Registry Based Studies and Familial Recurrence Risks

Rolv T. Lie

Important central health registers in Norway

Register	«Dataansvarlig»
Dødsårsaksregisteret	Folkehelseinstituttet
Kreftregisteret	Oslo universitetssykehus
Medisinsk fødselsregister	Folkehelseinstituttet
Meldingssystem for smittsomme sykdommer (MSIS)	Folkehelseinstituttet
Nasjonalt vaksinasjonsregister (SYSVAK)	Folkehelseinstituttet
Forsvarets helseregister	Forsvaret
Norsk pasientregister	Helsedirektoratet
Nasjonalt register over hjerte- og karlidelser	Folkehelseinstituttet
Kommunalt pasient- og brukerregister	Helsedirektoratet
Individbasert pleie- og omsorgsstatistikk (IPLOS)	Helsedirektoratet
Reseptbasert legemiddelregister	Folkehelseinstituttet

Health Analysis Platform (HAP) 2021-

Funksjonalitet på plattformen

Søknadstjenester

- Felles søknadsskjema på tvers av registre
- Redaksjonell veiledningstjeneste
- Felles saksbehandlingsløsning på tvers av registrene
- Metadatakatalog med informasjon om tilgjengelige datakilder og variabler

Analysetjenester

- Eksplorative analysetjenester som kohorteksplorer og anonymisert analyse
- Definerte sett med analyseverktøy
- Definerte sett med visualiserings- og BI- verktøy
- Økoystem av eksterne analysetjenester

Dataplattform

- Virtuell infrastruktur for nye helseregistre
- Sikre prosjektrom på plattformen
- Åpne data, både aggregerte og syntetiske
- Harmonisert datamodell
- API for innbyggerdata

Datakilder på plattformen

Datakilder som skal dupliseres på plattformen

- Sentrale helseregistre
- Nasjonale medisinske kvalitetsregistre
- Befolkningsbaserte helseundersøkelser
- Sosioøkonomiske data fra SSB

Datakilder som på sikt inngår på plattformen

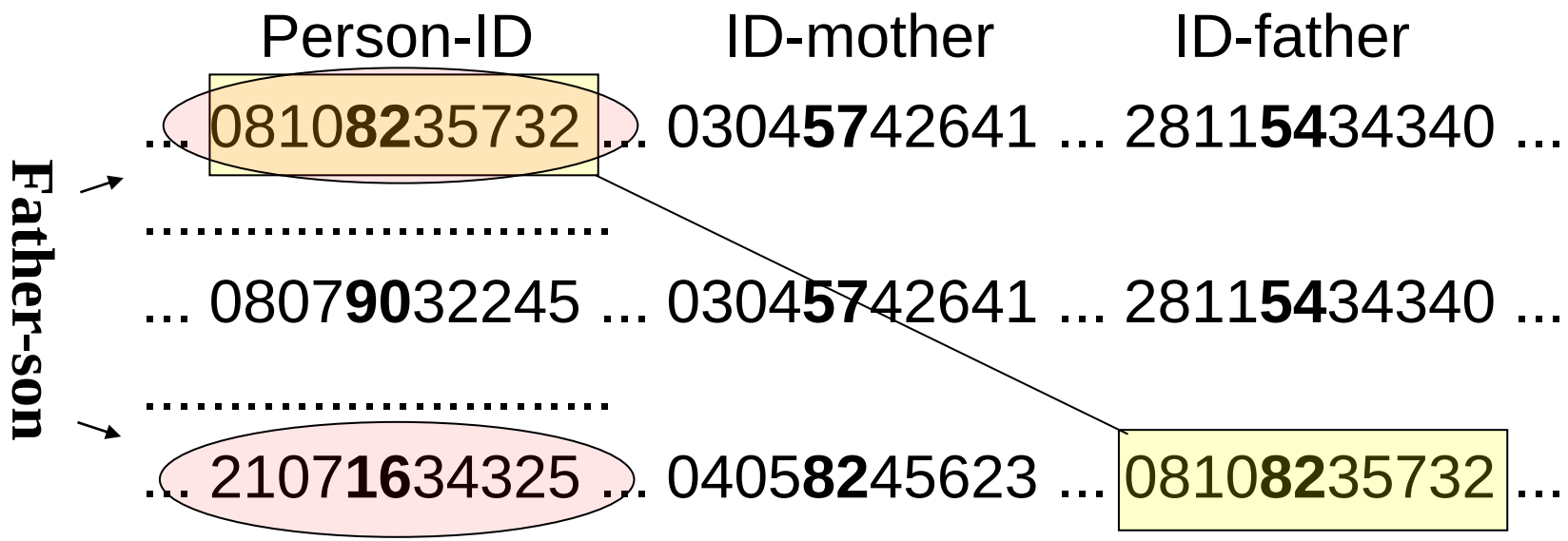
- PAS/EPJ-data
- Data fra tredjepartsaktører, som næringsliv og innbyggere
- Virtuelle helseregistre (masterdata)

How to access data: https://helsedata.no/

GWAS-data in Norway (soon available on HAP?)

MoBa	~240 000 individuals
HUSK	~36 000 individuals
HUNT	~90 000 individuals
Tromsø	~35 000 individuals
Total	~401 000 individuals

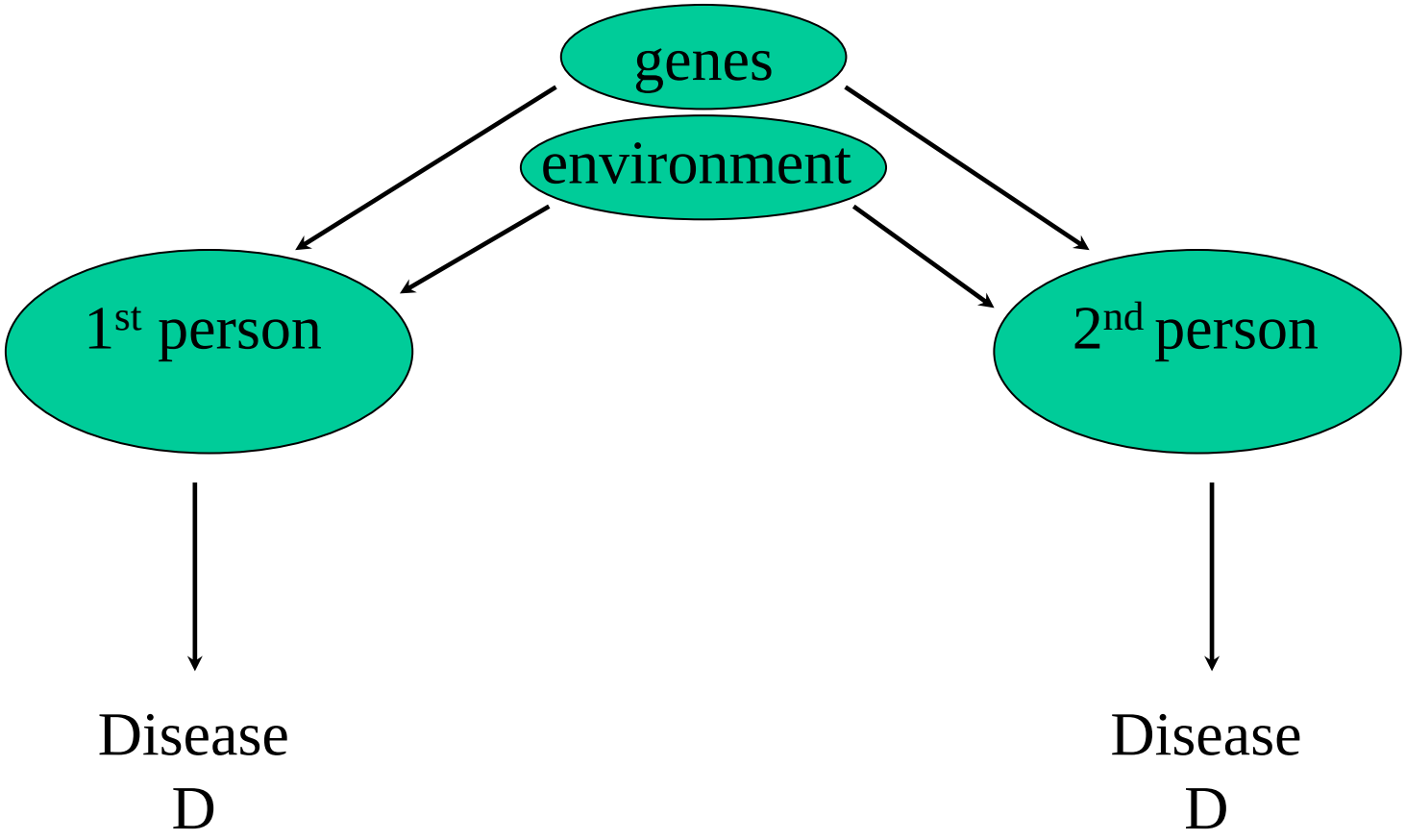
Record linkage, registers



- ID-numbers of most persons born in Norway since 1953 are registered with the ID of theirs parents in DSF*
- In reality we use pseudonym numbers in stead of real ID-numbers for privacy protection

* Det sentrale folkeregister, Skattedirektoratet

Recurrence risks in families are indirect effects



Recurrence risk, genetic contribution

Absolute risk of recurrence from person 1 to person 2

G_i, genotype of person i

D_i, disease in person i

Genetic relationship — Product of penetrances

$$P(D_2 \mid D_1)=[\sum_{G_1,G_2} P(G_1 \cap G_2)\, P(D_2|G_2)\, P(D_1|G_1)]/P(D_1)$$

Relative risk of recurrence (recurrence risk ratio):

$RR = P(D_2 \mid D_1) / P(D_2 \mid \text{not } D_1)$

Some studies use the population prevalence $P(D_2)$ as numerator or *OR* in stead of *RR*

Calculation of relative risk of recurrence

A prospective measure of risk relevant for counselling families

$$RR = \frac{\text{child's risk of } D \text{ given } parent \text{ had } D}{\text{child's risk of } D \text{ given } parent \text{ did not have } D}$$

Disease *D* among older relatives is treated as exposure

just as in other epidemiological studies

Adjustment for confounding not so relevant (indirect effects)

Adjustment for correlation within families may be important

Organize data-files to predict future relatives, keep family-ID

Regression models like logistic (*OR*), log-bionomial and Cox (*RR*) may be used for estimation

Examples of relative risk of recurrence for siblings

Table 6.2 Sibling recurrence risk ratio λ_S estimates as reported for dichotomous traits.

Phenotype	λ_S	Reference	Phenotype	λ_S	Reference
AITD[a]	16.9	[678]	Hypertension	4	[20]
Alcohol dependence	4	[20]	IDDM[c]	15	[560]
Asthma	3	[20]	Multiple sclerosis	20–30	[20, 560]
Autism	75	[20]	Open-angle glaucoma	8	[20]
Bipolar[b]	15	[20]	Osteoarthritis	23	[20]
Crohn disease	25–35	[20]	Prostate cancer	2.3–3	[20, 736]
Hemochromatosis[c]	41	[560]	Psoriasis	7	[20]
Hemochromatosis[d]	65	[560]	Rheumatoid arthritis	5–8	[20]
High myopia	4.9	[202]	Schizophrenia	9–10	[20, 561]
Hodgkin's disease	7	[560]	Tuberculoid leprosy	2.4	[560]

[a] Autoimmune thyroid disease, [b] Bipolar affective disorder, [c] Recurrence risk ratio for idiopathic hemochromatosis in males, [d] Recurrence risk ratio for idiopathic hemochromatosis in females, [e] Insulin-dependent diabetes mellitus.

Ziegler & König, 2012, Ch. 6

Pregnancy outcomes

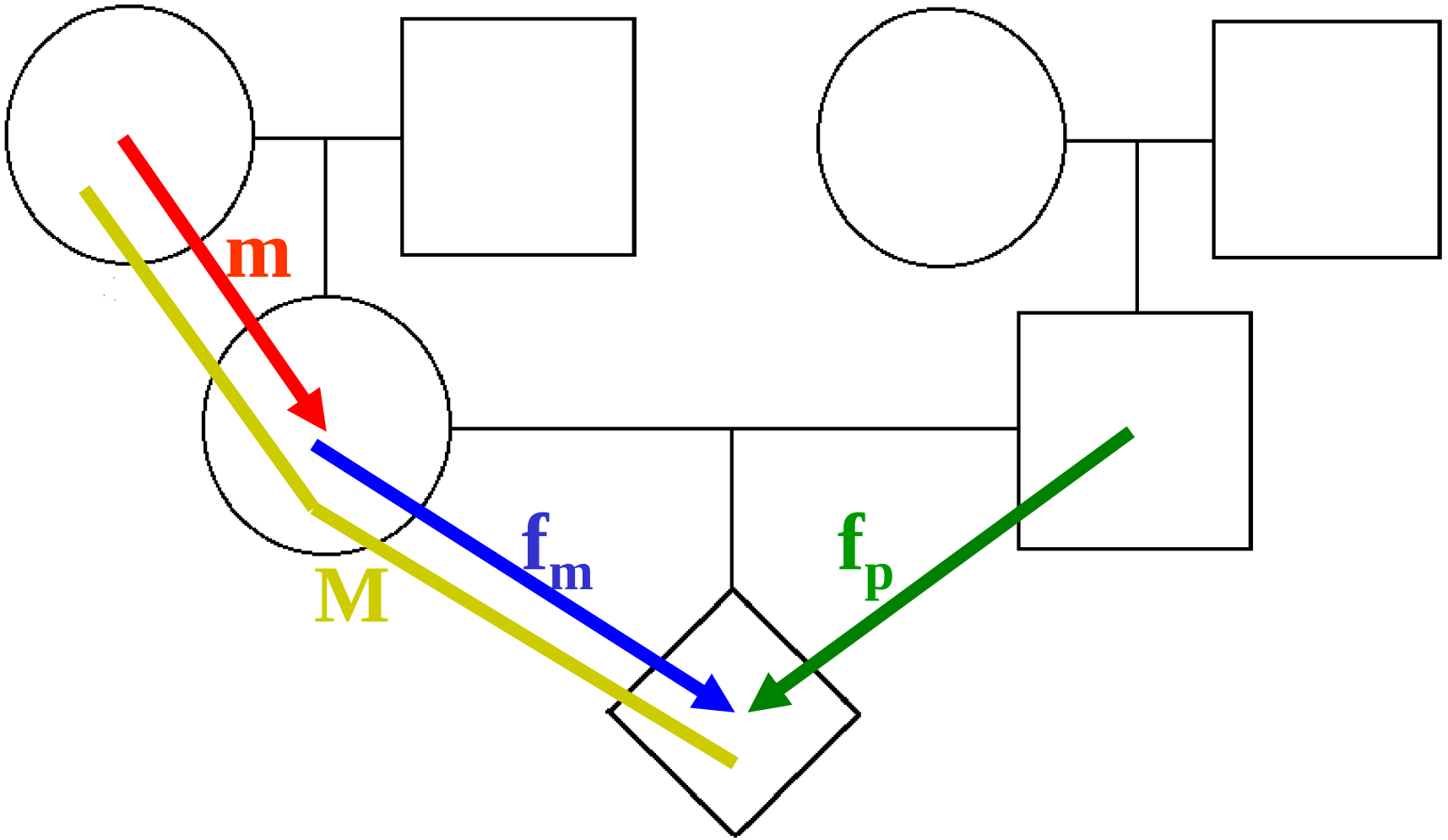
- Genetic influences during pregnancy
- Fetal and maternal genes
- Maternal could still be important for adult disease (fetal origins of adult disease etc.)

Main (genetic) components of transmission of risk in families

- Maternal genes (m)
- Fetal autosomal genes (f)
  - Paternal (f_p)
  - Maternal (f_m)
  - Imprinting or PoO ($f_p \neq f_m$)
- Mitochondrial DNA (M)
- + shared environment

Lie, PPE suppl. 2007

Main genetic components for parent-offspring recurrence



Genetic contributions to parent-offspring recurrence

	Mother-Child	Father-Child
Genes		
Foetal	½	½
Maternal	½	0
Mitochondrial	1	0

Proportion of genome shared by two individuals

Gjessing&Lie, Statistical Methods in Medical Research, 2008
(models for a continuous phenotype)

Questions we could address with recurrence risks

1. Are fetal (regular autosomal) genes involved?
2. Are maternal genes involved (through the fetal or childhood environment)?
3. Are PoO-effects likely?
4. Are mitochondria involved?

Recurrence risks may tell us where to look for genetic effects

Here are some examples:

Table 1| **Recurrence of cerebral palsy (CP) among relatives. Singletons and twins born in Norway 1967-2002 surviving first three years of life**

Relatives	Prevalence of CP (per 1000)	Relative risk (95% CI)	
		Crude	Adjusted
Twins			
Prevalence in twin population	228/45 116 (5.1)	1 (reference)	—
Proband-wise concordance rate	18/228 (78.9)	15.6 (9.8 to 24.8)	—
First degree			
Full siblings:			
Sibling without CP	1929/1 226 413 (1.6)	1 (reference)	1 (reference)
Sibling with CP	30/2014 (14.9)	9.5 (6.6 to 13.5)	9.2 (6.4 to 13.1)*
Parent-offspring:			
Parent without CP	813/622 480 (1.3)	1 (reference)	—
Parent with CP	2/237 (8.5)	6.5 (1.6 to 25.6)	—
Second degree			
Half siblings:			
Half sibling without CP	762/354 163 (2.2)	1 (reference)	1 (reference)
Half sibling with CP	5/774 (6.5)	3.0 (1.2 to 7.2)	3.0 (1.1 to 8.6)†
Aunt/uncle-niece/nephew:			
Aunt/uncle without CP	1930/1 342 559 (1.4)	1 (reference)	—
Aunt/uncle with CP	3/2360 (1.3)	0.9 (0.3 to 2.7)	—
Third degree			
First cousin with CP	8472/5 156 811 (1.6)	1 (reference)	—
First cousin without CP	23/9157 (2.5)	1.5 (0.9 to 2.7)	—

*Adjusted for maternal age at birth of older sibling (<20, 20-24, 25-29, 30-34, ≥35), maternal educational level (below high school, high school, above high school), and period of first birth (1967-71, 1972-77, 1978-84, 1985-91, 1992-2002).

†Adjusted for parental age at birth of older sibling (<20, 20-24, 25-29, 30-34, ≥35 for mothers, extended to 35-39, 40-44, ≥45 for fathers), parental educational level (below high school, high school, above high school), and period of first birth (1967-71, 1972-77, 1978-84, 1985-91, 1992-2002).

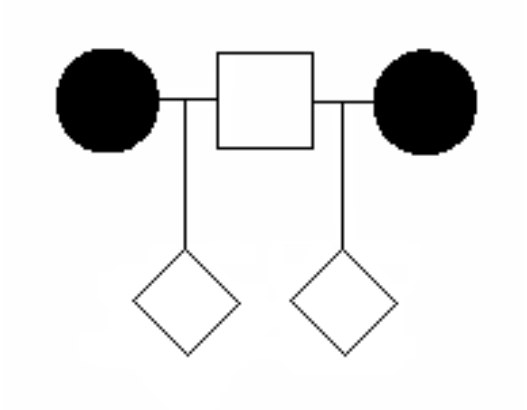
Example: Recurrence of CP

Tollånes et al. BMJ 2014

Example: Pre-eclampsia recurrence

Maternal disease, but only with a pregnancy

Children are paternal half-sibs:



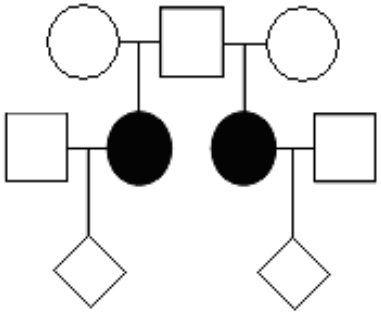
OR = 1.8 (1.2 – 2.6)

(Mostly from f_p)

(Lie et al. BMJ 1998)

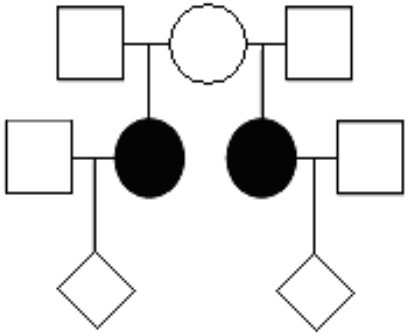
Preeklampsia recurrence, mitochondrial effect?

Mothers are paternal half-sisters



OR = 1.8 (0.9 - 2.6)
(Mostly from *m*)

Mothers are maternal half-sisters

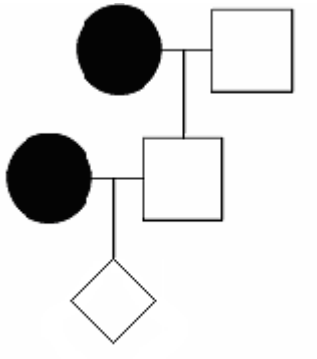


OR = 1.6 (1.01 – 2.9)
(Mostly from *m* and *M*)

(Lie et al. BMJ 1998)

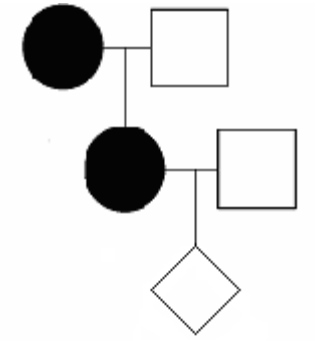
Pre-eclampsia recurrence

Through son (father-child):



OR = 1.4 (1.2-1.7) f_p +..

Through daughter (mother-child):



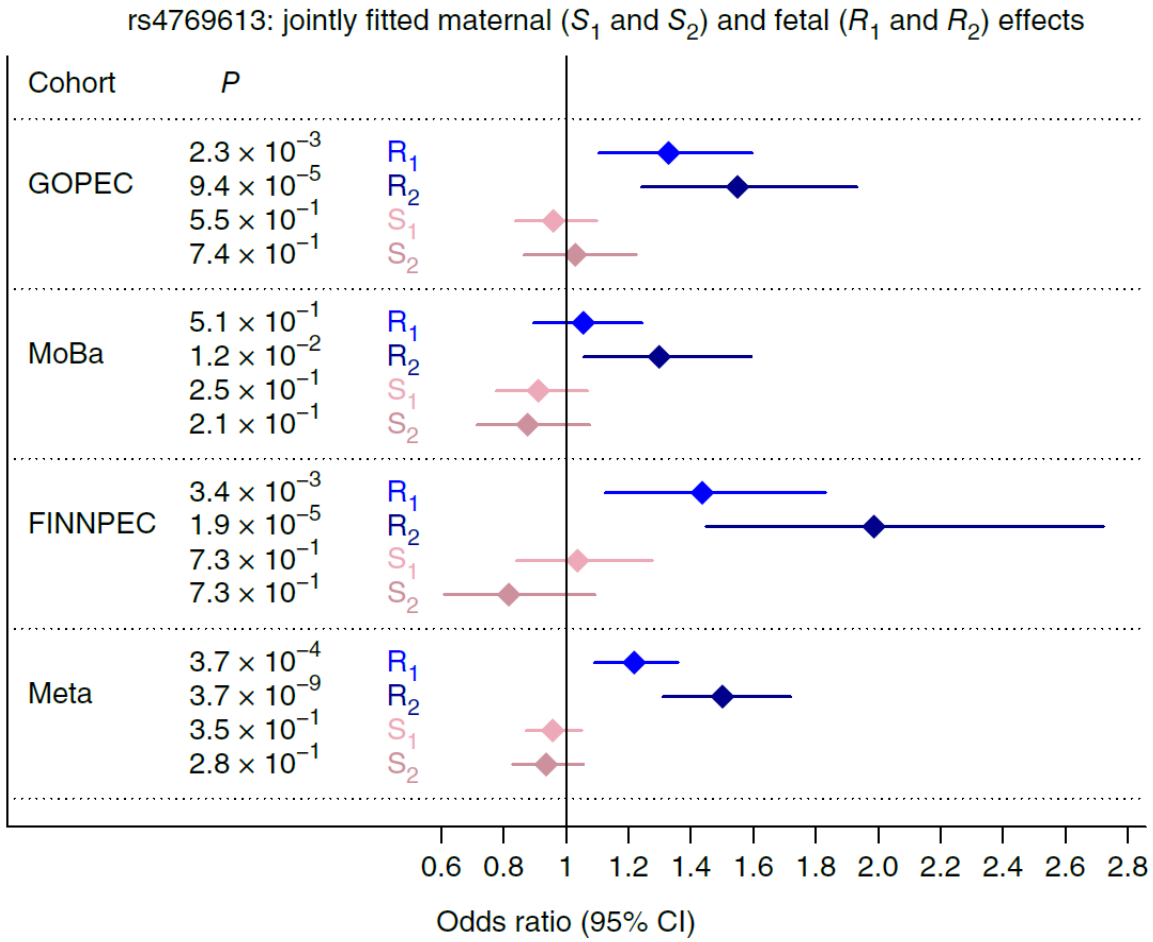
OR = 2.2 (1.9-2.4) $m + f_m + M$ +..

(Skjærven et al. BMJ 2005)

nature genetics

McGinnis R et al. Nat Genet. 2017 Aug;49(8):1255-1260.

Variants in the fetal genome near *FLT1* are associated with risk of preeclampsia



American Journal of Epidemiology
Published by the Johns Hopkins Bloomberg School of Public Health

Example: Preterm birth recurrence

Familial Patterns of Preterm Delivery: Maternal and Fetal Contributions

Allen J. Wilcox[1], Rolv Skjærven[2], and Rolv Terje Lie[2]

Am J Epidemiol 2008;167:474–479

Mother-offspring RR=1.5

Father-offspring RR=1.1

Maternal Effects for Preterm Birth: A Genetic Epidemiologic Study of 630,000 Families

Anna C. Svensson*, Sven Sandin, Sven Cnattingius, Marie Reilly, Yudi Pawitan, Christina M. Hultman, and Paul Lichtenstein

Am J Epidemiol 2009;170:1365–1372

Offspring of sisters RR=1.8

Offspring of sister and brother RR=1.1

Maternal Contributions to Preterm Delivery

Heather A. Boyd*, Gry Poulsen, Jan Wohlfahrt, Jeffrey C. Murray, Bjarke Feenstra, and Mads Melbye

Am J Epidemiol 2009;170:1358–1364

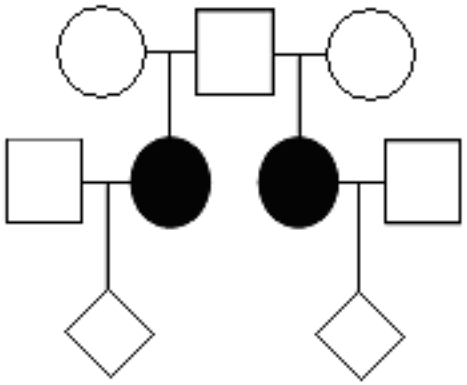
Mother-offspring RR=1.4

Offspring of sisters RR=1.6

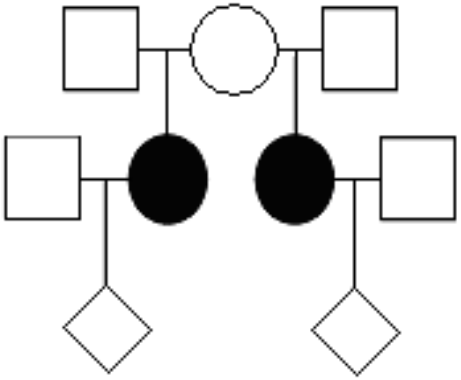
Father-offspring RR=1.2

Paternal half-siblings RR=1.1

Preterm birth recurrence, mitochondrial effect?



Paternal half-sisters
(Unrelated mitochondria)
RR=1.1 (0.9-1.2)



Maternal half-sisters
(Identical mitochondria)
RR=1.4 (1.2-1.6)

Boyd HA et al. Am J Epidemiol. 2009

The NEW ENGLAND JOURNAL of MEDICINE
N Engl J Med 2017;377:1156-67.

ORIGINAL ARTICLE

Genetic Associations with Gestational Duration and Spontaneous Preterm Birth

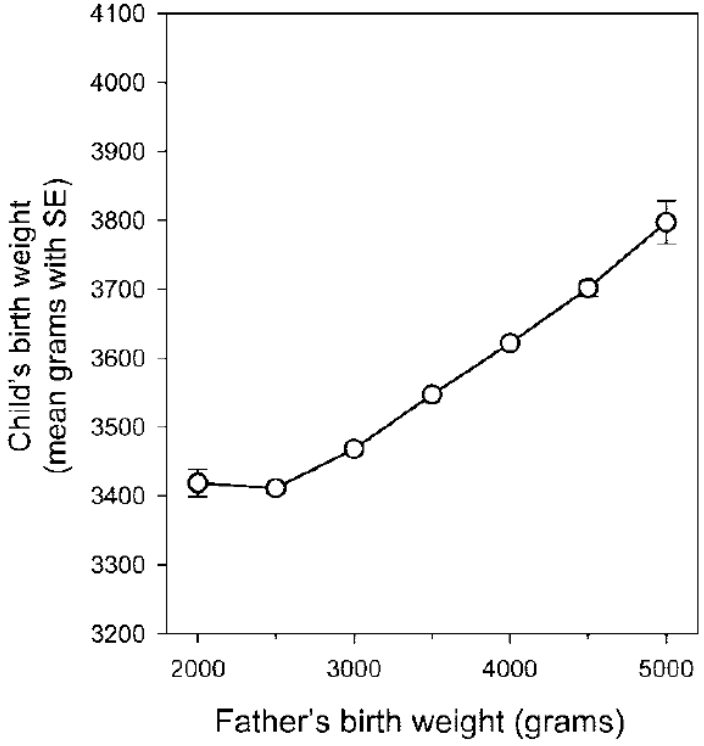
G. Zhang, B. Feenstra, J. Bacelis, X. Liu, L.M. Muglia, J. Juodakis, D.E. Miller, N. Litterman, P.-P. Jiang, L. Russell, D.A. Hinds, Y. Hu, M.T. Weirauch, X. Chen, A.R. Chavan, G.P. Wagner, M. Pavličev, M.C. Nnamani, J. Maziarz, M.K. Karjalainen, M. Rämet, V. Sengpiel, F. Geller, H.A. Boyd, A. Palotie, A. Momany, B. Bedell, K.K. Ryckman, J.M. Huusko, C.R. Forney, L.C. Kottyan, M. Hallman, K. Teramo, E.A. Nohr, G. Davey Smith, M. Melbye, B. Jacobsson, and L.J. Muglia

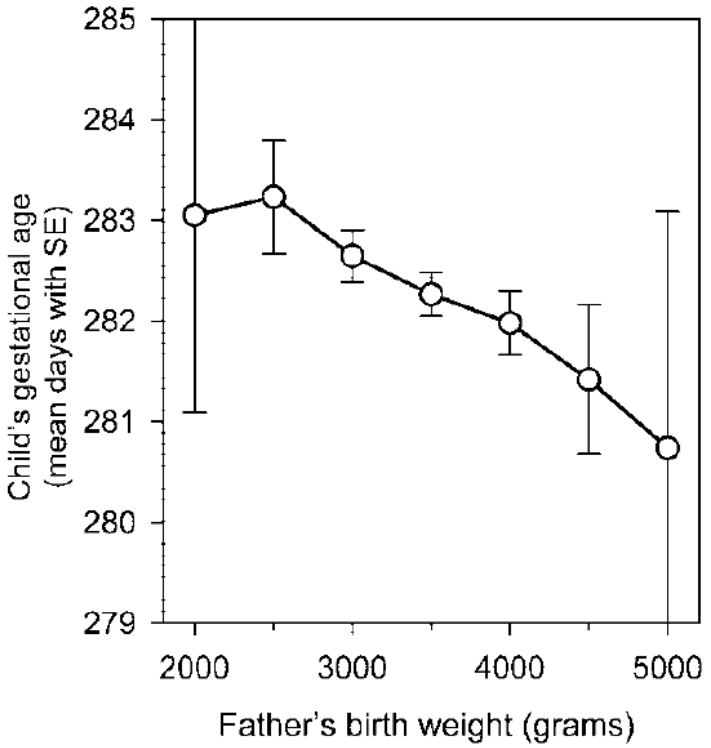
RESULTS

In the discovery and replication data sets, four loci (*EBF1*, *EEFSEC*, *AGTR2*, and *WNT4*) were significantly associated with gestational duration. Functional analysis showed that an implicated variant in *WNT4* alters the binding of the estrogen receptor. The association between variants in *ADCY5* and *RAP2C* and gestational duration had suggestive significance in the discovery set and significant evidence of association in the replication sets; these variants also showed genomewide significance in a joint analysis. Common variants in *EBF1*, *EEFSEC*, and *AGTR2* showed association with preterm birth with genomewide significance. An analysis of mother–infant dyads suggested that these variants act at the level of the maternal genome.

Maternal and Paternal Influences on Length of Pregnancy
Lie, Wilcox, Skjærven

VOL. 107, NO. 4, APRIL 2006 OBSTETRICS & GYNECOLOGY





Fathers BW → *Fetal growth* → Time of delivery

Paternal alleles → *Fetal growth* → Time of delivery

PLOS MEDICINE

Published: August 25, 2020

RESEARCH ARTICLE

Dissecting maternal and fetal genetic effects underlying the associations between maternal phenotypes, birth outcomes, and adult phenotypes: A mendelian-randomization and haplotype-based genetic score analysis in 10,734 mother–infant pairs

Jing Chen[1], Jonas Bacelis[2,3], Pol Sole-Navais[2], Amit Srivastava[4,5], Julius Juodakis[2], Amy Rouse[5], Mikko Hallman[6], Kari Teramo[7], Mads Melbye[8,9,10], Bjarke Feenstra[8], Rachel M. Freathy[11], George Davey Smith[12,13,14], Deborah A. Lawlor[12,13,14], Jeffrey C. Murray[15], Scott M. Williams[16], Bo Jacobsson[2,17], Louis J. Muglia[4,5]*, Ge Zhang[4,5]*

Table 3. Associations between birth weight genetic scores and birth outcomes and estimated causal effects per 1-SD change in gestational-age–adjusted birth weight.

Genetic Score Association and Causal Estimation[a]	Gestational Days			Preterm Birth (log[OR])			Birth Weight (g)			Birth Length (cm)		
	Beta	SE	*p*-Value	Beta	SE	*p*-Value	Beta	SE	*p*-Value	Beta	SE	*p*-Value
Maternal transmitted (β_{h1})	−0.0056	0.0027	0.035*	0.0024	0.00064	0.00018*	0.79	0.087	1.80×10^{-19}*	0.0018	0.00048	0.00022*
Maternal nontransmitted (β_{h2})	0.001	0.0027	0.7	−0.00058	0.00065	0.37	−0.19	0.088	0.029*	−0.00022	0.00049	0.66
Paternal transmitted (β_{h3})	−0.0099	0.0026	0.00018*	0.0019	0.00064	0.003*	1.3	0.087	1.70×10^{-48}*	0.0031	0.00048	2.30×10^{-10}*
Maternal effect (β_{MY})	0.0026	0.0023	0.26	-5.80×10^{-5}	0.00055	0.92	−0.33	0.076	1.10×10^{-5}*	−0.00073	0.00042	0.082
Fetal effect (β_{FY})	−0.0083	0.0023	0.00029*	0.0025	0.00056	9.20×10^{-6}*	1.1	0.075	1.00×10^{-50}*	0.0025	0.00042	1.30×10^{-9}*
Causal (TSLS)	−2.92	0.805	0.00029*	0.837	0.267	0.0017*	NA			1.05	0.124	2.40×10^{-17}*
Causal (ratio)	−3.24	0.889	0.00027*	0.624	0.215	0.0036*				1	0.172	5.10×10^{-9}*

[a]The effect size (beta) and SEs of genetic score association were given by per unit (g) change in genetic scores; the causal effect sizes were based on per 1-SD (1 SD = 426 g) change in gestational-age–adjusted birth weight. **Abbreviations:** NA, not applicable; SD, standard deviation; SE, standard error; TSLS, two-stage least-squares.

**p*-Values less than 0.05.

Genetic Score Association and Causal Estimation[a]	Gestational Days		
	Beta	SE	*p*-Value
Maternal transmitted (β_{h1})	−0.0056	0.0027	0.035*
Maternal nontransmitted (β_{h2})	0.001	0.0027	0.7
Paternal transmitted (β_{h3})	−0.0099	0.0026	0.00018*
Maternal effect (β_{MY})	0.0026	0.0023	0.26
Fetal effect (β_{FY})	−0.0083	0.0023	0.00029*
Causal (TSLS)	−2.92	0.805	0.00029*
Causal (ratio)	−3.24	0.889	0.00027*

STATISTICS IN MEDICINE
Statist. Med. 2004; **23**:449–465 (DOI: 10.1002/sim.1603)

Estimation of genetic and environmental factors for binary traits using family data

Y. Pawitan*,†, M. Reilly, E. Nilsson, S. Cnattingius and P. Lichtenstein

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, 17177 Stockholm, Sweden

Statistical Methods in Medical Research 2008; **17**: 75–96

Biometrical modelling in genetics: are complex traits too complex?

Håkon K. Gjessing Divison of Epidemiology, Norwegian Institute of Public Health, Norway and Section for Epidemiology and Medical Statistics, University of Bergen, Bergen, Norway and **Rolv Terje Lie** Section for Epidemiology and Medical Statistics, Department of Public Health and Primary Health Care, University of Bergen, Norway

References

Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal contributions to preterm delivery. Am J Epidemiol. 2009 Dec 1;170(11):1358-64.

Gjessing HK, Lie RT. Biometrical modelling in genetics: are complex traits too complex? Stat Methods Med Res. 2008 Feb;17(1):75-96.

Lie RT, Rasmussen S, Brunborg H, Gjessing HK, Lie-Nielsen E, Irgens LM. Fetal and maternal contributions to risk of pre-eclampsia: population based study. BMJ. 1998 May 2;316(7141):1343-7.

Lie RT. Intergenerational exchange and perinatal risks: a note on interpretation of generational recurrence risks. Paediatr Perinat Epidemiol. 2007 Jul;21 Suppl 1:13-8.

Pawitan Y, Reilly M, Nilsson E, Cnattingius S, Lichtenstein P. Estimation of genetic and environmental factors for binary traits using family data. Stat Med. 2004 Feb 15;23(3):449-65.

Skjærven R, Vatten LJ, Wilcox AJ, Rønning T, Irgens LM, Lie RT. Recurrence of pre-eclampsia across generations: exploring fetal and maternal genetic components in a population based cohort. BMJ. 2005 Oct 15;331(7521):877.

Svensson AC, Sandin S, Cnattingius S, Reilly M, Pawitan Y, Hultman CM, Lichtenstein P. Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. Am J Epidemiol. 2009 Dec 1;170(11):1365-72.

Tollånes MC, Wilcox AJ, Lie RT, Moster D. Familial risk of cerebral palsy: population based cohort study. BMJ. 2014 Jul 15;349:g4294.

Wilcox AJ, Skjærven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. Am J Epidemiol. 2008 Feb 15;167(4):474-9.