



Health from Cohorts and Biobanks (COHORT): Evaluation Report



Research Council of Finland

Table of Contents

1.	Summary	3
2.	Register-based research in Finland	4
3.	Health from Cohorts and Biobanks (COHORT) programme	6
4.	Programme evaluation: purpose of evaluation and methods used	7
5.	Call for applications and funded projects	7
5.1.	Early Determinants of Cardiovascular, Metabolic and Reproductive Health.....	7
5.2.	Finnish Psychiatric Birth Cohort Consortium	8
5.3.	Connecting Northern Finland Birth Cohorts with Borealis and THL Biobanks...8	
5.4.	Longitudinal Birth Cohort and Family-Based Studies of Morbidity, Mortality and Social Disadvantage: A Consortium Application on Sharing Data, Family Designs and Methods.....	9
5.5.	Finnish Platform of Birth Cohorts: Microbial Interactions and Health	9
5.6.	Sleep Patterns in Neuropsychological Development in Early Childhood – The Interplay of Genes and Environment	10
6.	Programme achievements.....	10
7.	Challenges.....	11
7.1.	Harmonisation of data.....	12
7.2.	Retrieving registry data.....	12
7.3.	Permits and legal issues	13
7.4.	Funding	13
7.5.	Application process.....	13
8.	Advice and suggestions for the future	14
8.1.	Ministerial-level suggestions.....	14
8.2.	General recommendations to the Research Council of Finland.....	14
8.3.	Programme-specific reflections and suggestions.....	14
8.4.	Recommendations to other stakeholders	15
9.	Concluding remarks	16
	Appendix A: Evaluation questions and material for the panel.....	17
	Appendix B: Projects funded in the Cohort call 2016	19

ISBN 978-951-715-944-9
Helsinki 2024

1. Summary

The Research Council of Finland’s research programme “Health from Cohorts and Biobanks (COHORT)” funded six research consortia between 2017 and 2020. The programme enhanced collaboration between different research studies on cohorts, registries and biobanks, with a particular focus on how to create joint research resources where different datasets could be combined to address complex questions that cannot be answered with single datasets. More specifically, the COHORT programme enabled structuring, harmonisation and combination of surveys, health data registries and biobanks. The infrastructure created can be used for questions that cannot be addressed with data from a single source only. Data standardisation methods developed during the programme may also be applied to other cohorts.

The lessons learned are likely to have long-term usefulness for a continued building and use of better research infrastructures based on scientific collaborations. It is of special note that as the COHORT programme was register-driven, the key outcomes are not scientific papers. This is a common feature of most other hypothesis-driven Research Council of Finland funding.

This evaluation report describes the results, opportunities and challenges faced by the researchers when implementing the projects and what can be learnt from their experiences to inform future initiatives. Based on these, the evaluation panel also makes suggestions and recommendations for the future, which are also highlighted in this report.

In general, all consortia were content with the programme funding, acknowledging the advantages of consortium type of funding for these types of projects, and that their projects would not have worked equally well as individual projects.

Two key challenges were highlighted:

- 1) how to obtain long-term funding for long-term initiatives
- 2) a difficult regulatory basis.

Although the European General Data Protection Regulation (GDPR) is the same throughout the EU, most consortia had encountered problems with the GDPR being interpreted more strictly in Finland than in other countries. A separate act on the [secondary use of health and social data](#) was also perceived as problematic. There were also very protracted delays in obtaining data from registers. Although there are regulations that stipulate the maximum time for handling data requests, there was a widespread experience that these regulations were not followed.

Possible improvements that were suggested included the following:

- rules on maximum waiting times with requirements to report waiting time statistics

- level of financing for registry holder(s) determined by their performance in meeting these requirements
- actions to promote faster evolving of widely recognised policies on interpretation of EU law, notably the GDPR.

In the current situation of unpredictable waiting times, rules that would allow no-cost extension of projects could ameliorate the problems.

Based on researcher interviews and final reports submitted by the consortia, the evaluation panel made several recommendations for the Research Council of Finland as well as other stakeholders involved.

The recommendations to the Research Council of Finland regarding research funding are as follows:

- 1) longer-term funding periods, and an opportunity to apply for extensions for particularly successful projects
- 2) funding instruments for research infrastructures, including parallel funding for research projects that use the infrastructure
- 3) funding instruments for improving resources (e.g. for expanding databases and cohorts with additional data or longer follow-up).

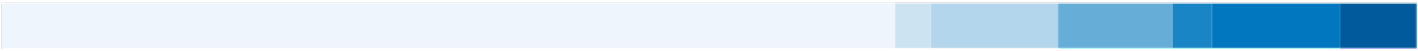
At programme management level, the panel suggests the following actions:

- 1) arranging national and international conferences and more programme events (e.g. kick-off seminar and meetings to interact with the Research Council of Finland and the other consortia)
- 2) requesting interim reports, preferably yearly reports, and providing more feedback from the funder to the consortia
- 3) continuation of the current type of evaluation with post-project interviews.

Finally, the organisations that are responsible for the research infrastructure could award sustainability grants for maintenance of existing infrastructures to ensure that these are sustained. Such grants could be quite limited in terms of amount but could save valuable resources for the future as well as help in developing and maintaining the expertise of the personnel involved.

2. Register-based research in Finland

Register-based studies are essential to study outcome, risk factors and protective factors for major adversities as well as health and social outcomes. The uniqueness with register-based research is to study rare risk factors and outcomes as well as interaction of different health and social factors. Register studies include possibilities to study these phenomena from pregnancy to adulthood. Another major advantage is complete datasets with low drop-out rates. Progressive legislation and a supportive public opinion have been



central to the development of key research infrastructures in Finland. However, despite an abundance of resources, their effective use is still hampered by a lack of cohesion and coordination: originally created in researcher-driven projects to address specific questions, the datasets available are narrowly focused on individual research interests. Translating cohorts, basic research and clinical data into population-level health effects will require closer integration of relevant infrastructures and research in this field, among other things by pooling cohorts (e.g. joint research questions) and complementing existing data (e.g. joint sampling) so that they can serve a wider range of users.

Register-based studies, especially in the Nordic setting, enable creation of datasets that would be extremely time-consuming, very expensive, and in most cases impossible to create by individual research groups. This field of research is traditionally championed by the Nordic countries, where comprehensive, population-based registries that are systematically linkable using the same personal identifiers have been an essential foundation of society for many decades, even centuries.

However, there is an international development of registry-based research that provides a challenge to the Nordic countries for keeping their traditional competitive advantage. Nonetheless, Finnish cohorts remain of the highest international standard, so this may be a potential strong point for Finnish science also in the future.

An opportunity that could be essential for Finland to exploit is to create datasets with a huge number of participants and long follow-up to detect, for instance, the development of chronic diseases, or study of less common or rare diseases, or diseases with a long latency period, or inter-generational effects. Such data do not exist and are very difficult to collect or obtain otherwise. An additional advantage of these datasets is that they are based on the total population with a large number of observations. They also allow for nested data structures. Use of data from population-based registries increases the efficiency, power, generalisability and reliability of data-based studies.

Major discoveries that can be made by crude designs have mostly already been done, and in the future more powerful approaches combining large sets of data in stronger designs is the only way to make progress. For example, studies on the outcome after treatment using registry- and cohort-based designs, which can have immense importance for better healthcare, are underused. Family studies, which make use of multigeneration registry data, can delineate whether associations are causal or indirect, and they are very important for public health.

3. Health from Cohorts and Biobanks (COHORT) programme

Finnish health promotion research has achieved a very high international standard in the past 10–15 years. A crucial asset has been its access to high-quality research datasets, including cohort data collected from different population groups. This has provided important evidence on the health effects of lifestyle factors. Progressive legislation and a supportive public opinion have been central to the development of key research infrastructures in Finland. However, despite an abundance of resources, their effective use is still hampered by a lack of cohesion and coordination: originally created in researcher-driven projects to address specific questions, the datasets available are narrowly focused on individual research interests.

There have been recent efforts to create a more coherent environment for the use of health data. The Act on Secondary Use of Health and Social Data (2019) is a special law applied only in Finland, on top of the GDPR. Analysis of data at individual level is permitted only in environments that meet the very strict requirements of the Act. The health data protection rules at the individual level have also been further updated in the modified Biobank Act (2023) to include GDPR regulations. Building a system to meet the additional guidelines above the GDPR is extremely expensive and time-consuming. Such resources should rather be used for research, following the GDPR, as in other EU countries and in the Nordics, with similar administrative register data. Most of the research community and members of all COHORT programme consortia have been extremely worried about the effects of these changes and new plans (European health data space), as they could dramatically reduce the opportunities for register-based research, making it more difficult and considerably more expensive to perform.

At the beginning of 2016, the Research Council of Finland announced the research programme “Health from Cohorts and Biobanks (COHORT)”. The programme was aimed at promoting initiatives combining cohorts, biobanks and associated datasets by funding several consortia of scientists from different disciplines that would integrate population cohorts, registries and biobank samples in a universally applicable way that would ultimately be used to gain knowledge on population-level health effects. In particular, the purpose of the funding was to facilitate the integration of projects with basic funding already in place and create greater cohesiveness and synergy among researcher-driven, multidisciplinary consortia working with the same sets of cohorts which, through scientific regeneration, would contribute to enhancing the impact of research based on registers, cohorts and biobanks.

The programme provided an opportunity to broaden understanding of the complexity of different types of data, such as wet data (e.g. sample analyses) and dry data (e.g. questionnaire data), as well as data from different electronic systems (e.g. medical records).

4. Programme evaluation: purpose of evaluation and methods used

The purpose of the present report is to evaluate the COHORT programme, to describe the main results achieved by the research consortia and to address questions such as what the main challenges faced by the consortia were and what kind of opportunities the programme enabled. A further purpose is to highlight strengths and shortcomings of the programme and identify possible areas of improvement. Was there a need for more support for the researchers, and if so what kind of support and by whom and at which stage? What is the legacy of the programme? Suggestions and ideas for future calls and investments were also asked for (see general interview structure in Annex 1).

This evaluation report is based first on the final reports prepared by the leaders of the consortia after the funding period, and second on a series of one-hour interviews with each of the consortium leaders. The interviews were conducted in August and September 2022 by Joakim Dillner, Professor of Infectious Epidemiology at the Karolinska Institutet, and Tea Lallukka, Professor of Medical Sociology at the University of Helsinki.

5. Call for applications and funded projects

The applications submitted to the COHORT programme call were reviewed by an international panel of experts within the field. The major review criteria were as follows:

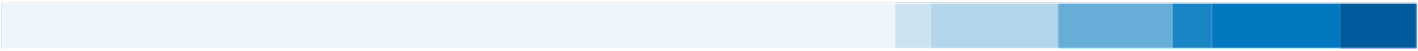
- innovativeness and feasibility of the research plan
- quality of the research teams and the research environment
- fit with the COHORT programme objectives.

The call attracted 15 applications, out of which six consortia received funding. This yielded a 40% success rate. The total programme volume was 5 million euros, and the funding period was four years. The funding period started in 2017. As COHORT was a small programme, the programme events were tied to the larger pHealth programme until the latter concluded at the end of 2019.

The six consortia selected for funding are presented below.

5.1. Early Determinants of Cardiovascular, Metabolic and Reproductive Health

The project was led by Harri Niinikoski from the University of Turku, and the consortium was granted 797,033 euros in funding. The other partners were from the Universities of Tampere and Oulu.



The project included rich data from the Special Turku Coronary Risk Factor Intervention Project (STRIP), the diabetes prediction and prevention study (DIPP) and Boy cohorts with follow-up of several decades up to date, from birth to midlife. The project set to examine the contributions of early nutrition and environment to cardiovascular and metabolic health, asthma and allergy, beta cell autoimmunity, and the effects of early testicular function on metabolic morbidity. An additional focus was on the role of overweight in sex hormone production.

The data used consisted of questionnaire data, laboratory analyses and physical measurements, pubertal development, laboratory data and clinical data. All research data were pseudonymised. As all of the data are planned to be used long-term, they have been securely stored to be used during the decades to come. The most important finding of the project was the development of a model that estimates the onset of puberty using growth data only. That is, pubertal staging is not needed. The project has continued after the funding period, addressing the remaining aims.

5.2. Finnish Psychiatric Birth Cohort Consortium

The project was led by Andre Sourander from the University of Turku. The consortium was granted 795,685 euros in funding. The other partners were from the University of Oulu and the Finnish Institute for Health and Welfare.

The consortium successfully harmonised seven Finnish birth cohorts: the Northern Finland Birth Cohorts, the Finnish 1981 Birth Cohort Study, Finnish Prenatal Studies, the 1987 and 1997 Birth Cohorts, the SSRI pregnancy cohort and the Southwest Finland Birth Cohort.

The main strength of the consortium was that the harmonisation of the cohorts enabled examining the same psychiatric phenomenon across all of the cohorts. More specifically, it also enabled addressing life-course development of psychiatric disorders and identifying issues related to such development. Data were available from 1960s onwards, which allowed for an examination of time trends between different cohorts and between time points.

5.3. Connecting Northern Finland Birth Cohorts with Borealis and THL Biobanks

The project was led by Minna Ruddock from the University of Oulu. The consortium was granted 744,389 euros in funding. The other partners were from Oulu University Hospital and the Finnish Institute for Health and Welfare.

The project used data from the Northern Finland Birth Cohorts (NFBCs), the THL Biobank and Biobank Borealis of Northern Finland. The harmonised data provided an opportunity to study healthy ageing, risk factors and preventive means to various diseases related to ageing. The datasets are advertised in public catalogues, at <https://site.fingenious.fi/en/all-partner-biobanks> and <https://directory.bbmri-eric.eu/#/>. While the datasets

themselves cannot be made publicly available, the information about them and the work done to harmonise them can attract future research projects to continue utilising them.

5.4. Longitudinal Birth Cohort and Family-Based Studies of Morbidity, Mortality and Social Disadvantage: A Consortium Application on Sharing Data, Family Designs and Methods

The project was led by Pekka Martikainen and Jaakko Kaprio from the University of Helsinki. The consortium was granted 799,983 euros in funding.

The project used several datasets. Firstly, the project used EKS70 data, which comprise national administrative register data from various sources and enable linking individuals to their biological parents and children and to other family or household members. Secondly, the project used the TwinRegistry, which is the older Finnish Twin Cohort. These data comprise twins who were born before 1958 and survey data from 1975, 1981, 1990 and 2011. Thirdly, the project used data that enabled use of family designs from a molecular genetic perspective. The data were from the FINRISK surveys in 1992, 1997, 2002, 2007 and 2012, Health2000/2011 as well as FinHealth 2017.

A highlight of the results produced by the time of the interview was confirming the multigenerational associations of health and social disadvantage. Thus, a socially disadvantaged background in the childhood family has been linked to later mental health problems. Additionally, the results highlight both the genetic and environmental determinants of cardiovascular diseases, largely independent of each other. Of special note is that the project successfully built data structures and that the PI of the project has been able to use them to apply for further extensive funding to continue this work.

5.5. Finnish Platform of Birth Cohorts: Microbial Interactions and Health

The project was led by Juha Pekkanen from the University of Helsinki. The consortium was granted 799,562 euros in funding. The other partners were from Helsinki University Hospital, the University of Oulu, the Finnish Institute for Health and Welfare, and Kuopio University Hospital.

The consortium generated a nationwide collaboration and database network to work long-term. This was achieved by harmonising several birth cohort datasets from both Southern and Northern Finland.

Moreover, new samples were collected using common protocols for future research to use. The network further made a platform that makes it possible to examine the contributions of early-life and environmental exposures to the development of allergic diseases.

Sequence data are available upon publication. Raw Illumina data are available at <https://www.ebi.ac.uk/ena/browser/view/PRJEB39137>.

Additionally, codes and programs are available regarding the metadata set of birth cohorts at <https://github.com/katrikorpela/mare>. The data repository is openly available at <https://github.com/katrikorpela/mare> and at www.ebi.ac.uk/ena/browser/view/PRJEB39137.

5.6. Sleep Patterns in Neuropsychological Development in Early Childhood – The Interplay of Genes and Environment

The project was led by Juulia Paavonen from the Finnish Institute for Health and Welfare. The consortium was granted 712,187 euros in funding. The consortium partner was from the University of Turku.

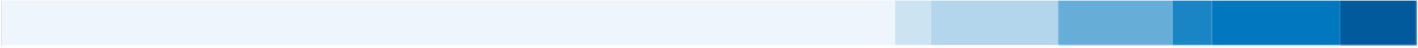
The project set to examine the significance of sleep in the development of self-regulation and cognitive and social skills in early childhood, also considering the role of environmental and genetic risk factors. Two cohort studies were pooled to enable cross-cohort comparisons and examination of similar phenomena. The cohorts were two pregnancy cohorts: the CHILD SLEEP cohort and the FinnBrain Birth Cohort Study. These provided large numbers for the complex statistical analyses. The data are not publicly available. The key results of the project highlight that it is important to aim for early detection of childhood sleep problems and to target interventions to tackle them. The project has received further funding to continue its work.

6. Programme achievements

There are many research groups studying common diseases in Finland using different registers, but there could be even more cooperation between them. Different studies have been initiated but the data collection is not uniform. There are data about similar questions, but if they are not collected in the same way, they cannot be directly combined.

The COHORT programme was intended to pilot a new programme concept. Rather than awarding funding to separate or individual projects built around new research hypotheses, the purpose was to facilitate the **integration** of research projects with adequate basic research funding already in place. The aim was to increase cooperation between different research partners building new networks of researchers and clusters of harmonised datasets and cohorts for register-based research.

COHORT enabled the structuring, harmonisation and combination of several datasets from different sources such as surveys, health data registries, biobanks and hospitals. This created an infrastructure that can be used to address questions that would not be possible to answer using individual datasets. The methods and protocols on how to standardise data that were developed within the programme can be applied to other cohorts to create compatible, homogenised data. The networks formed within and between the consortia create opportunities for current and future collaborations. The collaborations within a consortium enabled the use of more advanced



statistical methods and applying machine learning and predictive risk models from other research fields. This in turn increased the visibility and reputation of the research and attracted new PhD candidates and postdoctoral researchers to take the research forward. The joint funding of the consortia, compared to individual research group funding, enabled collaborations that likely would not have happened otherwise.

The funding made it possible to attract other funding for some of the consortia, for instance from the NIH and ERC Advanced Grants.

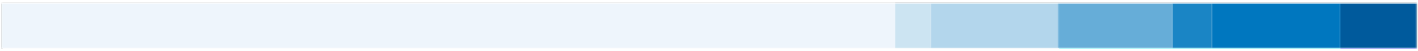
Some projects recognised that the funding period was too short to start conducting research using the collaborative infrastructure they had created. Although the creation of new collaborations based on a joint infrastructure was not the main intention of the programme, this is definitely a bonus. Overall, it can be concluded that the intention of the programme, to create an infrastructure of overarching and sustainable datasets for future health research by amalgamating and harmonising existing datasets nationwide, was met.

A major achievement expressed by most consortia was the **harmonisation of different datasets**. For example, it enabled cross-cohort comparisons and examination of the same phenomena across cohorts, which improved the generalisability. One example was the harmonisation of data from a large number of psychiatric cohort samples to address key questions about longitudinal outcomes, the most important being prenatal epidemiology and findings on possible roots of psychiatric conditions. Another example was the harmonisation of Finnish cohorts in combination with hospital data, as well as national register data and questionnaire data from participants, creating very large datasets where all individuals have the same pseudonyms throughout.

Other achievements included **creating harmonised datasets, standardised ways of sample collection and collaborative networks**. All consortia reported that the projects were successful and that funding them as a consortium instead of individual projects, enabled them to achieve more and to do work that would not have been possible otherwise. Thus, they could build stronger networks with stronger, harmonised data. Many of these new and **harmonised datasets are resources that can be used in future projects** also by other researchers.

7. Challenges

Register-based studies require profound understanding in advanced study designs, statistical methods, confounders, moderators and mediators. These are essential to discover possible mechanisms of adversities, wellbeing, help-seeking, time trend changes, etc. Understanding these aspects is essential for interventions, early detection, prevention, and development of effective health/social/educational services for families, children and adolescents.



The danger is that weak study designs or inadequate data may introduce a number of biases that require substantial experience to manage. Register-based studies rely on institutions with long-term experience of performing such studies. Therefore, the idea to create multicentre consortia is excellent, but this kind of cohort funding needs to have a plan for long-term commitment.

7.1. Harmonisation of data

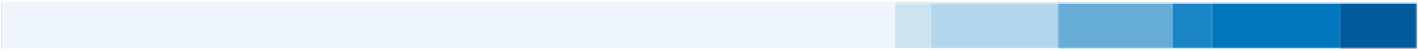
Harmonising data took much longer than most consortia initially anticipated. Although there are data about similar questions, the datasets could not be directly combined as the measures were not identical or the collected measures did not actually measure the same thing. For instance, study samples may have been collected at different timepoints, or from different subgroups, context, etc.

7.2. Retrieving registry data

All consortia that used register-based data witnessed that retrieving the administrative data for research purposes was a very slow process, and the researchers could do very little or nothing to speed it up. The process has become even more difficult after the programme. All consortia were very aware of how challenging it is under current law to continue any register-based research or, in particular, to apply for funding for it. Some projects were “delayed” by 1–2 years, as they waited for register data they had applied for from register data holders for the purposes of their project. It is very rare to apply for data from separate register data holders, as they can only grant permission to use their own data, which cannot be linked to any other register data sources.

Most projects used register data from several register data holders. All such applications have to be made to the Finnish Social and Health Data Permit Authority Findata (<https://findata.fi/en>). The application is complex and detailed. For example, all variables must be known exactly, and it takes a long time to search for them in the catalogues. Also, a lot of other documentation is needed. Drafting this complex application and revising it, if needed, and waiting for its processing, reduces the time available to work on the actual project.

The fee for applying for a decision and providing the data is expensive, and many worthwhile projects may not have sufficient funding for it. After the data is finally delivered, the project must also have funding to pay for storage and for placing the analyses on a secure server (<https://findata.fi/en/kapseli/>). One can imagine how little time there then actually is for the actual conduction of the project, when such a long time needs to be spent making another application to receive the data, and pay for that, before the research can even begin.



Under current legislation, the data should be provided within two months, but it altogether takes a lot longer to get the data application approved and access to the data (Act on Secondary Use of Personal Data in Social and Health Care 2019/552, section 5, subsection 48). This is also due to that the data should be processed on a highly secure server, and the rules are so strict that most universities and research institutes do not currently fulfil them. For example, they cannot provide 24/7 security surveillance.

7.3. Permits and legal issues

There is confusion and even justified anxiety in Finland around how to interpret the GDPR, which affects the combining of register data from different sources. This has led to delays in getting agreements in place. One COHORT consortium had to have multiple separate agreements, one for each data source. This is a severe problem that hampers Finnish register-based research, especially longitudinal studies.

One problem with combining consented data from two different cohorts is that it may be interpreted as a new cohort which requires a new consent, which is usually impossible to obtain (the original datasets can have been collected a long time ago, and the initial participants cannot be reached). One solution is to use the datasets in parallel, repeating the analyses in each cohort and combining only the results as aggregated estimates.

7.4. Funding

Generally, the opinion of the consortia was that the funding period was too short. Continued funding would likely have helped them all to continue the collaboration and produce more results. To apply for continued funding is difficult since there is not much funding for these kinds of collaborative research infrastructures. New funding usually also requires formulation of a new research question. Some, but not all, consortia had been successful in applying for and receiving funding to continue their work.

7.5. Application process

There were several specific complaints about the Research Council of Finland online system for applications that appear to already have been rectified. Specifying the budget for each year was considered difficult, since it is hard to estimate when certain costs will take place, in particular because of unpredictable waiting times. Writing a report after four years was also considered difficult. It would have been easier if the Research Council had requested interim reports.

8. Advice and suggestions for the future

8.1. Ministerial-level suggestions

- Make performance-based financing available for register holders to decrease data processing times, and so that data applications can be processed within the two months stipulated by law.
- Introduce research funding rules that would allow for a possibility to apply for a no-cost extension of a project. This is an important point due to the unpredictable waiting times to receive permissions and data.

8.2. General recommendations to the Research Council of Finland

- Introduce longer funding periods and an opportunity to apply for a continuation of funding for particularly successful projects, especially if the established resource is likely to be useful also for others and would benefit from a longer follow-up.
- In infrastructure calls, include parallel funding for research projects that use the infrastructure.
- Establish funding mechanisms for so called “soft infrastructure”, for example for expanding databases and cohorts with additional data or longer follow-ups.
- Arrange national and international conferences that could promote research interactions and initiate international cohort collaboration.

8.3. Programme-specific reflections and suggestions

- It would have been useful to arrange more COHORT-specific events, such as a kick-off seminar, interim meetings and an in-person final meeting to interact with the Research Council of Finland and the other consortia. Now, most programme events were tied to the larger pHealth programme, and the only COHORT-specific event was the final meeting, which was held online due to Covid-19 restrictions.
- Continuing the kind of evaluation process introduced in this call, with post-project interviews, was recommended.
- More feed-back on the application would be helpful. A suggested format is constructive feed-back from the evaluation panel, with an added comment from the Research Council of Finland.
- Requesting interim reports, preferably yearly reports, would facilitate writing the final report at the end of the project.

8.4. Recommendations to other stakeholders

- There should be specified maximum handling times for all procedures involved in obtaining data and permissions. The knowledge in the scientific community of such rules, including how to file complaints and where to find statistics on monitoring of handling times, should be promoted.
- Self-regulation by the registry holder(s) is suggested.
- GDPR-related problems appear to be mostly caused not by the law itself but by insecurity of how the rules should be interpreted. An established interpretation of GDPR in line with the international interpretation would greatly further Finnish register-based research. The process to arrive at an established policy of interpretation could be promoted if important, prejudicial cases were collected and promoted.
- This could be accomplished by scientific networks or institutes (e.g. university departments in medical law) or possibly by government agencies. Research funding by the Research Council of Finland might precipitate initiation of such a process.
- Introduce sustainability (“safety net”) grants for the maintenance of existing infrastructures, for example servers and security, to update data and employ, for instance, a permit and data manager to ensure that the infrastructure is not destroyed. Such grants could be quite limited in terms of amounts but could save valuable resources for the future. These could be launched by the institutes that are responsible for the research infrastructure. The Research Council of Finland could promote a process to establish local sustainability grants, if co-financing by applying universities (that could be used for sustainability grants) was mandated.

9. Concluding remarks

Finland could be an “El Dorado” for register-based studies, but this requires specific and long-term financing including, but not limited to, sustainability grants for maintenance of existing soft infrastructures for register-based research. Intervention trials using well documented, standardised transgenerational data and well-designed cohorts could be a future specialty of Finnish register-based research. Intervention studies based on cohorts that are stringently designed, population-based and use modern technologies should be promoted.

Large-scale cohort studies, with opportunities for decades of follow-up, pinpointing risk groups, studying (rarer) diseases with a long latency, and questions where it is not possible to conduct an intervention due to ethical or moral reasons or studying also intergenerational effects, and using very large datasets enabling more complex analyses will also be very important for the continued generation of new knowledge.

Appendix A: Evaluation questions and material for the panel

COHORT Programme evaluation

Autumn 2022

INTERVIEW QUESTIONS:

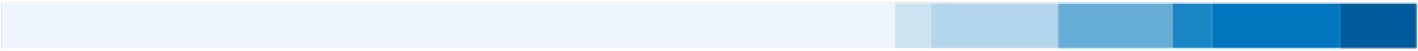
GENERAL NOTE: While it is also important to document what was done during the programme, for the Academy the most interesting question is what we can learn from the programme going forward.

POTENTIAL OF REGISTER-BASED STUDIES

- what opportunities do the register-based studies offer for research A) nationally B) internationally?
- which factors (apart from the funding itself) enhance OR make possible to conduct register-based studies? What factors hinder it?

COHORT PROGRAMME: ACHIEVEMENTS, CHALLENGES, IMPORTANCE & INTERACTIONS

- what were the major new achievements / results / breakthroughs in your consortium? What surprised you, positively or negatively, in your project? What was left undone and why?
 - NOTE: we would like to include 1-2 success stories into the report for human interest angle; in contrast, any misfortune/adversity (coronavirus epidemics for example) we discuss in the report should be written in such fashion that consortia/groups/individuals are not identified
- did following your research plan increase the integration between your groups / projects / data / institution as expected? What aspect of integration worked especially well / did not work as well as expected?
- Were there problems regarding the integration of different data sets? Was this sufficiently taken into account at the original planning stage?
- What was the added value of consortium funding? i.e. compared with giving individual funding to each research group separately, what were the advantages / disadvantages of COHORT programme funding? Would something have not been achieved with individual funding? Would



something extra have been achieved by funding individual groups instead?

- Was there any interaction between separate COHORT consortiums during your funding period? Would it have been beneficial if the Academy facilitated such interaction and if so, how?

MOVING FORWARD FROM THE PROGRAMME

- Did COHORT funding open new research or funding pathways for you?
- In which other ways could the integration between projects / groups be improved?
- What other ways do you envision to fund register-based studies? (both Academy and non-Academy funding, domestic and international)
 - from which sources have you / the groups received funding?
- apart from funding, what other ways do you envision to enhance / promote register-based studies?
- anything else you would like to bring up?

Appendix B: Projects funded in the Cohort call 2016

Funding period: 2017–2020 (excluding extensions due to Covid-19 epidemic)

Total funding: € 4.6 million

Early determinants of cardiovascular, metabolic and reproductive health (EDCar)

- consortium funding 797 033 €
 - Harri Niinikoski, University of Turku
 - Helena Virtanen, University of Turku
 - Suvi Virtanen, Tampere University
 - Riitta Veijola, University of Oulu

Finnish Psychiatric Birth Cohort Consortium (PSYCOHORTS)

- consortium funding 795 685 €
 - Andre Sourander, University of Turku
 - Juha Veijola, University of Oulu
 - Mika Gissler, The Finnish Institute for Health and Welfare (THL)

Identifying trajectories of healthy aging via integration of birth cohorts and biobank data (CoCoBi)

- consortium funding 744 389 €
 - Minna Ruddock, University of Oulu
 - Pia Nyberg, Oulu University Hospital
 - Kaisa Silander (previously Tiina Wahlfors, originally Anu Jalanko), The Finnish Institute for Health and Welfare (THL)

Longitudinal birth cohort and family based studies of morbidity, mortality and social disadvantage (BIRTHFAM)

- consortium funding 799 983 €
 - Pekka Martikainen, University of Helsinki
 - Jaakko Kaprio, University of Helsinki

Finnish Platform of Birth Cohorts: Microbial Interactions and Health (FINMIC)

- consortium funding 799 562 €
 - Juha Pekkanen, University of Helsinki
 - Pirkka Kirjavainen, The Finnish Institute for Health and Welfare (THL)
 - Willem de Vos, University of Helsinki
 - Juha Auvinen, University of Oulu
 - Seppo Heinonen, Helsinki University Hospital
 - Heikki Lukkarinen, Turku University Hospital
 - Katri Backman, Kuopio University Hospital

Sleep patterns in neuropsychological development in early childhood – the interplay of genes and environment (Sleep in Development)

- consortium funding 712 187€
 - Juulia Paavonen, The Finnish Institute for Health and Welfare (THL)
 - Marja Karlsson, University of Turku