

Swedish National  
Breast Cancer Study

---

**Karma**

## Karma Core variables

The Karma core variables are central variables picked from the entire Karma data collection and refer to the date at baseline, i.e. first assessment in Karma, if not stated otherwise. Karma holds the following core variables.

### Demographics

StudieID	ID of person in study
study_entrydate	Entry date of woman in study, the first date of survey or blood drawing.
age	Age at assessment. First date of survey interview or blood drawing.
yearofbirth	Year of birth.
interview_end	Survey finish time for complete response.
bmi	Body Mass Index
length	Length in cm
weight	Weight in kg

(Survey based variable at baseline in Karma)

## Medication

painkillers	Use of OTC painkillers in last year. See question sb9 in questionnaire, i.e. any of paracetamol, Ibuprofen, Diklofenak, acetylsalicylic acid.
-------------	---

(Survey based variable at baseline in Karma)

## Cancer and benign breast disease

benign_breastdisorder	Benign breast disorder
breastcancer	Breast cancer
othercancer	Other cancer than breast

(Survey based variable at baseline in Karma)

## Cancer in family

breastcancer_in_family_nbr	Number of breast cancer cases in family (1 <sup>st</sup> degree relatives, including children)
breastcancer_mother_age	Age at breast cancer onset for mother
breastcancer_mother_sister1_age	Age at breast cancer onset for (oldest) sister

breastcancer_mother_sister2_age	Age at breast cancer onset for (second oldest) sister
breastcancer_in_relative_nbr	Number of breast cancer cases in relatives (2 <sup>nd</sup> degree relatives)
ovariancancer_in_family	Ovarian cancer in family (1 <sup>st</sup> degree relatives, including children)

(Survey based variables at baseline in Karma)

### **BRCA mutation**

brca_mutation	Gene test for BRCA mutation detection.
---------------	--

(Survey based variable at baseline in Karma)

### **Education**

education_level	Education level, see question bg5 in questionnaire for interpretation
-----------------	---

### **Ancestry**

european_ancestry_iso	European ancestry, i.e. study participant born in Europe and parents born in Europe
born_country	Birth country of study participant
immigration_date	Date of immigration (first?)
emigration_date	Date of emigration (last?)

(All are survey based variables at baseline in Karma)

### **Reproductive health**

menarche_age	Age at menarche
cycles_irregular	Ever had irregular menstrual cycles
contraception_ever	Ever used oral contraceptives
pregnant_times	Number of pregnancies
birth_times	Number of births
age_at_firstchild	Age at first child birth
wom_pregn_tried1yr	Infertility status (tried to get pregnant for one year unsuccessfully)
menopause_status	1 is pre, 2 is peri, 3 is postmenopausal
postmenopausal	Menopause due to no period in last year.
postmenopausal_before_bc	Menopausal at least 1 year before BC. (a missing menopause age for a woman with BC over age 50 is set to 1)
menopause_age	Age at menopause

hrt_ever	Ever used hormonal treatment
hrt_status	2 is current user, 1 is previous user, 0 is never user. Variable applies to systemic HRT use only
hrt_status_at_bc	HRT status at time of BC
wom_pregn_tried1yr	Woman tried to get pregnant for at least one year without success.

(Survey based variables at baseline in Karma)

### **Living habits**

alcohol_gram_week	Grams of alcohol per week (based on last year consumption).
smoking_packyears	Number of pack years (1.0 means 20 cigarettes per day for 1 year).
smo_present	Deprecated.
smoking_status	2 is current user, 1 is previous user, 0 is never user.

(Survey based variables at baseline in Karma)

### **Physical activity**

pa_daily_total_controlled	MET (metabolic equivalent of task) of daily physical activity as estimated during last month since interview
---------------------------	--

(Survey based variables at baseline in Karma)

### **Quality of Life**

ql	Overall quality of life score (QLQ-C30).
----	--

(Survey based variable at baseline in Karma)

### **Breast cancer and breast disease**

bc_1stdiagdate	Breast cancer onset date
bc_1stdiagside	Breast cancer laterality (1=right)
bc_1stdiagage	Age at breast cancer onset
bc_2nddiagdate	Recurrence, second date
bc_2nddiagside	Recurrence, second laterality
bc_2nddiagage	Age at 2 <sup>nd</sup> breast cancer onset
icd10_“code”	ICD10 breast cancer codes
bc_malign_1stdiagdate	Malign invasive breast cancer onset

bc_case	Had a breast cancer
bc_prevalent	Had a breast cancer prior to
bc_malign	Breast cancer is invasive malign
bc_incident	Had a breast cancer since Karma study started
bc_incident_1stcancer	Had 1st breast cancer after joined Karma study (blood drawing date vs. BC date) If blood date is missing, survey date is used.
bc_case_1yr	Had a breast cancer in last year
bc_case_Xyr	Had a breast cancer in last X years
bc_prevalent	Had a breast cancer before entered Karma study
bc_case_karmastarted	Had a breast cancer since the start of the Karma study
bc_invasive_since2007	Invasive breast cancer since INCA started in 2007
othercancer_1stdiagdate	Onset date for non-BC cancer
othercancer_1stdiagage	Onset age for non-BC cancer
othercancer_malign_1stdiagdate	Onset date for non-BC invasive malign cancer
othercancer_malign_1stdiagage	Onset age for non-BC invasive malign cancer
othercancer_case	Had another cancer than breast cancer
othercancer_prevalent	Prevalent other cancer

othercancer_malign	At least one other cancer is invasively malign
cancer_nbr_swecan	Number of cancers as registered in Swedish cancer register (2 year lag)
cancer_malign_nbr_swecan	Number of invasive malign cancers as registered in Swedish cancer register (2 year lag)
nonmelanom_skincancer_malign	Non-melanoma skin cancer, malignant
nonmelanom_skincancer_case	Non-melanoma skin cancer, malignant or benign
oc_case	Ovarian cancer case

(Above variables are current cancer status based on Swedish cancer register, the RCC STHLM breast cancer register, and the INCA breast cancer register)

### **Tumor characteristics**

t	Tumor size
n	Lymph node metastases
m	Distant metastases
er	ER receptor status
her2	HER2 status



(Information from Swedish Cancer Register for worst tumor in case woman had more occurrences)

### **Cause of death**

deathcause	Cause of death (ICD code)
deathdate	Date of death
deathcause_bc	Death caused by breast cancer
deathcause_any	Death not caused by any reason

### **Karma unit**

ipt_karma_unit	Unit performing the blood sampling
ipt_bloodsampling_date	Date of blood sampling
ipt_bloodpressure_diastolic	Diastolic blood pressure at assessment
ipt_bloodpressure_systolic	Systolic blood pressure at assessment

(Variables are at the time of baseline and was registered by nurses)

### **Biobank**

bbk_alquot_saldo	Current number of aliquots remaining in biobank.
------------------	--

## Mammogram measures

volpara_mammography_date	Mammography date for Volpara measurement closest in time to baseline for controls and closest prior to BC diagnosis date for cases.
volpara_viewposition	Measurement performed on CC or MLO view.
VolumetricBreastDensity	Volumetric breast density (%), mean of right/left breast, MLO-view
BreastVolumeCm3	Total breast volume, mean of right/left breast, MLO-view
stratus_mammography_date	Mammography date for Stratus measurement closest in time to baseline for controls and closest in time prior to BC diagnosis date for cases.
stratus_viewposition	Measurement performed on CC or MLO view.
cbirads	Categorization of stratus_pd into four BI-RADS similar categories
stratus_pd	Mammographic density measurement based on Stratus area based method. Mean value of left/right mammogram, MLO view.
stratus_densearea_cm2	Measured dense area of breast (cm <sup>2</sup> ). Mean value of left/right mammogram, MLO view.

stratus\_breastarea\_cm2      Measured breast area of breast (cm<sup>2</sup>).  
Mean value of left/right mammogram,  
MLO view.

cbirads\_gt3mnt      Same as above, but with mammogram  
taken at least 90 days prior to breast  
cancer diagnosis date.

stratus\_pd\_gt3mnt

stratus\_densearea\_cm2\_gt3mnt

stratus\_breastarea\_cm2\_gt3mnt

### **Breast cancer risk**

age\_riskassessed      Age of women when risk is assessed.  
Survey interview date for controls and age  
at breast cancer for cases.

risk\_breastcancer\_own\_10yr      10-year risk for developing breast cancer.

risk\_breastcancer\_own\_lifetime      Lifetime risk (remaining lifetime from  
current age) for developing breast cancer.

(Tyrer-Cuzick risk model assessment. Variables are at the time of baseline  
and based on survey variables)

### **Genotype data**

Published SNPs related to breast density

rs10995190, rs10509168 (ZNF365 gene)

rs10483813 (Rad5 1L1 gene)

rs3817198 (LSP1 gene)

rs1520220 (IGF1 gene)

rs2946834 (IGF1 gene)

## **Cause of death**

dodsdatn	Date of death
ulorsak	Main cause of death, ICD-10.

(Variables on current status based on the Swedish case of death register)

## **Mammogram occasions**

rounds_image_prestudy	Rounds of mammograms available for woman prior to date of study entry
rounds_image_prestudy_raw	Rounds of raw mammograms available for woman prior to date of study entry
rounds_image_instudy	Rounds of mammograms available for woman in study, i.e. at study entry or after study entry
rounds_image_instudy_raw	Rounds of raw mammograms available for woman in study, i.e. at study entry or after study entry
rounds_image_prediag	Rounds of mammograms available for woman prior to (and up to same date of) her date of 1 <sup>st</sup> BC diagnosis date

rounds_image_prediag_raw	Rounds of raw mammograms available for woman prior to (and up to same date of) her date of 1 <sup>st</sup> BC diagnosis date
rounds_image_postbc	Rounds of mammograms available for woman after her date of 1 <sup>st</sup> BC diagnosis date
rounds_image_postbc_raw	Rounds of raw mammograms available for woman after her date of 1 <sup>st</sup> BC diagnosis date

(Variables based on KARMA mammogram repository)

### **Available data sources**

source_image	Image data available
source_image_analog	Analog image data available
source_image_digital	Digital image data available
source_image_prediag	Image data available prior to (and up to same day of) BC diagnosis
source_image_postdiag	Image data available after BC diagnosis
source_genotype	Genotype data available
source_genotype_cogs	Genotype data available from COGS
source_genotype_onco	Genotype data available from Oncoarray

source\_survey

Survey data available