



## WORK-PACKAGE 4

Taking stock of emerging and new data collection technologies and their potential impact on the development of future cohort studies and the need to optimise the integration of data.

### D4.2

Repository of European/international communication technologies (including social media) used in population-based cohorts, patients' cohorts, clinical trials, and exposure cohorts ordered by region

**SYnergies for Cohorts in Health: integrating the Role of all Stakeholders**

Grant Agreement No. 825884

Start Date: 01/01/2019

Duration: 36 months





## DOCUMENT INFORMATION

<b>Authors</b>	Kinga Polanska
<b>Contributors</b>	Bonaventure Ikediashi, Teresa Torres, Andreas Weser, Agnieszka Janowska, Wojciech Hanke
<b>Reviewer</b>	WP4 partners (PSSJD, NTNU, THE SYNERGIST) Ellen Vorstenbosch
<b>Responsible Partner</b>	Nofer Institute of Occupational Medicine (NIOM) / The SYNERGIST
<b>Dissemination Level</b>	Public
<b>Nature</b>	Other
<b>Keywords</b>	New communication technologies; repository; cohort; clinical trials
<b>Due Date</b>	30 <sup>th</sup> of June 2020
<b>Actual Submission Date</b>	29 <sup>th</sup> of June 2020
<b>Version</b>	Version 1.0

Disclaimer:

*This document has been produced in the context of the SYNCHROS Project. The SYNCHROS Project has received funding from the European Union's H2020 Programme under grant agreement N<sup>o</sup> 825884. For the avoidance of all doubts, the opinions expressed in this document reflect only the author's view and reflects in no way the European Commission's opinions. The European Commission has no liability in respect to this document and is not responsible for any use that may be made of the information it contains.*





## TABLE OF CONTENTS

DOCUMENT INFORMATION.....	1
TABLE OF CONTENTS .....	3
EXECUTIVE SUMMARY .....	4
1. INTRODUCTION .....	5
2. METHODOLOGY.....	5
2.1 Inclusion/exclusion criteria for mapping new and emerging data and communication technologies – developed under task 4.1.....	6
2.2 Results of the searching process – studies included in the public repository <a href="https://repository.synchros.eu">https://repository.synchros.eu</a> .....	9
3. REPOSITORY OF THE STUDIES USING NEW COMMUNICATION TECHNOLOGIES.....	9
4. EUROPEAN/INTERNATIONAL POPULATION COHORTS, PATIENTS' COHORTS AND CLINICAL TRIALS USING NEW COMMUNICATION TECHNOLOGIES .....	16
5. LIMITATIONS .....	17





## EXECUTIVE SUMMARY

This report contains a concise description of the online repository of European/international technologies used in population cohorts, patients' cohorts and clinical trials oriented by region (Deliverable 4.2).

The document contains the following information:

1. Description of methodology of mapping population cohorts, patients' cohorts and clinical trials using social media and new communication technologies as well as overall structure and content of the repository.
2. An overview of the public repository of European/international technologies used in population cohorts, patients' cohorts and clinical trials.
3. Conclusion drawn from the search of the repository including analysis by region, type of study and type of new communication technologies.

Most of the studies that use new communication technologies were conducted in US (69 studies) and in Europe (47 studies). Mobile data collection was the technology most frequently used in the studies (in 119 studies) of which smartphones and/or smart apps were used in 112 studies and mobile phones in 42 studies. Nineteen studies used social media for data collection and among them Facebook and/or Twitter was the most popular. Geospatial technology was used in 30 studies (with GPS and remote sensing used in most of the studies) and smart sensors in 9 studies (including 5 studies using ingestible sensors). There were no studies using artificial intelligence, crowdsourcing and crowd seeding techniques or passive data collection.

The information collected under WP4 (task 4.2) will be presented during the 32nd Annual Conference of the International Society for Environmental Epidemiology (August 24-27, 2020) – the presentation title: “New communication and data collection technologies used in population and patient cohorts”.





## 1. INTRODUCTION

The project aims to establish a sustainable European strategy for the development of the next generation of integrated population, patient and clinical trial cohorts, thereby contributing to an international strategic agenda for enhanced coordination of cohorts globally. To support the delivery of the overall objective, WP4 is centered on taking stock of the emergence of the use of digital technology for real time data collection – such as social media, wearables and smart phone sensors that will be particularly relevant for the design of future cohort studies involving patients, including clinical trials, and populations. The emergence of new communication technologies brings different new possibilities for data collection for cohorts and clinical trials. Work package 4 conducted a mapping of the different new communication technologies relevant to digital data collection (Deliverable 4.1). Task 4.2 builds on the findings of D4.1 to map existing population cohorts, patient cohorts and clinical trials, in Europe and beyond, using new and emerging data and communication technologies with the ultimate project goal to evaluate if and how they could play a role in the optimization of cohort studies.

## 2. METHODOLOGY

To ascertain the types of technologies that have been used and the type of studies in which they have been implemented, the following approach was taken:

- I. Conduct a mapping using the ascendant (e.g., electronic records, repositories, initiatives/studies provided by WP1 and partners within the project) and descendant (e.g., review of scientific papers) search approaches to map the main cohorts and clinical trials conducted in Europe as well as internationally.
- II. Request technical data for the data collection from the identified studies on: a) the study type; b) the study main objective; c) region, country; d) the design of the study/ population included in the study/number of participants; e) wave number and years in which they were carried out; f) data collection source; g) institution responsible for the study/principal investigator/contact person; h) type of data collection technology.





To achieve the aim of mapping of existing population cohorts, patients’ cohorts and clinical trials using social media and new communication technologies two approaches for searching strategies have been selected:

- I. WP1 identified initiatives of harmonization and integration studies through MEDLINE search. The results from WP 1 were evaluated to determine population cohorts, patient cohorts and clinical trials using new and emerging data communication technologies for data collection.
- II. A separate search strategy has been developed which involved the review of the manuscripts obtained in the search in MEDLINE/PUBMED/COCHRANE/EBSCO databases or reference list of selected papers or review papers.

## 2.1 Inclusion/exclusion criteria for mapping new and emerging data and communication technologies – developed under task 4.1.

*Table 1 Inclusion and exclusion criteria for mapping new and emerging data and communication technologies*

Inclusion	Exclusion
Existing data collection technologies	Publications/articles/reports with a publication date before 2012
Innovative data collection technologies	Blogs
Data collection technologies from unrelated fields	Articles not in English
Active & passive data collection technologies	New communication technologies not used for data collection
Mobile and smart devices	
Personal monitoring devices	
Disease/public health surveillance technology	
Social media	
Crowdsourcing/crowd seeding	
Geospatial technology	





Inclusion	Exclusion
Technologies that enable storing, managing, and querying data and sharing data among devices and databases	

The following keywords were used:

*Table 2 Keywords used for search in MEDLINE, PUBMED, COCHRANE and EBSCO databases*

Related to study design	(AND) Related to new and emerging data and communication technologies
Cohort	Mobile data collection
Population based cohort	Smartphone
Birth cohort	Cellular phone
Patient cohort	Mobile phone
Clinical trial	Mobile apps
	Tablet
	Laptop
	Wearable device
	Smartwatch
	Digital watch
	Social media
	Social network sites
	Social platforms
	Twitter
	Facebook
	Google
	Geospatial technology
	Global navigation satellite systems
	Geographical Information System
	Remote sensing
	Smart sensors
	Ingestible electronic devices
	Bio-sensors





Related to study design	(AND) Related to new and emerging data and communication technologies
	Crowdsourcing
	Crowd seeding
	Artificial intelligence
	Passive data collection

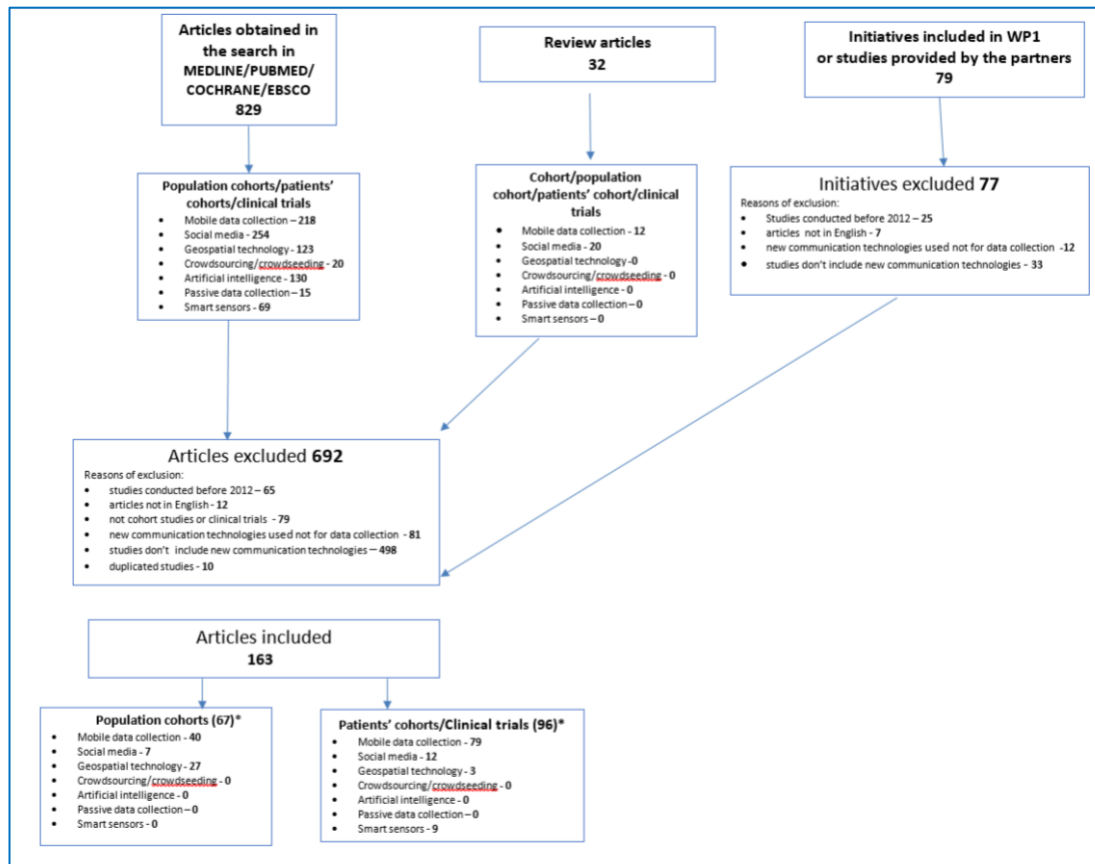
The result of the search process is presented in the diagram below. As a first step, the initiatives of harmonization and integration of health studies identified by WP1 partners and MEDLINE search under WP1 or provided by WP4 partners were evaluated for the identification of those that use new communication technologies. However, using this methodology only resulted in the identification of two studies: The Nord-Trøndelag Health Study (HUNT4) and Avon Longitudinal Study of Parents and Children (ALSPAC). Therefore, we decided to develop a separate search strategy which involved the review of the manuscripts obtained in the search in MEDLINE/PUBMED/COCHRANE/EBSCO databases or reference list of selected papers or review papers. All the studies listed according to the established keywords were evaluated by two researchers (based on the title/abstract and if needed full text of the paper). Finally, 163 studies met the inclusion criteria for that search strategy. The most important reasons for exclusion of the studies were as follows: 1) studies were conducted before 2012, 2) articles were not in English language, 3) there were no cohort studies or clinical trials, 4) new communication technologies were not used for data collection or communication, 5) studies have not included selected new communication technologies and 6) some of them were published in several papers (i.e. duplicated studies).





## 2.2 Results of the searching process – studies included in the public repository <https://repository.synchros.eu>

Figure 1 PRISMA diagram of mapping search results



\*In one study more than one new communication technologies can be used.

## 3. REPOSITORY OF THE STUDIES USING NEW COMMUNICATION TECHNOLOGIES

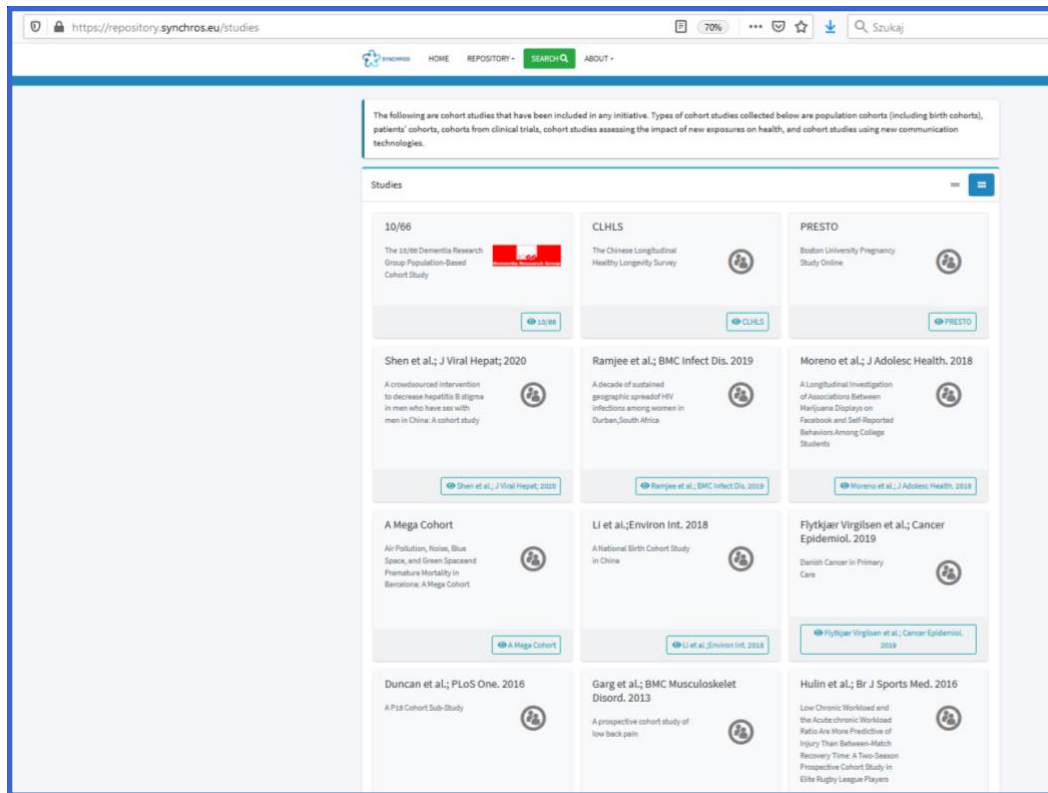
The repository itself (<https://repository.synchros.eu/>) has been developed under the WP 1 (D1.2). It offers a public collection of information about the main European and international efforts to harmonize and integrate cohort data including details about harmonization and integration methodology, communication technologies used in cohorts and exposure and health risk cohorts. It represents the digital platform of compiled European and international initiatives and cohorts mapped by WP1, WP2 and WP4 of the SYNCHROS project accessible for the public. The SYNCHROS repository has been developed with the objective of having a long-term solution that can continuously be enriched, even after the project has ended. The final studies identified from





the search conducted, have been mapped in the repository. The repository contains 163 studies (population cohorts, patients' cohorts and clinical trials) that use new communication technologies.

Figure 2 SYNCHROS Repository – overview of cohort studies

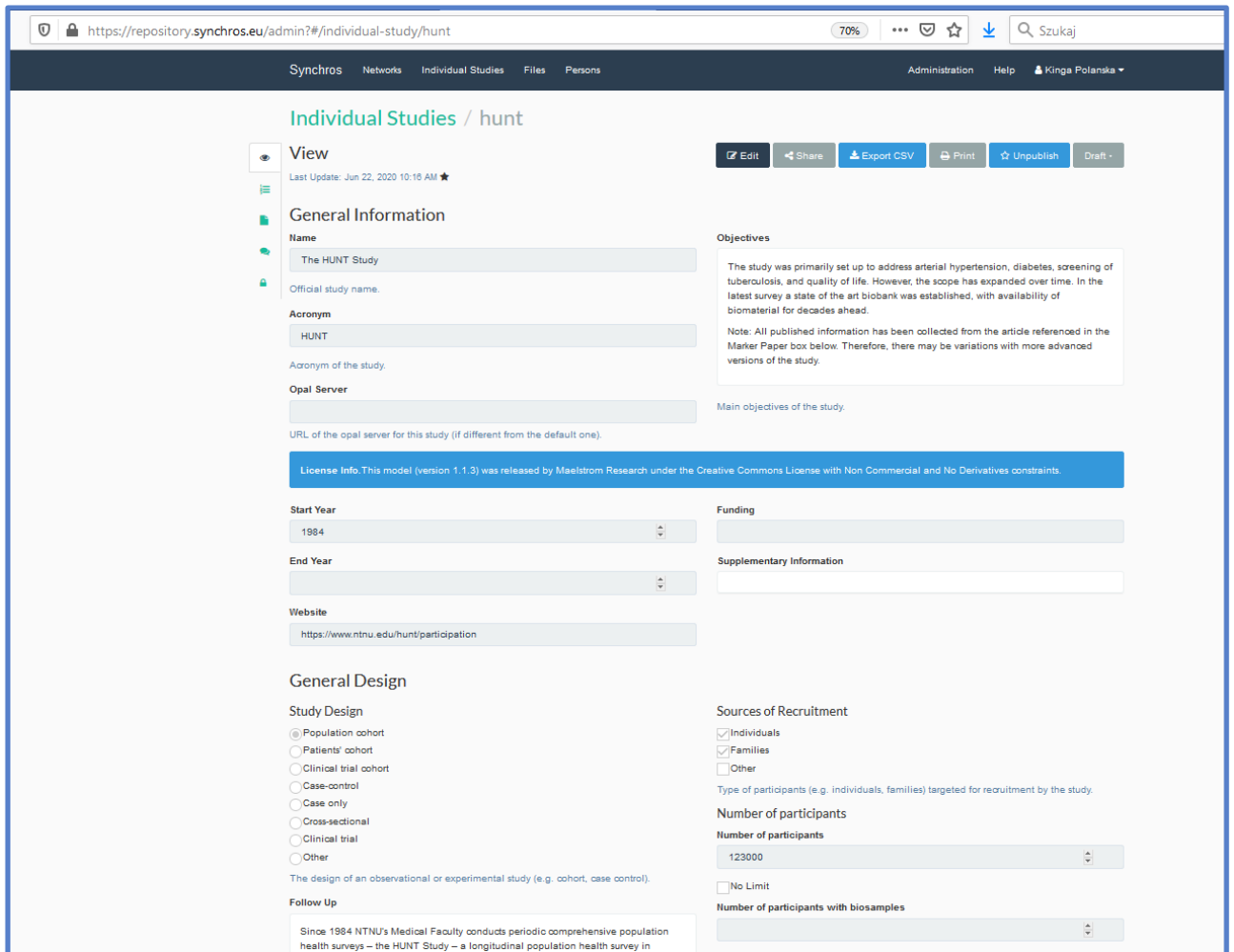


For each study included in the repository the following information has been added:

- 1) General information: study name and acronym (if available), study objectives, study design and funding sources.



Figure 3 – SYNCHROS repository – data of individual study



The screenshot shows the 'View' page for the 'hunt' study in the SYNCHROS repository. The page is organized into several sections:

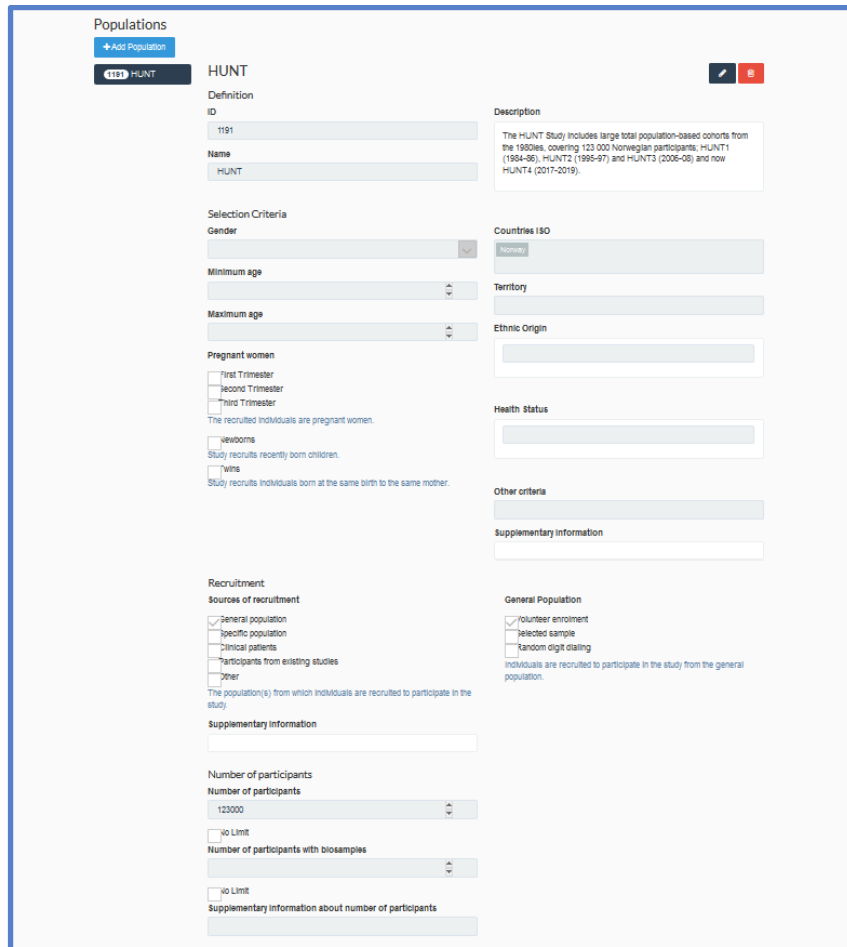
- General Information:**
  - Name:** The HUNT Study
  - Acronym:** HUNT
  - Opal Server:** (Empty field)
  - Start Year:** 1984
  - End Year:** (Empty field)
  - Website:** <https://www.ntnu.edu/hunt/participation>
- General Design:**
  - Study Design:** Population cohort (selected), Patients' cohort, Clinical trial cohort, Case-control, Case only, Cross-sectional, Clinical trial, Other.
  - Sources of Recruitment:** Individuals (checked), Families (checked), Other (unchecked).
  - Number of participants:** 123000
  - Number of participants with biosamples:** (Empty field)
- Objectives:**
  - The study was primarily set up to address arterial hypertension, diabetes, screening of tuberculosis, and quality of life. However, the scope has expanded over time. In the latest survey a state of the art biobank was established, with availability of biomaterial for decades ahead.
  - Note: All published information has been collected from the article referenced in the Marker Paper box below. Therefore, there may be variations with more advanced versions of the study.

2) Population description including type of population and number of participants.



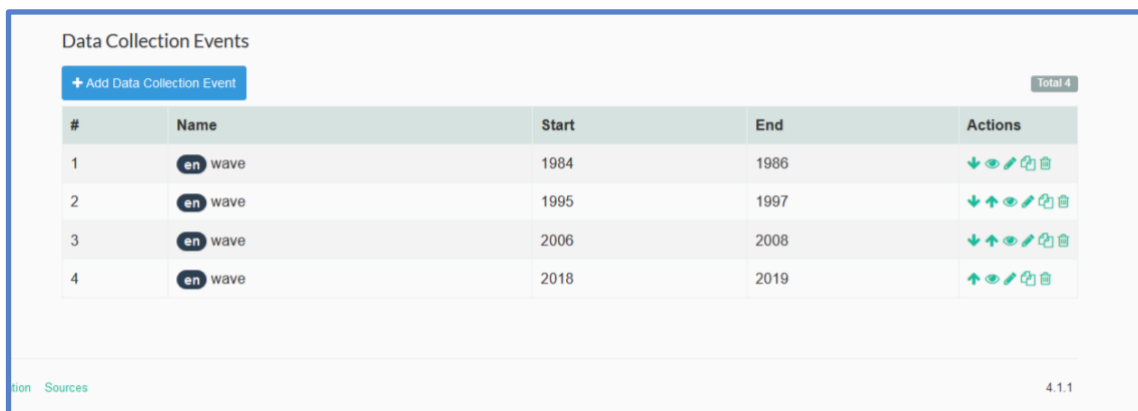


Figure 4 SYNCHROS repository - Population description including type of population and number of participants



- 3) Data collection event describing each wave of the study and new communication technologies used in the study.

Figure 5 SYNCHROS repository – description of data collection event



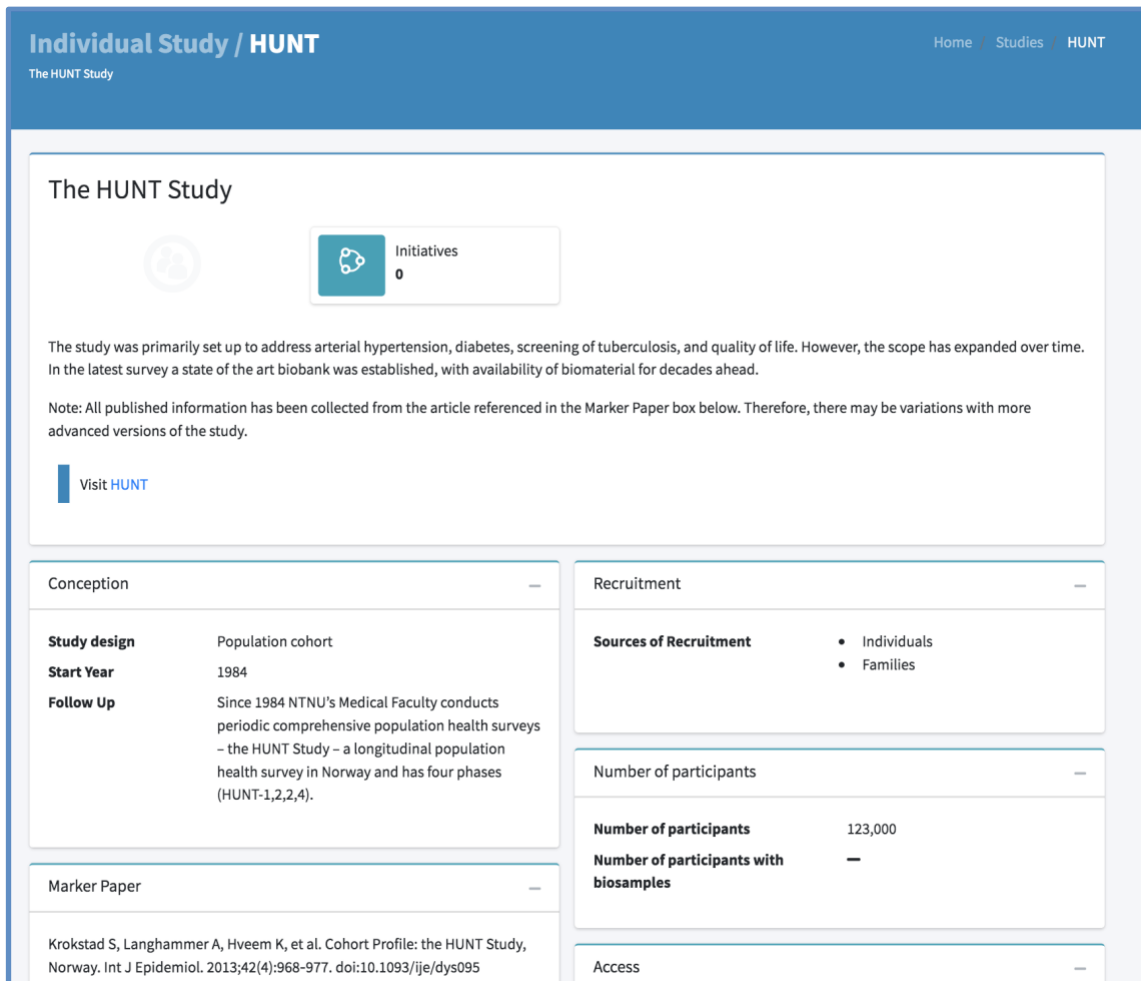
#	Name	Start	End	Actions
1	en wave	1984	1986	↓ 👁 ✎ 🗑
2	en wave	1995	1997	↓ ↑ 👁 ✎ 🗑
3	en wave	2006	2008	↓ ↑ 👁 ✎ 🗑
4	en wave	2018	2019	↑ 👁 ✎ 🗑





All entered studies (163 studies using new communication technologies) are available to the public (<https://repository.synchros.eu>) in the searchable repository. The detailed information for each individual study can be generated as seen in the example presented below:

Figure 6 SYNCHROS Repository – detailed information of individual studies



**Individual Study / HUNT** Home / Studies / HUNT

The HUNT Study

**Initiatives**  
0

The study was primarily set up to address arterial hypertension, diabetes, screening of tuberculosis, and quality of life. However, the scope has expanded over time. In the latest survey a state of the art biobank was established, with availability of biomaterial for decades ahead.

Note: All published information has been collected from the article referenced in the Marker Paper box below. Therefore, there may be variations with more advanced versions of the study.

[Visit HUNT](#)

<p><b>Conception</b></p> <p><b>Study design</b> Population cohort</p> <p><b>Start Year</b> 1984</p> <p><b>Follow Up</b> Since 1984 NTNU's Medical Faculty conducts periodic comprehensive population health surveys – the HUNT Study – a longitudinal population health survey in Norway and has four phases (HUNT-1,2,2,4).</p>	<p><b>Recruitment</b></p> <p><b>Sources of Recruitment</b></p> <ul style="list-style-type: none"> <li>• Individuals</li> <li>• Families</li> </ul>
<p><b>Marker Paper</b></p> <p>Krokstad S, Langhammer A, Hveem K, et al. Cohort Profile: the HUNT Study, Norway. <i>Int J Epidemiol.</i> 2013;42(4):968-977. doi:10.1093/ije/dys095</p>	<p><b>Number of participants</b> 123,000</p> <p><b>Number of participants with biosamples</b> –</p>
	<p><b>Access</b></p>

\*Image continues in below





Marker Paper

Krokstad S, Langhammer A, Hveem K, et al. Cohort Profile: the HUNT Study, Norway. *Int J Epidemiol.* 2013;42(4):968-977. doi:10.1093/ije/dys095

[PUBMED 22879362](#)

**Number of participants with biosamples** —

---

Access

General Information

**Availability of data and biosamples**

Data	✓
Biosamples	✓
Other	—

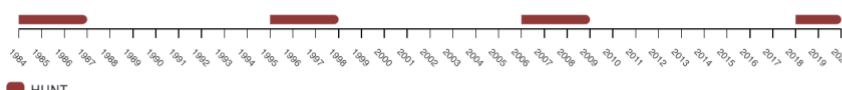
**Supplementary Information**

The HUNT Study includes data from questionnaires, interviews, clinical measurements and biological samples (blood and urine). Data from the HUNT Study are available for researchers who satisfy some basic requirements ([www.ntnu.edu/hunt](http://www.ntnu.edu/hunt)), whether affiliated in Norway or abroad. HUNT Biobank is an established and modern research biobank with high-technology equipment for storage, analysis, sample handling and delivery of samples. Our samples shall satisfy high quality standards and are stored in accordance with the Data Inspectorates laws and regulations. HUNT Biobank engages in sample handling from The Nord-Trøndelag Health Study (HUNT), Cohort of Norway (CONOR), and can receive samples from other researchers and research projects for storage, analysis and processing of DNA. We do not store samples from private individuals.

Yes   
  No   
  Not applicable   
  Don't know

---

**Timeline**



■ HUNT

---

**Population**

The HUNT Study includes large total population-based cohorts from the 1980ies, covering 123 000 Norwegian participants; HUNT1 (1984-86), HUNT2 (1995-97) and HUNT3 (2006-08) and now HUNT4 (2017-2019).

**Selection Criteria**

Newborns	✗
Twins	✗
Countries	• Norway
Ethnic Origin	•
Health Status	•

**Recruitment**

<b>Sources of recruitment</b>	• General population
<b>General Population</b>	• Volunteer enrolment

**Number of participants**

<b>Number of participants</b>	123,000
<b>Number of participants with biosamples</b>	—

---

**Data Collection Events**

Show  entries Search:

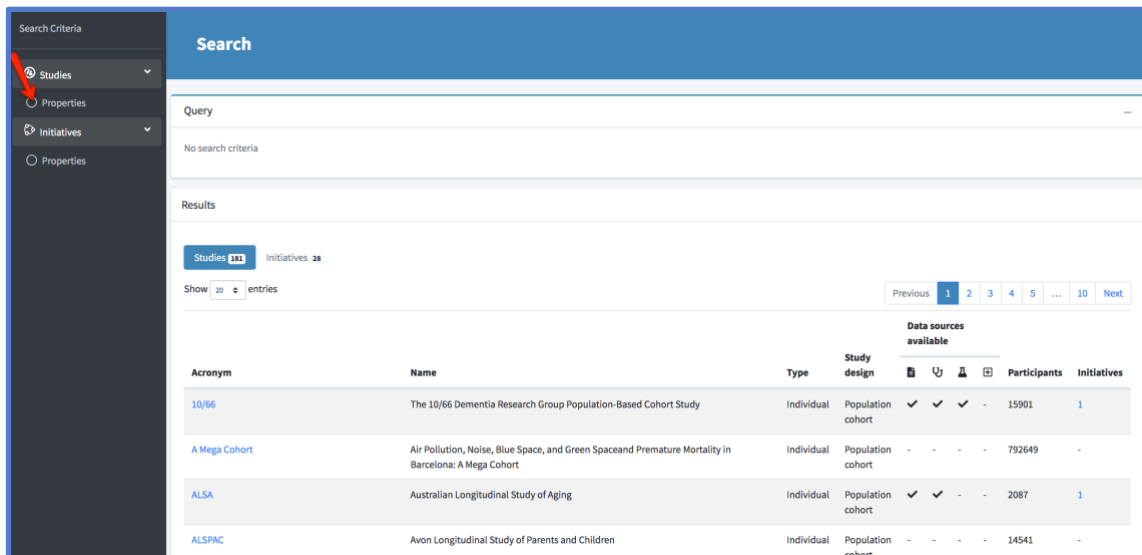
Name	Description	Start	End
wave		1984	1986
wave		1995	1997
wave		2006	2008
wave	Wearable device for measuring physical activity: AX3 (Axivity, Newcastle, UK)	2018	2019





In addition, the search can be carried out by selecting or writing the desired initiative/study properties through the left menu of the general search page. This can be done by country/region, population, type of study and type of new communication technologies used in the studies.

Figure 7 SYNCHROS Repository – Search engine



The screenshot shows the SYNCHROS search engine interface. On the left, there is a 'Search Criteria' sidebar with options for 'Studies', 'Properties', 'Initiatives', and another 'Properties' option. The main area is titled 'Search' and contains a 'Query' input field with the text 'No search criteria'. Below this, the 'Results' section shows 'Studies 11' and 'Initiatives 28'. A 'Show 20 entries' dropdown is present, along with pagination controls (Previous, 1, 2, 3, 4, 5, ..., 10, Next). A table of results is displayed with the following columns: Acronym, Name, Type, Study design, Data sources available (with icons for a folder, a lock, and a document), Participants, and Initiatives.

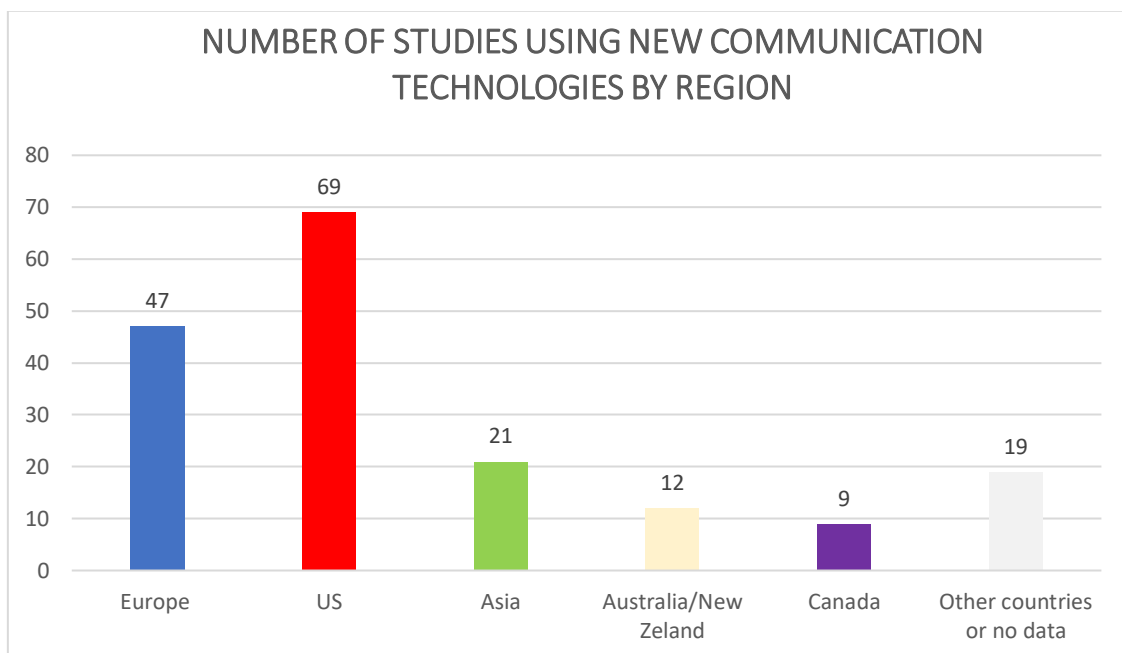
Acronym	Name	Type	Study design	Data sources available	Participants	Initiatives
10/66	The 10/66 Dementia Research Group Population-Based Cohort Study	Individual	Population cohort	✓ ✓ ✓ -	15901	1
A Mega Cohort	Air Pollution, Noise, Blue Space, and Green Space and Premature Mortality in Barcelona: A Mega Cohort	Individual	Population cohort	- - - -	792649	-
ALSA	Australian Longitudinal Study of Aging	Individual	Population cohort	✓ ✓ - -	2087	1
ALSPAC	Avon Longitudinal Study of Parents and Children	Individual	Population cohort	- - - -	14541	-



## 4. EUROPEAN/INTERNATIONAL POPULATION COHORTS, PATIENTS' COHORTS AND CLINICAL TRIALS USING NEW COMMUNICATION TECHNOLOGIES

Our findings show that most of the studies that use new communication technologies were conducted in the US (69 studies) and in Europe (47 studies). Looking at the studies conducted in Europe most of them used mobile data collection, geospatial technology and smart sensors. Some of the studies are based on international collaboration.

Figure 8 – Number of identified studies using new communication technologies by region



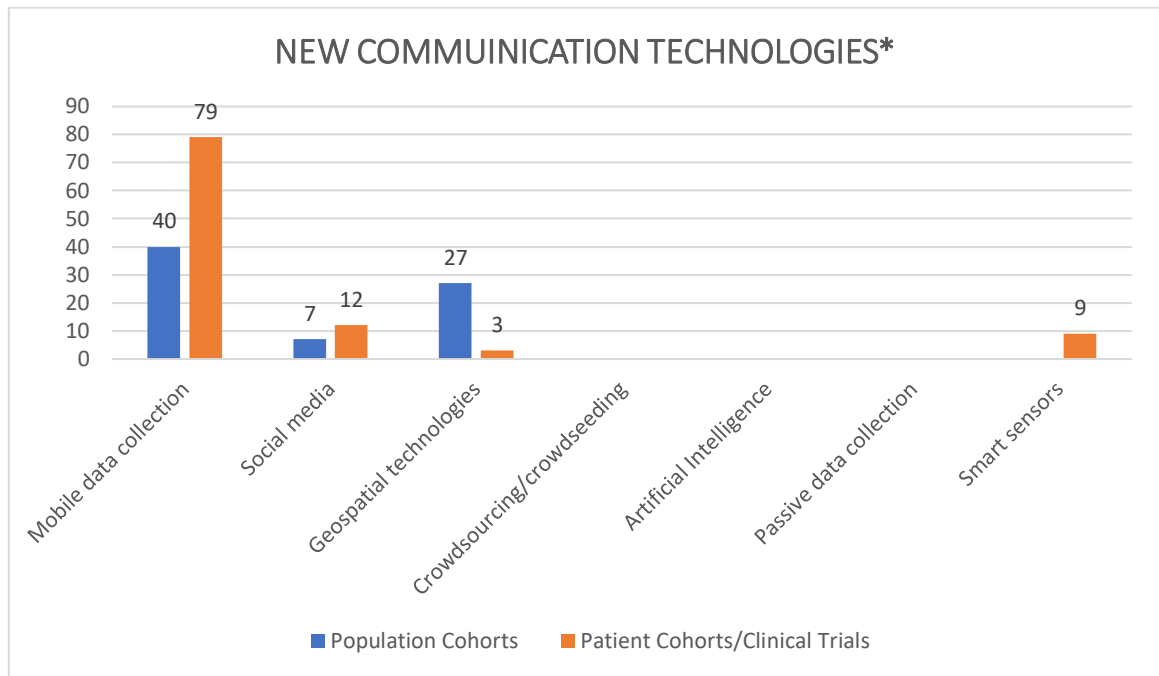
Mobile data collection was the technology that was most frequently used in the studies (in 119 studies) and among this type of technology smartphones and/or smart apps were used in 112 studies and mobile phones in 42 studies (the other studies collect the data using smartwatches and wearables or tablets/laptops). Nineteen studies used social media for data collection and among them Facebook (15 studies) or Twitter (8 studies) were the most popular. Geospatial technology was used in 30 studies (with GPS and remote sensing used in most of the studies) and smart sensors in 9 studies (including 5 studies using ingestible sensors). There were no studies using artificial intelligence, crowdsourcing and crowd seeding techniques or passive data collection.







Figure 9 Number of identified studies that use new communication technologies by type of technology



\*Number of studies in which new communication technologies were used are presented (in one study more than one new communication technologies can be used).

## 5. LIMITATIONS

While the methodology we employed for the search adopted was comprehensive, we are aware that our finding is not exhaustive. There may be other cohorts/ clinical trials conducted that may have used different new communication technologies that are yet to be published.