 435 This project is funded by the European Union's Horizon 2020 Research and Innovation Programme
35
36 436 under Grant Agreement No 101016216.
37
38

39 437

40
41 438 **DATA AVAILABILITY**
42

43 439 unCoVer data will be available via a tiered-access web-based interface and complying with data
44
45 440 protection requirements.
46
47

48 441

49
50 442 **COMPETING INTERESTS**
51

52 443 The authors declare no conflicts of interest.
53
54

55 444

56
57 445 **COLLABORATORS**
58
59
60

1
2
3 446 The unCoVer network: Institute of Tropical Medicine and Antwerp University Hospital, Belgium (José
4
5 447 L. Peñalvo, Elly Mertens, Marianne Van der Sande, Patrick Soentjens, Diana Sagastume, James Cottam,
6
7 448 Gloria Soriano, Sabrina Van Ierssel, and Hanne Van Tiggelen); Fundación Investigación HM Hospitales,
8
9 449 Spain (José M. Castellano, Paula Villares, José Barberán, Mercedes Villareal, Justo Menéndez, Nerea
10
11 450 Ruiz del Árbol, and Alberto Estirado); Universidad Politécnica de Madrid, Spain (Ernestina Menasalvas,
12
13 451 Enrique Javier Gómez, Alberto Blázquez Herranz, David Fernandez Lobón, and Paloma Chausa);
14
15 452 Universidad de Navarra, Spain (Paul Nguewa); Universidade do Porto, Portugal (David M. Pereira, and
16
17 453 Morteza Hosseini); Technological University Dublin, Ireland (Paul Hynds, John Kelleher, and Elizabeth
18
19 454 Hunter); University College Cork, Ireland (Zubair Kabir, Ella Arensman, Brendan Palmer, and Georgie
20
21 455 O’Sullivan); Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca, Romania (Milena Man,
22
23 456 Lucia Maria Lotrean, Mihaela Lupse, and Mira Florea); Universitatea De Medicina Si Farmacie Grigore
24
25 457 T Popa Din Iasi, Romania (Antigona Carmen Trofor, Andrei Tudor Cernomaz, Radu Adrian Crisan-Dabija,
26
27 458 and Cristina Grigorescu); Luxembourg Institute of Health (Michel Vaillant, and Guy Fagherazzi);
28
29 459 Universidade Católica Portuguesa, Portugal (João C Fernandes, and João Silva); Trnava University,
30
31 460 Slovakia (Marek Majdan, Daria Rabarova, Adriana Krsakova, Jaroslava Brnova, Janka Prnova, Jaroslav
32
33 461 Slany and Dominika Plancikova); Instituto Politécnico de Coimbra, Portugal (Ana Lucía Baltazar);
34
35 462 Hospital Universitario de La Princesa, Spain (Joan B Soriano, Julio Ancochea, Nisa Boukichou
36
37 463 Abdelkader, Adrián Peláez, Elena Ávalos, and Gorane Iturricastillo); Instituto Investigación Sanitaria
38
39 464 Fundación Jiménez Díaz, Spain (Lucía Llanos, Miguel Górgolas, Olga Sánchez-Pernaute, Arnoldo Santos
40
41 465 Oviedo, Sergio Luis Lima, Antonio Herrero, and Pablo Minguez), Panepistimio Thessalias, Greece
42
43 466 (Polychronis Kostoulas, Olympia Lioupi, Eleftherios Meletis, Konstantinos Pateras, and Costas Tsiamis);
44
45 467 Universitetet I Sorost-Norge, Norway (Jörn Klein, and Mustafa Asfari); Istituto Don Calabria, Italy (Dora
46
47 468 Buonfrate, Tamara Ursini, and Nicoletta De Santis); Sciensano, Belgium (Brecht Devleesschauwer,
48
49 469 Petronille Bogaert, Koen Blot, Miriam Saso, and Mathil Vandromme); Croatian Institute of Public
50
51 470 Health, Croatia (Jakov Vukovic, Ivan Pristas, Tamara Poljicanin, Jelena Dimnjakovic, Marko Brkic, and
52
53 471 Marija Svajda); Institut Za Antropologiju, Croatia (Miran Čoklo, Saša Missoni, Jelena Šarac, Natalija
54
55
56
57
58
59
60

1
2
3 472 Novokmet, Luka Bočkor, Ivan Dolanc, Antonija Jonjić, and Iva Šunić); Baskent Universitesi Vakfi, Turkey
4
5 473 (Seval Akgun, Tugba Gürgen Erdogan, Süleyman Çetinküner, Cenk Belibağlı, Kübra Demir, Mustafa
6
7 474 Görür, Turgut Bulut, and K.R. Nayar); St Mary's University Twickenham, United Kingdom (Silvia Riva);
8
9
10 475 Azienda ULSS6 Euganea, Italy (Mary Elizabeth Tamang, Carlo Giordani, and Petra Golin); Korea
11
12 476 University, South Korea (In-Hwan OH, and Seok Jun Yoon); University of South Florida, United States
13
14 477 (Miguel Reina Ortiz); Universidad de Antioquía, Colombia (Paula Andrea Diaz Valencia, Lina Ruíz, Juan
15
16 478 Pablo Pérez Bedoya, Oscar Ignacio Mendoza, Camilo Hincapie, Boris Rodriguez, and Noël Barengo);
17
18 479 Feevale University, Associacao Pro Ensino Superior Em Novo Hamburgo, Brazil (Fernando Spilki, Juliane
19
20 480 Deise Fleck, and Matheus Nunes Weber); Univerzitet U Sarajevu, Bosnia and Herzegovina (Enisa
21
22 481 Ademović, Lejla Burnazović-Ristić, Semra Čavaljuga, Džan Ahmed Jesenković, and Lejla Džananović).
23
24
25
26 482

483 REFERENCES

- 484
485 1. World Health Organization (WHO). Expanding our understanding of post COVID-19 condition: report
486 of a WHO webinar, 9 February 2021., 2021.
- 487 2. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-
488 19 in Wuhan, China: a retrospective cohort study. *The Lancet* 2020;395(10229):1054-62. doi:
489 10.1016/S0140-6736(20)30566-3
- 490 3. Shi Y, Yu X, Zhao H, et al. Host susceptibility to severe COVID-19 and establishment of a host risk
491 score: findings of 487 cases outside Wuhan. *Critical Care* 2020;24(1):108. doi: 10.1186/s13054-
492 020-2833-7
- 493 4. World Health Organization (WHO). COVID-19 Clinical management: living guidance, 2021.
- 494 5. Jin Y, Ji W, Yang H, et al. Endothelial activation and dysfunction in COVID-19: from basic mechanisms
495 to potential therapeutic approaches. *Signal Transduction and Targeted Therapy* 2020;5(1):293.
496 doi: 10.1038/s41392-020-00454-7

- 1
2
3 497 6. Wan Y, Shang J, Graham R, et al. Receptor Recognition by the Novel Coronavirus from Wuhan: an
4
5 498 Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology*
6
7 499 2020;94(7):e00127-20. doi: 10.1128/jvi.00127-20
8
9
10 500 7. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased
11
12 501 risk for COVID-19 infection? *Lancet Respir Med* 2020;8(4):e21. doi: 10.1016/S2213-
13
14 502 2600(20)30116-8 [published Online First: 2020/03/15]
15
16 503 8. Kassir R. Risk of COVID-19 for patients with obesity. *Obesity Reviews*;21(6):e13034. doi:
17
18 504 10.1111/obr.13034
19
20
21 505 9. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness
22
23 506 among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study.
24
25 507 *BMJ* 2020;360:m1966. doi: 10.1101/2020.04.08.20057794
26
27
28 508 10. Doiron D, Marcon Y, Fortier I, et al. Software Application Profile: Opal and Mica: open-source
29
30 509 software solutions for epidemiological data management, harmonization and dissemination.
31
32 510 *Int J Epidemiol* 2017;46(5):1372-78. doi: 10.1093/ije/dyx180 [published Online First:
33
34 511 2017/10/13]
35
36
37 512
38
39 513
40
41 514
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Overview of the COVID-19 real-world data among the unCoVer network

Country	Institution ¹	Num. centres ²	Start date	End date	Num. patients ²	Study Population ³	Num. time points	Variable type																
								Demographics	Clinical/Epidemiology	Human OMICs	Pathogen OMICs	Bio-specimens	Imaging	Social network	Movement-related	Mental health	Economic	Diet	Screening					
Patient data from Electronic Medical Records																								
BA	UNSA	10	17/03/2020	20/05/2021	2,000	1, 2, 3	≥2	X	X				X											
BE	ITM-UZA	1	22/04/2020	Ongoing	187		≥2	X	X				X				X							
BR	ASPEUR	40	26/03/2020	Ongoing	2,000	1, 3	15	X	X		X	X						X						
ES	FIHM	12	01/03/2020	15/02/2021	4,480		≥2	X	X				X											
ES	HULPr	1	01/01/2020	29/09/2020	2,217		≥2	X	X															
ES	IIS-FJD	1	07/03/2020	31/05/2020	1,861		≥2	X	X				X											
ES	UNAV	2	01/04/2020	01/01/2021	100		≥1	X	X															
HR	INANTRO	2	01/08/2020	Ongoing	200		≥1	X	X			X	X				X							
IT	IRCCS - DB1	1	03/03/2020	Ongoing	200		≥1	X	X				X											
IT	IRCCS - DB4	1	01/03/2020	09/05/2020	355		≥1	X	X															
IT	ULSS6	4	01/03/2020	Ongoing	1,000		≥3	X	X				X											
RO	UMF Cluj	1			100		≥4	X	X				X											
RO	UMF IASI	1	01/03/2020	Ongoing	150		≥2	X	X				X											
SK	TU	1	01/10/2020	Ongoing	800		≥1	X	X															
TR	BU - DB1	6	01/03/2020	Ongoing	7,000		4	X	X															
Public Health Surveillance data and Registers																								
SK	TU	240	01/03/2020	Ongoing	776,000	2	≥1	X	X												X			
BE	Sciensano	98	14/03/2020	Ongoing	50,000	1	≥2	X	X				X											
CO	UdeA	1,314	14/03/2020	Ongoing	3,997,021	2	1	X	X															
HR	CIPH	NA	25/02/2020	Ongoing	220,000	2	≥1	X	X															
IE	TUDublin	48	29/02/2020	30/11/2020	74,000	2	≥1	X	X					X				X						
NO	USN	40	31/03/2020	Ongoing	2,313	1	10	X	X								X							
PT	UPORTO	NA	03/03/2020	Ongoing	830,000	2	≥1	X	X															

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Country	Institution ¹	Num. centres ²	Start date	End date	Num. patients ²	Study Population ³	Num. time points	Variable type																
								Demographics	Clinical/Epidemiology	Human OMICs	Pathogen OMICs	Bio-specimens	Imaging	Social network	Movement-related	Mental health	Economic	Diet	Screening					
Observational Research data																								
IT	IRCCS - DB3	1	22/04/2020	25/05/2020	1,515	5	≥1	X	X															
LU	LIH	1	15/04/2020	01/04/2021	900	2	5	X	X			X												
PT	IPC	1	01/03/2020	01/07/2020	550	5	≥1	X												X				
UK	SMUC	1	01/07/2020	01/01/2021	200	2	1	X						X										
Screening data																								
IT	IRCCS - DB2	1	31/03/2020	12/05/2020	1,635	6	1	X	X												X			
TR	BU – DB2	1	01/03/2020	Ongoing	18,000	5,7	1	X													X			

¹Institution’s acronyms: ITM-UZA, Institute of Tropical Medicine and Antwerp University Hospital; FIHM, Fundación Investigación HM Hospitales; UNAV, Universidad de Navarra; UPORTO, Universidade Do Porto; TUDublin, Technological University Dublin; UMF Cluj, Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca; UMF IASI, Universitatea De Medicina Si Farmacie Grigore T Popa Din Iasi; LIH, Luxembourg Institute of Health; TU, Trnavska Univerzita V Trnave; IPC, Instituto Politécnico de Coimbra; HULPr, Hospital Universitario de La Princesa; IIS-FJD, Instituto Investigación Sanitaria Fundación Jiménez Díaz; USN, Universitetet I Sorost-Norge; IRCCS, Istituto Don Calabria; CIPH, Croatian Institute of Public Health; INANTRO, Institut Za Antropologiju; BU, Baskent Universitesi Vakfi; SMUC, St Mary’s University Twickenham; ULSS6, Azienda ULSS6 Euganea; UdeA, Universidad de Antioquia; ASPEUR, Associacao Pro Ensino Superior Em Novo Hamburgo; UNSA, Univerzitet U Sarajevu. ² Number of centres providing information, and number of patients/individuals in the datasets could change due to the continuous update of the information; NA, Not available at the moment. ³ Study population refers to, 1) COVID-19 hospitalized patients, 2) COVID-19 cases, 3) COVID-19 patients attending primary care services, 4) COVID-19 patients attending emergency services, 5) general population, 6) health care workers, and 7) tourists.

1
2
3
4
5 **Figure legend**
6
7
8

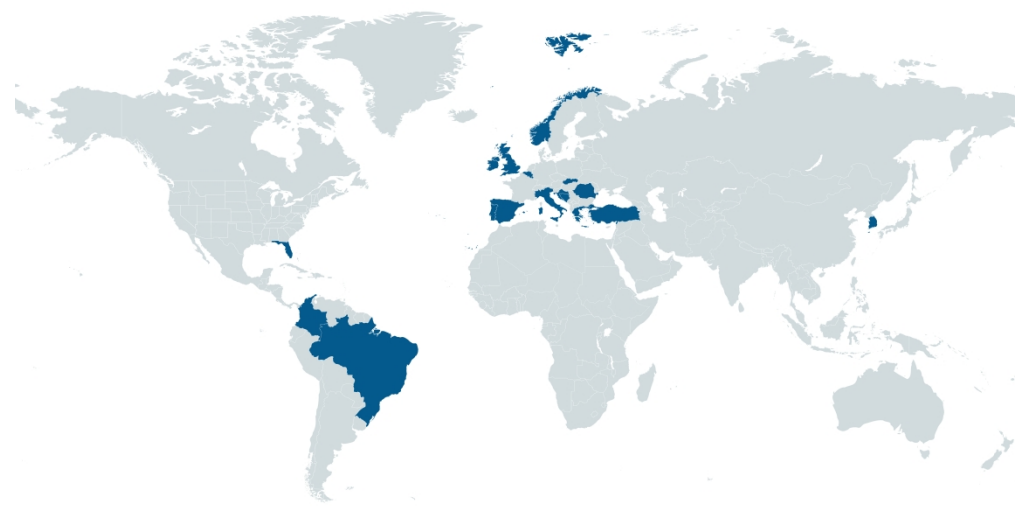
9 **Figure 1.** Geographical distribution of the unCoVer network (Belgium, Bosnia and Herzegovina, Brazil,
10 Colombia, Croatia, Greece, Ireland, Italy, Luxembourg, Norway, Portugal, Romania, Slovakia, South
11 Korea, Spain, Turkey, United Kingdom, and the United States of America)
12

13 **Figure 2.** Management structure of the unCoVer network
14

15
16 **Figure 3.** Secure multi-party computation of unCoVer data based on Opal/DataShield infrastructure
17

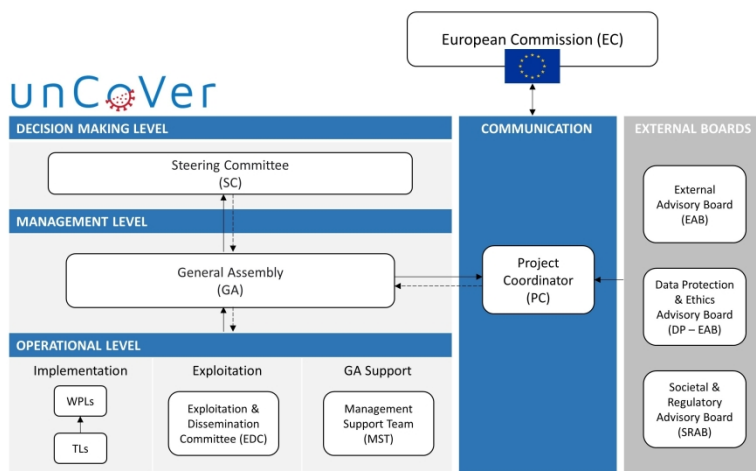
18 **Figure 4.** Data management process of the COVID-19 related data available within the unCoVer
19 network
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Geographical distribution of the unCoVer network (Belgium, Bosnia and Herzegovina, Brazil, Colombia, Croatia, Greece, Ireland, Italy, Luxembourg, Norway, Portugal, Romania, Slovakia, South Korea, Spain, Turkey, United Kingdom, and the United States of America)

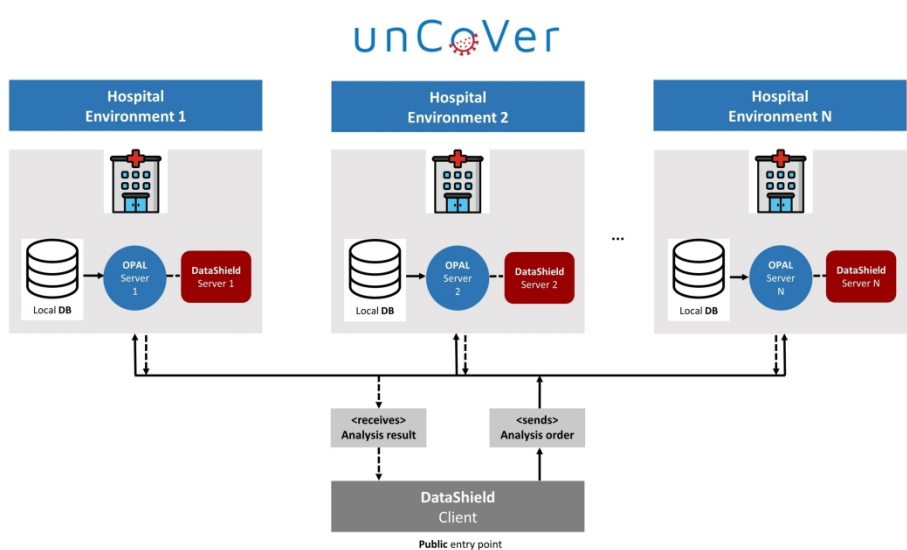
1382x704mm (118 x 118 DPI)



Management structure of the unCoVer network

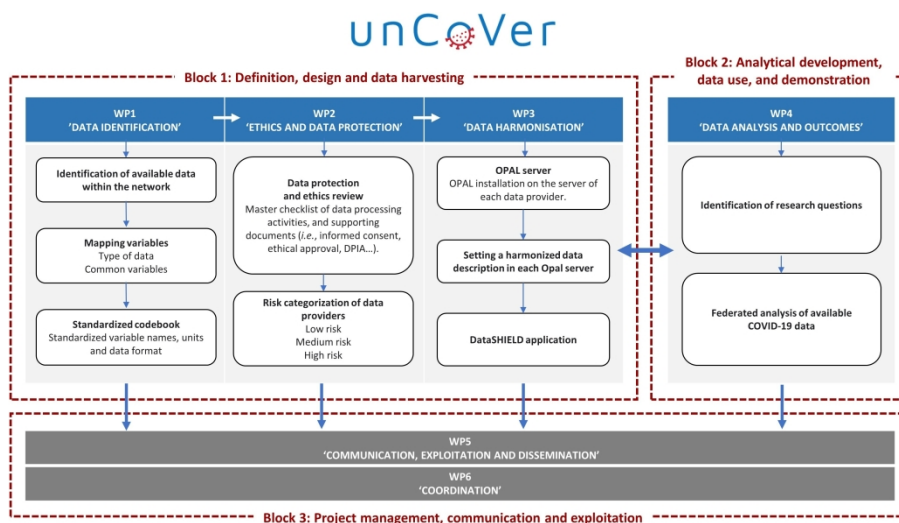
338x190mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Secure multi-party computation of unCoVer data based on Opal/DataShield infrastructure

304x171mm (300 x 300 DPI)



Data management process of the COVID-19 related data available within the unCoVer network

338x190mm (300 x 300 DPI)

Supplemental Table 1. Clinical and epidemiological data available in unCoVer network databases containing hospitalized patient data.

Country	Institution ¹	Patient admission CFR						Daily CFR during hospitalization						Patient outcome CRF			
								At the ward			In ICU						
		Demographics	Onset and/or Admission date	Signs and symptoms	Laboratory results	Pre-admission medication	Co-morbidities & risk factors	Signs and symptoms	Laboratory results	Supportive Treatment	ICU admission date	Signs and symptoms	Laboratory results	Medication treatment	Complications	Discharge date	Outcome status
Patient data from Electronic Medical Records																	
BA	UNSA	X	X	X	X	X	X			X	X		X			X	
BE	ITM-UZA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
BR	ASPEUR	X	X	X	X	X	X	X	X		X					X	
ES	FIHM	X	X	X ²	X ²	X ²	X ²	X		X ²	X		X ²		X	X	
ES	HULPr	X	X	X ²			X		X ²	X ³			X		X ⁴	X	
ES	IIS-FJD	X	X	X	X	X	X	X		X ⁴	X		X	X	X	X	
ES	UNAV	X	X		X												
HR	INANTRO	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
IT	IRCCS - DB1	X	X	X	X	X	X	X		X ³	X		X	X	X	X	
IT	IRCCS - DB4	X	X	X	X		X ³										
IT	ULSS6	X	X	X ²	X	X	X	X	X	X	X	X	X		X	X	
RO	UMFCluj	X	X	X			X		X	X			X		X ⁴	X	
RO	UMFIASI	X	X	X		X	X		X	X	X		X		X	X	
SK	TU	X	X	X	X	X	X	X	X	X	X	X	X		X	X	
TR	BU – DB1	X	X	X			X						X				
TR	BU – DB2	X	X	X			X		X	X	X		X		X	X	
Patient data from Public Health Surveillance and National Registers																	
SK	TU	X	X	X	X	X	X	X	X	X	X	X	X		X	X	
BE	Sciensano	X	X	X	X		X		X		X	X	X	X	X	X	
CO	UdeA	X													X	X	
HR	CIPH	X	X				X								X	X	
IE	TUDublin	X	X				X										
NO	USN	X	X	X	X		X						X	X	X	X	
PT	UPORTO	X	X				X ²		X ³	X						X	
Patient data from Observational Research																	

IT	IRCCS - DB2	X		X			X ³									
IT	IRCCS - DB3	X		X			X ³									
LU	LIH	X	X	X	X	X	X	X	X	X	X ³	X	X	X	X	X

¹Acronyms: ITM-UZA, Institute of Tropical Medicine and Antwerp University Hospital; FIHM, Fundación Investigación HM Hospitales; UNAV, Universidad de Navarra; UPORTO, Universidade Do Porto; TUDublin, Technological University Dublin; UMF Cluj, Universitatea De Medicina Si Farmacie Iuliu Hatieganu Cluj-Napoca; UMF IASI, Universitatea De Medicina Si Farmacie Grigore T Popa Din Iasi; LIH, Luxembourg Institute of Health; TU, Trnavska Univerzita V Trnave; IPC, Instituto Politécnico de Coimbra; HULPr, Hospital Universitario de La Princesa; IIS-FJD, Instituto Investigación Sanitaria Fundación Jiménez Díaz; USN, Universitetet I Sorost-Norge; IRCCS, Istituto Don Calabria; CIPH, Croatian Institute of Public Health; INANTRO, Institut Za Antropologiju; BU, Baskent Universitesi Vakfi; SMUC, St Mary's University Twickenham; ULSS6, Azienda ULSS6 Euganea; UdeA, Universidad de Antioquia; ASPEUR, Associacao Pro Ensino Superior Em Novo Hamburgo; UNSA, Univerzitet U Sarajevu.²Data needs to be extracted from text variables; ³Coded as yes/no without any further specifications; ⁴Length of stay in days instead of exact date of admission and/or discharge from hospital and/or ICU

Supplemental Table 2. Data protection Risk Assessment checklist

> Was personal data/sensitive collected?		
Yes: What permissions and safeguarding are in place for the individual (evaluated through question below)?		
No: What data was collected and what are the risks?		
> Was Explicit Consent or Assent received (if necessary)?		
Yes: Method of requesting individual consent or assuming assent must be provided to unCoVer		
No: Should it have been collected?		
> Is Ethical Approval necessary?		
Yes: Ethical Approval must be provided to unCoVer		
No: If Ethical Approval is not necessary, personal data must not be identifiable (GDPR does not apply) and the data should be 'public' to some extent		
> Has the partner received Ethical Approval?		
Yes: Does unCoVer need any additional information following Ethical Approval?		
No: Hold until Ethical Approval is provided and assess		
> Anonymisation / Pseudonymization: Is the process used to anonymise or pseudonymise the data to be shared with uncover known? Who undertook this process?		
Yes: Transparent and appropriate methods used were clear and acceptable. Parties with data access are identified. Still ensure no personal data is passed onto unCoVer		
No: if these processes were undertaken, is there follow-up required?		
> Is there a letter from the Data Protection Officer (DPO) confirming the standardisation of Ethical Approval?		
Yes: Letter has been signed and returned by local DPO		
No: Follow-up with partner is needed		
Low risk	Medium risk (a combination of some of the below)	High risk (a combination of some of the below)
Ethical approval received or not applicable	Ethical approval received or pending	Ethical approval not received
Explicit consent & assent received or not applicable	Explicit consent & assent received or waiting confirmation	Unclear details of consent / assent
Anonymisation and/or pseudonymization transparent & appropriate	Waiting confirmation of anonymization/ pseudonymization process(es)	Anonymization or pseudonymization incomplete. Or process unexplained
Data transfer agreement in place if necessary	Data transfer pending or in place if necessary	Non-EU partners without appropriate Data transfer agreements
DPO letter received	DPO letter pending/received	DPO letter pending

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Abstract, page 4
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported			Pag 6, 7
Objectives	3	State specific objectives, including any prespecified hypotheses			Pag 7, 8
Methods					
Study Design	4	Present key elements of study design early in the paper			Pag 8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection			Pag 8, 9, and Figure 1
Participants	6	(a) <i>Cohort study</i> - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.	Pag 8, 9, 10, Table 1 and Supplemental Table 1

		<p><i>Case-control study</i> - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</p> <p><i>Cross-sectional study</i> - Give the eligibility criteria, and the sources and methods of selection of participants</p> <p><i>(b) Cohort study</i> - For matched studies, give matching criteria and number of exposed and unexposed</p> <p><i>Case-control study</i> - For matched studies, give matching criteria and the number of controls per case</p>		<p>RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.</p> <p>RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.</p>	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Table 1 and Supplemental Table 1
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group			Table 1 and Supplemental Table 1
Bias	9	Describe any efforts to address potential sources of bias			Pag 10, 11, 12
Study size	10	Explain how the study size was arrived at			Pag 10, 11, 12
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why			Pag 10, 11, 12, and Figure 3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> - If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> - If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses		Pag 10, 11, 12, and Figure 3
20 21 22 23 24 25 26 27 28	Data access and cleaning methods		..	RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Figure 3
29 30 31 32 33 34 35	Linkage		..	RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Figure 3
36	Results				
37 38 39 40 41 42 43 44	Participants	13	(a) Report the numbers of individuals at each stage of the study (<i>e.g.</i> , numbers potentially eligible, examined for eligibility, confirmed eligible, included in the	RECORD 13.1: Describe in detail the selection of the persons included in the study (<i>i.e.</i> , study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in	NA (protocol paper not including results)

45
46
47

		<p>study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram</p>		<p>the text and/or by means of the study flow diagram.</p>	
Descriptive data	14	<p>(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) <i>Cohort study</i> - summarise follow-up time (e.g., average and total amount)</p>			NA (protocol paper not including results)
Outcome data	15	<p><i>Cohort study</i> - Report numbers of outcome events or summary measures over time <i>Case-control study</i> - Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> - Report numbers of outcome events or summary measures</p>			NA (protocol paper not including results)
Main results	16	<p>(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period</p>			NA (protocol paper not including results)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1 2 3 4	Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses			NA (protocol paper not including results)
5	Discussion					
6 7 8	Key results	18	Summarise key results with reference to study objectives			NA (protocol paper not including results)
9 10 11 12 13 14 15 16 17	Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Pag 16, 17
18 19 20 21 22 23 24	Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			Pag 16, 17
25 26 27	Generalisability	21	Discuss the generalisability (external validity) of the study results			Pag 16, 17
28	Other Information					
29 30 31 32 33 34	Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based			Pag 18
35 36 37 38 39 40	Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Pag 18

*Reference: Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1 *Checklist is protected under Creative Commons Attribution ([CC BY](#)) license.
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

For peer review only