Karisma Core variables

Version History

Version	Date	Comments
1.0	2020-06-23	Initial version
1.1	2021-02-23	Add new variables: cbirads_0m, cbirads_6m, cbirads_exit

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

Demographics

studieid   ID of person in study

age_studyentry   Age at assessment. First date of survey interview and blood drawing.

dateofbirth   Date of birth.

started_time   Baseline survey started time for response.

finished_time   Baseline survey finish time for complete response.

rand_group   Randomization group

randomization_nr   Randomization number

firstpill_date   Date of the first pill

(Variables at the time of baseline)

visitdate_0m	Entry date of the woman in the study, first date of the survey and blood drawing.
bmi_0m	Body Mass Index at study entry
length_0m	Length in cm at study entry
weight_0m	Weight in kg at study entry
bloodpressure_systolic_0m	Systolic blood pressure at study entry
bloodpressure_diastolic_0m	Diastolic blood pressure at study entry
imp_out_0m	Number of pills out at study entry
survey_status_0m	Baseline survey status:
	0 = not answered, 1 = complete, 2 = incomplete

(Variables at 1 month)

survey_status_1m	Follow-up survey status at 1 month:
	0 = not answered, 1 = complete, 2 = incomplete

(Variables at 3 months)

survey_status_3m	Follow-up survey status at 3 months:
	0 = not answered, 1 = complete, 2 = incomplete

(Variables at 6 months)

visitdate_6m	Visit date at 6 months
bmi_6m	Body Mass Index at 6 months
length_6m	Length in cm at 6 months
weight_6m	Weight in kg at 6 months
imp_in_6m	Number of returned pills out at 6 months
imp_in_forgot_6m	Forgot to bring remaining pills at 6 months
survey_status_6m	Follow-up survey status at 6 months:
	0 = not answered, 1 = complete, 2 = incomplete

(Variables at 12 months)

visitdate_12m	Visit date at 12 months
bmi_12m	Body Mass Index at 12 months

length_12m	Length in cm at 12 months
weight_12m	Weight in kg at 12 months
survey_status_12m	Follow-up survey status at 12 months:
	0 = not answered, 1 = complete, 2 = incomplete

Study Interruption

interrupt_cause	Cause of study interruption:
	100 = the participant wishes to interrupt her participation in the study
	200 = Unacceptable side effects
	300 = Interrupt due to violation to protocol
	400 = Severe adverse events (SAE).
imp_stop_decision	Decision to interrupt study:
	1 = PI decision
	2 = Participant's decision.
interrupt_date	Date of study interruption
imp_stop_date	Date of stop taking pills
study_stop	Stop study participation

(Variables at date of exit examination)

visitdate_exit	Visit date at exit examination
bmi_exit	Body Mass Index at date of exit examination
length_exit	Length in cm at date of exit examination
weight_exit	Weight in kg at date of exit examination
survey_status_exit	Follow-up survey status at date of exit examination:
	0 = not answered, 1 = complete, 2 = incomplete
imp_in_exit	Number of returned pills out at date of exit examination
imp_in_forgot_exit	Forgot to bring remaining pills at date of exit examination

Medication

(Survey based variable at baseline in Karisma)

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

Cancer and benign breast disease

(Survey based variables at baseline in Karisma)

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

Cancer in family

(Survey based variables at baseline in Karisma)

breastcancer_in_family_nbr	Number of breast cancer cases in family (1st degree relatives, including children)
breastcancer_mother	Mother has breast cancer
breastcancer_sisters	Sister(s) has breast cancer, nbr
breastcancer_daughter	Daughter has breast cancer
breastcancer_mother_age	Age at breast cancer onset for mother
breastcancer_sister1_age	Age at breast cancer onset for (oldest) sister
breastcancer_sister2_age	Age at breast cancer onset for (second oldest) sister
breastcancer_in_family	Breast cancer in family (1st degree relatives, including children)
ovariancancer_in_family	Ovarian cancer in family (1st degree relatives, including children)
breastcancer_in_relative_nbr	Number of breast cancer cases in relatives (2nd degree relatives)

BRCA mutation

(Survey based variable at baseline in Karisma)

brca_mutation   Gene test for BRCA mutation detection.

Education

(Survey based variable at baseline in Karisma)

education_level   Education level, see question bg5 in questionnaire for interpretation

Ancestry

(Survey based variables at baseline in Karisma)

european_ancestry_iso   European ancestry, i.e. studyparticipant born in Europe and parents born in Europe

born_country   Birth country of study participant

Reproductive health

(Survey based variables at baseline in Karisma)

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_firstbirth   Age at first child birth

wom_pregn_tried1yr   Infertility status (tried to get pregnant for one year unsuccessfully)

menopause_status   1 = pre, 2 = peri, 3 = postmenopausal (age cut-off for missings: pre - peri age 46, peri - post age 55)

postmenopausal   Menopause due to no period in last year (age cut-off for missings: pre - post age 50)

menopause_age	Age at menopause
hrt_ever	Ever used hormonal treatment
hrt_status	2 = current user, 1 = previous user, 0 = never user. Variable applies to systemic HRT use only

Alcohol and tobacco use

(Survey based variables at baseline in Karisma)

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).
smoking_status	2 = current user, 1 = previous user, 0 = never user

Quality of Life

(Survey based variable at baseline in Karisma)

ql	Overall quality of life score (QLQ-C30)

Sense of Coherence

(Survey based variable at baseline in Karisma)

comprehensibility	Comprehensibility means that everyday life can be predicated, makes sense and is not confusing
manageability	Manageability means that everyday life is within the control to handled and take care of with sufficient resources
meaningfulness	Meaningfulness means that everyday life is a source of satisfaction, being interesting and exciting, and is really worth doing
kasam	kasam scoring is the sum of the three sub-scales "comprehensibility", "manageability" and "meaningfulness"

Karma unit

(Variables are at the time of baseline)

ipt_karma_unit	Unit performing the blood sampling
ipt_registration_date	Date of study inclusion
ipt_consent_k2	Karisma study consent

FACT-ES (Functional Assessment of Cancer Therapy)

(Survey based variables at baseline/follow-up in Karisma)

ess_0m	FACT-ES score baseline
ess_1m	FACT-ES score 1 month
ess_3m	FACT-ES score 3 months
ess_6m	FACT-ES score 6 months
ess_exit	FACT-ES score exit
ess_12m	FACT-ES score 12 months

MMAS- 8 (Morisky Medication Adherence Scales)

(Follow-up survey based variables in Karisma)

mmas_1m	MMAS score 1 month
mmas_3m	MMAS score 3 months
mmas_6m	MMAS score 6 months
mmas_exit	MMAS score exit
mmas_12m	MMAS score 12 months

Imaging measurement of volumetric breast density

density_0m	Volpara density at study entry
mammography_date_0m	Mammography date baseline
mammography_date_6m	Mammography date 6 months

mammography_date_exit	Mammography date exit
mammography_date_12m	Mammography date 12 months
stratus_pd_0m	Mammographic density measurement based on Stratus baseline
stratus_pd_6m	Mammographic density measurement based on Stratus 6 months
stratus_pd_exit	Mammographic density measurement based on Stratus at date of exit examination
stratus_pd_12m	Mammographic density measurement based on Stratus 12 months
stratus_da_0m	Dense area (cm2) baseline
stratus_da_6m	Dense area (cm2) 6 months
stratus_da_exit	Dense area (cm2) exit study
stratus_da_12m	Dense area (cm2) 12 months
cbirads_0m	Computer generated BI-RADS breast composition classification, baseline
cbirads_6m	Computer generated BI-RADS breast composition classification, 6 months
cbirads_exit	Computer generated BI-RADS breast composition classification, exit

Analysis variables

analysis_itt	Indication of the full population that is used for some analyses
analysis_mitt	Indication of the intention to treat population
analysis_pp	Indication of the per-protocol analyses population
responder_density	Mammographic density responder to tamoxifen as defined in the study protocol, i.e. median relative density decrease in the 20 mg arm in the per-protocol population at 6 months compard with baseline (-9.7% decrease).

Tamoxifen and metabolite concentrations

Z_endoxifen_ng_mL	Endoxifen (ng/mL), below detection limit = 0.05

Z_endoxifen_nmol_L	Endoxifen (nmol/L), below detection limit = 0.134
four_OH_TAM_ng_mL	Afimoxifene (ng/mL), below detection limit = 0.1
four_OH_TAM_nmol_L	Afimoxifene (nmol/L), below detection limit = 0.258
N_DM_TAM_ng_mL	N-desmethyltamoxifen (ng/mL), below detection limit = 1.0
N_DM_TAM_nmol_L	N-desmethyltamoxifen (nmol/L), below detection limit = 2.797
TAM_ng_mL	Tamoxifen (ng/mL), below detection limit = 1.0
TAM_nmol_L	Tamoxifen (nmol/L), below detection limit = 2.692
tamoxifen_analysis_date	Date of tamoxifen and metabolite analysis

CYP2D6

CYP2D6_score	Cyp2D6 gene activity score (Goetz)
	https://pubmed.ncbi.nlm.nih.gov/29385237/
	https://ascpt.onlinelibrary.wiley.com/doi/epdf/10.1002/cpt.1007
	> 2.0 = CYP2D6 ultrarapid metabolizer
	1.5 and 2.0 = CYP2D6 normal metabolizer
	1.0 = CYP2D6 normal metabolizer or intermediate metabolizer (controversy remains)
	0.5 = CYP2D6 intermediate metabolizer
	0 = CYP2D6 poor metabolizer
CYP2D6_group	Cyp2D6 gene activity group:
	3 = CYP2D6 score > 2.0
	2 = CYP2D6 score 1.5 and 2.0
	1 = CYP2D6 score 0.5 and 1.0
	0 = CYP2D6 score 0

Available data sources

source_sourvey	Survey data available