Data quality report

Report created: 13:08:46 14 Jul 2023, dqrep version 1.34

Report content

Dataset overview

Integrity issues and notes

Descriptive variable overview

Missing values (Item missingness)

Range violations

Univariate outliers

Variance proportion overview

Overview for single variable

Change-log for modified variables

Dataset overview

The dataset contains in total 29 variables based on the files: SHIP_study with 2154 observations. Out of these variables, in total 28 were selected for further analyses.

Variables of the following categories are available: Primary variables: 7; Secondary variables: 12; Control variables: 2; Time variables: 1; Observer variables: 3; Device variables: 3;
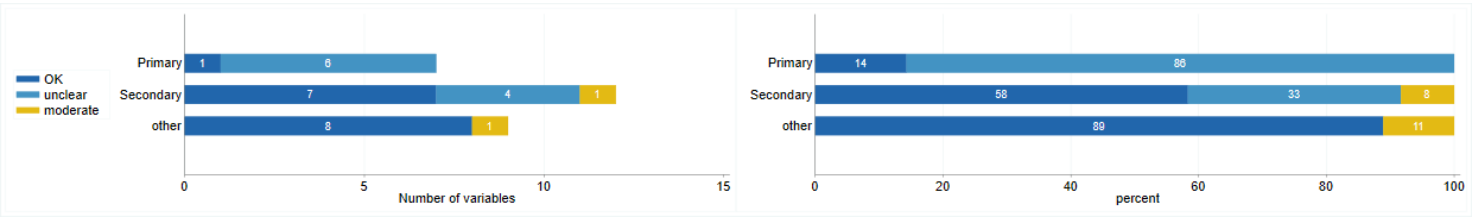
The following metadataset has been used for the provision of attributes on study data variables: SHIP_metadata.xlsx

An ID variable is specified based on the variable selection: id
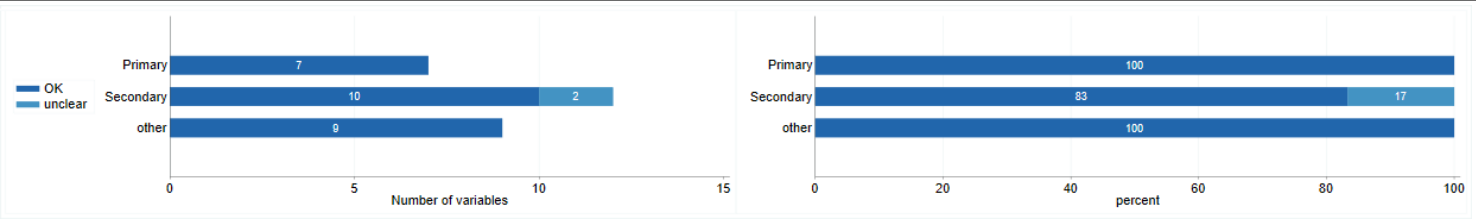
Integrity issues and notes

There are notes or issues related to data integrity or the provided metadata. Please see tables in the appendix.
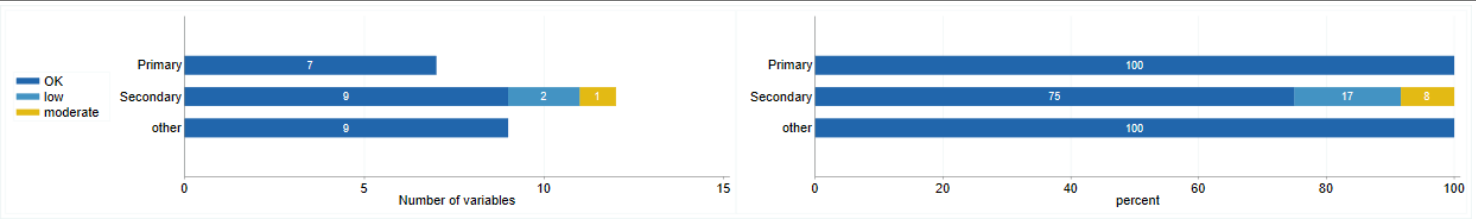
Variables with data quality issues

This graph shows the maximum problem category per variable.

Crude item missingness

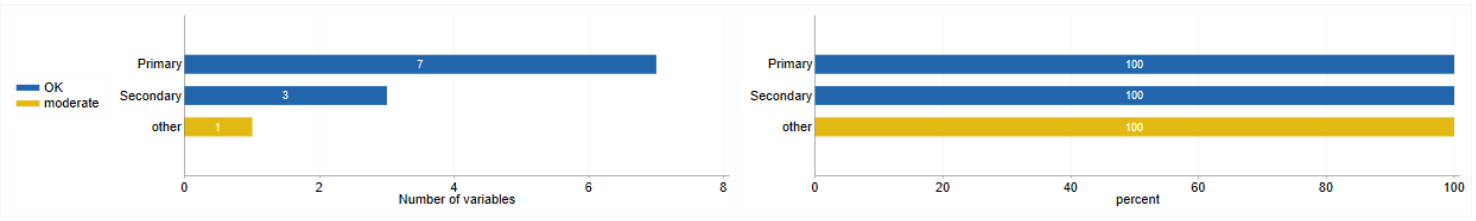Categories for Item missing %: OK: <1; unclear: >=1;

Response rate - Item

Categories for Item Response %: OK: <1; low: <5; moderate: <10; important: <25; critical: >=25;
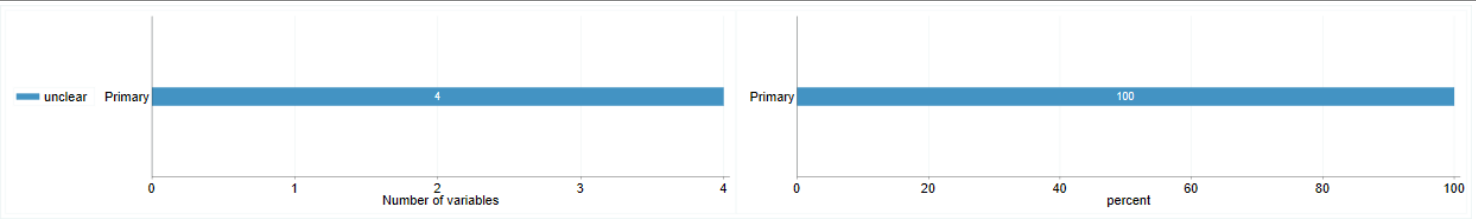
Inadmissible values

Categories for Inadmissible values %: OK: ==0; moderate: <2; important: <5; critical: >=5;

Uncertain values

Categories for Implausible values %: OK: ==0; unclear: <=100;

Univariate outliers
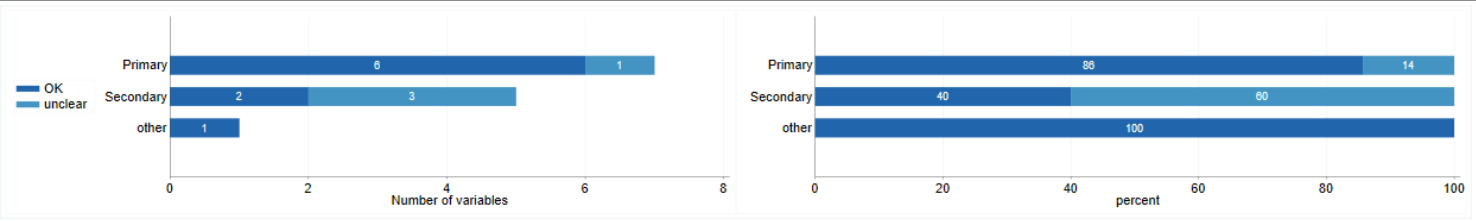
Categories for Implausible values %: OK: ==0; unclear: <2; moderate: <5; important: <10; critical: >=10;

Unexpected effects by Observer variables
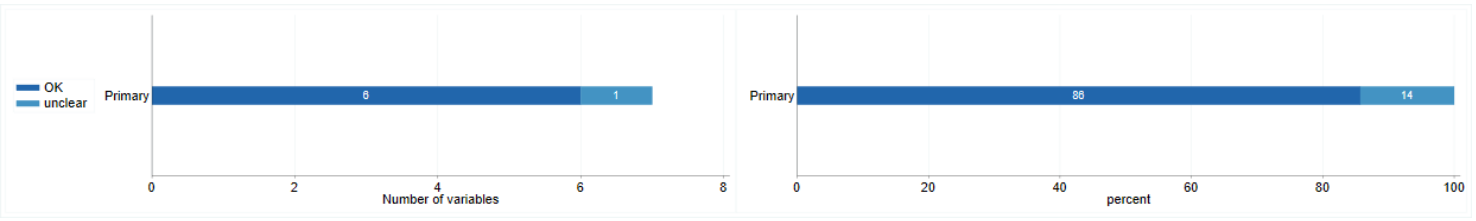
Unexpected effects by Device variables
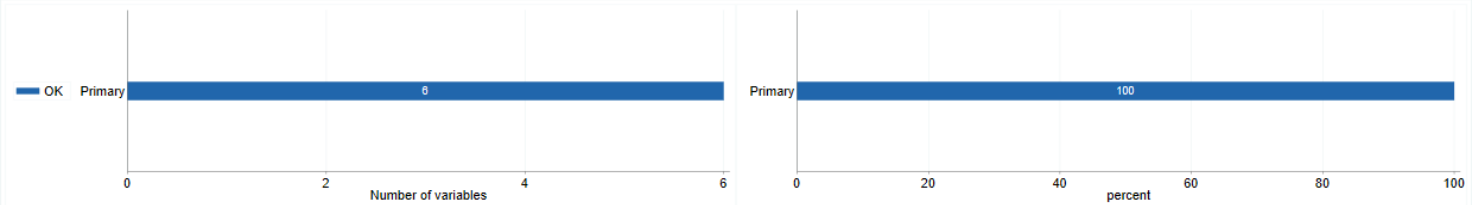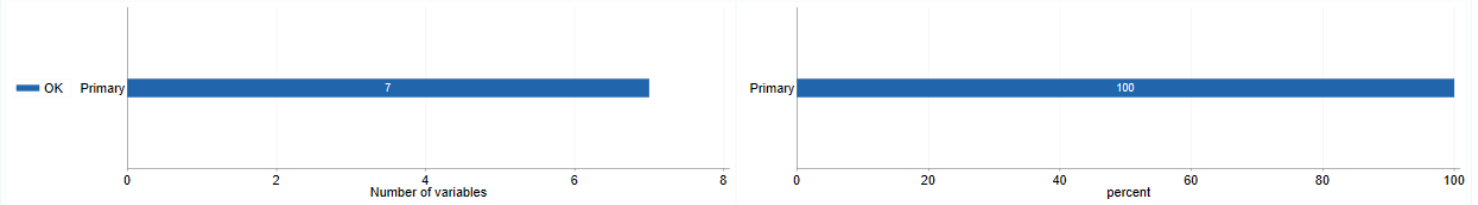
Time trend

Categories for Time trend variance proportion: OK: <.1; unclear: <.2; moderate: <.3; important: <.4; critical: >=.4;

Data quality measures overview 1

Primary variables	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
sbp1 Systolic blood pressure 1. measurement (mmHg)	.09%	.09%	0%	.23%	0%	.01	.01
sbp2 Systolic blood pressure 2. measurement (mmHg)	.28%	.28%	0%	.28%	0%	.02	.02
dbp1 Diastolic blood pressure 1. measurement (mmHg)	.09%	.09%	0%	.46%	0%	.01	<.01
dbp2 Diastolic blood pressure 2. measurement (mmHg)	.28%	.28%	0%	.33%	0%	.01	<.01
height Heigth (cm)	.14%	.14%	0%	n.a.	0%	<.01	.01
weight Body weight (kg)	.19%	.19%	0%	n.a.	.05%	.01	<.01
waist Waist circumference	.14%	.14%	0%	n.a.	0%	n.a.	.03

Secondary variables	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
school Educational level	.32%	.32%	n.a.	n.a.	n.a.	n.a.	n.a.
family Marital status	.32%	.32%	n.a.	n.a.	n.a.	n.a.	n.a.
smoking Smoking	.37%	.37%	n.a.	n.a.	n.a.	n.a.	n.a.
stroke Stroke ever	.46%	.46%	n.a.	n.a.	n.a.	n.a.	n.a.
myocard Myocardial infarction ever	.65%	.65%	n.a.	n.a.	n.a.	n.a.	n.a.
diab_known Diabetes	.32%	.32%	n.a.	n.a.	n.a.	n.a.	n.a.
diab_age Diabetes Age onset	.32%	3.89%	n.a.	n.a.	0%	n.a.	n.a.
contraception Contraceptives Ever	.42%	.82%	n.a.	n.a.	n.a.	n.a.	n.a.
income Houshold income	5.39%	5.39%	n.a.	n.a.	0%	n.a.	n.a.
hdl HDL coholesterol	.74%	.74%	0%	n.a.	.42%	n.a.	n.a.
ldl LDL coholesterol	1.3%	1.3%	0%	n.a.	.05%	n.a.	n.a.
cholesterol Cholesterol	.7%	.7%	0%	n.a.	.09%	n.a.	n.a.

Control variables	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
sex Sex	0%	0%	n.a.	n.a.	n.a.	n.a.	n.a.
age Age (years)	0%	0%	.05%	n.a.	0%	n.a.	n.a.

Time variables	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
exdate Examination date	0%	0%	n.a.	n.a.	n.a.	n.a.	n.a.

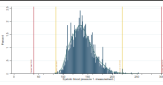| Observer variables	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
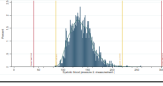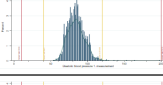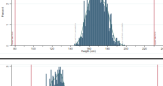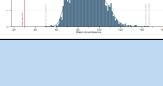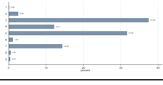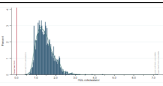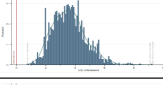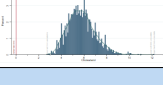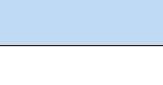
	Missings all %	Nonresponse rate - Item %	Inadmissible values %	Uncertain values %	Univariate outliers %	VP DEV	VP EXA
obs_int Observer Interview	.05%	.05%	n.a.	n.a.	n.a.	n.a.	n.a.
obs_soma Observer Somatometry	.09%	.09%	n.a.	n.a.	n.a.	n.a.	n.a.
obs_bp Observer blood pressure	.14%	.14%	n.a.	n.a.	n.a.	n.a.	n.a.
Device variables	**Missings all %**	**Nonresponse rate - Item %**	**Inadmissible values %**	**Uncertain values %**	**Univariate outliers %**	**VP DEV**	**VP EXA**
dev_weight Device weight	.09%	.09%	n.a.	n.a.	n.a.	n.a.	n.a.
dev_length Device heigth	.09%	.09%	n.a.	n.a.	n.a.	n.a.	n.a.
dev_bp Device blood pressure	.14%	.14%	n.a.	n.a.	n.a.	n.a.	n.a.
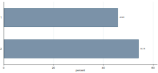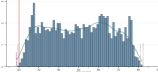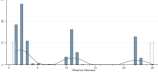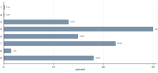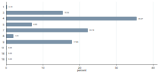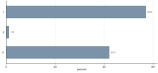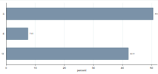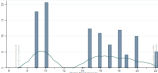
Data quality measures overview 2

Primary variables	Time Trend Variance proportion
sbp1 Systolic blood pressure 1. measurement (mmHg)	.07
sbp2 Systolic blood pressure 2. measurement (mmHg)	.07
dbp1 Diastolic blood pressure 1. measurement (mmHg)	.07
dbp2 Diastolic blood pressure 2. measurement (mmHg)	.07
height Heigth (cm)	.02
weight Body weight (kg)	.07
waist Waist circumference	.06

Secondary variables	Time Trend Variance proportion
school Educational level	n.a.
family Marital status	n.a.
smoking Smoking	n.a.
stroke Stroke ever	n.a.
myocard Myocardial infarction ever	n.a.
diab_known Diabetes	n.a.
diab_age Diabetes Age onset	n.a.
contraception Contraceptives Ever	n.a.
income Houshold income	n.a.
hdl HDL coholesterol	n.a.
ldl LDL coholesterol	n.a.
cholesterol Cholesterol	n.a.

Control variables	Time Trend Variance proportion
sex Sex	n.a.
age Age (years)	n.a.

Time variables	Time Trend Variance proportion
exdate Examination date	n.a.

Observer variables	Time Trend Variance proportion
obs_int Observer Interview	n.a.
obs_soma Observer Somatometry	n.a.
obs_bp Observer blood pressure	n.a.

Device variables	Time Trend Variance proportion
dev_weight Device weight	n.a.
dev_length Device heigth	n.a.
dev_bp Device blood pressure	n.a.

Descriptive variable overview

Primary variables	Levels		Mean [SD]	Min Max	Missings %	N
sbp1 Systolic blood pressure 1. measurement (mmHg)	125		139.01 [22.5]	83 253	.09%	2152
sbp2 Systolic blood pressure 2. measurement (mmHg)	123		136.4 [21.86]	83 258	.28%	2148
dbp1 Diastolic blood pressure 1. measurement (mmHg)	75		84.52 [11.81]	53 198	.09%	2152
dbp2 Diastolic blood pressure 2. measurement (mmHg)	71		83.52 [11.52]	55 151	.28%	2148
height Heigth (cm)	1606		168.71 [9.25]	144.05 198.13	.14%	2151
weight Body weight (kg)	1763		77.63 [15.08]	42.6 144.44	.19%	2150
waist Waist circumference	2105		89.2 [13.82]	49.27 145.77	.14%	2151

Secondary variables	Levels		Mean [SD]	Min Max	Missings %	N
school Educational level	9			1 9	.32%	2147
family Marital status	5			1 5	.32%	2147
smoking Smoking	3			0 2	.37%	2146
stroke Stroke ever	2			1 2	.46%	2144
myocard Myocardial infarction ever	2			1 2	.65%	2140
diab_known Diabetes	2			0 1	.32%	2147
diab_age Diabetes Age onset	47		53.68 [13.33]	12 79	.32%	173
contraception Contraceptives Ever	2			1 2	.42%	1089
income Houshold income	3		1.92 [.71]	1 3	5.39%	2038
hdl HDL coholesterol	1111		1.45 [.44]	.42 7.2	.74%	2138
ldl LDL coholesterol	2125		3.58 [1.13]	.7 9.24	1.3%	2126
cholesterol Cholesterol	1668		5.76 [1.2]	2.67 12.12	.7%	2139

Control variables	Levels		Mean [SD]	Min Max	Missings %	N

	Levels		Mean [SD]	Min Max	Missings %	N
sex Sex	2			1 2	0%	2154
age Age (years)	64		49.86 [16.19]	19 82	0%	2154

Time variables	Levels		Mean [SD]	Min Max	Missings %	N
exdate Examination date	688		07apr1999	16oct1997 19may2001	0%	2154

Observer variables	Levels		Mean [SD]	Min Max	Missings %	N
obs_int Observer Interview	13		n.a.	1 25	.05%	2153
obs_soma Observer Somatometry	8		n.a.	1 9	.09%	2152
obs_bp Observer blood pressure	10		n.a.	1 18	.14%	2151

Device variables	Levels		Mean [SD]	Min Max	Missings %	N
dev_weight Device weight	3		n.a.	1 11	.09%	2152
dev_length Device heigh	3		n.a.	3 11	.09%	2152
dev_bp Device blood pressure	11		n.a.	7 22	.14%	2151

Missing values (Item missingness)

Primary variables	Permitted jump N (%)	Missing spec. N (%)	Missing system N (%)	Missings N (%)	Available %	Response rate - Item %	N
sbp1 Systolic blood pressure 1. measurement (mmHg)	0 (0%)	0 (0%)	2 (.09%)	2 (.09%)	99.91%	99.91%	2152
sbp2 Systolic blood pressure 2. measurement (mmHg)	0 (0%)	0 (0%)	6 (.28%)	6 (.28%)	99.72%	99.72%	2148
dbp1 Diastolic blood pressure 1. measurement (mmHg)	0 (0%)	0 (0%)	2 (.09%)	2 (.09%)	99.91%	99.91%	2152
dbp2 Diastolic blood pressure 2. measurement (mmHg)	0 (0%)	0 (0%)	6 (.28%)	6 (.28%)	99.72%	99.72%	2148
height Heigth (cm)	0 (0%)	0 (0%)	3 (.14%)	3 (.14%)	99.86%	99.86%	2151
weight Body weight (kg)	0 (0%)	0 (0%)	4 (.19%)	4 (.19%)	99.81%	99.81%	2150
waist Waist circumference	0 (0%)	0 (0%)	3 (.14%)	3 (.14%)	99.86%	99.86%	2151

Secondary variables	Permitted jump N (%)	Missing spec. N (%)	Missing system N (%)	Missings N (%)	Available %	Response rate - Item %	N
school Educational level	0 (0%)	7 (.32%)	0 (0%)	7 (.32%)	99.68%	99.68%	2147
family Marital status	0 (0%)	7 (.32%)	0 (0%)	7 (.32%)	99.68%	99.68%	2147
smoking Smoking	0 (0%)	7 (.32%)	0 (0%)	8 (.37%)	99.67%	99.63%	2146
stroke Stroke ever	0 (0%)	10 (.46%)	0 (0%)	10 (.46%)	99.54%	99.54%	2144
myocard Myocardial infarction ever	0 (0%)	14 (.65%)	0 (0%)	14 (.65%)	99.35%	99.35%	2140
diab_known Diabetes	0 (0%)	7 (.32%)	0 (0%)	7 (.32%)	99.68%	99.68%	2147
diab_age Diabetes Age onset	1974 (91.64%)	7 (.32%)	0 (0%)	7 (.32%)	8.03%	96.11%	173
contraception Contraceptives Ever	1056 (49.03%)	9 (.42%)	0 (0%)	9 (.42%)	50.56%	99.18%	1089
income Houshold income	0 (0%)	116 (5.39%)	0 (0%)	116 (5.39%)	94.61%	94.61%	2038
hdl HDL coholesterol	0 (0%)	0 (0%)	16 (.74%)	16 (.74%)	99.26%	99.26%	2138
ldl LDL coholesterol	0 (0%)	0 (0%)	28 (1.3%)	28 (1.3%)	98.7%	98.7%	2126
cholesterol Cholesterol	0 (0%)	0 (0%)	15 (.7%)	15 (.7%)	99.3%	99.3%	2139

Control variables	Permitted	Missing spec.	Missing	Missings N	Available %	Response rate	N

	Permitted jump N (%)	Missing spec. N (%)	Missing system N (%)	Missings N (%)	Available %	Response rate - Item %	N
sex Sex	0 (0%)	0 (0%)	0 (0%)	0 (0%)	100%	100%	2154
age Age (years)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	100%	100%	2154
Time variables	**Permitted jump N (%)**	**Missing spec. N (%)**	**Missing system N (%)**	**Missings N (%)**	**Available %**	**Response rate - Item %**	**N**
exdate Examination date	0 (0%)	0 (0%)	0 (0%)	0 (0%)	100%	100%	2154
Observer variables	**Permitted jump N (%)**	**Missing spec. N (%)**	**Missing system N (%)**	**Missings N (%)**	**Available %**	**Response rate - Item %**	**N**
obs_int Observer Interview	0 (0%)	1 (.05%)	0 (0%)	1 (.05%)	99.95%	99.95%	2153
obs_soma Observer Somatometry	0 (0%)	2 (.09%)	0 (0%)	2 (.09%)	99.91%	99.91%	2152
obs_bp Observer blood pressure	0 (0%)	3 (.14%)	0 (0%)	3 (.14%)	99.86%	99.86%	2151

This table differentiates between: (a) allowed jumps, (b) defined missing values of any kind, and (c) missing values that have no coding. Permitted jumps are not considered to be missing values. Therefore, these are not included in the column for all missing values. The last three columns refer to available data values and out of them, the first to the proportion of available values amongst the entire sample, and the second to the proportion of all expected observations after exclusion of permitted jumps.

Range violations

Primary variables	Lower inadm. limit (**N**)	Upper inadm. limit (**N**)	Lower uncertainty limit (**N**)	Upper soft limit (**N**)	Inadmissible values %	Uncertain values %
sbp1 Systolic blood pressure 1. measurement (mmHg)	40 (0)	300 (0)	85 (1)	220 (4)	0%	.23%
sbp2 Systolic blood pressure 2. measurement (mmHg)	40 (0)	300 (0)	85 (2)	220 (4)	0%	.28%
dbp1 Diastolic blood pressure 1. measurement (mmHg)	10 (0)	200 (0)	40 (0)	120 (10)	0%	.46%
dbp2 Diastolic blood pressure 2. measurement (mmHg)	10 (0)	200 (0)	40 (0)	120 (7)	0%	.33%
height Heigth (cm)	80 (0)	230 (0)	n.a.	n.a.	0%	n.a.
weight Body weight (kg)	30 (0)	250 (0)	n.a.	n.a.	0%	n.a.
waist Waist circumference	30 (0)	n.a.	n.a.	n.a.	0%	n.a.

Secondary variables	Lower inadm. limit (**N**)	Upper inadm. limit (**N**)	Lower uncertainty limit (**N**)	Upper soft limit (**N**)	Inadmissible values %	Uncertain values %
diab_age Diabetes Age onset	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
income Houshold income	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
hdl HDL coholesterol	0 (0)	n.a.	n.a.	n.a.	0%	n.a.
ldl LDL coholesterol	0 (0)	n.a.	n.a.	n.a.	0%	n.a.
cholesterol Cholesterol	0 (0)	n.a.	n.a.	n.a.	0%	n.a.

Control variables	Lower inadm. limit (**N**)	Upper inadm. limit (**N**)	Lower uncertainty limit (**N**)	Upper soft limit (**N**)	Inadmissible values %	Uncertain values %
age Age (years)	20 (1)	n.a.	n.a.	n.a.	.05%	n.a.

Time variables	Lower inadm. limit (**N**)	Upper inadm. limit (**N**)	Lower uncertainty limit (**N**)	Upper soft limit (**N**)	Inadmissible values %	Uncertain values %
exdate Examination date	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Univariate outliers

Primary variables		Grubbs N N	Medc. LB(N) Medc. UB(N)	Outliers N (%)
msbp1 Systolic blood pressure 1. measurement (mmHg) mod		0 0	88.34(3) 213.65(6)	0 (0%)
msbp2 Systolic blood pressure 2. measurement (mmHg) mod		0 0	88.07(1) 216.61(1)	0 (0%)
mdbp1 Diastolic blood pressure 1. measurement (mmHg) mod		0 0	53.5(1) 113.5(19)	0 (0%)
mdbp2 Diastolic blood pressure 2. measurement (mmHg) mod		0 0	53.73(0) 118.55(3)	0 (0%)
height Heigth (cm)		0 0	144.23(1) 199.84(0)	0 (0%)
weight Body weight (kg)		0 1	35.57(0) 120.61(14)	1 (.05%)
waist Waist circumference		0 0	40.53(0) 124.67(12)	0 (0%)

Secondary variables		Grubbs N N	Medc. LB(N) Medc. UB(N)	Outliers N (%)
diab_age Diabetes Age onset		0 0	19.5(5) 87.5(0)	0 (0%)
income Houshold income		0 0	-.5(0) 3.5(0)	0 (0%)
hdl HDL coholesterol		0 5	.61(4) 3.1(6)	9 (.42%)
ldl LDL coholesterol		0 1	.67(0) 6.67(18)	1 (.05%)
cholesterol Cholesterol		0 2	2.97(4) 9.44(12)	2 (.09%)

Control variables		Grubbs N N	Medc. LB(N) Medc. UB(N)	Outliers N (%)
mage Age (years) mod		0 0	-7.67(0) 100.91(0)	0 (0%)

Outliers: (LB= lower boundary; UB= upper boundary); Medcouple (mc) Brown (2005) P25/75 -/+1.5*exp(-3.5/-4/4/-3.5*mc)*IQR; Grubbs Test: (CI 97.5); Decision rule: If skewness <-1;>1 use only medcouple; if skewness [-1;1] use both rules.

The prefix 'm' to the variable name indicates that properties of a modified variable have been displayed. Please check the change-log table. The changes concern range violations.

Cluster effects for: Observer variables obs_soma

Primary variables	Variance prop.	# Clusters used	# Clusters	Min Cluster size N	Mean Cluster size N	Max Cluster size N	N in analysis
height Heigth (cm)	.01	6	8	2	268	645	2144
mweight Body weight (kg) mod	<.01	6	8	2	268	644	2142
waist Waist circumference	.03	6	8	2	268	645	2145

Variables default recoded. Only Clusters with at least 10 observations are included. Results are based on fixed effects regression models.

The prefix 'm' to the variable name indicates that properties of a modified variable have been displayed. Please check the change-log table.

Cluster effects for: Observer variables obs_bp

Primary variables	Variance prop.	# Clusters used	# Clusters	Min Cluster size N	Mean Cluster size N	Max Cluster size N	N in analysis
msbp1 Systolic blood pressure 1. measurement (mmHg) mod	.01	6	10	1	214	760	2138
msbp2 Systolic blood pressure 2. measurement (mmHg) mod	.02	6	10	1	214	760	2133
mdbp1 Diastolic blood pressure 1. measurement (mmHg) mod	<.01	6	10	1	214	759	2133
mdbp2 Diastolic blood pressure 2. measurement (mmHg) mod	<.01	6	10	1	214	762	2132

Variables default recoded. Only Clusters with at least 10 observations are included. Results are based on fixed effects regression models.

The prefix 'm' to the variable name indicates that properties of a modified variable have been displayed. Please check the change-log table.

Cluster effects for: Device variables dev_weight

Primary variables	Variance prop.	# Clusters used	# Clusters	Min Cluster size N	Mean Cluster size N	Max Cluster size N	N in analysis
mweight Body weight (kg) mod	.01	3	3	22	716	1225	2147

Variables default recoded. Only Clusters with at least 10 observations are included. Results are based on fixed effects regression models.

The prefix 'm' to the variable name indicates that properties of a modified variable have been displayed. Please check the change-log table.

Cluster effects for: Device variables dev_length

Primary variables	Variance prop.	# Clusters used	# Clusters	Min Cluster size N	Mean Cluster size N	Max Cluster size N	N in analysis
height Heigth (cm)	<.01	3	3	160	716	1088	2149

Variables default recoded. Only Clusters with at least 10 observations are included. Results are based on fixed effects regression models.

Cluster effects for: Device variables dev_bp

Primary variables	Variance prop.	# Clusters used	# Clusters	Min Cluster size N	Mean Cluster size N	Max Cluster size N	N in analysis
msbp1 Systolic blood pressure 1. measurement (mmHg) mod	.01	9	11	1	195	442	2143
msbp2 Systolic blood pressure 2. measurement (mmHg) mod	.02	9	11	1	194	442	2138

mdbp1 Diastolic blood pressure 1. measurement (mmHg) mod	.01	9	11	1	194	443	2138
mdbp2 Diastolic blood pressure 2. measurement (mmHg) mod	.01	9	11	1	194	444	2137

Variables default recoded. Only Clusters with at least 10 observations are included. Results are based on fixed effects regression models.

The prefix 'm' to the variable name indicates that properties of a modified variable have been displayed. Please check the change-log table.

Results for variable: sbp1

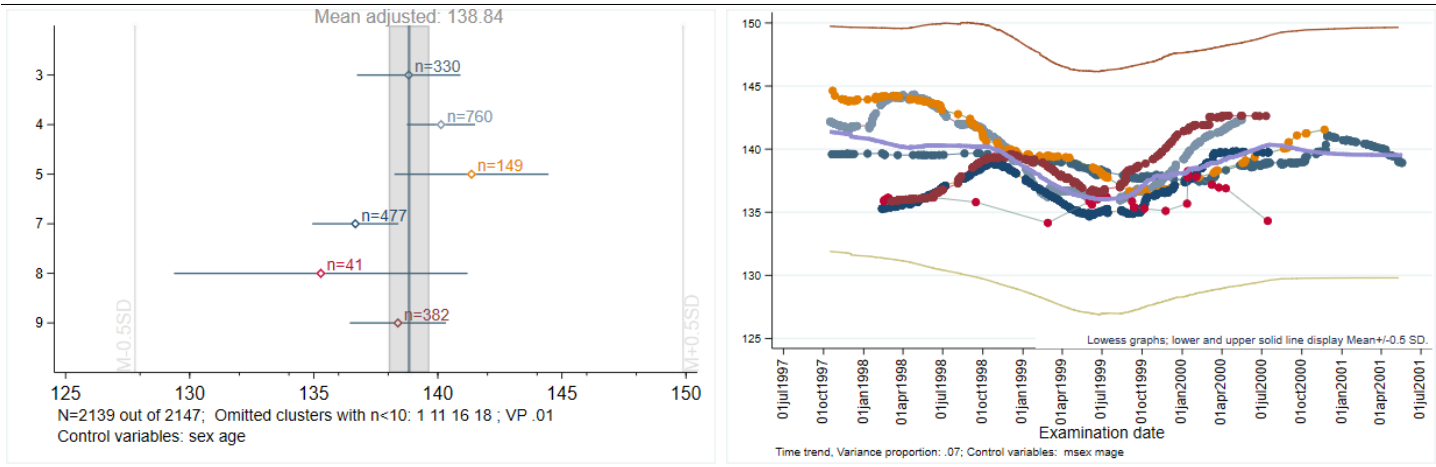Primary variable: Systolic blood pressure 1. measurement (mmHg)

Data type: float / Stata-Format: float %8.0g converted: int %8.0g / Scale level: ratio

Measures	Original variable	Modified variable
N	2152	2147
Missing values	2	7
Mean	139.01	138.85
Standard deviation	22.50	22.10
Skewness	0.64	0.55
Minimum	83.00	87.00
Percentile 1	98.00	99.00
Percentile 50	137.00	137.00
Percentile 99	200.00	198.00
Maximum	253.00	220.00
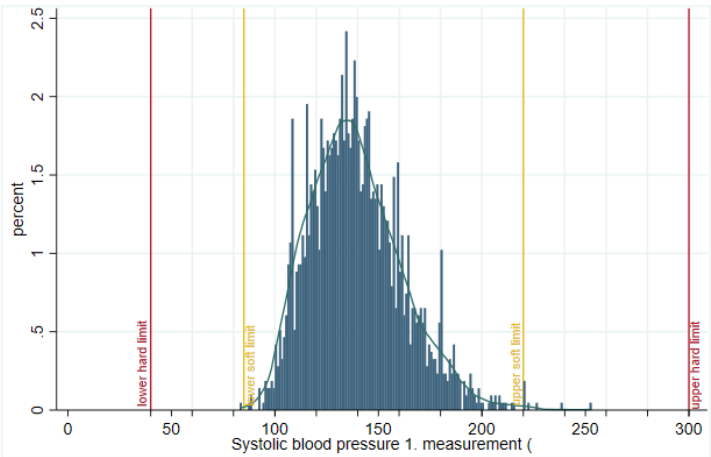
For information on modifications see table: Change-log for modified variables.

Variable: sbp1   Cluster effects for: Observer variables obs_bp

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Variable: sbp1   Cluster effects for: Device variables dev_bp

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Results for variable: sbp2

Primary variable: Systolic blood pressure 2. measurement (mmHg)

Data type: float / Stata-Format: float %8.0g converted: int %8.0g / Scale level: ratio

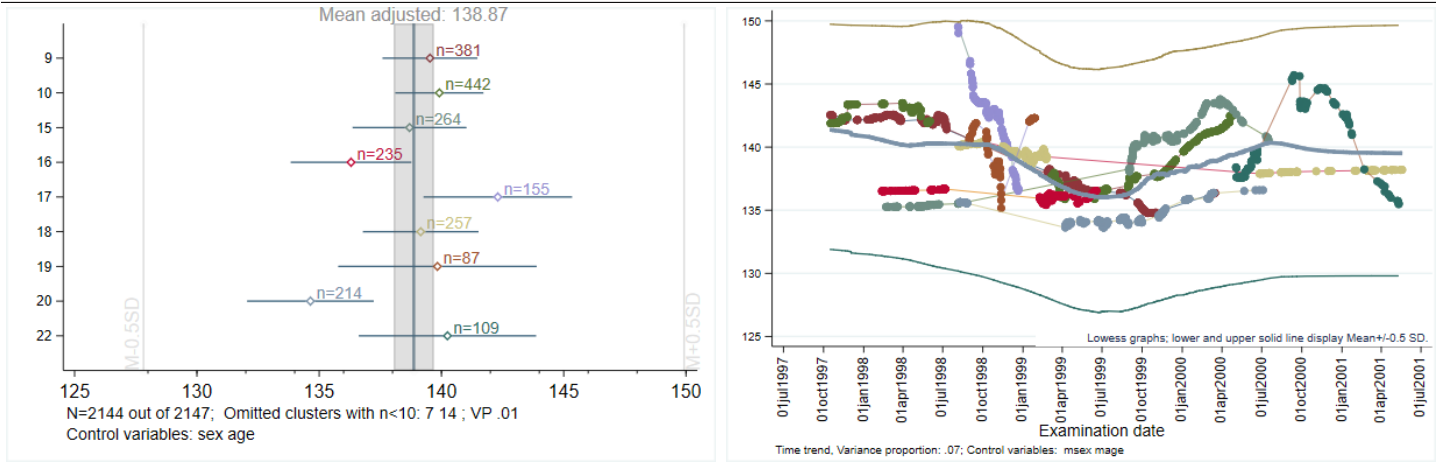| Measures	Original variable	Modified variable
N	2148	2142
Missing values	6	12
Mean	136.40	136.26
Standard deviation	21.86	21.39
Skewness	0.63	0.51
Minimum	83.00	87.00
Percentile 1	96.00	97.00
Percentile 50	134.00	134.00
Percentile 99	197.00	194.00
Maximum	258.00	218.00

For information on modifications see table: Change-log for modified variables.

Variable: sbp2   Cluster effects for: Observer variables obs_bp

N=2134 out of 2142; Omitted clusters with n<10: 1 11 16 18 ; VP .02
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.

Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

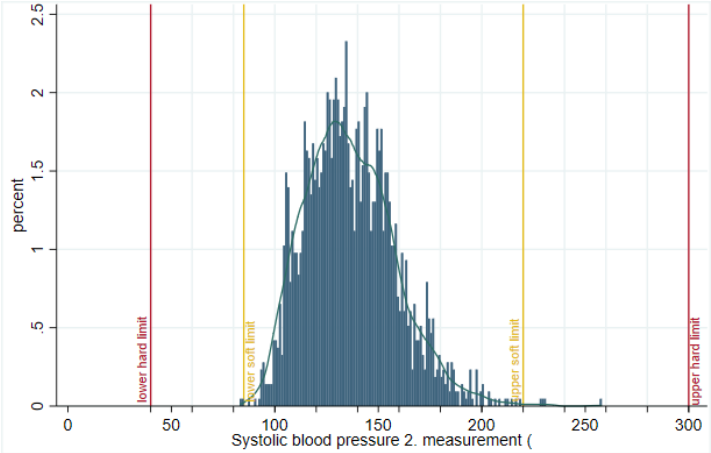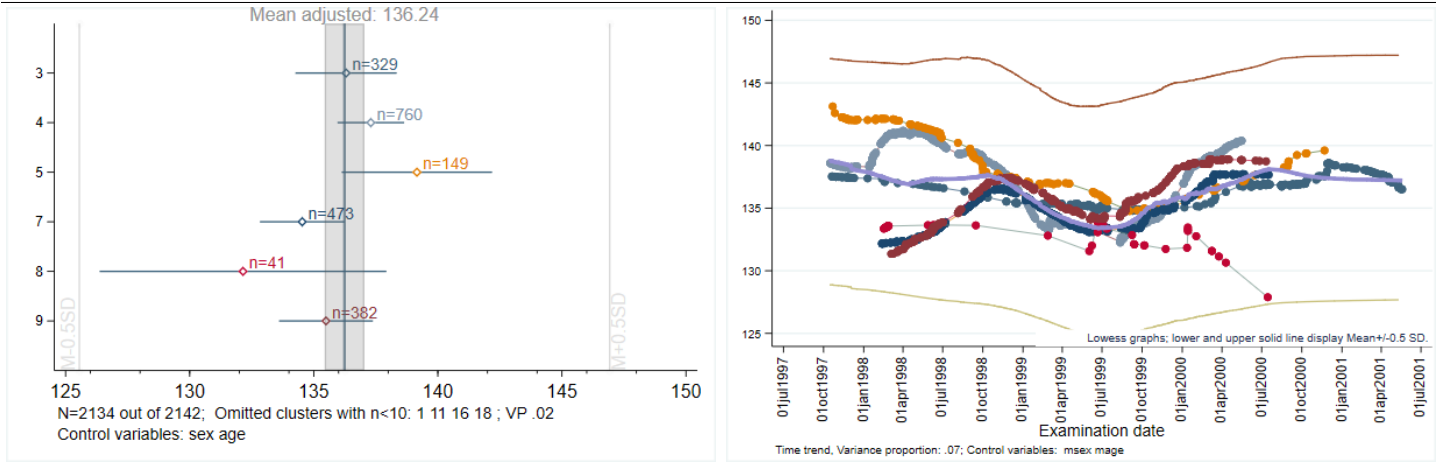## Variable: sbp2   Cluster effects for: Device variables dev_bp

N=2139 out of 2142; Omitted clusters with n<10: 7 14 ; VP .02
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.

Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Results for variable: dbp1

Primary variable: Diastolic blood pressure 1. measurement (mmHg)

Data type: float / Stata-Format: float %8.0g converted: int %8.0g / Scale level: ratio

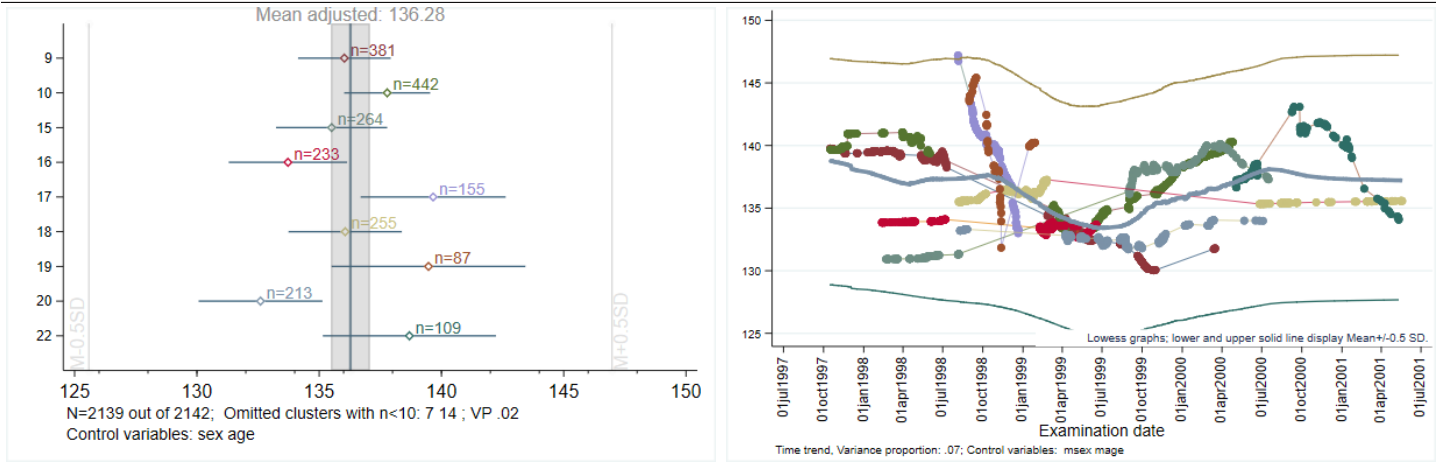| Measures	Original variable	Modified variable
N	2152	2142
Missing values	2	12
Mean	84.52	84.25
Standard deviation	11.81	11.06
Skewness	1.01	0.30
Minimum	53.00	53.00
Percentile 1	62.00	62.00
Percentile 50	84.00	84.00
Percentile 99	115.00	113.00
Maximum	198.00	120.00

For information on modifications see table: Change-log for modified variables.

Variable: dbp1   Cluster effects for: Observer variables obs_bp

N=2134 out of 2142; Omitted clusters with n<10: 1 11 16 18 ; VP <.01
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.
Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Variable: dbp1   Cluster effects for: Device variables dev_bp

N=2139 out of 2142; Omitted clusters with n<10: 7 14 ; VP .01
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.
Time trend, Variance proportion: .07; Control variables: msex mage
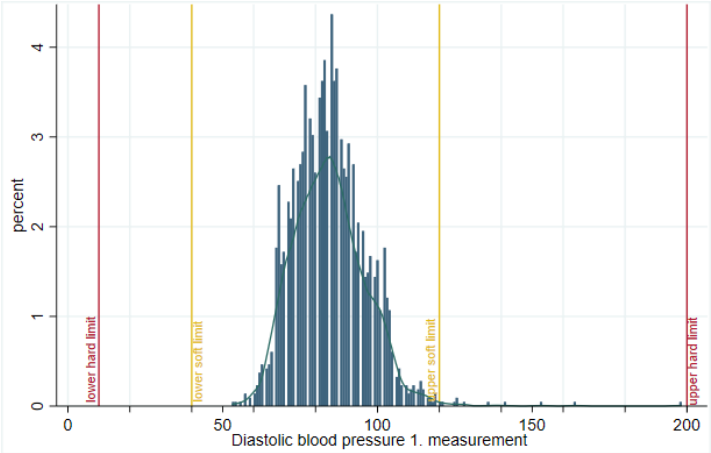
Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Results for variable: dbp2

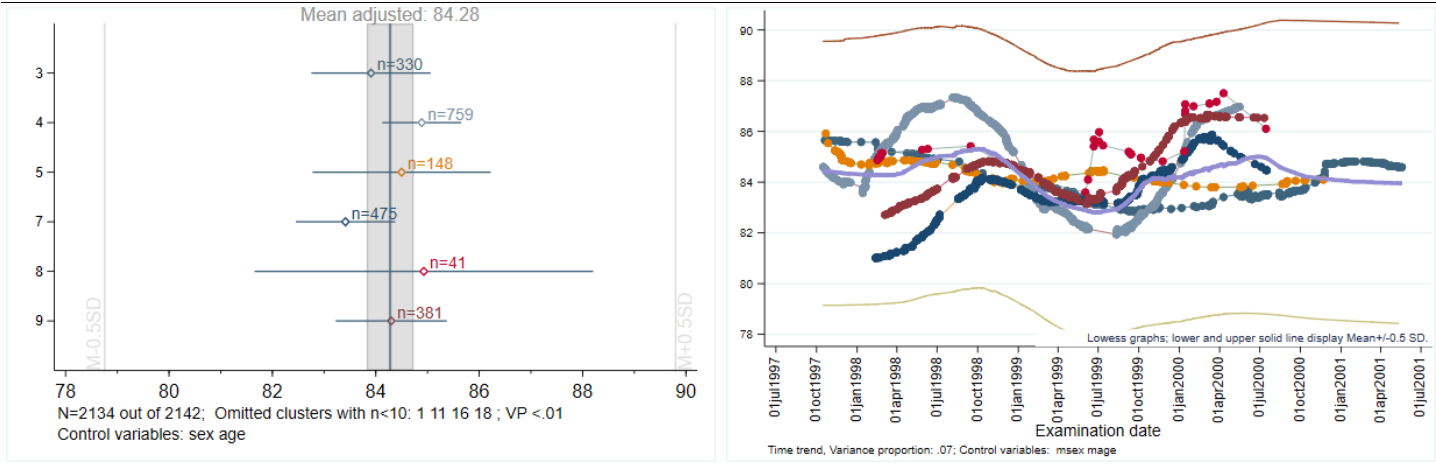Primary variable: Diastolic blood pressure 2. measurement (mmHg)

Data type: float / Stata-Format: float %8.0g converted: int %8.0g / Scale level: ratio

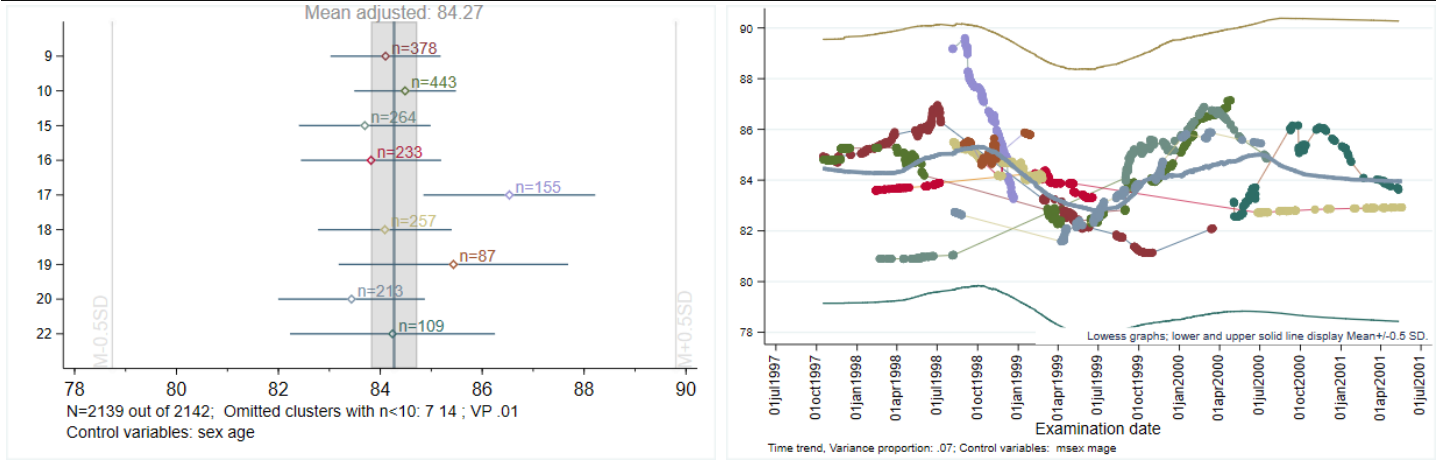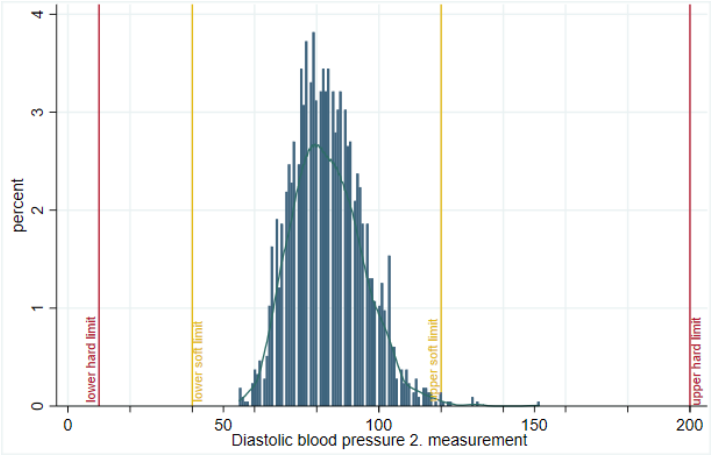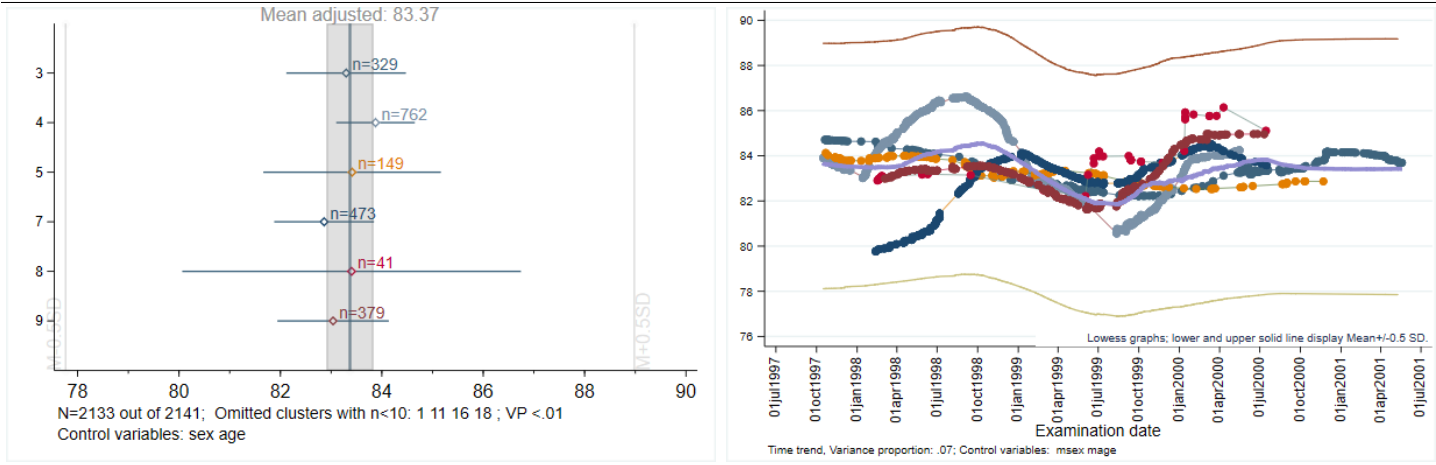| Measures	Original variable	Modified variable
N	2148	2141
Missing values	6	13
Mean	83.52	83.37
Standard deviation	11.52	11.21
Skewness	0.51	0.33
Minimum	55.00	55.00
Percentile 1	61.00	61.00
Percentile 50	83.00	83.00
Percentile 99	114.00	112.00
Maximum	151.00	120.00

For information on modifications see table: Change-log for modified variables.

Variable: dbp2   Cluster effects for: Observer variables obs_bp

N=2133 out of 2141; Omitted clusters with n<10: 1 11 16 18 ; VP <.01
Control variables: sex age

Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Variable: dbp2   Cluster effects for: Device variables dev_bp

N=2138 out of 2141; Omitted clusters with n<10: 7 14 ; VP .01
Control variables: sex age

Time trend, Variance proportion: .07; Control variables: msex mage
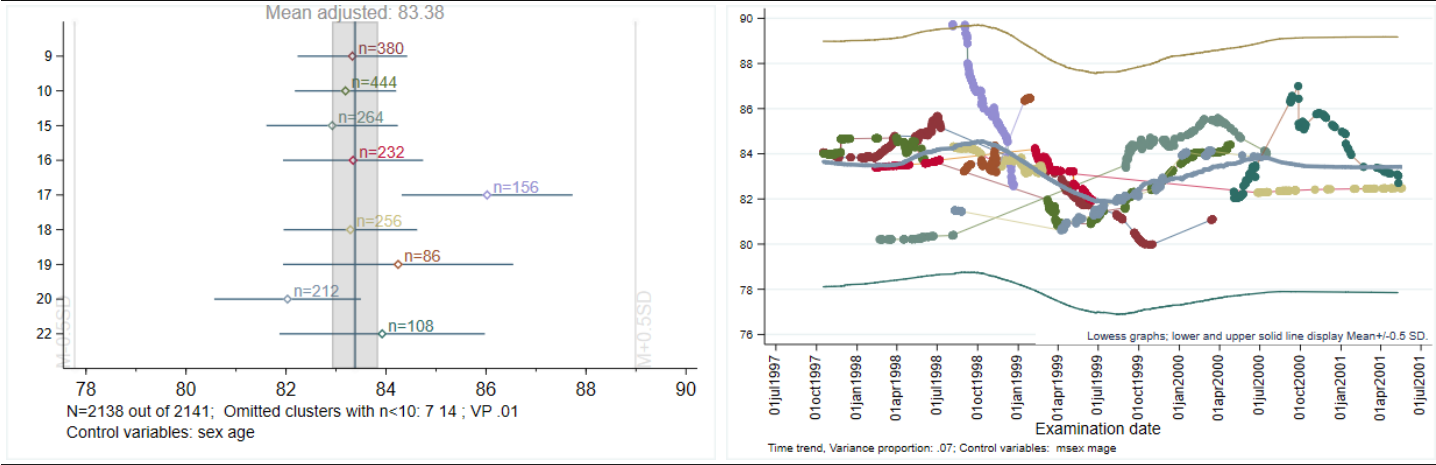
Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Results for variable: height

Primary variable: Heigth (cm)

Data type: float / Stata-Format: double %12.0g / Scale level: ratio

Measures	Original variable	Modified variable
N	2151	
Missing values	3	Measurement values
Mean	168.71	weren't modified.
Standard deviation	9.25	
Skewness	0.11	
Minimum	144.05	
Percentile 1	149.26	
Percentile 50	168.38	
Percentile 99	190.10	
Maximum	198.13	

Variable: height   Cluster effects for: Observer variables obs_soma

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Variable: height   Cluster effects for: Device variables dev_length
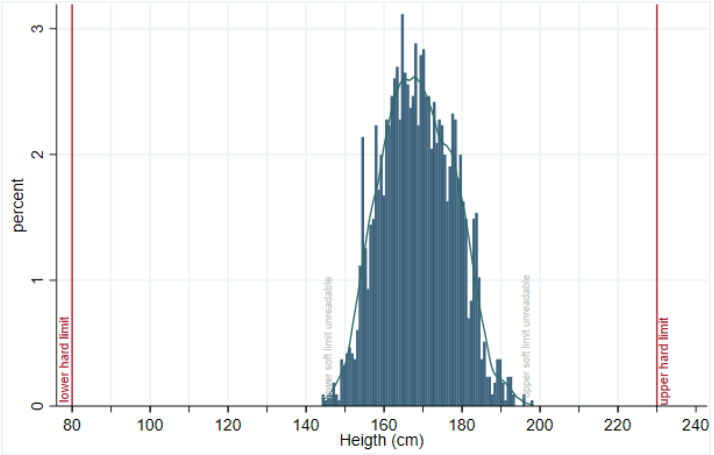

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate
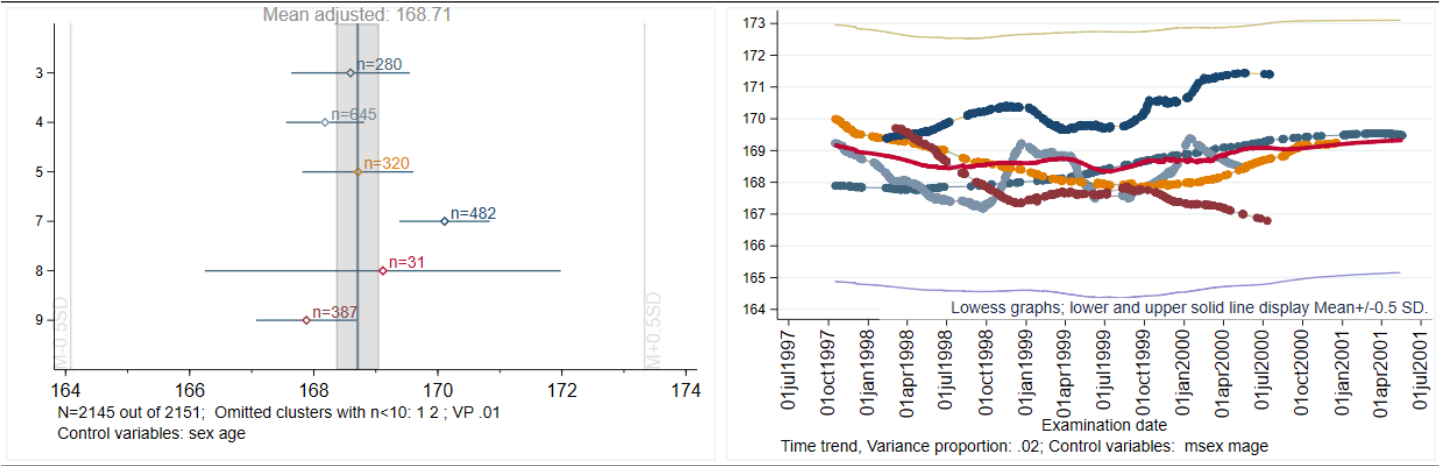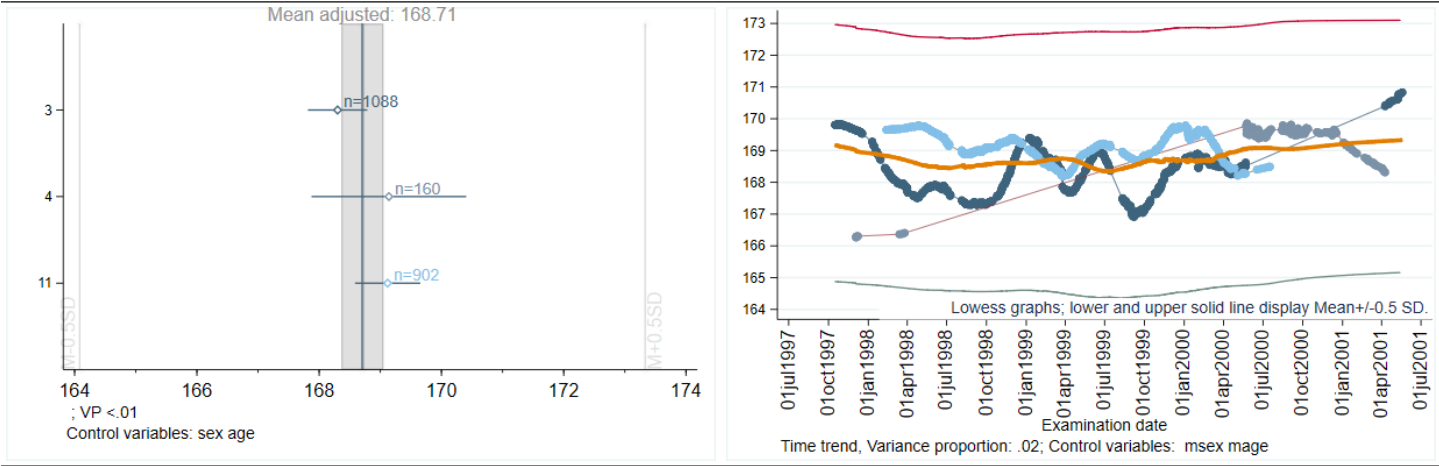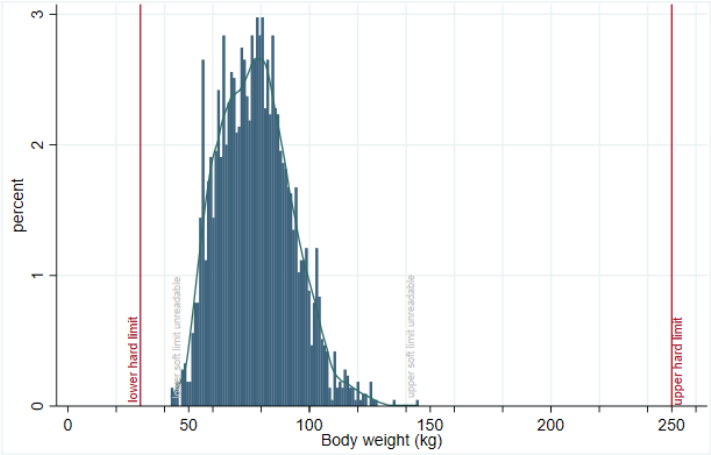
Results for variable: weight
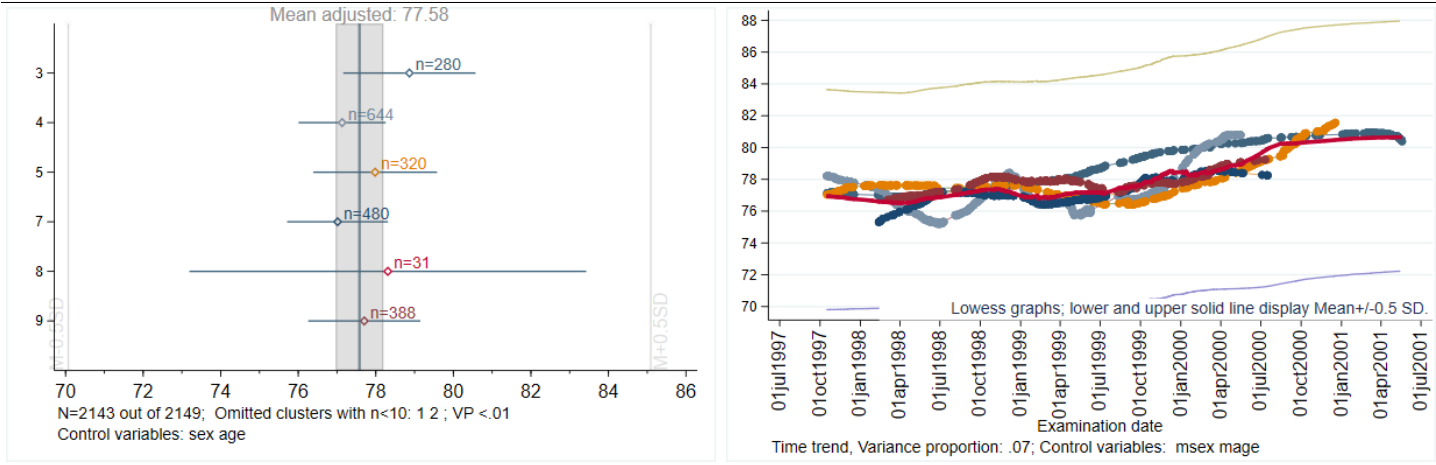
Primary variable: Body weight (kg)

Data type: float / Stata-Format: double %12.0g / Scale level: ratio

Measures	Original variable	Modified variable
N	2150	2149
Missing values	4	5
Mean	77.63	77.60
Standard deviation	15.08	15.01
Skewness	0.41	0.38
Minimum	42.60	42.60
Percentile 1	48.62	48.62
Percentile 50	77.04	77.03
Percentile 99	117.22	116.49
Maximum	144.44	135.48

For information on modifications see table: Change-log for modified variables.

Variable: weight   Cluster effects for: Observer variables obs_soma

N=2143 out of 2149; Omitted clusters with n<10: 1 2 ; VP <.01
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.

Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=1) Variable for time trend: exdate

Variable: weight   Cluster effects for: Device variables dev_weight
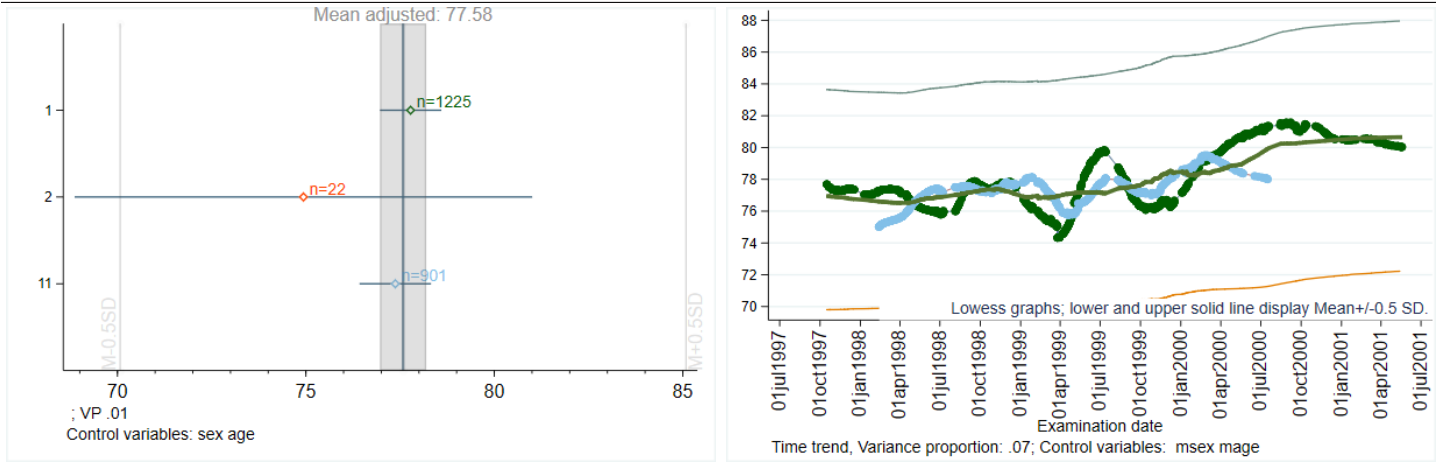
; VP .01
Control variables: sex age

Lowess graphs; lower and upper solid line display Mean+/-0.5 SD.

Time trend, Variance proportion: .07; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=1) Variable for time trend: exdate

Results for variable: waist

Primary variable: Waist circumference

Data type: float / Stata-Format: double %12.0g / Scale level: ratio

Measures	Original variable	Modified variable
N	2151	
Missing values	3	Measurement values
Mean	89.20	weren't modified.
Standard deviation	13.82	
Skewness	0.18	
Minimum	49.27	
Percentile 1	63.33	
Percentile 50	89.52	
Percentile 99	120.48	
Maximum	145.77	
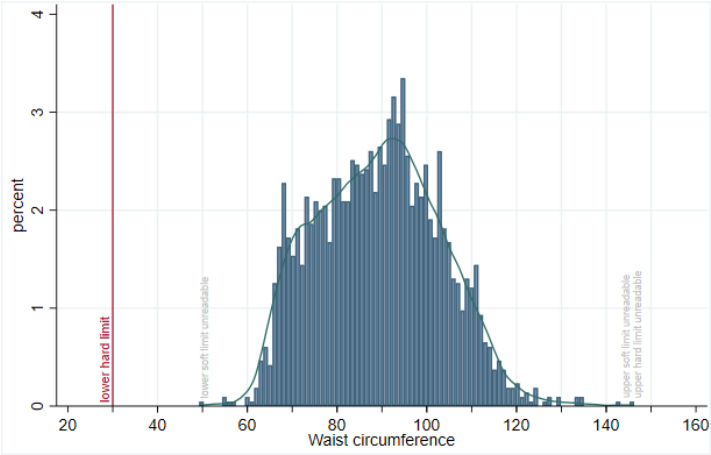
Variable: waist   Cluster effects for: Observer variables obs_soma
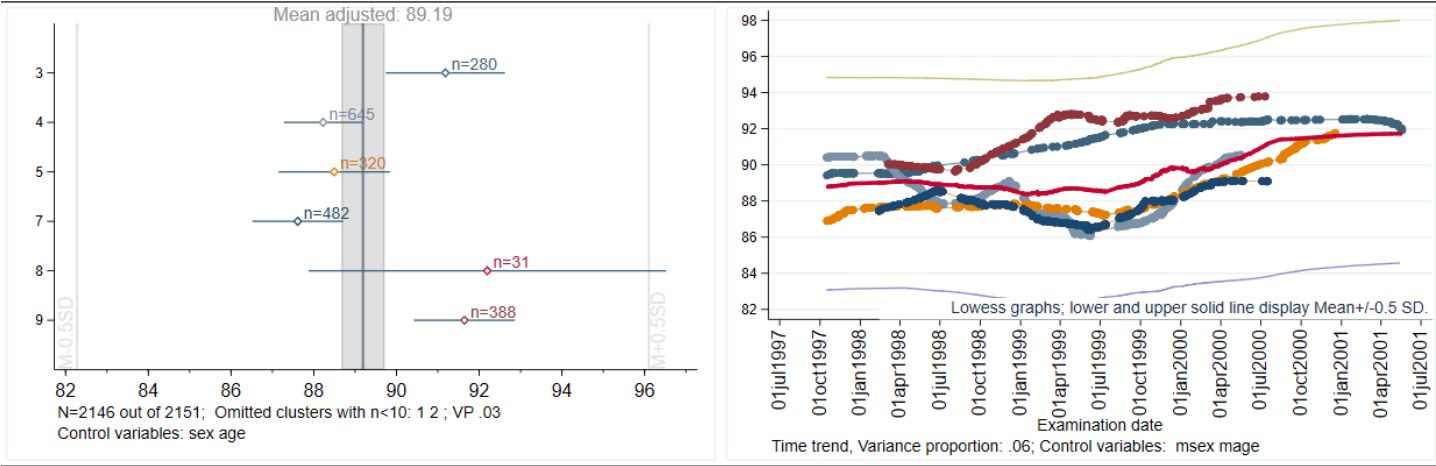
N=2146 out of 2151; Omitted clusters with n<10: 1 2 ; VP .03
Control variables: sex age

Time trend, Variance proportion: .06; Control variables: msex mage

Automatic recodings: Missings recoded (n=0) ; Outliers as missings (n=0) Variable for time trend: exdate

Change-log for modified variables

Primary variables	Missings recoded	Categories changed	Range violations as missings	Outliers as missings	# Any changes	N
sbp1 Systolic blood pressure 1. measurement (mmHg)	0	0	5	0	5	2147
sbp2 Systolic blood pressure 2. measurement (mmHg)	0	0	6	0	6	2142
dbp1 Diastolic blood pressure 1. measurement (mmHg)	0	0	10	0	10	2142
dbp2 Diastolic blood pressure 2. measurement (mmHg)	0	0	7	0	7	2141
height Heigth (cm)	0	0	0	0	0	2151
weight Body weight (kg)	0	0	0	1	1	2149
waist Waist circumference	0	0	0	0	0	2151

Secondary variables	Missings recoded	Categories changed	Range violations as missings	Outliers as missings	# Any changes	N
school Educational level	14	2146	0	0	2160	2147
family Marital status	14	2147	0	0	2161	2147
smoking Smoking	14	668	0	0	682	2146
stroke Stroke ever	20	2144	0	0	2164	2144
myocard Myocardial infarction ever	28	2140	0	0	2168	2140
diab_known Diabetes	14	0	0	0	14	2147
diab_age Diabetes Age onset	14	0	0	0	14	173
contraception Contraceptives Ever	18	1089	0	0	1107	1089
income Houshold income	232	0	0	0	232	2038
hdl HDL coholesterol	0	0	0	9	9	2129
ldl LDL coholesterol	0	0	0	1	1	2125
cholesterol Cholesterol	0	0	0	2	2	2137

Control variables	Missings recoded	Categories changed	Range violations as missings	Outliers as missings	# Any changes	N
sex Sex	0	2154	0	0	2154	2154
age Age (years)	0	0	1	0	1	2153

Time variables	Missings recoded	Categories changed	Range violations as missings	Outliers as missings	# Any changes	N
exdate	0	0	0	0	0	2154

Examination date

Observer variables	Missings recoded	Categories changed	Range violations as missings	Outliers as missings	# Any changes	N
obs_int Observer Interview	2	0	0	0	2	2153
obs_soma Observer Somatometry	4	0	0	0	4	2152
obs_bp Observer blood pressure	6	0	0	0	6	2151
Device variables	**Missings recoded**	**Categories changed**	**Range violations as missings**	**Outliers as missings**	**# Any changes**	**N**
dev_weight Device weight	4	0	0	0	4	2152
dev_length Device heigth	4	0	0	0	4	2152
dev_bp Device blood pressure	6	0	0	0	6	2151

This table provides an overview for all variable changes when computing modified variables.

Study data integrity issues and notes

Message	Object	Data quality classification and comment
WARNING	Metavariable: varorder	Unexpected data type
WARNING	Primary variables sbp1	Unexpected data type: expected: float observed: int %8.0g
Note	Primary variables sbp1	The format of this variable was changed. float %8.0g converted: int %8.0g
WARNING	Primary variables sbp2	Unexpected data type: expected: float observed: int %8.0g
Note	Primary variables sbp2	The format of this variable was changed. float %8.0g converted: int %8.0g
WARNING	Primary variables dbp1	Unexpected data type: expected: float observed: int %8.0g
Note	Primary variables dbp1	The format of this variable was changed. float %8.0g converted: int %8.0g
WARNING	Primary variables dbp2	Unexpected data type: expected: float observed: int %8.0g
Note	Primary variables dbp2	The format of this variable was changed. float %8.0g converted: int %8.0g
Note	Control variables age	The format of this variable was changed. float %8.0g converted: byte %8.0g
Note	Observer variables obs_int	The format of this variable was changed. float %9.0g converted: byte %9.0g
Note	Observer variables obs_soma	The format of this variable was changed. float %9.0g converted: byte %9.0g
Note	Observer variables obs_bp	The format of this variable was changed. float %9.0g converted: byte %9.0g
Note	Device variables dev_length	The format of this variable was changed. int %8.0g converted: byte %8.0g