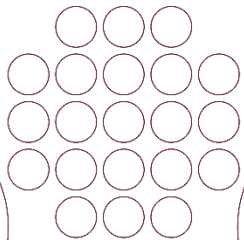


Nationwide enrichment of biological sample data with data from health registers – the Danish Biobank Register search system

Dr. Bartłomiej Wilkowski

Head of the Danish Biobank Register

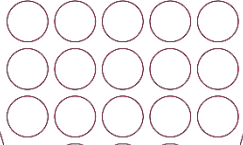


About me

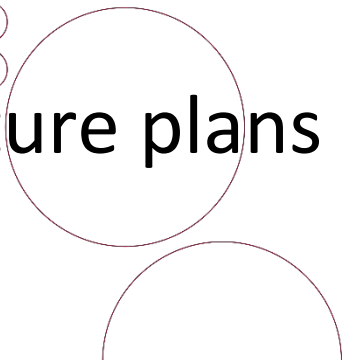
- Bart Wilkowski:
 - Employed at the Danish National Biobank since June 2011
 - Head of the Danish Biobank Register system
 - Ph.D. in Biomedical Informatics (Technical University of Denmark)
 - M.Sc. in Telecommunications and Computer Science (Technical University of Lodz, Poland)

- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- Related work
- Future plans

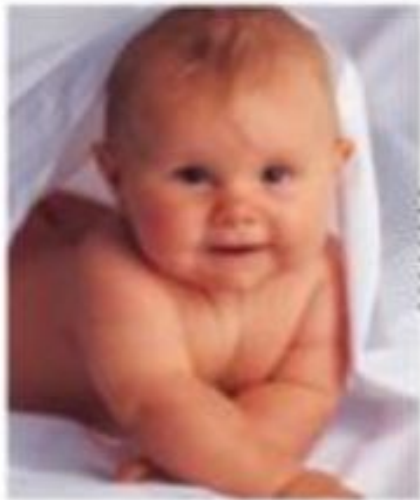
➤ Related work



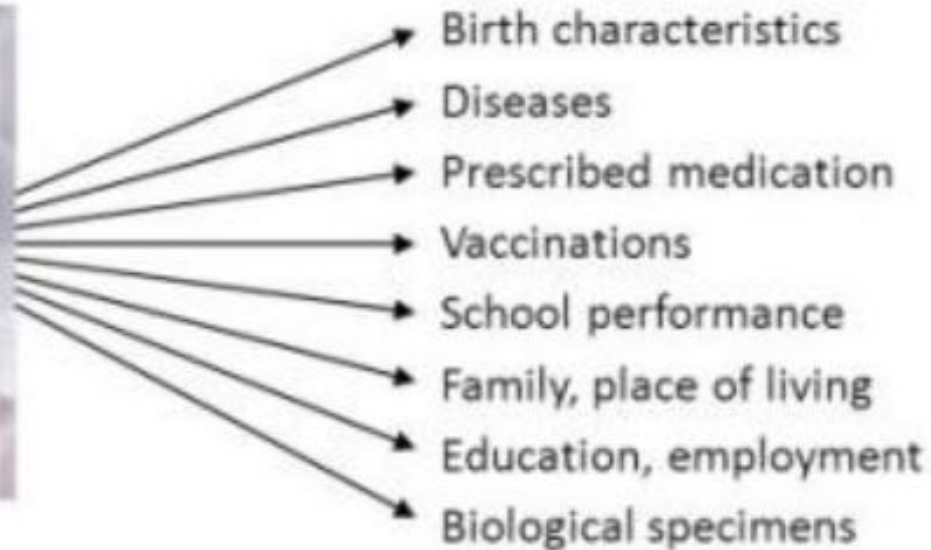
➤ Future plans



Denmark: one key, many registers



CPR number



- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- Related work
- Future plans

➤ Related work

➤ Future plans

Background

- The Danish National Biobank is placed at Statens Serum Institut, a national laboratory for more than 100 years
 - Collections with many (millions) of biological samples
 - Increasing demand from researchers
 - Overview and structure was needed

Funds were raised to establish

- National Biobank that offers state-of-the art and well monitored freezing units and laboratory facilities
- Danish Biobank Register (SSI samples and other joining biobanks)

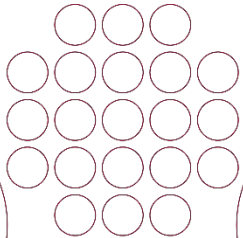
ново nordisk fonden

Biobank construction 2011-12

DANMARKS
NATIONALE
BIOBANK



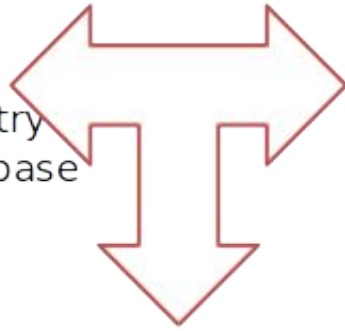
- 3. level – Epidemiological department
- 2. level – Biobank laboratories and offices
- 1. level – Sample storage
- Basement – Storage rooms og walk-in-freezers



The Biobank research potential

REGISTRIES

- Civil Registrations System
- Danish National Patient Registry
- Danish Family Relations Database
- Danish Medical Birth Registry
- National Pathology Registry
- Register of Medicinal Product Statistics
- Vaccination Registry
- Demographic Registries
- And many more...



BIOLOGICAL SAMPLES

- Serum/ Plasma
- Buffy Coat
- Whole Blood
- Filter paper blood spots
- Spinal fluid
- Amniotic fluid

Research in health

EPIDEMIOLOGY

When an Entire Country Is a Cohort

Denmark has gathered more data on its citizens than any other country. Now scientists are pushing to make this vast array of statistics even more useful.

Science. 2000 Mar 31;287(5462):2398-9

The concept of the Danish National Biobank

Danish National Biobank



www.nationalbiobank.dk

Danish Biobank Register

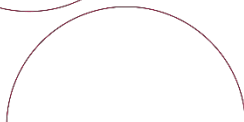
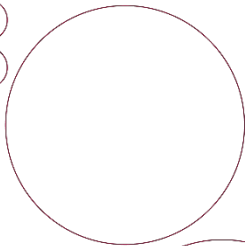
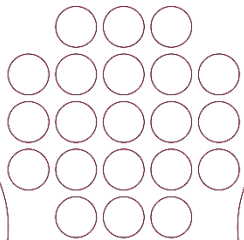


www.biobanks.dk

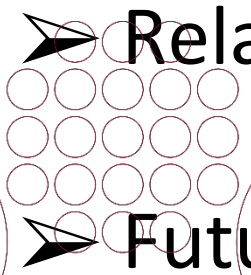
The Coordinating Centre



DANMARKS
NATIONALE
BIOBANK



- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- Related work
- Future plans

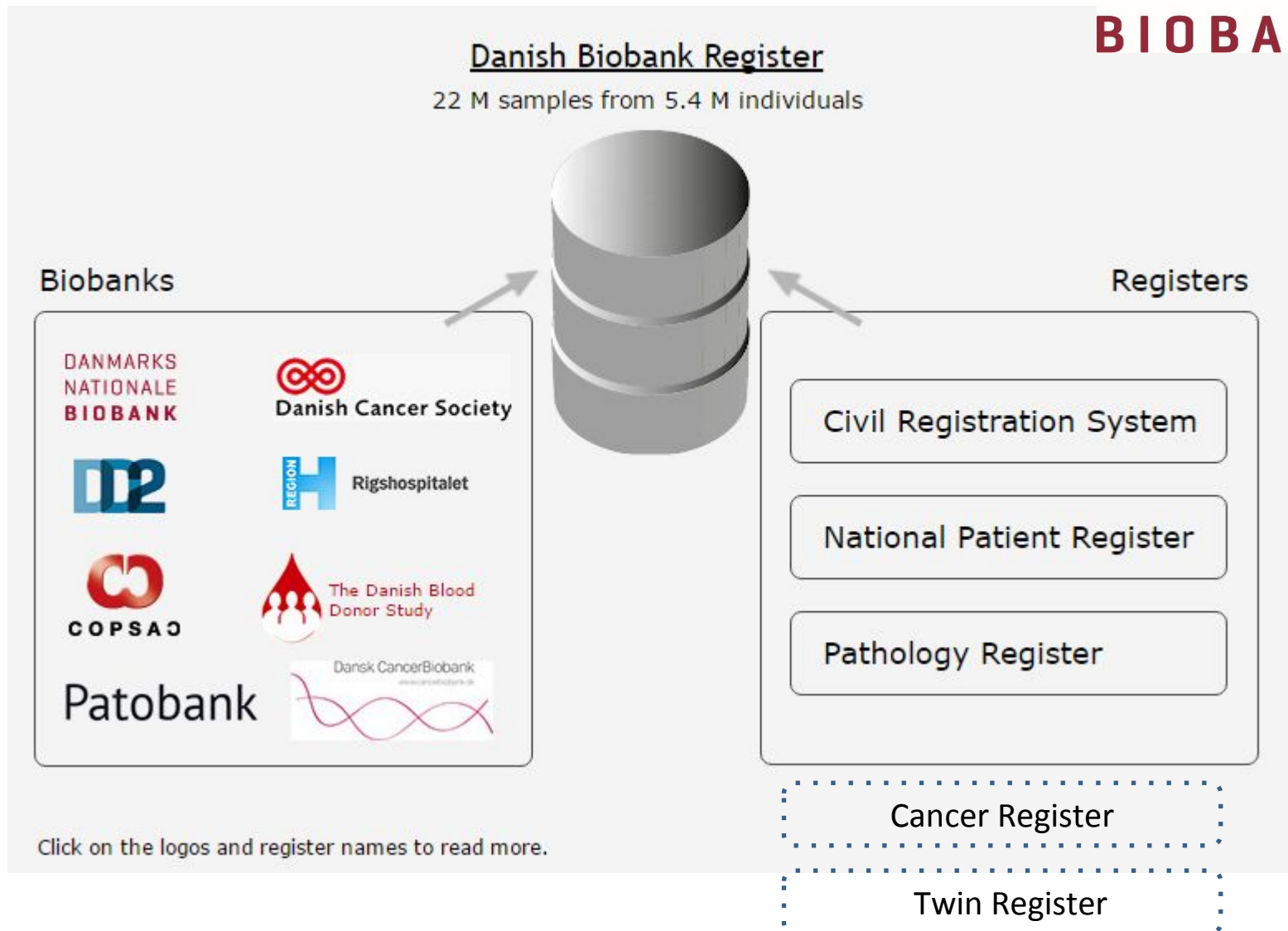


Danish Biobank Register

- Web-based search system (released on March 21, 2012)
 - Quick and easy overview of samples in Denmark
- Support for register-based research
 - Efficient linkage of
 - administrative registers (CPR)
 - diagnosis registers (LPR, PAT)
 - samples
- Results: pseudo anonymized, aggregate form
- Shortened from months to minutes the process of identifying enough pre-existing samples to support new research projects
- Part of Nordic Center of Excellence in Health Related e-Sciences (NIASC) – the Nordic Biobank Register initiative

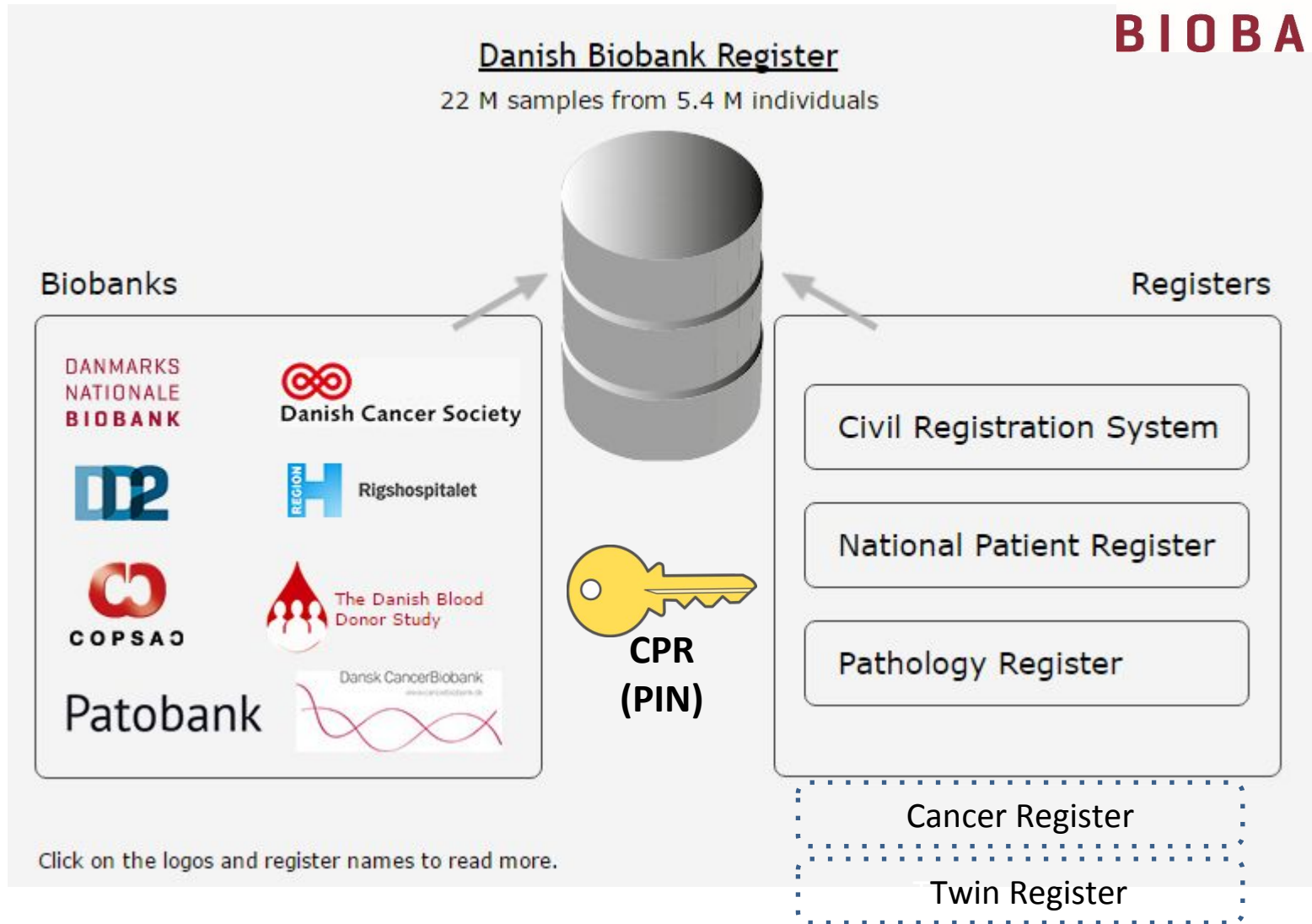
Danish Biobank Register

DANMARKS
NATIONALE
BIOBANK



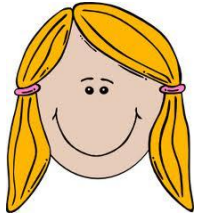
Danish Biobank Register

DANMARKS
NATIONALE
BIOBANK

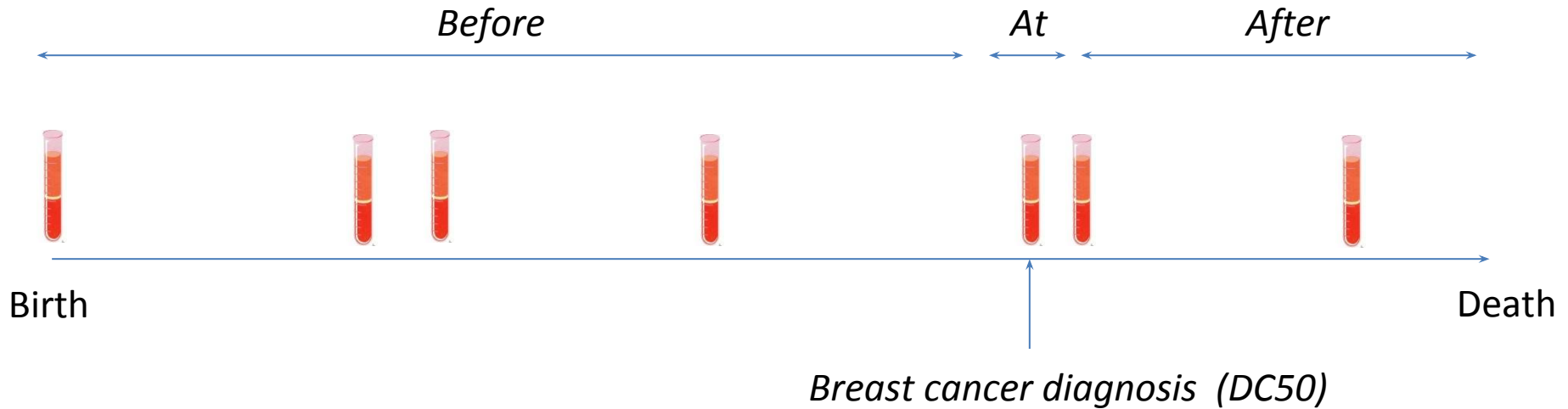


~ 160 million diagnoses
(ICD 8,10 and SNOMED)

Person



DANMARKS
NATIONALE
BIOBANK



The Danish Biobank Register system gives an overview of samples and diagnoses from birth to death. It links samples and diagnoses of individuals, and highlights which samples were taken before, at or after a diagnosis.

Participating biobanks (22 million samples from 5.4 million people)

7.500 samples



Danish centre for strategic research in type 2 diabetes

3.5 million samples

DANMARKS
NATIONALE
BIOBANK



STATENS
SERUM
INSTITUT



Bedre Sundhed for mor og barn

230.000 samples

Dansk CancerBiobank
www.cancerbiobank.dk



17 mill. samples

Patobank



170.000 samples

Danish Cancer Society



900.000 samples

The Danish Blood Donor Study

10.000 samples

COPSAC
Copenhagen Studies on Asthma in Childhood

220.000 samples



Rigshospitalet

Copenhagen Hospital Biobank

Person

Diagnosis

Sample

Report Definition

Finish

General info

Gender



Any

Country of birth

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Country of birth (mother)

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Country of birth (father)

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Year of birth

From

1850

To

2015

Age at sample taken (expressed in years)

From

0

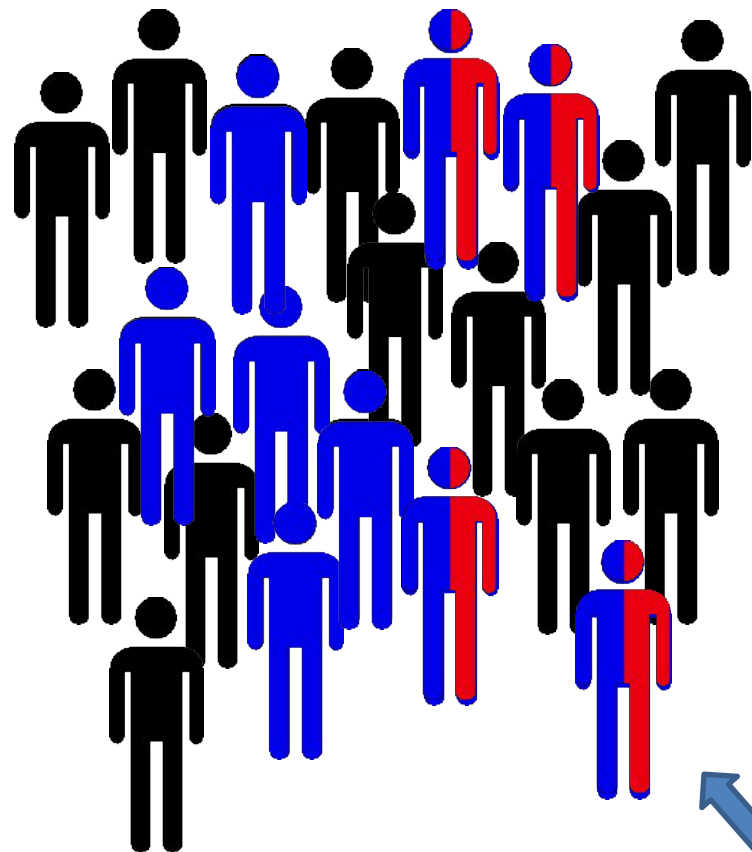
To

150

www.nationalbiobank.dk
www.biobankdenmark.dk

– DNB homepage

www.biobanks.dk – direct access



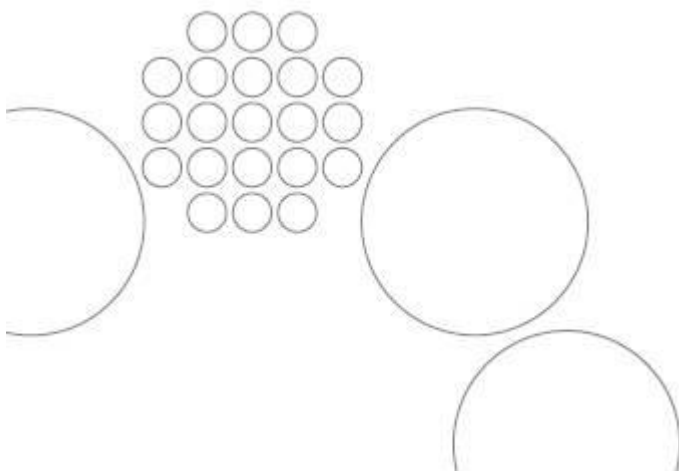
Diagnosis of breast cancer

ICD10: DC50

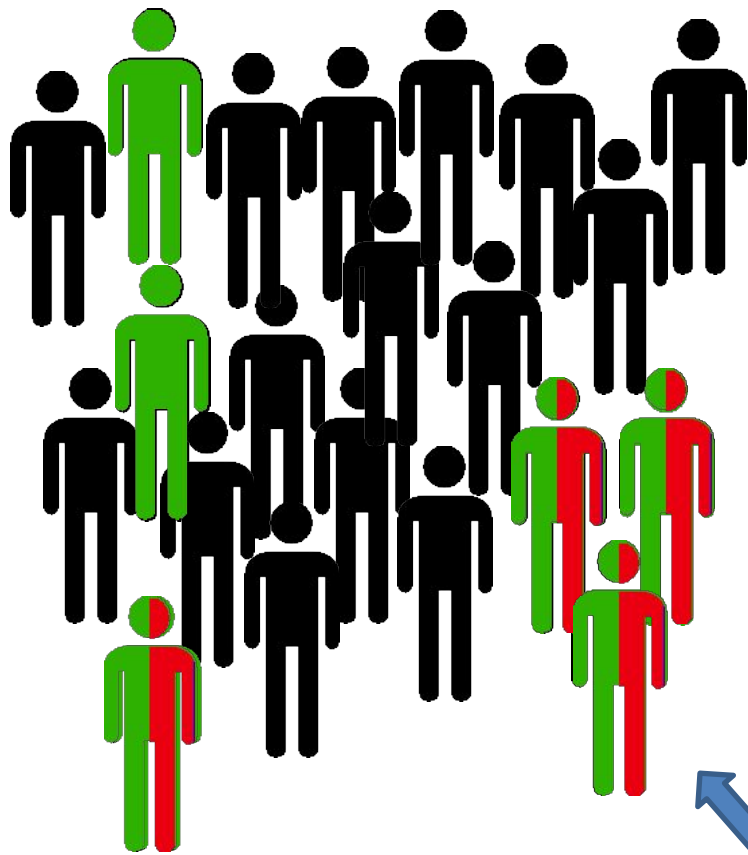
ICD8: 174

Plasma samples taken before the
diagnosis of breast cancer
(≥ 365 days before the diagnosis)

CASE group



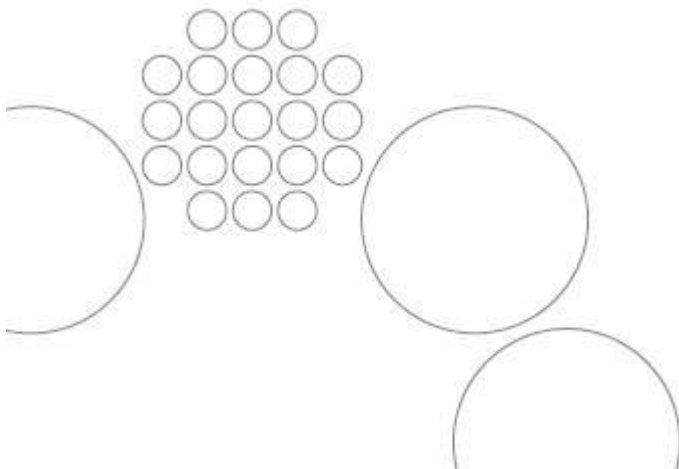
DANMARKS
NATIONALE
BIOBANK



Persons without a history of
breast cancer
(except ICD10: DC50*,
ICD8: 174*
)

Serum samples from the Danish
National Biobank

CONTROL group



Get samples out

Obtain approval from a Health
Research Ethics Committee



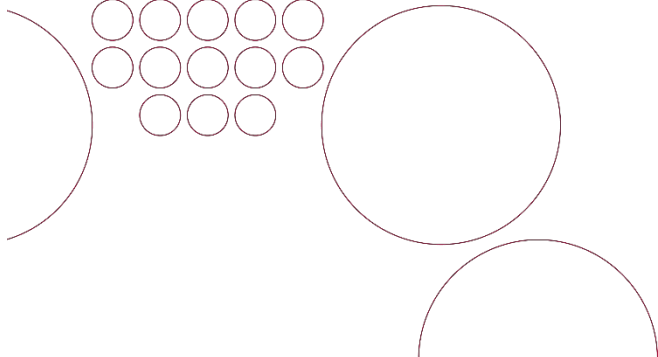
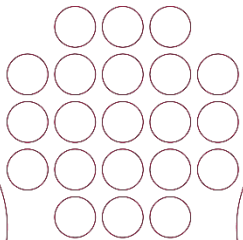
Apply to Scientific Services at
Sundhedsdatastyrelsen



Applications are evaluated by the
Scientific Board



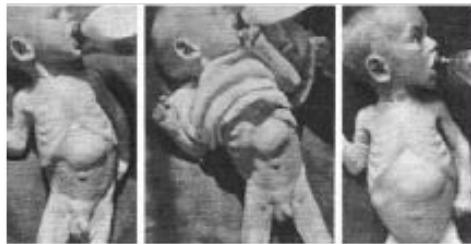
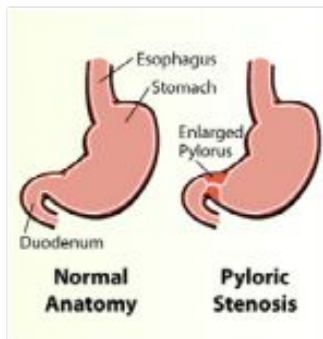
Samples are handed
out



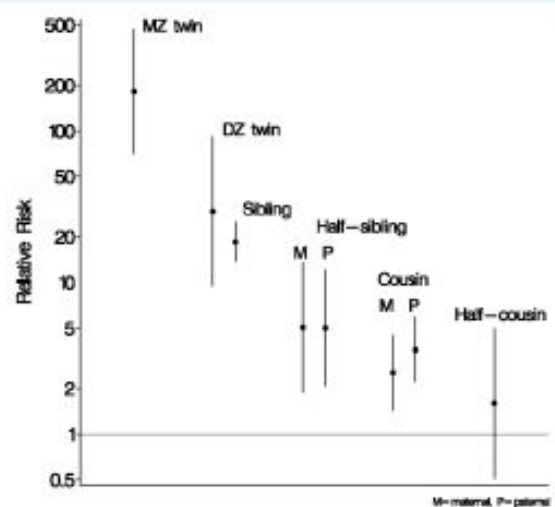
Pyloric stenosis: An example

1. Background

- Disease in newborns
- Muscle enlargement:
- Food cannot pass.
- Most surgeries in newborns



2. National registries - linkage study



National registries info on:

- Family-linkage info on all Danes (up to 4 generations)
- All surgery-confirmed cases in Denmark

3. Biobank study

A. Biobank registry

Identify blood sample on 1,001 pyloric stenosis cases and 2,401 controls

B. Biobank

In 2 months: identify samples and extract DNA

C. Genetic analyses

Identified 3 important gene variants involved in disease development

D. Published

Nature Genetics 2012

Bjarke Feenstra et al.

- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- [Related work](#)
- Future plans

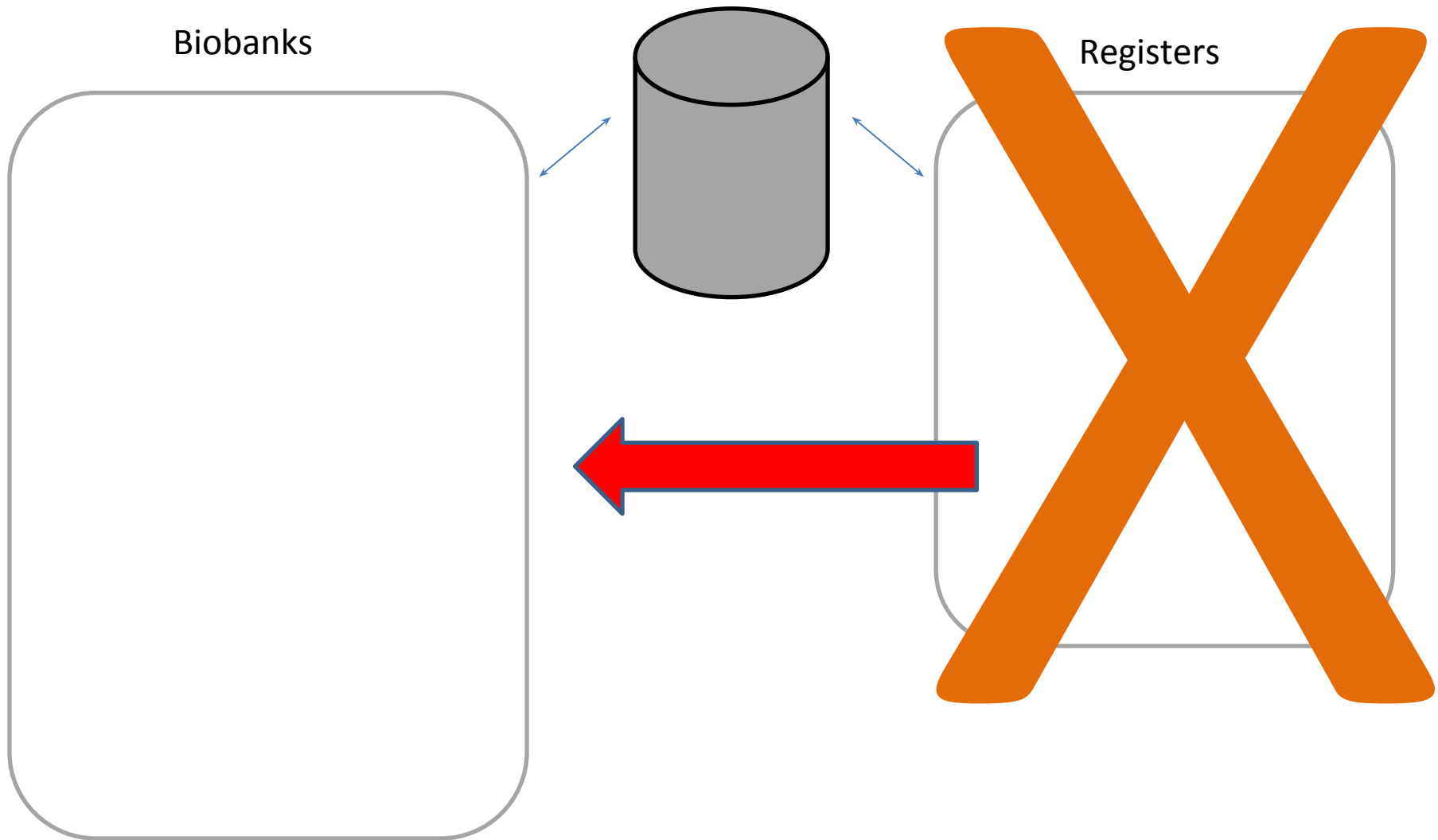
- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- Related work
(clone:Swedish Biobank Register)
- Future plans

Swedish Biobank Register

Nordic Center of Excellence
in Health-Related e-Sciences



Swedish Biobank Register



Swedish Biobank Register

Nordic Center of Excellence
in Health-Related e-Sciences

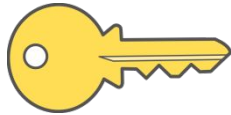


Swedish Biobank Register

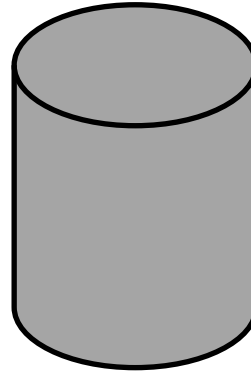
Biobanks

STHLM2

prostate cancer study
(samples + diagnoses)



Local UUID



<http://biobanks.se>

Swedish Biobank Register

Nordic Center of Excellence
in Health-Related e-Sciences



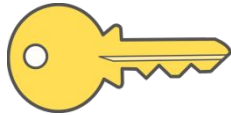
Swedish Biobank Register

~ 174,000 samples from 78,500 individuals

Biobanks

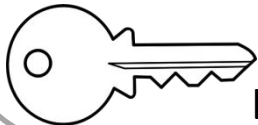
STHLM2

prostate cancer study
(samples + diagnoses)

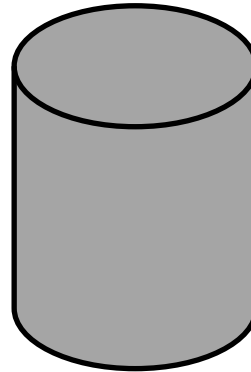


Local UUID

Svenska
Tvillingregistret
(samples +
diagnoses)



Local UUID 2



<http://biobanks.se>

Swedish Biobank Register

Nordic Center of Excellence
in Health-Related e-Sciences



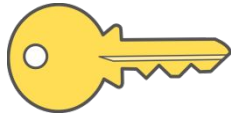
Swedish Biobank Register

~ 174,000 samples from 78,500 individuals

Biobanks

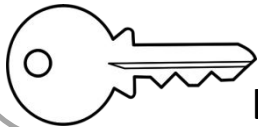
STHLM2

prostate cancer study
(samples + diagnoses)

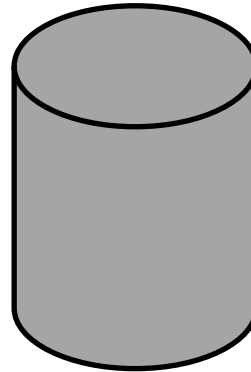


Local UUID

Svenska
Tvillingregistret
(samples +
diagnoses)



Local UUID 2



<http://biobanks.se>

107,000 samples; 24,500 individuals

67,000 samples; 54,000 individuals

Swedish Biobank Register

Nordic Center of Excellence
in Health-Related e-Sciences



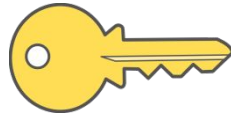
Swedish Biobank Register

~ 174,000 samples from 78,500 individuals

Biobanks

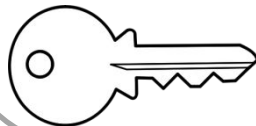
STHLM2

prostate cancer study
(samples + diagnoses)

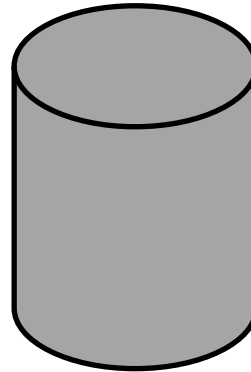


Local UUID

Svenska
Tvillingregistret
(samples +
diagnoses)



Local UUID 2



<http://biobanks.se>

107,000 samples; 24,500 individuals

1.2 million diagnoses
(ICD7,8,9,10 and SNOMED)

67,000 samples; 54,000 individuals

Prostate cancer search example

Swedish Twin Register - overview

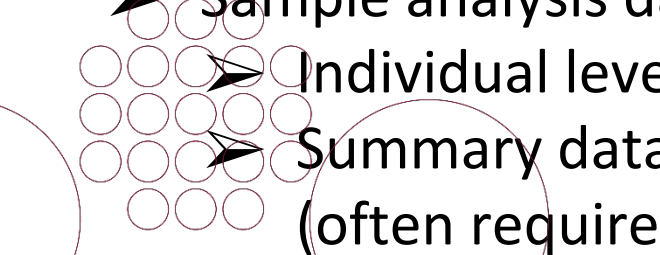
		Complete pairs		Incomplete pairs
		One diagnosed	Both diagnosed	
Monozygotic twins	Female	1	0	0
	Male	113	40	42
Dizygotic twins - same sex	Female	0	0	0
	Male	199	29	92

	with diagnosis	Complete pairs	Incomplete pairs
Dizygotic twins - opposite sex	only female twin	0	0
	only male twin	219	97
	both twins	0	N/A

Zygosity	Gender	No diagnosis (controls)	
		Complete pairs	Incomplete pairs
Monozygotic twins	F	4795	1116
	M	3379	1125
Dizygotic twins - same sex	F	4141	1686
	M	3271	1693
Dizygotic twins - opposite sex	F	5731	2787
	M		1748

- Introduction: registers in Denmark
- Introduction: Danish National Biobank
- Danish Biobank Register
- Related work
- **Future plans**

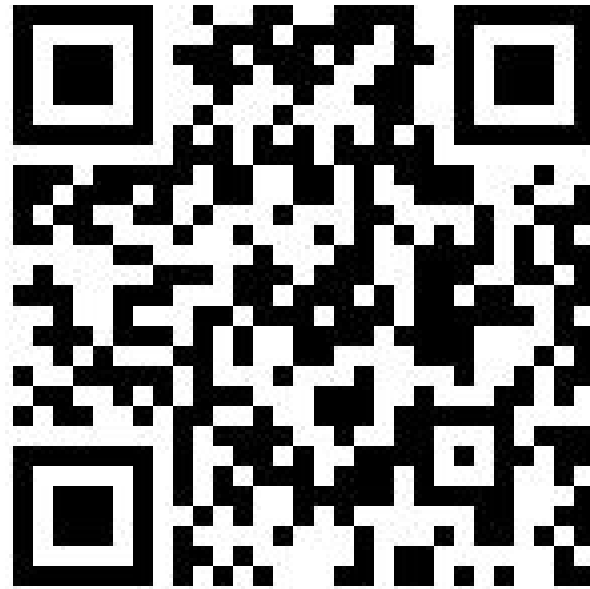
Future

- Adding more data (biobanks & registers) to the existing systems (Danish and Swedish Biobank Register)
 - Cloning the system in other countries (Faroe Islands, Norway, Finland?, Estonia?, etc.)
 - Expanding current Danish Biobank Register functionalities:
 - Sample's body location - a criteria for tissue samples
 - Sample analysis data from a future databank, aka local EGA
 - Individual level data
 - Summary data from the completed research projects (often required by funding agencies)
- 

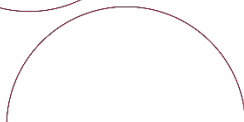
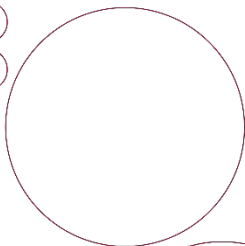
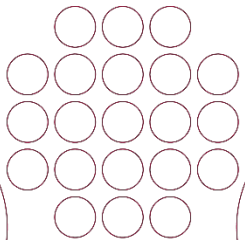
Thank you!

DANMARKS
NATIONALE
BIOBANK

BAW@ssi.dk



www.danishnationalbiobank.com



Research projects

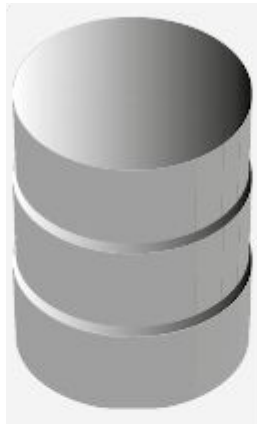
DANMARKS
NATIONALE
BIOBANK

Parkinson's
disease



Inflammatory
bowel disease

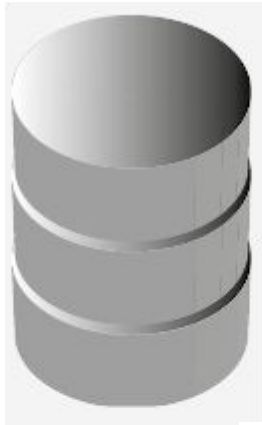
Danish Biobank Register



1. Parkinson's disease (ICD10: G20 eller ICD8: 342)
2. Inflammatory bowel disease (Mb. Crohn eller colitis ulcerosa) (ICD10: DK50, DK51; ICD8: 563.00–563.09, 563.19).

Research projects

Danish Biobank Register



1. Parkinson's disease (ICD10: G20 eller ICD8: 342)
2. Inflammatory bowel disease (Mb. Crohn eller colitis ulcerosa) (ICD10: DK50, DK51; ICD8: 563.00–563.09, 563.19).

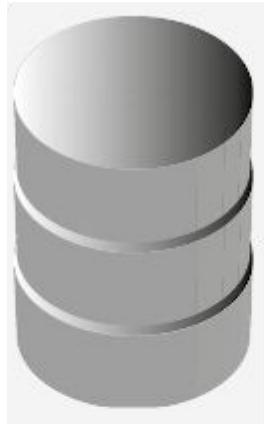
1. [Biobank] - [Prøvetype] - [Antal individer]

		Biobank				
		<i>Copenhagen Hospital Biobank</i>	<i>Danish Cancer Society Biobank</i>	<i>Danish National Biobank</i>	<i>Dansk CancerBiobank</i>	<i>Patobanken</i>
Prøvetype	<i>Biopsy</i>	0	0	0	0	346
	<i>Buffycoat</i>	0	0	0	*	0
	<i>Clot</i>	0	0	*	0	0
	<i>DNA</i>	0	22	*	0	0
	<i>Filter paper</i>	0	0	*	0	0
	<i>Whole blood</i>	43	0	0	*	0

Research projects

DANMARKS
NATIONALE
BIOBANK

Danish Biobank Register



Data



43 Samples



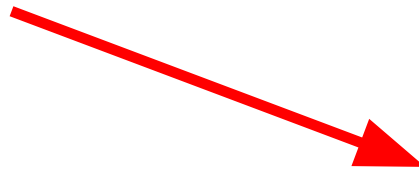
Without the register: All cases of Parkinson's disease (7300) and inflammatory bowel disease (45.000) would have to be extracted from health registers and the list of potential donors cross referenced with all relevant biobanks one by one.

Research projects

DANMARKS
NATIONALE
BIOBANK

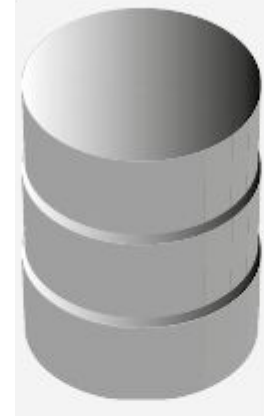


Rare neurological
disease

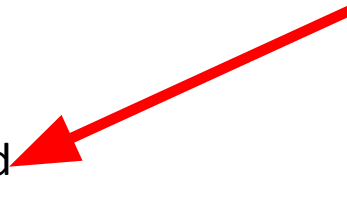


Danish Biobank Register

22 M samples from 5.4 M individuals



11 Spinal fluid
samples



Person

Diagnosis

Sample

Report Definition

Finish

General info

Gender



Any ▾

Country of birth

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Country of birth (mother)

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Country of birth (father)

[Unselect all](#)

Denmark Greenland Faroe Islands Norway Sweden Finland Iceland Other Unknown

Year of birth

From

1850

To

2016

Age at sample taken (expressed in years)

From

0

To

150

Danish Biobank Register

Person

Diagnosis

Sample

Report Definition

Finish

General info

Search for diagnosis codes

Choose code type

ICD10

Search codes (e.g. DJ46, 493, asthma, KMAA-KMAD)

dc50

Select codes

- ICD10/DC50 (Brystkræft)
- ICD10/DC500 (Kræft i brystvorte el areola mammae)
- ICD10/DC500A (Kræft i areolae mammae)
- ICD10/DC500B (Kræft i brystvorte)
- ICD10/DC500C (Paget's sygdom i mamma)
- ICD10/DC500D (Paget's disease of the nipple)
- ICD10/DC501 (Brystkræft i den centrale del af mamma)
- ICD10/DC502 (Brystkræft i den øvre mediale kvadrant af mamma)
- ICD10/DC503 (Brystkræft i den nedre mediale kvadrant af mamma)

Select all

Deselect all

Add as primary search criteria

Add as secondary search criteria

Add as omitted search criteria

Primary search criteria (any of)

- ICD10/DC50
- ICD10/DC500
- ICD10/DC500A
- ICD10/DC500B
- ICD10/DC500C
- ICD10/DC500D
- ICD10/DC501
- ICD10/DC502

Delete all

Delete selected

Secondary search criteria (any of)

AND

Delete all

Delete selected

Omitted search criteria (any of)

BUT NOT

Delete all

Delete selected

Diagnosis period (applies only to the primary search criteria)

Person

Diagnosis

Sample

Report Definition

Finish

General info

Select Biobank



[Unselect all](#)

Danish National Biobank	<input checked="" type="checkbox"/> BSMB	Selected projects: All
Copenhagen Hospital Biobank	<input checked="" type="checkbox"/> PKU	
Danish Cancer Society Biobank	<input checked="" type="checkbox"/> Greenland samples	
Patobanken	<input checked="" type="checkbox"/> SSI diagnostic samples	
Genetic Biobank of the Faroes		
The Danish Blood Donor Study		
COPSAC		
DD2		
Dansk CancerBiobank		

Sample types

Blood

[Unselect all](#)

Buffycoat Serum Plasma Filter paper Whole blood Clot PBMC

Tissue

[Unselect all](#)

Dry frozen FFPE OCT embedded RNAlater treated

Other types

[Unselect all](#)

DNA Spinal fluid Amniotic fluid Urine Feces Nasopharyngeal swab Bone marrow Fine needle aspiration Cervical cytology Other cytologies Breast milk RNA

Advanced sample types

All samples (excluding umbilical cord samples) All samples (including umbilical cord samples) Only umbilical cord blood samples

Person

Diagnosis

Sample

Report Definition

Finish

General info

Report definition



Primary variable

Biobank

Secondary variable

Sample type

Counting unit

Number of samples

(Optional) Classify samples into 3 categories:

Sample taken date

- 365 days **before** the "incident diagnosis" date and earlier
- **at** the "incident diagnosis" date
- 7 days **after** the "incident diagnosis" date and later

For more details, use help.

Add report

X-value	Y-value	Counting unit	Before interval in days	After interval in days	
Biobank	Sample type	Number of individuals			Delete
Biobank	Sample type	Number of samples	365	7	Delete

Version: 1.1.17

2. [Biobank] - [Sample type] - [Number of samples]

With before/at/after - ("at diagnosis" range: -365 days, +7 days)

			Biobank		
			<i>Copenhagen Hospital Biobank</i>	<i>Danish Cancer Society</i>	<i>Danish National Biobank</i>
Sample type	<i>Buffycoat</i>	Before	0	0	1416
		At(-365,+7)	0	0	17
		After	0	0	57
	<i>DNA</i>	Before	0	2200	0
		At(-365,+7)	0	116	0
		After	0	100	0
	<i>Filter paper</i>	Before	0	0	1656
		At(-365,+7)	0	0	18
		After	0	0	66
	<i>Plasma</i>	Before	0	2200	1789
		At(-365,+7)	0	116	37
		After	0	100	274
	<i>Serum</i>	Before	0	2200	4184
		At(-365,+7)	0	116	977
		After	0	100	10190
<i>Whole blood</i>	Before	149	0	275	
	At(-365,+7)	328	0	93	
	After	5684	0	848	