

The biobank of the Norwegian Mother and Child Cohort Study – present status

Liv Paltiel, Anita Haugan, Trine Skjerden, Kari Harbak, Siri Bækken, Nina Kristin Stensrud, Gun Peggy Knudsen and Per Magnus

Norwegian Institute of Public Health

E-mail: liv.paltiel@fhi.no

This is an open access article distributed under the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The Norwegian Mother and Child Cohort Study (MoBa) is a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health (NIPH) (1). The main aim is to find causes of diseases and explain trajectories and variability of health-related traits over a life-course span. Additional aims are to detect early signs of disease and to describe the development of disease (1,2). MoBa provides a data set with questionnaires and biological samples together with the possibility, if approved by the right authorities, for linkage to other Norwegian registries. The pregnant women and their partners were recruited at 50 hospitals in Norway between 1999 and 2008 with a 40.6% response rate. Almost 113 000 pregnancies are included in the study, counting more than 114 000 children including about 1950 pairs of twins, and more than 75 000 fathers. About 16 700 women have participated with more than one pregnancy; hence the cohort is suitable also for sibling studies. The selection to the sample has been described (3), and suggests that although some prevalence estimates are biased, the association measures between exposures and outcomes are not.

The biological samples were collected from both parents during pregnancy and from mothers and children (umbilical cord blood) at birth. All biological samples were sent to the NIPH Biobank for processing and storage. The MoBa biobank constitutes the major part of the NIPH biobank. The methods initially used for collecting, processing and storing the biological material, as well as the quality programs have been described previously (4,5).

This follow-up article provides a comprehensive update of the methods used for collecting and storing biological samples in MoBa together with a description of the status of the biobank in 2014. First, we describe changes made in sampling and processing during the last years of the recruitment period. Second, we describe and discuss sample integrity. Third, we describe the biological samples available for research. Finally, we briefly summarize the use of the biological material and discuss the value of the MoBa biobank.

The MoBa primary teeth biobank (6) and biological samples collected in sub-studies are not included in this paper.

SAMPLING AND PROCESSING

Sampling and sample processing

In MoBa, the nationwide recruitment of pregnant women was conducted between 1999 and 2008. Biological samples were received during a 10-year period from 1999 until the last child in the study was born (2009). Biological samples were collected twice from the mothers; first at the ultrasound appointment at 17–18 weeks of gestation and later after delivery. Also, a sample from the cord blood was drawn at delivery. From the fall of 2000, the fathers were also invited to give a blood sample at the ultrasound appointment (4).

Over the course of the study, the specimens collected changed as research interests emerged. The updated timeline of the specimen collection activities in the main project is presented in Figure 1. In addition, some of the sub-projects, like the ABC study (7), BraPust (Asthma) and The Language and Learning Study (SOL) (www.fhi.no) collect biological samples.

Umbilical cord samples

Until the spring of 2005, umbilical cord blood samples in MoBa were collected in two 7 ml EDTA-tubes (Becton-Dickinson (BD), Plymouth, UK). In the spring of 2005, a strong interest in RNA samples led to changes in the collection protocol. In MoBa, we started to collect 3 ml cord blood in a 9 ml RNA-preservation tube with buffer (Tempus, Applied Biosystems, Foster City, CA). At the biobank, the Tempus tubes were frozen at -20°C overnight to prevent tube damages, and then transferred to -80°C for long time storage. The procedures and testing to establish a practical protocol for sampling, handling, transportation and storage of blood RNA is described elsewhere (8). Recently, also the effects of long-term storage of blood RNA collected in RNA stabilizing Tempus tubes was reported (9).

Before adding the Tempus tube, the filling levels of the two 7 ml EDTA tubes was evaluated. The results showed that few samples had maximum blood volume; hence, we simultaneously reduced the EDTA blood volume and switched to one 10 ml EDTA tube (Becton-Dickinson (BD), Plymouth, UK).

All blood collected from the cord or capillary blood drawn from the child a few days after birth was sent to the MoBa Biobank for processing. Plasma was first

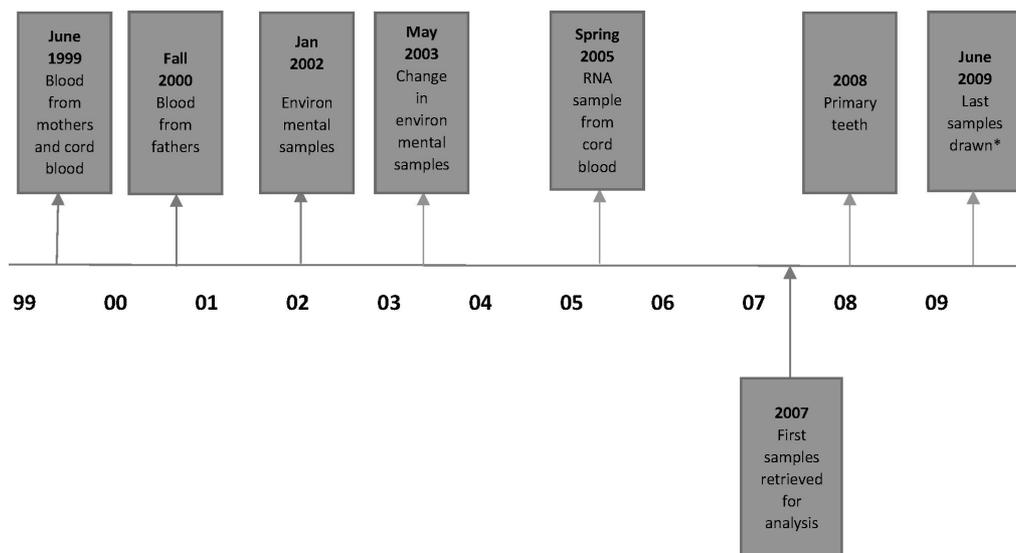


Figure 1. Timeline of sample collection in MoBa.
* Last samples collected in the main study.

pulled off in the biobank after shipping, as opposed to the parental samples where plasma was pulled off at the hospital (4). Thus, the cord blood samples were processed differently from the parental MoBa samples, as concerns the timing of centrifugation and plasma pull off from EDTA tubes. Figure 2 presents the collection and processing procedures for cord blood sample sets.

DNA extraction

DNA is isolated from all MoBa participants, and procedures for normalization and quality control have been described previously (4,5). During the 10 year collection period, four different DNA isolation kits have been used in the biobank. Most often the Flexigene kit (Qiagen, Hilden, Germany) was used, but during some periods PureGene (Gentra, Minneapolis, MN), SQ DNA kit E.Z.N.A (Omega bio-tek, Norcross, GA) and ArchivePure (5'Prime, Hamburg, Germany) kits were used. Each time a new kit was implemented in the laboratory, thorough testing was performed to assure that DNA quality and quantity met the requirements.

As discussed by Rønningen et al (4), it was a priority to isolate DNA and normalize DNA concentration from fresh samples for all participants at the time of collection to facilitate DNA-related research. The intention has been to make DNA readily available, reducing delivery times to scientists and analytical laboratories, especially for the larger studies. To date, DNA from more than 45 000 different participants has been retrieved for analysis.

Storage

Table 1 gives an overview over stored specimens in MoBa, and includes aliquote sizes and storage temperatures. After processing in the biobank, all aliquoted sample types were stored in two different chest

freezers (4). From 2013, one DNA aliquot (two if a participant has seven or more aliquots) has been stored in a backup storage in another building 10 min walk from the main Biobank location. The rest of the DNA aliquots are stored at the main location in an automated storage facility (RTS, Manchester, UK). The Biobank also harbors an automated -80°C storage facility (Hamilton, Bonaduz, Switzerland) and started transferring plasma and urine samples from chest freezers in May 2014. In addition to facilitating sample retrieval, the use of automated storages reduces icing of tubes, plates and racks and ensures a more stable storage temperature.

Sample retrieval

The MoBa Cohort welcomes researchers that satisfy some basic requirements to apply for access to answer specific research questions within the overall aim of the MoBa study. Terms and conditions for access to biological samples and associated data from MoBa can be found on the NIPH website (www.fhi.no/studier/den-norske-mor-og-barn-undersokelsen). Samples will only be retrieved and sent to researchers after approval from the MoBa scientific management group, approval from the applicable authorities for the handling and use of the samples for the research purpose in question, following a signed agreement between MoBa and the PI and her/his research institution.

The retrieval process in the MoBa Biobank consists of four main steps after a retrieval list is received from MoBa statisticians or IT-personnel; 1) Retrieving samples from freezers, 2) If necessary, the correct sample volume is pipetted and transferred to a new tube, 3) Retrieval controls to ensure that the correct material is sent, 4) Shipment of samples and sample documentation.

The use of automated freezers has facilitated the retrieving of samples. A retrieval list is sent from the

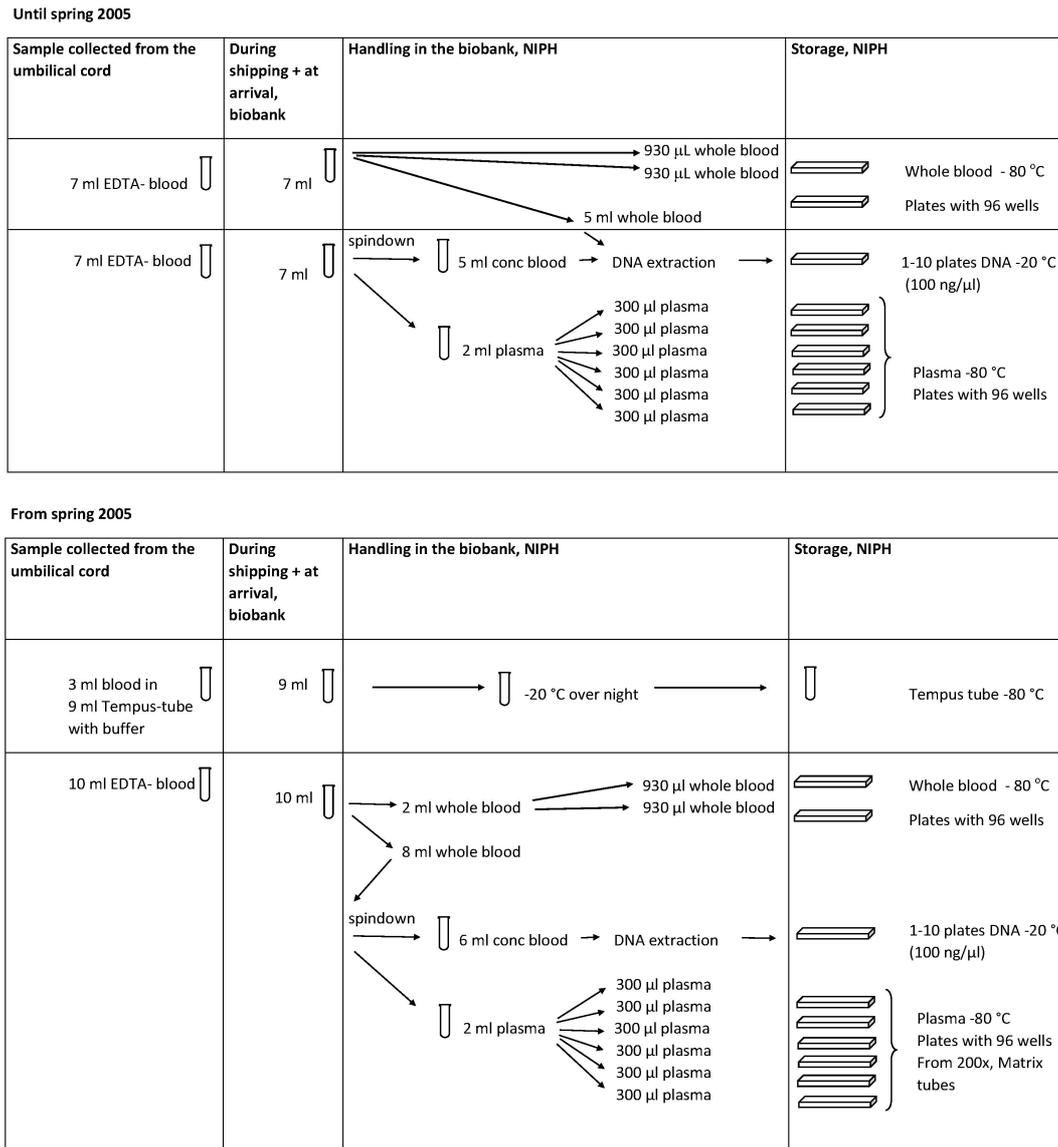


Figure 2. Sample sets collected from the umbilical cord.

MoBa Laboratory Information Management System (MoBa LIMS) to the storage unit, and the correct plates or tubes are retrieved in short time. Manual retrieving of samples from chest freezers in the Biobank is more time consuming. To ensure that the correct plates or tubes are picked, a personal digital assistant (PDA) with a control program is used to scan all plates and barcodes on single tubes.

If it is necessary to pipette and transfer material before delivery, the Biobank has thorough routines both for automated and manually performed retrievals. For automated retrievals, a retrieval file from the MoBa LIMS is transferred to one of our retrieval pipetting robots. The retrieved volume from a certain position and the number of freeze thaw cycles are tracked in the MoBa LIMS.

Before shipment, controls are performed to ensure that the volume and concentration (for DNA) are correct.

SAMPLE INTEGRITY

Background

Sample mix-up in the pre-analytical phase can occur at any time of the collection and processing procedures, and in the MoBa Biobank, action is taken at every step to minimize the frequency of sample mix-up, through standardized operation procedures and several control steps managed by the MoBa LIMS (4). Based on observations in studies using the material, the sample mix-up frequency in MoBa has been suggested to be about 1%.

Method

Based on mother-child duos, a study was set up to estimate the proportion of samples that were mixed-up. To calculate the sample size needed for a reasonably precise estimate of the proportion of the samples that could have been mixed-up, we examined variances in

Table 1. Overview over available biological specimens in The Norwegian Mother and Child Cohort Study (MoBa) to be used for research purposes stored in the biobank.

Participants	Biological Material	Time of Sampling*	Sample	Sample volume	Aliquotes (n)***	Storage material	Storage temp. (°C)
Father	Whole Blood (peripheral; EDTA)	UL	DNA**	Up to 9.3 ml	1-10*930 µl	Deep well plates	-20
	Whole Blood (peripheral; EDTA)	UL	Whole blood	Up to 1860 µl	2*930 µl	Deep well plates	-80
	Whole Blood (peripheral; EDTA)	UL	Plasma	Up to 1.8 ml	6*300 µl	Microtiter plates, single tubes with 2D barcode	-80
Mother	Whole Blood (peripheral; EDTA)	UL,K2	DNA**	Up to 9.3 ml	1-10*930 µl	Deep well plates	-20
	Whole Blood (peripheral; EDTA)	UL, K2	Whole blood	Up to 1860 µl	2*930 µl	Deep well plates	-80
	Whole Blood (peripheral; EDTA)	UL, K2	Plasma	Up to 1.8 ml	6*300 µl	Microtiter plates, single tubes with 2D barcode	-80
	Whole Blood (peripheral; EDTA)	UL	Plasma (PMI****)	Up to 2.79 ml	3*750/930 µl	Single tubes 2D barcode	-80
	Urine (chlorhexidine)	UL	Urine	Up to 5.58 ml	6*750/930 µl	Single tubes 2D barcode	-80
	Whole Blood (peripheral; EDTA)	UL	Whole blood	Up to 3000 µl	1 tube each individual	Single vacutainer tube	-20
Child	Umbilical cord blood (EDTA)	at birth	DNA**	Up to 9.3 ml	1-10*930 µl	Deep well plates	-20
	Umbilical cord blood (EDTA)	at birth	RNA	1 TEMPUS 9 ml tube	1 tube containing 3 ml whole blood	Single vacutainer tubes	-80
	Umbilical cord blood (EDTA)	at birth	Whole blood	Up to 1860 µl	2*930 µl	Deep well plates	-80
	Umbilical cord blood (EDTA)	at birth	Plasma	Up to 1.8 ml	6*300 µl	Microtitre plates, Single tubes with 2 D barcode	-80
	Teeth	Pre-school (Age 6)	Tooth	One milk tooth		RT	RT

Abbreviations & Notes:

EDTA = Ethylenediaminetetraacetic acid

* UL = sample collection at ultrasound appointment around pregnancy week 17-18.

K2 = sample collection during the stay in maternal ward (the majority of samples collected within 2 days post-partum)

** Stored in a mean concentration of 100 ng/µl in each aliquot

*** Numbers of available aliquotes depends on volume of whole blood obtained during blood drawing

**** PMI = plasma EDTA sample collected for environmental purposes

the binomial distribution (10) under the assumption that the estimate would be low, around 1%. We decided that a sample size of 380 would give a precise enough estimate. Randomly collected blood sample sets from 99 women, both during pregnancy and after delivery, and their children (umbilical cord blood) were analyzed for both gender and relatedness. DNA from 83 unrelated fathers was analyzed for gender only.

The samples were analyzed at the Department of Family Genetics at NIPH. All samples were tested for a minimum of 24 genetic markers including gender identification, using the PowerPlex16 kit (Promega, Madison, WI). The sample mix-up frequency was calculated based on relatedness and gender by comparing these results to the recorded information in the MoBa LIMS.

Results

Overall, 99.2% of the samples analyzed showed correct gender and relatedness. Of the 380 samples analyzed, more than one DNA profile was observed in three samples (0.8%, 95% CI 0.77-0.83). The three samples

consisted of at least 30% contribution of DNA from each individual DNA-profile found in the mixture.

Discussion

The 0.8% sample mix-up frequency found in this study corresponds to previous observations.

The Biobank uses barcode labelling and scanning in all sample handling, including all steps in the DNA isolation procedure, to prevent sample mix-up. Thus, any observed sample mix-up was expected to be caused by misidentification of sample sets during sample collection and/or during sample registration in the Biobank and only one DNA profile would have been found in each sample. Surprisingly, more than one DNA profile was found in three of the tested samples. It suggests that sample contamination, probably due to mixture of samples from different individuals during DNA isolation or at the time of registration, is likely to be the cause.

The three samples that were mixed up were collected and handled during the peak time in the MoBa collection period. To minimize errors that can cause sample mix-up in any future biological sample

Table 2. Number of samples in the MoBa Biobank available for retrieval as of August 2014.

Sample type	No of sample sets	No of DNA samples	No of plasma (PE ^c) samples	No of plasma (PMI ^f) samples	No of plasma heparin (PH) samples	No of urine samples without chlorhexidine	No of urine samples with chlorhexidine	No of RNA samples	No of whole blood (FE3 ^g) samples	No of whole blood (FE2 ^h) samples
K1 ^a	100,276	96,450	98,207	69,668	9,884	9,737	68,393	-	77,940	96,973
F ^b	73,113	71,171	72,482	-	-	-	-	-	-	72,163
C ^c	93,327	89,475	91,473	-	-	-	-	45,446	-	51,018
K2 ^d	87,215	80,123	86,426	-	-	-	-	-	-	85,646

a) Mother sample collected at ultrasound

b) Father sample collected at ultrasound

c) Blood sample from the umbilical cord or capillary sample from the baby

d) Mother sample collected at delivery

e) Plasma EDTA samples for all purposes

f) Plasma EDTA samples collected especially for environmental purposes

g) 3 ml EDTA tube with whole blood collected for environmental purposes

h) Whole blood collected for all purposes

collections, staffing should be given high priority, as well as the importance of even stronger focus on safe and reliable routines.

As opposed to the potential harm that sample mix-up of patient samples in a clinical setting can do, a sample mix-up rate below 1% is assumed to have limited effect when studying the strengths of associations (e.g OR, RR) between genotypes and various health outcomes. In genetic analyses, sample contamination will be detected and the contaminated samples will not be used in the statistical analyses. Thus, it will not affect this kind of analyses.

Handling of sample mix-up in the MoBa biobank

Indications of sample-mix up may come from researchers that have received DNA samples for analyses. Questions regarding the gender of sampled subjects or regarding the supposed relatedness between two samples may be raised. On any suspicion of sample mix-up, the Biobank always retrieves DNA and sends it to an independent analysis. If the new analysis shows that we are dealing with sample mix-up, the sample will not be used in further analysis.

Some of the retrievals from the last years included thorough QC analysis by analytical laboratories. The returned results to MoBa have shown less than 0.8% sample-mix up. The samples that have shown contamination or misidentification have been retrieved and analyzed once more, resulting in three equally sized groups of classification; 1) samples with more than one DNA profile, 2) samples where the gender and/or relatedness were wrong and 3) samples with no errors found.

THE COLLECTION

Recruitment to the MoBa cohort study ended December 2008, and the last samples in the main study were collected June 2009 – ten years after the first sample arrived at the Biobank. The numbers of available samples are presented in Table 2. The overall response rates for blood collection are good. Sample sets are

available for 89, 77, 84 and 81% of the eligible mothers at ultrasound (K1 sample), mothers after delivery (K2), fathers and children. The larger number of missing samples for mothers after delivery is mainly due to the sampling routines. In some hospitals, these samples were taken during weekdays only and short hospital stays resulted in a certain amount of missing values.

From January 2002, two additional blood tubes and a urine sample, the so-called "environmental samples", were collected at the ultrasound appointment. The purpose was to measure the levels of environmental exposures that the pregnant women and her unborn child could have been exposed to. From May 2003, there was a change in the protocol regarding type of tube for urine and plasma collection for these samples (4). In total there are almost 81 000 sample sets collected from the mothers at ultrasound that includes one or more of the environmental samples. The numbers of the different available plasma and urine environmental samples are lower than for whole blood due to the changes in the protocol.

Looking at the children, the number of RNA samples is low because the collection of these samples first started in 2005. Whole blood was collected during the entire period, and the lower number of these samples has two reasons – only samples with 7 ml or more were considered for whole blood storage and the samples should not be coagulated. If these criteria were not met, whole blood was used for DNA isolation only. Characteristics of children with and without available biological material are presented in Table 3. The number of missing samples is higher for children with low birth weight and preterm birth, probably because of the priorities during these deliveries.

Table 4 presents the number of available DNA mother and child duos and mother, father and child trios. There are more than 90 000 sample sets from mother-child duos in the MoBa, and as of August 2014, DNA is available for retrieval for more than 86 000 of these duos. Likewise, there are more than 62 000 samples sets from mother-father-child trios, and DNA is available for more than 58 000 of these

Table 3. Characteristics of children with and without available biological material.

	Available biological material	Not available biological material
Female, %	48.7	48.4
Birth weight, g	3,600 (553)	3,360 (802)
Birth weight <2500 g, %	3.0	11.3
Gestational age, weeks	39.5 (1.7)	38.6 (3.4)
Gestational age <37 weeks, %	4.9	14.1

Values are percentages and means (standard deviation)

Table 4. Available DNA duos and trios in the MoBa as of August 2014.

Type	Number of received sample sets in Moba	Number available for retrieval of DNA in MoBa
Mother-child duos	90 623	86 274
Mother-father-child trios	62 776	58 557

Table 5. Time from collection till the sample is received in the NIPH Biobank.

	Same day	1 day	2 days	3 days	4 days	5+ days
Total	11 767 (3%)	191 293 (54%)	51 029 (14%)	62 515 (18%)	25 522 (7%)	12 105 (3%)
K1	1 393 (1%)	63 682 (63%)	10 705 (11%)	15 850 (16%)	6 202 (6%)	2 469 (3%)
F	965 (1%)	46 797 (64%)	7 343 (10%)	12 014 (16%)	4 179 (6%)	1 889 (3%)
C	5 022 (5%)	32 208 (32%)	23 323 (25%)	16 947 (18%)	10 844 (12%)	5 156 (6%)
K2	4 387 (5%)	48 606 (56%)	9 658 (11%)	17 704 (20%)	4 297 (5%)	2 591 (3%)

trios. If we have biological material from the mother, father and child, and we still do not have available DNA trios, it is usually the DNA from the father or the child that is missing.

Temperature and duration of sample transport is a critical pre-analytical factor (11,12), and as discussed by Rønningen et al, most of the MoBa samples were shipped by ordinary mail. Samples from the three closest hospitals were delivered on a daily basis (except weekends) directly to the Biobank with the hospital couriers. All samples were processed on the day of receipt. Table 5 shows the time from collection to the samples were received and processed in the biobank. A majority of the samples, approximately 57%, are received the same or the day after collection. For all samples, except samples from the umbilical cord, more of them were received three days after collection than two days after. The reason for this is that most samples drawn Fridays first were received on Mondays, and samples drawn Saturdays that were sent by ordinary mail first were received Tuesdays. As samples from the parents usually were drawn at daytime, samples from the umbilical cord naturally were taken both days and nights. Therefore the pattern is different for these samples.

Both the date of sample collection and sample receipt are registered in the MoBa Laboratory Informa-

tion Management System (called MoBaStudy), and it is possible to request a maximum difference in days between collection and receipt, when applying for samples from MoBa.

USE OF THE BANKED MOBA SAMPLES

Since the first retrieval in 2006, the number of national and international research groups that base their research on the MoBa biobank and its comprehensive dataset linkable to national databases, has steadily increased. So far, 47 publications based on MoBa data have utilized the banked samples, including one Ph.D. thesis.

As of August 2014, more than 100 000 samples have been retrieved for research purposes and analyses of genetic, epigenetic and environmental biomarkers have been performed. The number of samples shipped annually from the biobank for research purposes is given in Table 6. As of today, the retrievals of DNA and plasma samples constitute 58 and 40% of the approved research applications, respectively. Urine and RNA samples represent the remaining retrievals.

In addition to MoBa's size, it's banked biological material from families (mother, father, child) collected from mid-pregnancy, makes it a valuable research resource to study how genetic and environmental risk factors affect the health and development of children and their parents. The collected biological material is suitable for genetic analyses, transcript profiling and gene expression studies, biomarker analyses and toxicology analyses, and examples of the research projects utilizing the biological material collected through MoBa and its rich dataset, will be given in separate contributions within this special issue.

Table 6. Number of samples retrieved and shipped to sub-projects.

Year	No of retrievals	No of retrieved samples	No of DNA	No of plasma	No of urine	No of RNA
2006	1	53		53		
2007	5	1,847	1,676	171		
2008	9	13,685	6,876	6,809		
2009	6	5,684	3,874	1,000	810	
2010	7	3,024	2,023	1,001		
2011	4	3,173	710	2,313	150	
2012	9	14,808	4,155	10,653		
2013	10	29,846	12,922	15,601		1,323
2014*	8	32,326	28,047	3,741		538
In total	51	104,446	60,283	41,342	960	1,861

* Until August 25th, 2014

ACKNOWLEDGEMENT

The Norwegian Mother and Child Cohort Study is supported by the Norwegian Ministry of Health and the Ministry of Education and Research, NIH/NIEHS (contract no N01-ES-75558), NIH/NINDS (grant no.1 U01 NS 047537-01 and grant no.2 U01 NS 047537-06A1). Many have contributed

to the MoBa Biobank from the establishment until present, and we are grateful to everyone who are or have been involved in the Biobank Team. We are also grateful to all the participating families in Norway who take part in this ongoing cohort study.

REFERENCES

1. Magnus P, Irgens LM, Haug K, Nystad W, Skjærven R, Stoltenberg. Cohort Profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol* 2006; **35**: 1146-1150.
2. Norwegian Mother and Child Cohort Study. Revised protocol. 2012. <http://www.fhi.no>.
3. Nilsen, R, Vollset SE, Gjessing HK, Skjærven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol* 2009; **23**: 596-608.
4. Rønningen KS, Paltiel L, Meltzer HM, Nordhagen R, Lie KK, Hovengen R, et al. The biobank of the Norwegian mother and child cohort study: A resource for the next 100 years. *Eur J Epidemiol* 2006; **21**: 619-625.
5. Paltiel L, Aarem J, Bækken S, Stensrud NK, Harbak K. Biospecimen program in the biobank of the Norwegian Institute of Public Health. *Norsk Epidemiologi* 2012; **21** (2): 225-229.
6. Tvinnereim HM, Lygre GB, Haug K, Schreuder P, Klock K. A biobank of primary teeth within the Norwegian Mother and Child Cohort Study (MoBa): a resource for the future. *Paediatr Perinat Epidemiol* 2012; **26**: 264-271.
7. Stoltenberg C, Schjølberg S, Bresnahan M, Hornig M, Hirtz D, Dahl C, et al. The autism birth cohort (ABC): A paradigm for gene-environment-timing research. *Mol Psychiatry* 2010; **15** (7): 676-680.
8. Duale N, Brunborg G, Rønningen KS, Briese T, Aarem J, Aas KK, et al. Human blood RNA stabilization in samples collected and transported for a large biobank. *BMC Res Notes* 2012; **5**: 510.
9. Duale N, Lipkin WI, Briese T, Aarem J, Rønningen KS, Aas KK, et al. Long-term storage of blood RNA collected in RNA stabilizing Tempus tubes in a large biobank – evaluation of RNA quality and stability. *BMC Res Notes* 2014; **7**: 633.
10. Larsen RJ, Morris LM. An introduction to mathematical statistics and its applications, 2nd edn. New Jersey: Prentice-Hall, 1986: 276-282.
11. Campbell LD, Betsou F, Garcia DL, Giri JG, Pitt KE, Pugh RS, et al. International society for biological and environmental repositories (ISBER) 2012. Best practices for repositories: Collection, storage, retrieval, and distribution of biological materials for research. *Biopreserv Biobank* 2012; **10** (2): 79-161.
12. Tuck MK, Chan DW, Chia D, Godwin AK, Grizzle WE, Krueger KE, et al. Standard operating procedures for serum and plasma collection: Early Detection Research Network Consensus statement. *J Proteome Res* 2009; **8** (1): 113-117.