

CONOR (COhort NORway) – en oversikt over en unik forskningsdatabank

Anders Engeland og Anne Johanne Søgaard

Divisjon for Epidemiologi, Nasjonalt folkehelseinstitutt

Korrespondanse: Anders Engeland, Divisjon for Epidemiologi, Nasjonalt folkehelseinstitutt, Postboks 4404 Nydalen, 0403 Oslo

Telefon: 23 40 82 13 Telefaks: 23 40 82 52 E-post: anders.engeland@fhi.no

RESYMÉ

CONOR (COhort NORway) er navnet på et stort samarbeidsprosjekt mellom sentrale epidemiologiske forskningsmiljøer i Norge. Fra og med 1994 er det i store regionale helseundersøkelser inkludert et felles sett av spørsmål, og deltakerne har gitt blod som er lagret for senere analyser. Målsetningen er å inkludere 200 000 personer. Så langt har 167 000 personer deltatt og gitt samtykke til at blodet og resultatene kan brukes i medisinsk forskning. Den gjennomsnittlige alderen på disse var 50 år, og alderen på personene varierte fra 18 til 103 år. Ved utgangen av juni 2002 var 6 600 av disse deltakerne døde. CONOR-dataene er personidentifiserbare slik at de kan kobles til personidentifiserbare registre for forskningsformål.

Engeland A, Søgaard AJ. **CONOR (COhort NORway) – an overview of a unique research material.**

Nor J Epidemiol 2003; 13 (1): 73-77.

ENGLISH ABSTRACT

CONOR (COhort NORway) is a large collaborative project between epidemiological research centres in Norway. Since 1994, large regional health surveys have included a common set of questions, and blood samples from the participants has been stored for later analyses. The intention is to include 200,000 persons. So far 167,000 persons have participated. They have agreed that the information and their blood samples can be used in medical research. The mean age at survey was 50 years and the age ranged from 18 to 103 years. At the end of June 2002 6,600 of the participants had died. The data in the CONOR-project contain the Norwegian personal identification number which makes it possible to link these data with health registries including the personal identification number as e.g. the Cause of Death Registry and the Cancer Registry.

INNLEDNING

CONOR (COhort NORway) er navnet på et stort samarbeidsprosjekt mellom sentrale epidemiologiske forskningsmiljøer i Norge. Bakgrunnen for CONOR, samt målsetting og hvilke potensialer som ligger i dette datamaterialet er nærmere omtalt i en annen artikkel i dette nummeret av Norsk Epidemiologi (1).

De samarbeidende epidemiologiske miljøene ved Universitetet i Tromsø, HUNT forskningssenter i Nord-Trøndelag/NTNU, Universitetet i Bergen, Folkehelse og Statens helseundersøkelser (de to siste er nå deler av Nasjonalt folkehelseinstitutt), har gjennomført flere store befolkningsbaserte helseundersøkelser i løpet av de siste tiårene (2). Data fra disse undersøkelsene, både måledata og opplysninger innsamlet ved hjelp av spørreskjema, er blitt brukt i en rekke nasjonale og internasjonale vitenskapelige publikasjoner.

I CONOR vil dataene fra alle disse undersøkelsene brukes samlet – bl.a. for å finne årsaker til sjeldne sykdommer – ved å teste mulige miljøfaktorer, arvelige faktorer og livsstilsfaktorer – og å beskrive utbredelsen

av sykdommer i Norge etter tid, sted og sosiodemografiske faktorer (1). Med utgangspunkt i disse målsettingene, har de store regionale helseundersøkelsene fra og med 1994 derfor inkludert et felles sett av spørsmål (CONOR-spørsmål) (3), og tappet fullblod (serum i Tromsø IV) fra alle deltakerne. Blodet er frosset ned ved -80°C .

Selv om målet om å samle inn data og biologisk materiale fra 200 000 voksne personer i CONOR ennå ikke er nådd, er data fra de undersøkelsene som er gjennomført til nå, samlet i en datafil. Hensikten er at CONOR-dataene skal stilles til rådighet for forskere i de deltakende miljøer, og for forskere fra andre miljøer etter nærmere fastsatte retningslinjer.

DATAINNSAMLING

I CONOR inngår foreløpig data fra de store regionale helseundersøkelser gjennomført i perioden 1994-2002 – dvs. Tromsø IV-undersøkelsen, Helseundersøkelsen i Nord-Trøndelag (HUNT 2), Helseundersøkelsen i Hordaland (HUSK), Helseundersøkelsen i Oslo

(HUBRO), Oslo II-undersøkelsen, Helseundersøkelsen i Hedmark og Oppland (OPPHED), Tromsø V-undersøkelsen, Helseundersøkelsen i Troms og Helseundersøkelsen i Finnmark (de to siste omtales også som TROFINN). I HUBRO, OPPHED, Tromsø V og TROFINN inngår det også en ren spørreskjemaundersøkelse av 15-åringer, men disse undersøkelsene er ikke en del av CONOR.

Tromsø

Tromsø IV ble gjennomført i 1994-95 og alle innbyggere i Tromsø som var født i 1969 eller tidligere, ble invitert til å delta (4). De som møtte til fase II i Tromsø IV (alle som møtte i utvalgte aldersgrupper ble invitert til en mer omfattende fase II) ble invitert til å delta i en ny undersøkelse, Tromsø V, i 2001. I tillegg ble de som fylte 30, 40, 45, 60 og 75 år i 2001 invitert (3). Tromsø-undersøkelsene er gjennomført som et samarbeid mellom Universitetet i Tromsø, Statens helseundersøkelser/Nasjonalt folkehelseinstitutt, Regi-nsykehuset i Tromsø og Tromsø kommune.

Nord-Trøndelag

Alle innbyggere i Nord-Trøndelag som var 13 år eller eldre ble invitert til å delta i HUNT 2 (1995-97) (5). Det ble tatt blodprøver av personer over 20 år (og noen få 18- og 19-åringer). Bare disse personene inngår i CONOR. Datainnsamlingen ved HUNT 2 ble gjennomført som et samarbeid mellom Folkehelse (nå Folkehelseinstituttet), Statens helseundersøkelser (nå Folkehelseinstituttet), Nord-Trøndelag fylkeskommune og Norges Teknisk-naturvitenskapelige universitet (NTNU).

Hordaland

I 1997-99 ble alle personer bosatt i Hordaland og født i 1953-57 invitert til en helseundersøkelse (HUSK) i regi av Universitetet i Bergen, Statens helseundersøkelser og kommunehelsetjenesten i Hordaland (6). I tillegg til disse 5 årskullene ble også alle personer født i 1925-27 og 1950-51 som hadde deltatt i en tilsvarende helseundersøkelse i 1992-93, invitert til HUSK.

Oslo (HUBRO)

HUBRO ble gjennomført i 2000-01 og alle som var bosatt i Oslo og født i 1924-25, 1940-41, 1955, 1960 og 1970 ble invitert (7). Mot slutten av studien ble også personer født i 1954 og 1969 invitert, men disse fikk ingen parring. HUBRO ble gjennomført som et samarbeid mellom Universitetet i Oslo, Oslo kommune og Statens helseundersøkelser.

Oslo II

I forkant av HUBRO i 2000, ble alle gjenlevende personer bosatt i Oslo og Akershus som i sin tid var blitt invitert til Oslo-undersøkelsen (1972-73) (8),

invitert til en ny helseundersøkelse (Oslo II). Alle menn født i 1923-32, samt et 7% utvalg av menn født 1933-52 ble innkalt til den første Oslo-undersøkelsen. Oslo II ble gjennomført av Statens helseundersøkelser i samarbeid med forsker Lise Lund Håheim.

Oppland og Hedmark

Statens helseundersøkelser gjennomførte i 2000-2001 helseundersøkelsen OPPHED i samarbeid med kommunene i Oppland og Hedmark (9). Alle personer bosatt i disse to fylkene som var født i 1925, 1940, 1955, 1960 og 1970, ble invitert til å delta.

Troms og Finnmark

Helseundersøkelsene i Troms og Finnmark (TROFINN) ble startet i 2002 i regi av Folkehelseinstituttet. I begge fylker ble alle personer (unntatt de som var bosatt i Tromsø (undersøkt tidligere) – og de som var bosatt i 6 kommuner i Troms og 9 i Finnmark) født i 1927, 1942, 1957, 1962 og 1972, invitert. I tillegg ble deltakere fra tidligere helseundersøkelser i Finnmark, gjennomført i 1974, 1977 og 1987, invitert. De resterende kommunene i Finnmark – kommuner med samisk og norsk bosetting, vil bli undersøkt i 2002-2003 i forbindelse med "Helse- og levekårsundersøkelsen" – som har hovedfokus på kosthold, levekår og verdiorientering blant befolkningen i disse områdene. Denne undersøkelsen er et samarbeid mellom Senter for samisk helseforskning Karasjok, Institutt for samfunnsmedisin, Universitetet i Tromsø og Folkehelseinstituttet (10). Data fra denne undersøkelsen er foreløpig ikke inkludert i CONOR.

Totalt er det blitt sendt ut 308 017 innkallelser til disse helseundersøkelsene. Antall som møtte opp til undersøkelsene var 178 889 (58%). Noen personer fikk innkallelse til flere undersøkelser, slik at antall personer som ble innkalt var 295 570. Det var 170 116 personer (58%) som møtte til en eller flere helseundersøkelser. Vi vil i denne fremstillingen kun bruke svar/opplysninger fra den første undersøkelsen hver enkelt deltok i.

SAMTYKKEERKLÆRING

Alle undersøkelser som inngår i CONOR er såkalt samtykke-baserte. Dvs. at alle personer som enten møtte til helseundersøkelse og/eller svarte på spørsmål i forbindelse med undersøkelsen, ble bedt om å skrive under på en samtykkeerklæring. Samtykkeerklæringene gikk på hvordan data som samles inn fra den enkelte person kunne benyttes videre. Samtykkeerklæringene har variert noe, men kan grovt settes opp slik:

1. Samtykke til at resultater følges opp av helsepersonell
2. Samtykke til å få innkalling til lignende-/tilleggsundersøkelser (kun undersøkelser startet f.o.m. 1995)

3. Samtykke til at resultater kan brukes i medisinsk forskning (i undersøkelser før 2000 inneholdt dette også samtykke til å sammenholde opplysningene med opplysninger fra andre helse- og sykdomsregistre)
4. Samtykke til at blodprøve kan oppbevares
5. Samtykke til at resultater kan kobles til andre registre (kun brukt i undersøkelser startet f.o.m. 2000)
6. Samtykke til at analyse av arvestoff kan gjennomføres (kun brukt i undersøkelser startet f.o.m. 2001)

For hvert av samtykkene, skulle personene stryke det de ikke samtykket i. Deretter skulle de skrive under på samtykkeerklæringene.

Av de 170 116 personene som møtte til helseundersøkelsene var det relativt få som ikke ga de samtykkene det ble spurt om (tabell 1). Over 99% samtykket til at resultatene kunne følges opp av helsepersonell, til å få innkalling til lignende undersøkelser eller tilleggsundersøkelser, til at resultatene kunne brukes til medisinsk forskning og til at blodprøven ble oppbevart. Noen færre (97%) ga samtykke til at opplysningene fra undersøkelsen kunne kobles til andre registre og til at analyse av arvestoff kunne gjennomføres (98%). I alle de tre undersøkelsene der det ble spurt eksplisitt om samtykke til kobling til andre registre og om det kunne gjøres analyser av arvestoff, var det flere som samtykket til arvestoffanalyser enn til å koble opplysningene med andre registre.

Kun data fra personer som møtte til en helseundersøkelse, og som har gitt og skrevet under på samtykke om at opplysningene kan brukes i medisinsk forskning, at blodprøver kan oppbevares, at resultatene kan kobles til andre registre og at det kan gjøres arvestoff-

analyser (hvis det ble spurt om de ulike samtykkene), blir brukt i denne oversikten over CONOR-data. Det var 2 760 personer som ikke ga de nødvendige samtykkene, dvs at i alt 167 356 personer (98,4% av personene som møtte til undersøkelse) inngår i den omtalte CONOR-filen.

Hvis vi inkluderer alle dem som har samtykket i at data kan brukes til medisinsk forskning, består CONOR-materialet av totalt 168 884 individer. Der- som man bare ønsker å analysere spørreskjemadata, uten koblinger til andre registre, vil en slik datafil kunne brukes.

Som før nevnt var samtykke til å koble resultatene til andre registre, det som flest deltakere reserverte seg mot. Dette til tross for at dette samtykket (innført i 2000) inneholdt opplysninger om at slike koblinger må godkjennes av Datatilsynet i tillegg til at navn og personnummer skal fjernes etter kobling. Før år 2000 inneholdt samtykket til at resultatene kunne brukes i medisinsk forskning, også en formulering om at dette innebar kobling med andre registre. Det var en større andel som samtykket til kobling før år 2000 enn senere.

TILGJENGELIGE DATA OM CONOR-DELTAAGERNE

Alle personer som møtte til undersøkelse ble bedt om å fylle ut ett eller flere spørreskjema. Spørreskjemaene var forskjellige i de ulike undersøkelsene, men de fleste undersøkelsene inneholdt alle de såkalte CONOR-spørsmålene (tilgjengelig på CONORs hjemmeside (3)). Det er spørsmål om egen helse og familiens helse, fysisk aktivitet, røyking, kaffe, alkohol, skolegang, lokalmiljø, bolig, venner, arbeid, bruk av medisiner og,

Tabell 1. Andel av de som møtte som ga samtykke og skrev under på samtykkeerklæringen.

	Antall * personer	Oppf. av helsepers.	Tilleggs- undersøkelse	Medisinsk forskning	Lagring av blodprøve	Kobling til registre	Arvestoff- analyser**	Underskr. på erklæringen
Tromsø IV	27 169	99,6	–	99,3	99,5	–	–	99,9
HUNT II	65 331	99,7	99,8	99,5	99,5	–	–	99,9
HUSK	25 630	96,8	99,7	99,6	99,5	–	–	100,0
Oslo 2	7 309	99,7***	–	94,5	99,7***	94,5	–	94,5
HUBRO	22 070	99,7***	99,6	99,5	98,7***	96,7	–	99,9
OPPHED	12 441	99,8	99,8	99,5	98,9	98,5	–	100,0
Tromsø V	907	100,0	99,8	99,5	98,8	97,1	97,2	100,0
Troms§	2 496	99,6	99,8	99,4	98,3	96,5	98,0	100,0
Finnmark§	6 803	99,6	99,8	99,1	98,4	95,9	98,1	100,0
Totalt	170 116	99,2	99,8	99,3	99,3	96,7	98,0	99,7

* Hver person telles kun i første undersøkelse personene har møtt til. Dvs. at personer som f.eks. har møtt både til Tromsø IV og Tromsø V kun regnes med under Tromsø IV.

** Ved HUNT 2 ble det innhentet fornyet samtykke som omfattet analyse av arvestoff våren 2002. Det betyr at HUNT 2-biobanken i dag er basert på samtykke som omfatter genetisk forskning. Dette fornyede samtykket er det imidlertid ikke tatt hensyn til her.

*** I HUBRO og Oslo 2 er det inkludert henholdsvis 107 og 753 personer som ikke ble spurt om samtykke til å følges opp av helsepersonell og lagring av blodprøve. (Deltok i undersøkelsene HYRIM, DOIT og LIFE (7))

§ Del av TROFINN (10)

for kvinner, spørsmål om reproduksjon og barnefødsler. I tillegg til å svare på spørsmål, ble høyde, vekt, midje, hofte, blodtrykk og puls målt, og det ble tatt en blodprøve. Blodprøvene ble brukt til å måle kolesterol, triglyserider, glukose (unntatt i Tromsø IV) og HDL-kolesterol. I tillegg ble det frosset ned fullblod (serum i Tromsø IV) til senere analyser i CONOR.

Gjennomsnittsalderen ved undersøkelsestidspunktet for de 167 356 personene som er inkludert i CONOR-filen er 50 år. Den yngste er 18 år og den eldste 103 år. Personene er født i perioden 1893-1978 (tabell 2).

KOBLING AV CONOR-FILEN MED ULIKE REGISTRE

Det unike personnummeret (11) og personnummerbaserte person- (11), sykdoms- (12,13) og dødelighetsregistre (14) av høy kvalitet, gjør at sykdoms- og dødelighetsdata kan kobles sammen med resultater og opplysninger fra helseundersøkelsene på en sikker og relativt enkel måte. Også andre kilder med personidentifiserbare data, f.eks. data fra folke- og bolig-tellinger, kan utnyttes på samme måte. Det samme kan nye fremtidige sykdomsregistre – og et eventuelt nytt personidentifiserbart pasientregister (15).

Foreløpig er CONOR-filen bare koblet til personregisteret for å finne ut hvor mange som allerede er døde. Ved utgangen av juni 2002 var det i overkant av 6 600 dødsfall blant de 167 356 personene som møtte til en av undersøkelsene som inngår i CONOR. Den

samlede observasjonstiden (tiden fra undersøkelsesdato til emigrasjon, død eller 30. juni 2002) var 730 000 personår (tabell 3). Det observerte antallet dødsfall var noe mindre (72%) enn det antallet en kunne forvente hvis denne gruppen hadde hatt samme dødelighet som den generelle norske befolkningen.

Det forventede antall krefttilfeller i CONOR, beregnet ved bruk av kreftregister fra hele landet for 1995-99, er i overkant av 5 300, og forventet antall lungekrefttilfeller er i overkant av 500.

Påkoblet informasjon om f.eks. dødsårsaker eller krefttyper, kan deretter være utgangspunktet for bl.a. å gjennomføre nøstede kasus-kontroll studier hvor det gjøres nye analyser av blodprøvene.

Tabell 2. Antall personer i CONOR-filen etter fødselsår og kjønn (bare personer med alle de omtalte samtykkene er med).

	Menn	Kvinner
1893-99	11	14
1900-09	285	525
1910-19	2 363	3 481
1920-29	13 236	11 550
1930-39	9 550	7 288
1940-49	14 409	14 484
1950-59	23 822	27 395
1960-69	12 139	14 428
1970-78	5 465	6 911
	81 280	86 076

REFERANSER

- Magnus P, Arnesen E, Holmen J, Stoltenberg C, Søgaard AJ, Tell GS. CONOR – COhort NORway: historie, formål og potensiale. *Norsk Epidemiologi* 2003; **13**: 79-82.
- Bjartveit K. Statens helseundersøkelser: Fra tuberkulosekamp til mangesidig epidemiologisk virksomhet. *Norsk Epidemiologi* 1997; **7**: 157-74.
- Nasjonalt folkehelseinstitutt. CONOR. <http://www.fhi.no/tema/conor>. Juni 2003.
- Universitetet i Tromsø. Tromsøundersøkelsen. <http://www.ism.uit.no/tromso5/>. Juni 2003.
- Helseundersøkelsen i Nord-Trøndelag. Helseundersøkelsen i Nord-Trøndelag. <http://www.hunt.ntnu.no/>. Juni 2003.
- Universitetet i Bergen. Helseundersøkelsen i Hordaland (HUSK). <http://www.uib.no/isf/husk/>. Juni 2003.
- Nasjonalt folkehelseinstitutt. HUBRO. <http://www.fhi.no/tema/helseundersokelse/oslo/>. Juni 2003.
- Holme I, Håheim LL, Hjermann I. Prediksjonsfunksjoner for 21 års utvikling av hjerteinfarkt fra Osloundersøkelsen 1972/73. *Tidsskr Nor Lægeforen* 2003; **123**: 1050-1053.
- Nasjonalt folkehelseinstitutt. Helseundersøkelsen i Oppland og Hedmark. http://www.fhi.no/tema/helseundersokelse/oppland_hedmark/. Juni 2003.
- Nasjonalt folkehelseinstitutt. Helseundersøkelse i Troms og Finnmark. http://www.fhi.no/tema/helseundersokelse/troms_finnmark/. Juni 2003.
- Hammer H. Det sentrale folkeregister i medisinsk forskning. *Tidsskr Nor Lægeforen* 2002; **122**: 2550-2550.
- Irgens LM. Medisinsk fødselsregister – et sentralt utgangspunkt for perinatalmedisinsk forskning. *Tidsskr Nor Lægeforen* 2002; **122**: 2546-2549.
- Arnesen K. Opprettelsen av Kreftregisteret. *Tidsskr Nor Lægeforen* 2001; **121**: 1256-1257.
- Gjertsen F. Dødsårsaksregisteret – en viktig datakilde for medisinsk forskning. *Tidsskr Nor Lægeforen* 2002; **122**: 2551-2554.
- SINTEF. Norsk pasientregister. <http://www.npr.no>. Juni 2003.

Tabell 3. Observert antall personår og dødsfall samt forventet antall dødsfall, krefttilfeller og lungekrefttilfeller etter alder fram til 30. juni 2002.

	Alder	Personår	Dødsfall		Krefttilfeller		
			Observert	Forventet	Observert/ forventet	Forventet kreft	Forventet lungekreft
Menn	20-24	4 687	6	5	1,29	2	0
	25-29	15 102	18	16	1,12	9	0
	30-34	28 985	16	32	0,50	18	0
	35-39	27 686	18	38	0,47	23	1
	40-44	48 920	46	95	0,49	62	5
	45-49	56 823	86	167	0,51	122	14
	50-54	34 235	102	164	0,62	132	19
	55-59	26 258	129	193	0,67	171	23
	60-64	23 456	201	302	0,67	278	40
	65-69	21 287	354	462	0,77	395	57
	70-74	29 187	711	1 083	0,66	718	97
	75-79	20 582	863	1 260	0,68	630	74
	80-84	7 463	623	757	0,82	259	23
	≥85	3 877	683	692	0,99	141	8
Totalt	348 549	3 856	5 266	0,73	2 959	362	
Kvinner	20-24	6 089	4	2	2,27	2	0
	25-29	17 755	3	6	0,47	10	0
	30-34	34 375	7	16	0,42	31	0
	35-39	31 333	12	23	0,52	48	1
	40-44	55 660	37	63	0,58	141	5
	45-49	63 316	72	115	0,63	256	13
	50-54	35 597	56	110	0,51	206	15
	55-59	26 682	90	121	0,74	198	16
	60-64	24 555	110	173	0,63	231	21
	65-69	19 163	166	215	0,77	220	22
	70-74	23 936	304	459	0,66	324	30
	75-79	23 490	492	799	0,62	361	25
	80-84	11 961	563	766	0,74	210	9
	≥85	7 375	846	1 040	0,81	139	4
Totalt	381 289	2 762	3 910	0,71	2 376	161	