

Delstudie:

KARISMA 2

Karisma extended datasets

The Karisma extended datasets are variables derived from the raw variables collected in Karisma. Karisma holds currently extended datasets on reproductive health, smoking, alcohol, quality of life, and sense of coherence.

Version 1.1

Contents

| | |
|---|----|
| Reproductive and menstrual health | 2 |
| Living habits | 13 |
| Quality of life | 23 |
| Sense of coherence..... | 26 |
| Update history | 28 |

Reproductive and menstrual health

Version 2.1 - August 2020, contact: mikael.eriksson@ki.se

Introduction

This document describes the menstrual and reproductive variables that were originally collected for the Karma study and the data processing and recoding that was done for analysis purposes.

This document is organized in the following sections:

1. Menstruation
2. Menopause
3. Pregnancy and child
4. Contraceptive use
5. Use of menopausal hormones
6. Infertility

Within each section, a definition is given for each newly created variable.

Codebook

| Variable name | Description | Type/format | Value description |
|------------------|--------------------|-------------|---|
| Menarche_age | Age at menarche | Numeric | Continuous in years |
| Cycles_irregular | Cycle irregularity | Numeric | 0 = regular 1 = irregular |
| Menopause_status | Menopausal status | Numeric | 1 = premenopausal 2 = perimenopausal 3 = postmenopausal |
| Menopause_age | Age at menopause | Numeric | Continuous in years |

| | | | |
|-----------------------------|-------------------------------------|---------|---------------------|
| Pregnant_times | Number of pregnancies | Numeric | Ordinal (0-15) |
| Birth_times | Number of births | Numeric | Ordinal (0-11) |
| Age_at_firstbirth | Age at first birth | Numeric | Continuous in years |
| Contraception_oral_ever | Ever use of oral contraceptives | Numeric | 0 = no 1 = yes |
| Contraception_systemic_ever | Ever use of systemic contraceptives | Numeric | 0 = no 1 = yes |
| Contraception_ever | Ever use of contraceptives | Numeric | 0 = no 1 = yes |

1. MENSTRUATION

1A. Age at menarche

Variable:

menarche_age

Definition:

Age at menarche is defined as the age at first menstruation. Women who reported having no periods are classified as missing on this variable.

1B. Cycle irregularity

Variable:

cycles_irregular

Definition:

A woman is considered to have irregular cycle if the difference between the shortest and the longest cycle is ≥ 7 days.

2. MENOPAUSE

2A. Menopausal status

Variable 1:

menopause_status (pre, peri/unknown, post)

Definition:

Menopausal status is defined as a 3-category variable (pre, peri/unknown, post) based on menstrual history, reasons for having no periods and the interval between last reported period and date of questionnaire when still menstruating. The categories of menopausal status are as follows:

- pre: having menses over the past year and still menstruating in the 3 months prior to study entry
- peri: having menses over the past year but no longer menstruating during the 3 months prior to study entry
- post: having no menses over the past year or had oophorectomy

Age cutoff handling:

Age cutoffs for menopausal status are used for women with missing or incomplete questionnaire data on menstruation status and for those who had no menses due to gynecological surgery other than oophorectomy. These women are classified as follows:

- pre: < 46 years
- peri/unknown: 46-55 years
- post: > 55 years

Exception handling premenopausal reasons:

We can distinguish between 3 types of reasons for having no periods during the past year (Figure 1):

- 1: specified reasons for menopause: menopause, medication, intense exercise, anorexia and oophorectomy
- 2: non-specified reasons for menopause: other and 998
- 3: premenopausal reasons: pregnancy and contraceptive use

Women who reported a specified reason for menopause are classified as postmenopausal regardless of any other reasons reported. Women who reported a premenopausal reason only are classified as premenopausal if they were ≤ 55 years and as postmenopausal if they were > 55 years.

Women who reported a non-specified reason are classified as postmenopausal if not in combination with a premenopausal reason only.

Women who reported a non-specified reason in combination with a premenopausal reason are classified as premenopausal if they were ≤ 55 years and as postmenopausal if they were > 55 years.

Menopausal status definition for Karma papers for "menopause_status":

Women were considered premenopausal when they reported having menses over the past year and when they were still menstruating in the 3 months prior to study entry. Women were considered postmenopausal when they had no menses over the past year or when they reported oophorectomy. Women were considered perimenopausal if they had menses over the past year but were no longer menstruating during the last 3 months. Age cutoffs for menopausal status were used for women having missing or incomplete questionnaire data and for those who had no

menses due to gynecological surgery other than oophorectomy. These women were considered postmenopausal if they were > 55 years, perimenopausal/unknown if they were between 46 and 55 years of age and premenopausal if they were < 46 years.

Variable 2:

postmenopausal (pre, post)

Definition:

Postmenopausal status is defined as a 2-category variable (pre, post) based on menstrual history, reasons for having no periods and the interval between last reported period and date of questionnaire. The categories of menopausal status are as follows:

pre: having menses over the past year

post: having no menses over the past year or had oophorectomy

Women who had no menses in the past year due to pregnancy, breast feeding, or use of contraceptives are in addition defined as premenopausal.

Age cutoff handling:

Age cutoffs for menopausal status are used for women with missing or incomplete questionnaire data on menstruation status and for those who had no menses due to gynecological surgery other than oophorectomy. These women are classified as follows:

pre: <= 50 years

post: > 50 years

2B. Age at menopause

Variable:

menopause_age

Definition:

Age at menopause is defined as the age at last menstruation for women who are postmenopausal according to **menopause_status**.

3. PREGNANCY AND CHILD

3A. Number of pregnancies

Variable:

pregnant_times

3B. Number of births

Variable:

birth_times

3C. Age at first birth

Variable:

age_at_firstbirth

Definition:

Age at first birth is defined as the time interval between year of birth and birthyear of first child when birthyear of first child > age at menarche.

4. CONTRACEPTIVE USE

Variables:

contraception_oral_ever
contraception_systemic_ever
contraception_ever

Definition:

We can define contraceptive use at three different levels (see Figure 2):

- 1: oral contraceptives (OCs): this group includes pill users only: mini pill, combination pill and pill type unknown
- 2: systemic hormonal contraceptives: this group includes users of systemic hormonal contraceptives: mini pill, combination pill, pill type unknown, hormone patch, p-injection and implant.
- 3: hormonal contraceptives: this group includes all types of hormonal contraceptives, both local and systemic: mini pill, combination pill, pill type unknown, hormone patch, p-injection, implant, hormone IUD, vaginal ring, hormone spray and hormone gel.

Women are classified as oral contraceptive users if they reported any of the OCs listed above.

Women are classified as ever users of systemic hormonal contraceptives if they reported any of the systemic hormonal contraceptives listed above. Women are classified as ever hormonal

contraceptive users if they reported any of the hormonal contraceptives listed above. Women who reported 'copper IUD' only are classified as non-users of hormonal contraceptives. Women who reported 'IUD type unknown' only are classified as missing on hormonal contraceptive use.

Note: other types of contraceptives (i.e. pill type unknown, vaginal ring, hormone spray and hormone gel) were recovered from the 'other' category. See Excel file 'Recategorization of contraceptive use in Karisma' for details regarding the recoding of the 'other' category of contraceptive use.

6. INFERTILITY

Variables:

wom_pregn_tried_1yr

wom_infertility_treat

wom_infertility_treat_whc_1

wom_infertility_treat_whc_2

wom_infertility_treat_whc_3

wom_infertility_treat_whc_5

wom_infertility_treat_whc_6

wom_infertility_treat_whc_998

Note: other types of treatment were recovered from the 'other' category. See Excel file 'Recategorization of fertility treatment in Karma' for details regarding the recoding of the 'other' category (wom_infertility_treat_whc_6).

Living habits

Version 2 - August 2019, contact: mikael.eriksson@ki.se

Introduction

This document describes the lifestyle variables that were originally collected for the Karma study and the data processing and recoding that was done for analysis purposes.

This document is organized in the following sections:

1. Smoking
2. Alcohol

Within each section, a definition is given for each newly created variable. For some variables, additional figures or tables are presented that describe how the variable was created. The SAS scripts for the newly created variables can be found in the Appendix at the end of this documentation.

Codebook for smoking variables

| Variable name | Description | Type/format | Value description |
|------------------|---------------------------|-------------|--|
| Smoking_status | Smoking status | Numeric | 0 = never 1 = former 2 = current |
| Smoking_onsetage | Age at smoking initiation | Numeric | 1 = < 20 years 2 = 20-29 years 3 = 30-39 years 4 = 40-49 years 5 = 50-59 years |

| | | | |
|-------------------|----------------------|---------|-----------------------------------|
| | | | 6 = 60-69 years 7 = > 70 years |
| Smoking_duration | Smoking duration | Numeric | Continuous in years |
| Smoking_packyears | Packyears of smoking | Numeric | Continuous in packyears |

Codebook for alcohol variables

| Variable name | Description | Type/format | Value description |
|-------------------------|---|--------------------|--------------------------|
| alcohol_gram_week_beer | Amount of ethanol consumed from beer | Numeric | Continuous in gram |
| alcohol_gram_week_wine | Amount of ethanol consumed from wine | Numeric | Continuous in gram |
| alcohol_gram_week_spir | Amount of ethanol consumed from spirits | Numeric | Continuous in gram |
| alcohol_gram_week_total | Amount of ethanol consumed from all kinds of alcoholic beverage | Numeric | Continuous in gram |

1. SMOKING

1A. Smoking status

Variable:

smoking_status

Definition:

The variable smoking status is created to distinguish between current, former and never smokers. Ever smoking is defined as smoking regularly for more than 1 year or smoking at least 100 cigarettes in lifetime. Ever smokers are further subdivided into current and past users, based on the smoking status at time of study entry.

1B. Age at smoking onset

Variable:

Smoking_onsetage

Definition:

Age at smoking initiation is a categorical variable (< 20 years, 20-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, > 70 years) and derived from the smoking data in different life periods.

1C. Smoking duration

Variable:

Smoking_duration

Definition:

Smoking duration represents the total time of smoking during a participant's lifetime, and is the sum of smoking durations during different age periods. Smokers are assigned a 10-year duration if they reported smoking during, before and after a 10-year age-period. If they reported smoking only before or after a 10-year age-period they are assigned a 5-year smoking duration. Participants reporting smoking before age 20 or after age 70 are assigned a 5-year smoking duration by default. For example: The smoking duration for a specific period, for instance, 40-49 years, was considered to be 10 years if the participant also reported smoking during both periods 30-39 and 50-59 years; or to be 5 years if the participant reported smoking during only 30-39 years or 50-59 years.

1D. Packyears

Variable:

Smoking_packyears

Definition:

Packyears is a continuous variable representing the sum of the smoking intensity x smoking duration for each 10-year age period. 1 packyear is defined as smoking 20 cigarettes (= 1 pack)/day for 1 year

2. ALCOHOL CONSUMPTION

Method of calculating alcohol consumption

The total weekly alcohol consumption (gram/week) represents the amount of ethanol consumed per week, and is the sum of ethanol consumption from three kinds of alcoholic beverage including beer, wine and spirits.

To calculate the weekly consumption of ethanol from each kind of beverage, we multiplied the volume of each beverage per week by its ethanol index. The ethanol index is the gram of ethanol contained in 100 grams of beverage. Since 1 ml of alcoholic liquid is considered equal to 1 gram, the ethanol index represents the amount of ethanol included in 100ml of beverage (see Figure 1 – Ethanol index).

2A. Ethanol consumption from beer

Variable:

alcohol_gram_week_beer

Definition:

This variable represents the amount of ethanol consumed per week from six kinds of beer including folköl, mellanöl, strong beer, very strong beer, cider and alco-pops.

Note:

- Ethanol index of different kinds of beer

- The ethanol index of different kinds of beer was calculated based on the ethanol index of folköl 3.5%
- The concentration of strong beer (5-7%) was considered as 6%
- The concentration of very strong beer (> 7%) was considered as 8%
- The ethanol index of Alco-pops was considered equal to that of cider 5%
- Outlier handling: It is suggested to consider the amount of ethanol from each kind of beer as missing if the number of bottle of that beer is > 40 bottles/week.

2B. Ethanol consumption from wine

Variable:

alcohol_gram_week_wine

Definition:

This variable represents the amount of ethanol consumed per week from wine.

Handling of missingness

Some participants who drink wine miss information on the kind of wine, and thus miss the ethanol index for the wine they drink. The ethanol index for the wine drunk by these participants was assigned as the average ethanol index from five predefined kinds of wine in the questionnaire.

Note

- Quantity of wine:
 - The number of wine glass is 0.5 for option "Less than one glass" and is 5.5 for option "5 or more glasses"
 - The number of wine bottle is 3.5 for option "3 or 4 bottles" and is 5.5 for option "5 bottles or more"
- Unit of wine
 - 1 glass of wine contains 125 ml wine
 - 1 bottle of wine contains 750 ml wine
- Ethanol index of wine
 - The ethanol index of pure white wine was calculated as an average of the ethanol index of four kinds of white wine coded from 1908-1911 in the Nutrient Database (see Figure 1)
 - The ethanol index of different mixtures of wine was calculated based on the proportion of red and white wine in each mixture

2C. Ethanol consumption from spirits

Variable:

alcohol_gram_week_spir

Definition:

This variable represents the amount of ethanol consumed per week from spirits.

Note:

The ethanol index of spirits was calculated as an average of ten kinds of liquor coded from 1918-1927 in the Nutrient Database (see Figure 1).

2D. Total weekly ethanol consumption

Variable:

alcohol_gram_week_total

Definition:

This variable represents the amount of ethanol consumed per week from all kinds of alcoholic beverage.

Figure 1. Ethanol index of alcoholic beverage

The ethanol index (indicated in **column L**) is calculated using the Nutrient Database 2012 provided by the Swedish National Food Agency (image of the session of alcoholic beverages is enclosed below).

| | A | B | C | D | E | F | G | H | I | J | K | L |
|------|---|------|------|-----|------|-----|------|-----|-----|-----|------|------|
| 1674 | Öl lättöl vol. % 2,3 | 1901 | 120 | 29 | 0.2 | 0 | 3.8 | 0 | 0 | 0.2 | 94 | 1.8 |
| 1675 | Öl pilsner vol. % 3,5 | 1902 | 166 | 40 | 0.4 | 0 | 4.6 | 0 | 0 | 0.2 | 92 | 2.8 |
| 1676 | Cider vol. % 1 | 1903 | 195 | 47 | 0 | 0 | 10.1 | 0 | 0 | 0.1 | 89 | 0.8 |
| 1677 | Vitt vin vol. % 1 | 1904 | 71 | 17 | 0 | 0 | 2.8 | 0 | 0 | 0.3 | 96.1 | 0.8 |
| 1678 | Rödvin rosévin vol. % 1 | 1905 | 71 | 17 | 0 | 0 | 2.8 | 0 | 0 | 0.2 | 96.2 | 0.8 |
| 1679 | Öl starköl el exportöl vol. % 5,6 | 1906 | 195 | 47 | 0.6 | 0 | 3.2 | 0 | 0 | 0.2 | 91.5 | 4.5 |
| 1680 | Rödvin vol. % 12 | 1907 | 301 | 72 | 0 | 0 | 0.8 | 0 | 0 | 0.3 | 89 | 9.9 |
| 1681 | Vitt vin halvtorr el halvsött vol. % 12,5 | 1908 | 340 | 81 | 0.1 | 0 | 3 | 0 | 0 | 0.2 | 86.8 | 9.9 |
| 1682 | Vitt vin el rhenvin vol. % 10 | 1909 | 295 | 71 | 0.1 | 0 | 3.8 | 0 | 0 | 0.2 | 88 | 7.9 |
| 1683 | Vitt vin sött vol. % 13,5 | 1910 | 414 | 99 | 0.2 | 0 | 5.9 | 0 | 0 | 0.3 | 82.9 | 10.7 |
| 1684 | Vitt vin torrt vol. % 12 | 1911 | 282 | 68 | 0 | 0 | 0.4 | 0 | 0 | 0.3 | 89.8 | 9.5 |
| 1685 | Vinglög vol. % 10 | 1912 | 521 | 125 | 0 | 0 | 17 | 0 | 0 | 0.3 | 74.7 | 8 |
| 1686 | Madeira vol. % 18 | 1914 | 540 | 129 | 0 | 0 | 7.9 | 0 | 0 | 0.2 | 77.9 | 14 |
| 1687 | Portvin vitt rött vol. % 20 | 1915 | 626 | 150 | 0 | 0 | 9.5 | 0 | 0 | 0.2 | 74.3 | 16 |
| 1688 | Sherry halvtorr vol. % 17 | 1916 | 522 | 125 | 0 | 0 | 3.4 | 0 | 0 | 0.3 | 80.3 | 16 |
| 1689 | Starkvinslög vol. % 16 | 1917 | 750 | 179 | 0 | 0 | 22.3 | 0 | 0 | 0.3 | 64.6 | 12.8 |
| 1690 | Brännvin renat el vodka vol. % 40 | 1918 | 928 | 222 | 0 | 0 | 0 | 0 | 0 | 0 | 68 | 32 |
| 1691 | Likör söt vol. % 24 | 1919 | 1112 | 266 | 0 | 0 | 33 | 0 | 0 | 0 | 48 | 19 |
| 1692 | Likör Kaptenlöjtnant vol. % 38 | 1920 | 1344 | 321 | 0 | 0 | 27 | 0 | 0 | 0 | 42.5 | 30.5 |
| 1693 | Konjak el brandy vol. % 40 | 1921 | 942 | 225 | 0 | 0 | 0.8 | 0 | 0 | 0.1 | 67.1 | 32 |
| 1694 | Punsch vol. % 26 | 1922 | 1022 | 244 | 0 | 0 | 26 | 0 | 0 | 0 | 54 | 20 |
| 1695 | Rom vol. % 40 | 1923 | 933 | 223 | 0 | 0 | 0.3 | 0 | 0 | 0 | 67.7 | 32 |
| 1696 | Kryddat brännvin vol. % 40 | 1924 | 947 | 226 | 0 | 0 | 1.1 | 0 | 0 | 0 | 66.9 | 32 |
| 1697 | Gin vol. % 40 | 1925 | 928 | 222 | 0 | 0 | 0 | 0 | 0 | 0 | 68 | 32 |
| 1698 | Whisky vol. % 40 | 1926 | 928 | 222 | 0 | 0 | 0 | 0 | 0 | 0 | 68 | 32 |
| 1699 | Likör mindre söt vol. % 24 | 1927 | 971 | 232 | 0 | 0 | 24.7 | 0 | 0 | 0 | 56.3 | 19 |
| 1700 | Irish coffee | 1929 | 440 | 105 | 0.3 | 5.2 | 4.6 | 0 | 0 | 0.1 | 84.1 | 5.6 |
| 1701 | Bovetemjöl ljost | 1930 | 1493 | 357 | 6.6 | 2 | 76 | 1.9 | 0 | 0.9 | 12.6 | 0 |
| 1702 | Havremust pulver beräk | 1931 | 1400 | 358 | 15.8 | 5.8 | 56 | 8 | 1.6 | 1.0 | 9.5 | 0 |

Ethanol index of alcohol beverages

| Alcoholic beverage | Ethanol index (grams of ethanol per 100ml of alcoholic beverage) |
|---------------------------------|--|
| Beer | |
| Folköl (3.5%) | 2.8 |
| Mellanöl (4.5%) | 3.6 |
| Strong beer (5-7%) | 4.8 |
| Very strong beer (> 7%) | 6.4 |
| Cider (5%) | 4.0 |
| Alco-pops | 4.0 |
| Wine | |
| 100% red wine | 9.9 |
| 75% red wine and 25% white wine | 9.8 |

50% red wine and 50% white wine 9.7

25% red wine and 75% white wine 9.525

100% white wine 9.5

Spirits 28.05

Quality of life

Version 1 - October 2019, contact: mikael.eriksson@ki.se

Introduction

Quality of life is part of the Karisma survey assessment using the instrument EORTC QLQ-C30 instrument v3.

The instrument assesses the study participant quality of life as experienced in the last week. The instrument was developed in 1986 by an EORTC research program initiated to develop an integrated, modular approach for evaluating the QoL of patients participating in international cancer clinical trials. This research resulted in the development of a core questionnaire which is referred to as the EORTC QLQ-C30 (Aaronson et al., 1993). The EORTC QLQ-C30 incorporates nine multi-item scales: five functional scales (Physical, Role, Cognitive, Emotional and Social Functioning); three symptom scales (Fatigue, Pain and Nausea/Vomiting); and a Global Health Status/QoL scale. Six single item scales are also included (Dyspnoea, Insomnia, Appetite Loss, Constipation, Diarrhoea and Financial Difficulties). The psychometric properties of the questionnaire were tested and in conclusion it was found to possess the required standards such as validity (measuring what it is intended to measure), reliability (measuring with sufficient precision) and sensitivity (ability to detect changes) (Aaronson et al., 1993; Osoba et al., 1994; Kaasa et al., 1995). The questionnaire was initially tested in a population of lung cancer patients (Aaronson et al., 1993) and subsequently in a variety of cancer patient groups. A bibliography is contained in the EORTC QLQ-C30 Scoring Manual (Fayers et al., 2001). There is a continuing programme of development for the EORTC QLQ-C30.

QLQ-C30 in Karisma

Karma implemented the Swedish version of QLQ-C30 and used the standard EORTC coding of the 15 scales available as described in the manual SCManualQLQ-C30.pdf.

Karma generates one extended QLQ-C30 dataset for each survey (baseline, follow-up) with data collected from the study participants. Each dataset includes the following scales.

1. pf Physical functioning (i.e. problems with carrying things, taking a long/short walk, need to rest during the days, need help with hygiene or getting dressed)
2. rf Role functioning (i.e. able to work, have normal free time activities, good family life and social activities)
3. ef Emotional functioning (i.e. been feeling tense, worried, irritated, blue)
4. cf Cognitive functioning (i.e. difficulties concentrating e.g. when reading a paper or watching TV, difficulties remembering things)
5. sf Social functioning (i.e. interference in family life due to physical condition or medical treatment or interference in social activities due to same causes)
6. ql Global quality of life (i.e. general estimation of the persons health and quality of life)
7. fa Fatigue (i.e. need for rest, feeling weak, feeling tired)
8. nv Nausea/Vomiting
9. pa Pain (i.e. had pain, was affected in daily activities by pain)

10. dy Dyspnoea

11. sl Insomnia

12. ap Appetite loss

13. co Constipation

14. di Diarrhoea

15. fi Financial problems

The scales are defined as the mean value of the 1-4 scale (not at all--very much) responses to the questions within parentheses above with exception of the scale "general quality of life" which is based on two questions having 7-scale answer options (very poor--excellent).

Sense of coherence

Version 1 - October 2019, contact: mikael.eriksson@ki.se

Introduction

Karma assesses the study participant "sense of coherence", i.e. everyday life comprehensibility, manageability, and meaningfulness.

1. Comprehensibility means that everyday life can be predicated, makes sense and is not confusing.
2. Manageability means that everyday life is within the control to handled and take care of with sufficient resources.
3. Meaningfulness means that everyday life is a source of satisfaction, being interesting and exciting, and is really worth doing.

The main "kasam" scoring is the sum of the three sub-scales "comprehensibility", "manageability", and "meaningfulness".

Assesment and coding

The KASAM instrument used in Karma is based on KASAM-13 as described by Björvell/Langius, 1987. Adaption to modern survey technique was performed with the result of a single scale for answer options ranging from "very seldom or never" to "very often" in 7 scale steps, see the section quality of life / sense of coherence in the survey documentation at <https://karmastudy.org/contact/data-access/data-sources/>

The coding uses the coding shown in Björvell/Langius, 1987, adapted to the Karisma assessment.

Comprehensibility: sum of the answer options for questions 2,6,8,9,11

Manageability: sum of the answer options for questions 3,5,10,13

Meaningsfulness: sum of the answer options for questions 1,4,7,12

Update history

v1.1.

Update Reproductive and menstrual health to v2.1. Added description of variable postmenopausal.