Leids Universitair Medisch Centrum

Re-quantified Nightingale measurements compared to the old versions

Daniele Bizzarri

email: d.bizzarri@lumc.nl

Universiteit Leiden The Netherlands

MOLEPI molecular epidemiology

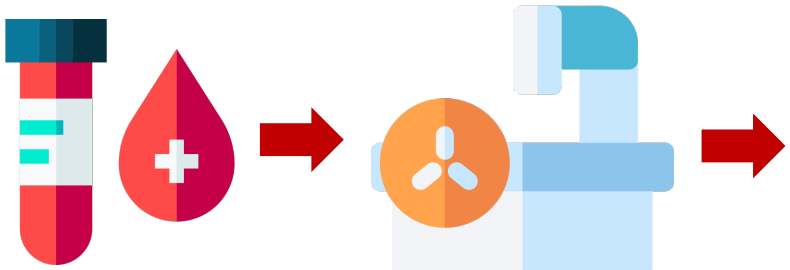
LCBC

The Delft Bioinformatics Lab

Metabolomics from Nightingale health

^{1}H-NMR Metabolomics in large epidemiological studies

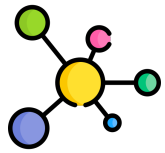
Nightingale



100 μL

NMR

Metabolic profile (>200 features) for:



routine lipids,
14 lipoprotein subclasses
particle sizes,
apoplipoproteins
fatty acids
omega-3 and 6

amino acids
ketons,
chronic inflammation
fluid balance
kidney function
glycolisis

BBMRI.nl
Biobanking and BioMolecular resources Research Infrastructure The Netherlands

~30.000 samples

biobank uk
Enabling scientific discoveries that improve human health

~ 120.000 samples

Nightingale platform versions

BBMRI-nl contains several versions of the platform

BBMRI.nl
Biobanking and BioMolecular resources Research Infrastructure The Netherlands

~35.000 samples

First wave:
~25,000 samples from 26 Biobanks → Nightingale data 2014

Second wave:
~9,900 samples from 10 biobanks (mostly repeated measures + HOF and STABILITEIT) → Nightingale data 2016

Requantification:
~35,000 samples from 28 biobanks → Nightingale data 2020

biobank uk
Enabling scientific discoveries that improve human health

~ 120.000 samples → Nightingale data 2020

To what extent are these measurements comparable?

What is this presentation about?

We used the Leiden Longevity Study to evaluate:

1) Comparison data 2014 and 2020
2) Comparison data 2016 and 2020
3) Comparisons of the ”clinically validated” markers by Nightingale with clinical chemistry 2014 and 2020
4) Comparisons mortality score and MetaboAge in 2014, 2016 and 2020

AIM: We want to inform you!

What data did we analyze?
First and second metabolomics

Leiden Longevity Study:

LLS_PARTOFFS (middle aged [30 - 79 y.o.])

Time-point	First wave	Recalibration	Common data after QC
IOP1	2014	2020	1863
IOP2	2016	2020	588
IOP3	2016	2020	388

LLS_SIBS (older individuals [89-103 y.o.])

Time-point	First wave	Recalibration	Common data after QC
IOP1	2014	2020	942

We focus on the 65 most uncorrelated measures.

What has changed in 2020?

2 discontinued, 25 new

Discontinued

name	description
HDL2-C	HDL2 cholesterol
HDL3-C	HDL3 cholesterol

New

name	description
non-HDL-C	Total cholesterol minus HDL-C
Clinical LDL-C	Clinical LDL cholesterol
Total-PL	Total phospholipids in lipoprotein particles
VLDL-PL	Phospholipids in VLDL
LDL-PL	Phospholipids in LDL
HDL-PL	Phospholipids in HDL
VLDL-CE	Cholesteryl esters in VLDL
LDL-CE	Cholesteryl esters in LDL
HDL-CE	Cholesteryl esters in HDL
VLDL-FC	Free cholesterol in VLDL
LDL-FC	Free cholesterol in LDL
HDL-FC	Free cholesterol in HDL
Total-L	Total lipids in lipoprotein particles
VLDL-L	Total lipids in VLDL
LDL-L	Total lipids in LDL
HDL-L	Total lipids in HDL
Total-P	Total concentration of lipoprotein particles
VLDL-P	Concentration of VLDL particles
LDL-P	Concentration of LDL particles
HDL-P	Concentration of HDL particles
PUFA/MUFA	Ratio of polyunsaturated fatty acids to monounsaturated fatty acids
Omega-6/Omega-3	Ratio of omega-6 fatty acids to omega-3 fatty acids
Total BCAA	Total concentration of branched-chain amino acids (leucine + isoleucine + valine)
Acetone	Acetone

What has changed in the content?

37 clinically validated metabolites

Cholesterol	
Total cholesterol	mmol/l
VLDL cholesterol	mmol/l
Clinical LDL cholesterol	mmol/l
HDL cholesterol	mmol/l
Triglycerides	
Total triglycerides	mmol/l
Apolipoproteins	
Apolipoprotein B	g/l
Apolipoprotein A1	g/l
Ratio of apolipoprotein B to apolipoprotein A1	
Fluid balance and inflammation	
Creatinine	mmol/l
Albumin	g/l
Glycoprotein acetyls	mmol/l
Glycolysis metabolites	
Glucose	mmol/l
Lactate	mmol/l

Fatty acids	
Total fatty acids	mmol/l
Omega-3 fatty acids	mmol/l
Omega-6 fatty acids	mmol/l
Polyunsaturated fatty acids	mmol/l
Monounsaturated fatty acids	mmol/l
Saturated fatty acids	mmol/l
Docosahexaenoic acid	mmol/l
Fatty acid ratios	
Omega-3 fatty acids to total fatty acids	%
Omega-6 fatty acids to total fatty acids	%
Polyunsaturated fatty acids to total fatty acids	%
Monounsaturated fatty acids to total fatty acids	%
Saturated fatty acids to total fatty acids	%
Docosahexaenoic acid to total fatty acids	%
Polyunsaturated fatty acids to monounsaturated fatty acids	%
Omega-6 fatty acids to omega-3 fatty acids	%

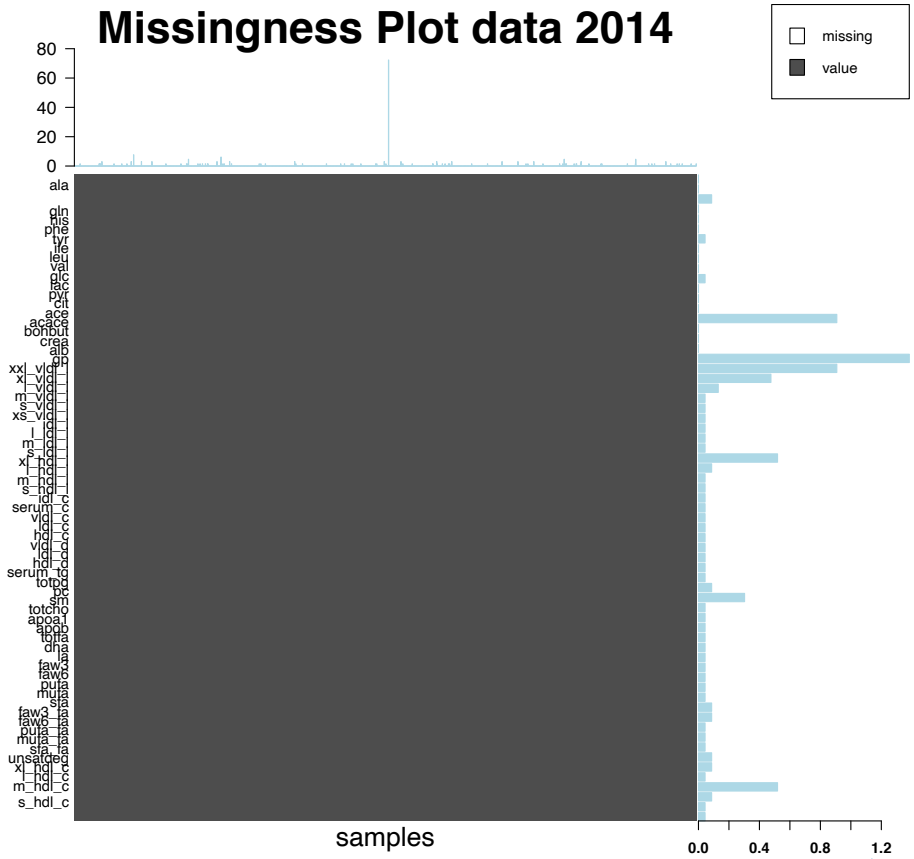
Branched-chain amino acids	
Total branched-chain amino acids	mmol/l
Isoleucine	mmol/l
Leucine	mmol/l
Valine	mmol/l
Aromatic amino acids	
Phenylalanine	mmol/l
Tyrosine	mmol/l
Other amino acids	
Alanine	mmol/l
Glycine	mmol/l
Histidine	mmol/l

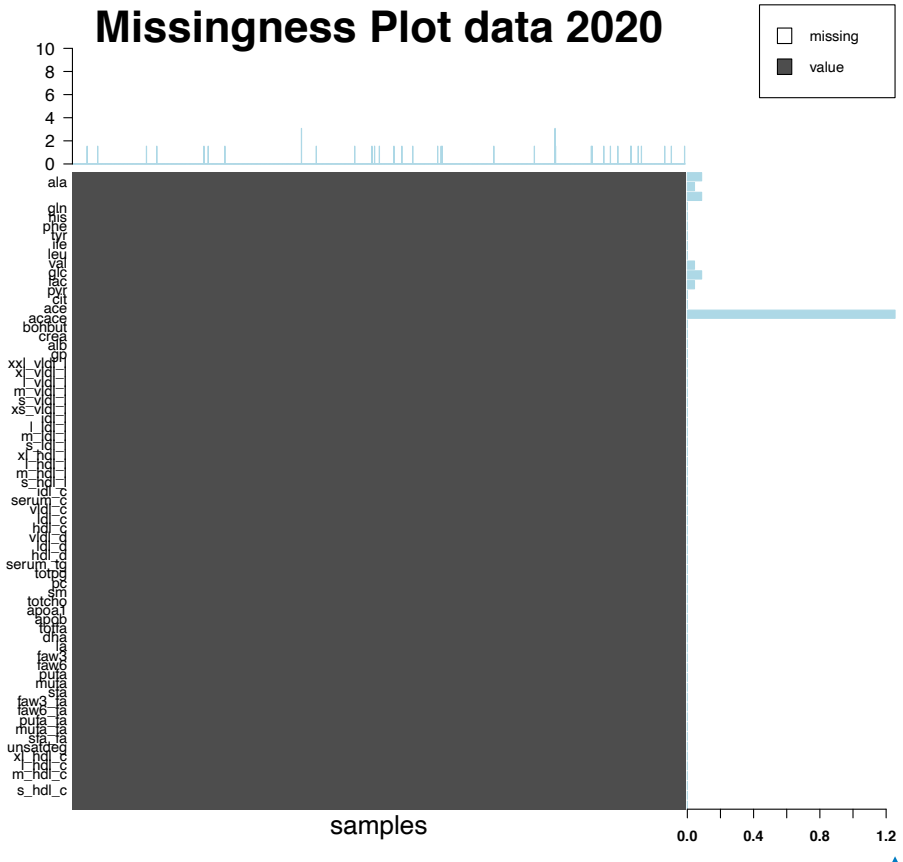
Link to site: https://research.nightingalehealth.com/clinically-validated-biomarkers

Missingness and 5-SD outliers

LLS_PAROFFS IOP1: From low to lower

Missingness in 2020 is lower



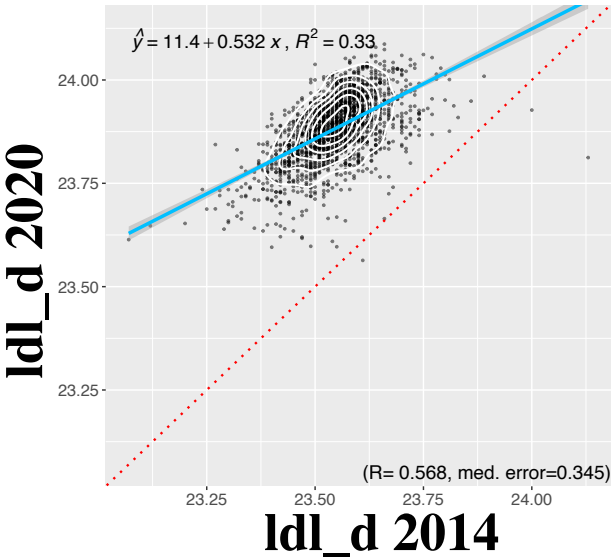


1%

LLS_PARTOFFS IOP1 (data 2014 and 2020)

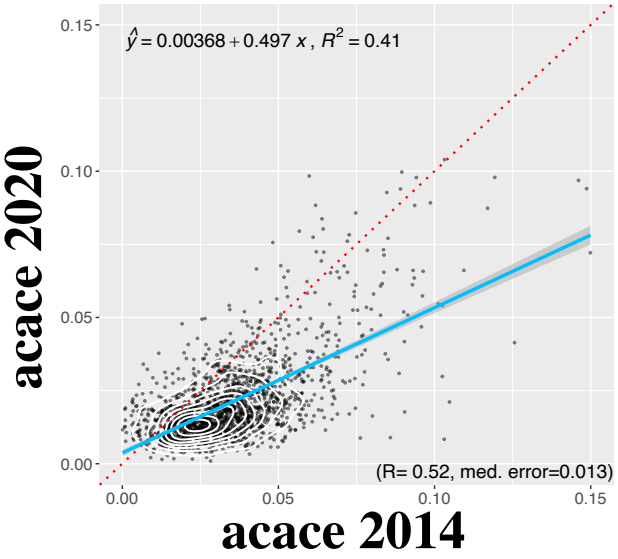
scatterplots: exemplify some differences in signal

ldl_d, corr=0.568



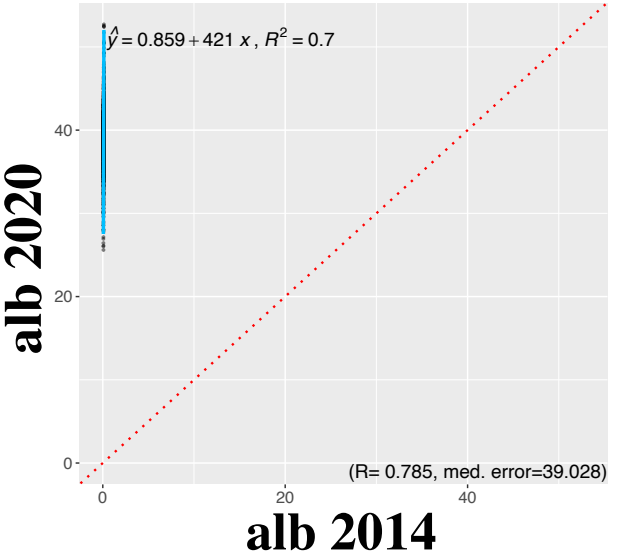
1. Poor correlation
2. Difference in mean

acace, corr=0.52



1. Poor correlation

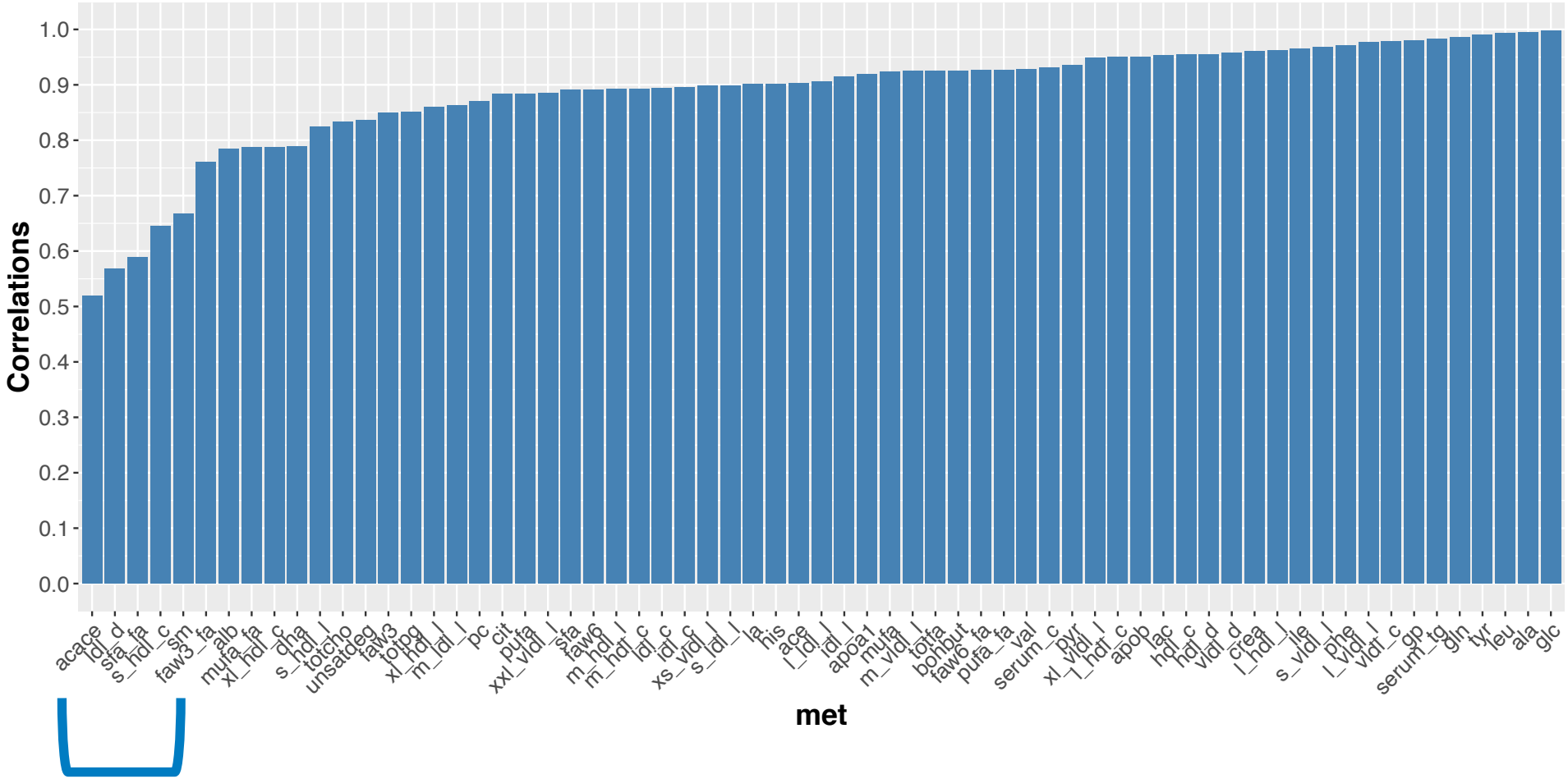
alb, corr=0.785



1. Ok correlation
2. Different unit

LLS_PARTOFFS IOP1

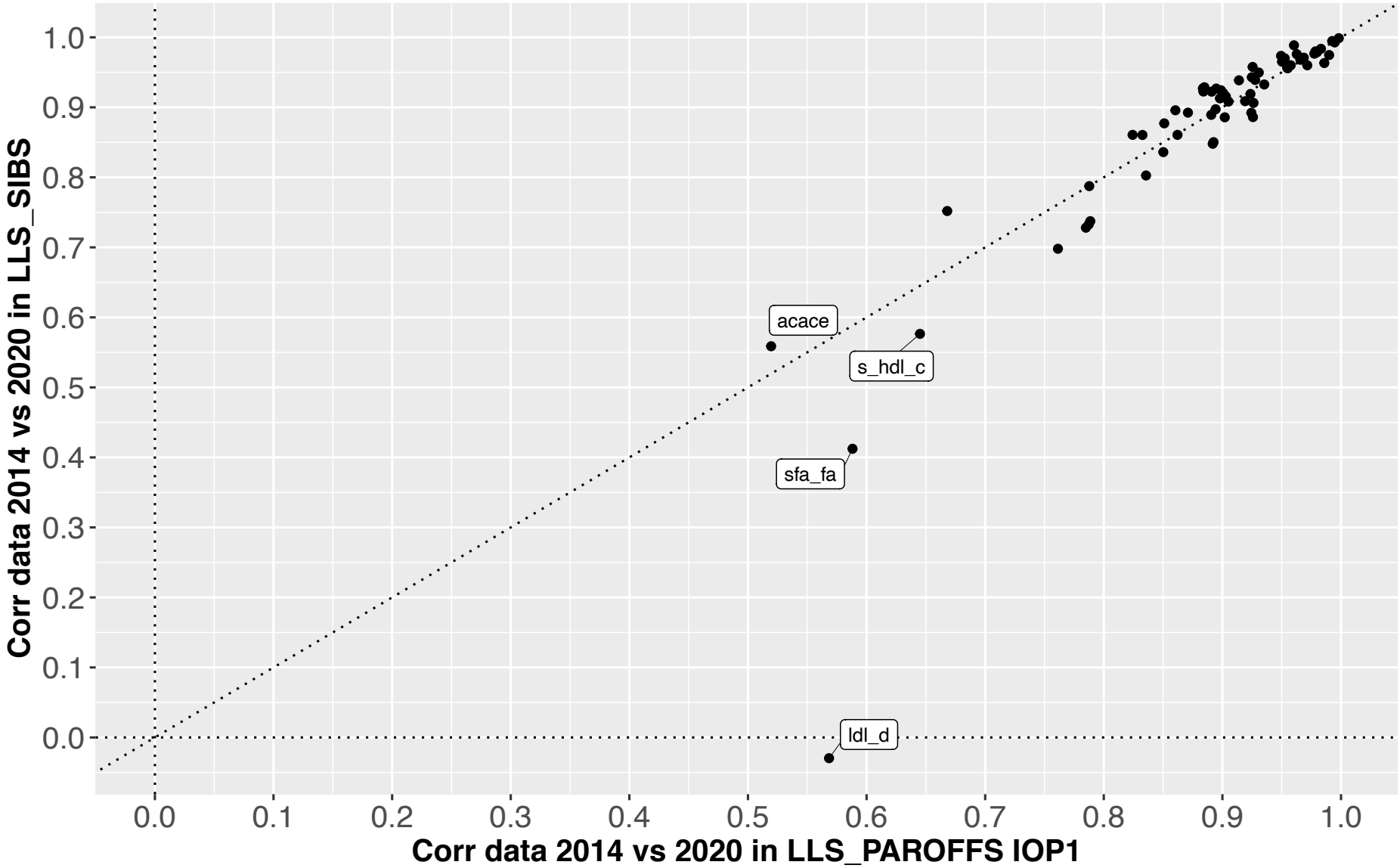
Spearman's Correlations 2014 vs 2020



5 poorly correlating met (cor≤0.7): acace, ldl_d, sfa_fa, s_hdl_c, sm!

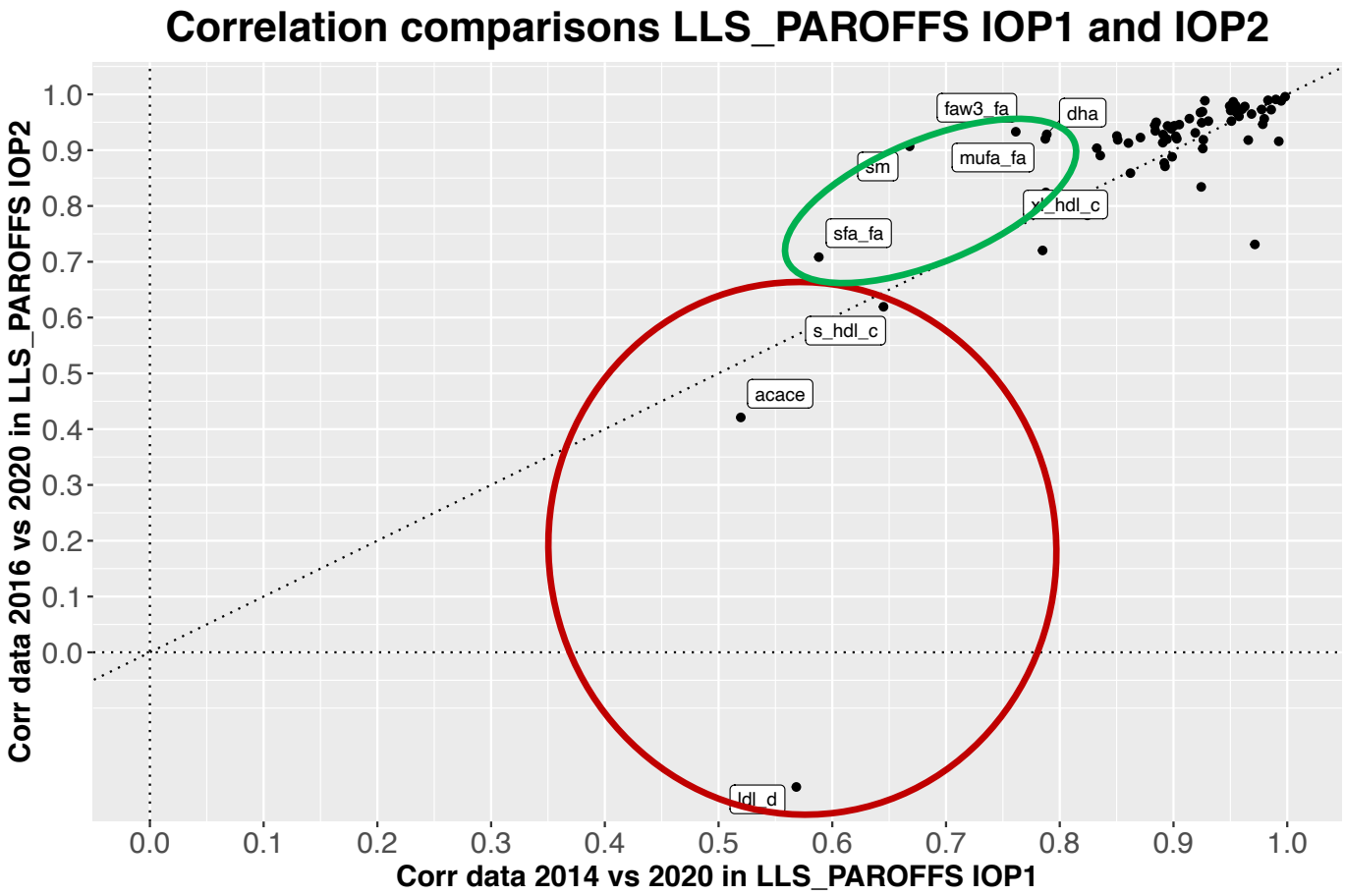
Comparing correlations 2014 vs 2020 in 2 cohort subsets

Consistent correlations



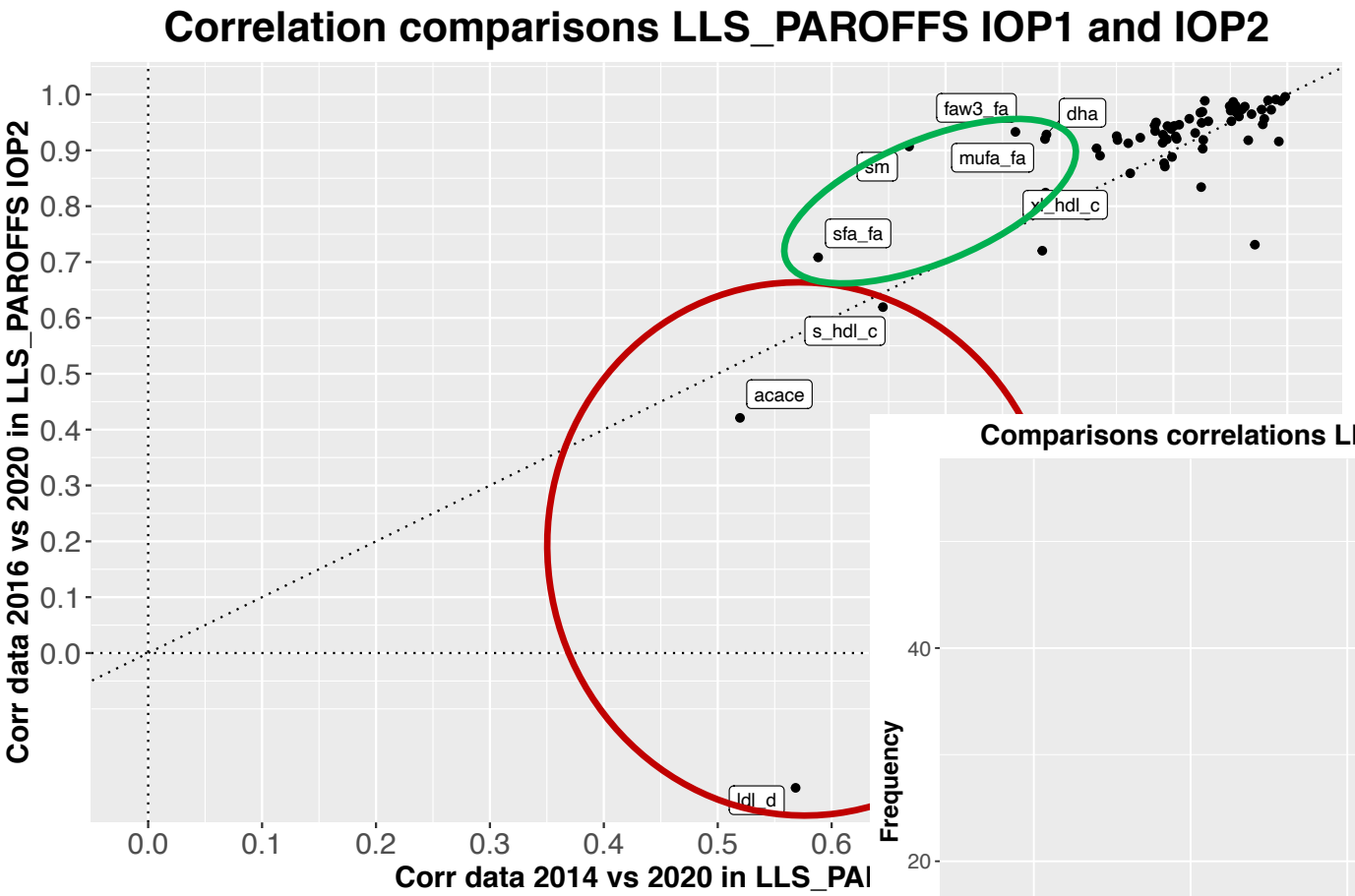
Comparing correlations wave 2014 and 2016 vs 2020

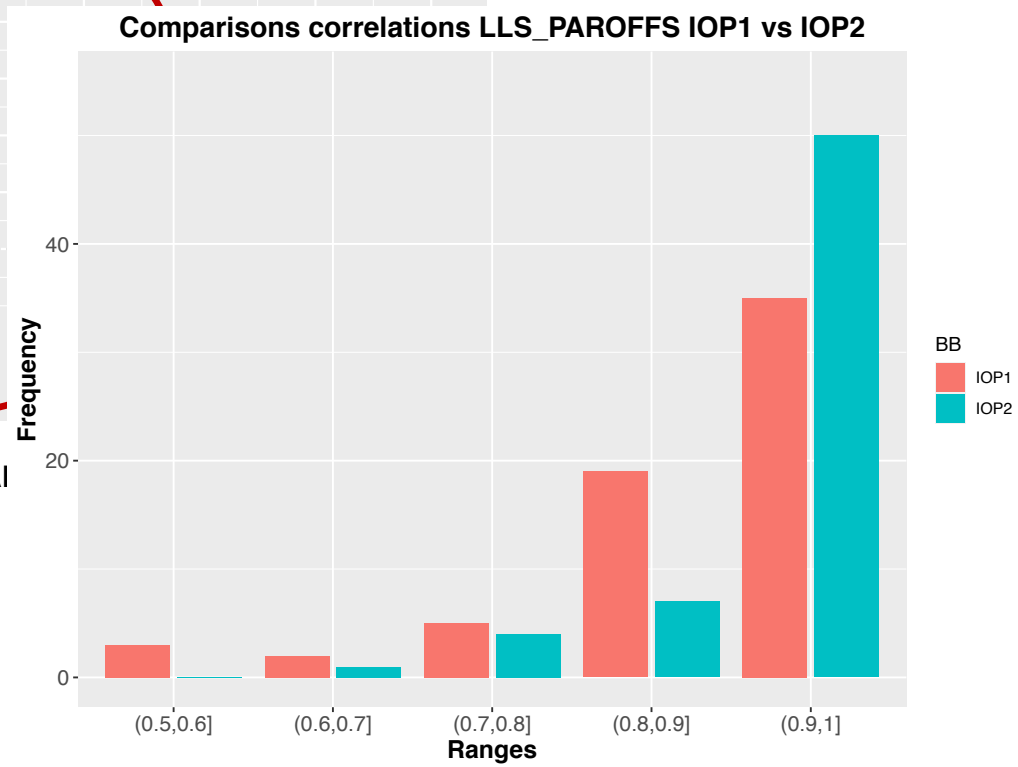
Better correlations with data 2016



Comparing correlations wave 2014 and 2016 vs 2020

Better correlations with data 2016

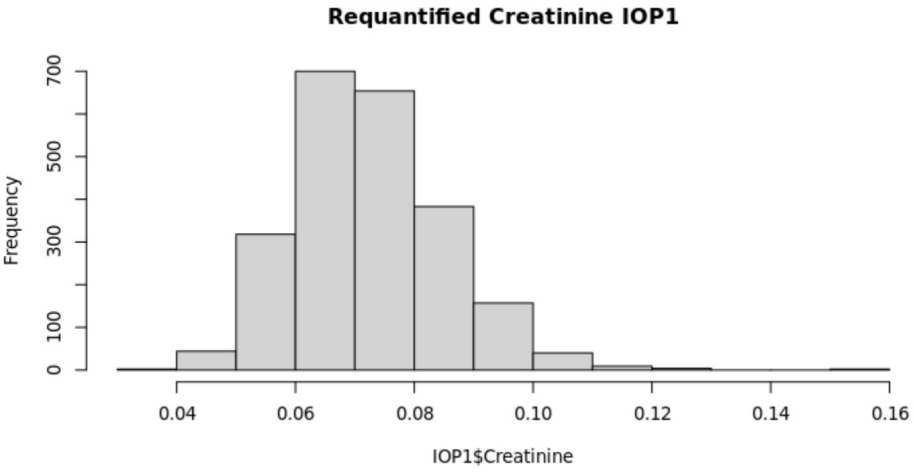


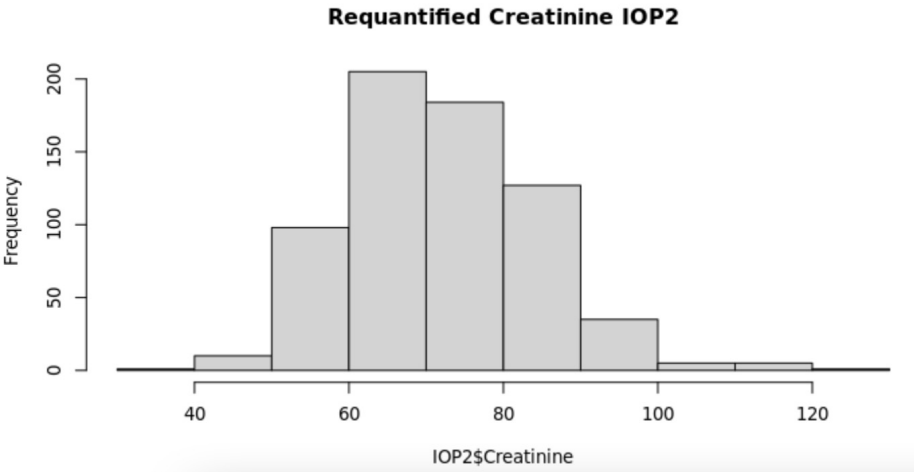


Re-quantification is sensitive to 2014 and 2016

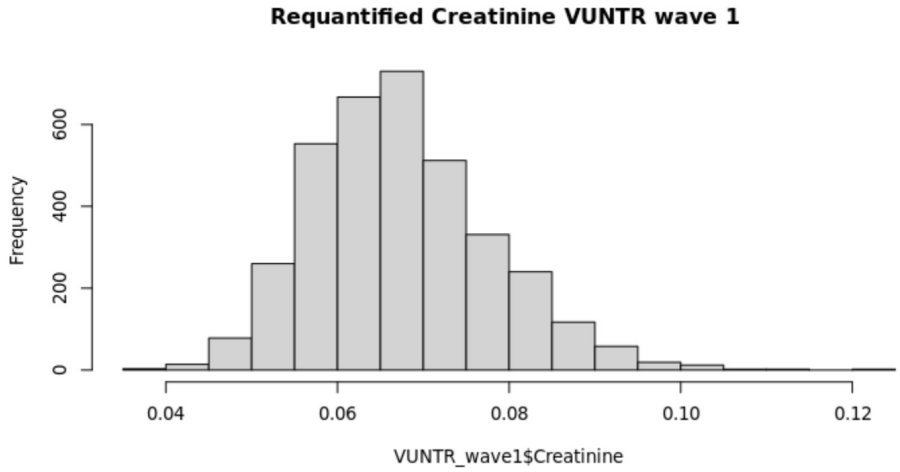
Re-calibrated creatinine is in different units when in 2014 or 2016

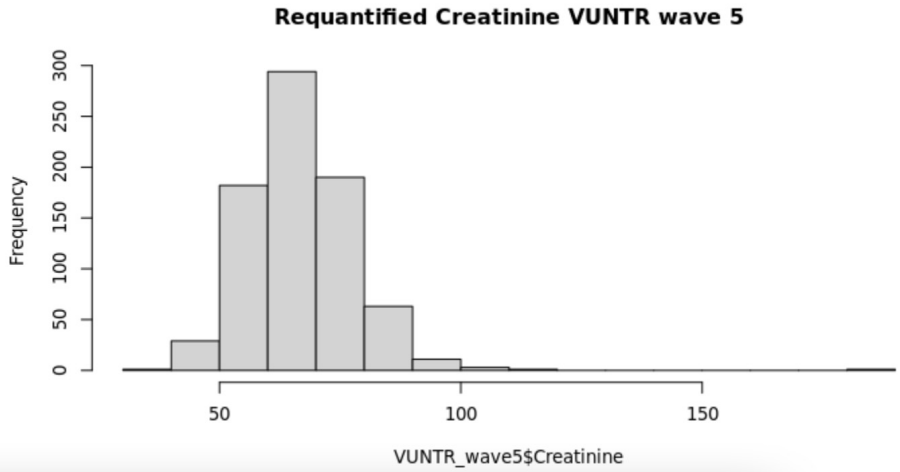
LLS



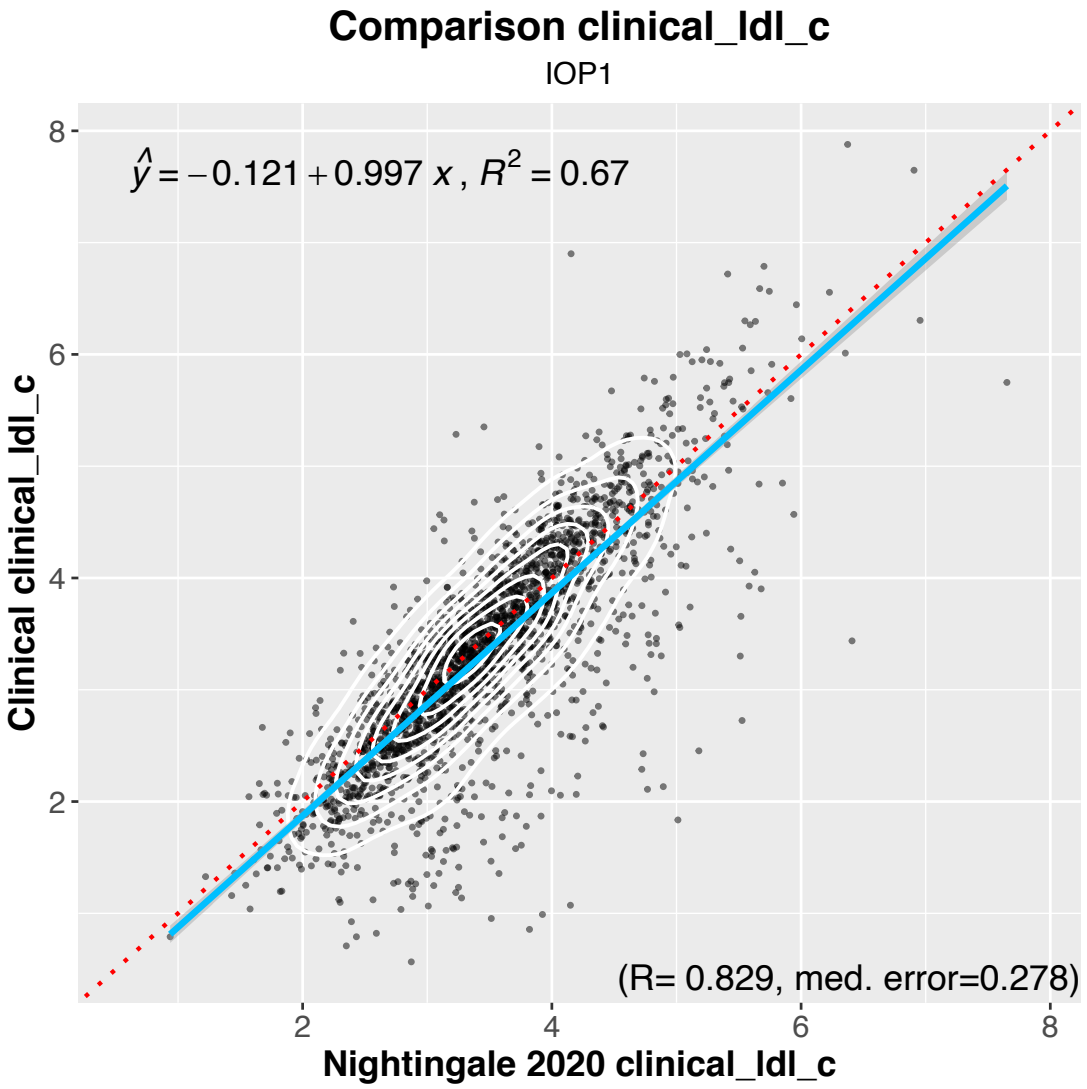


VUNTR



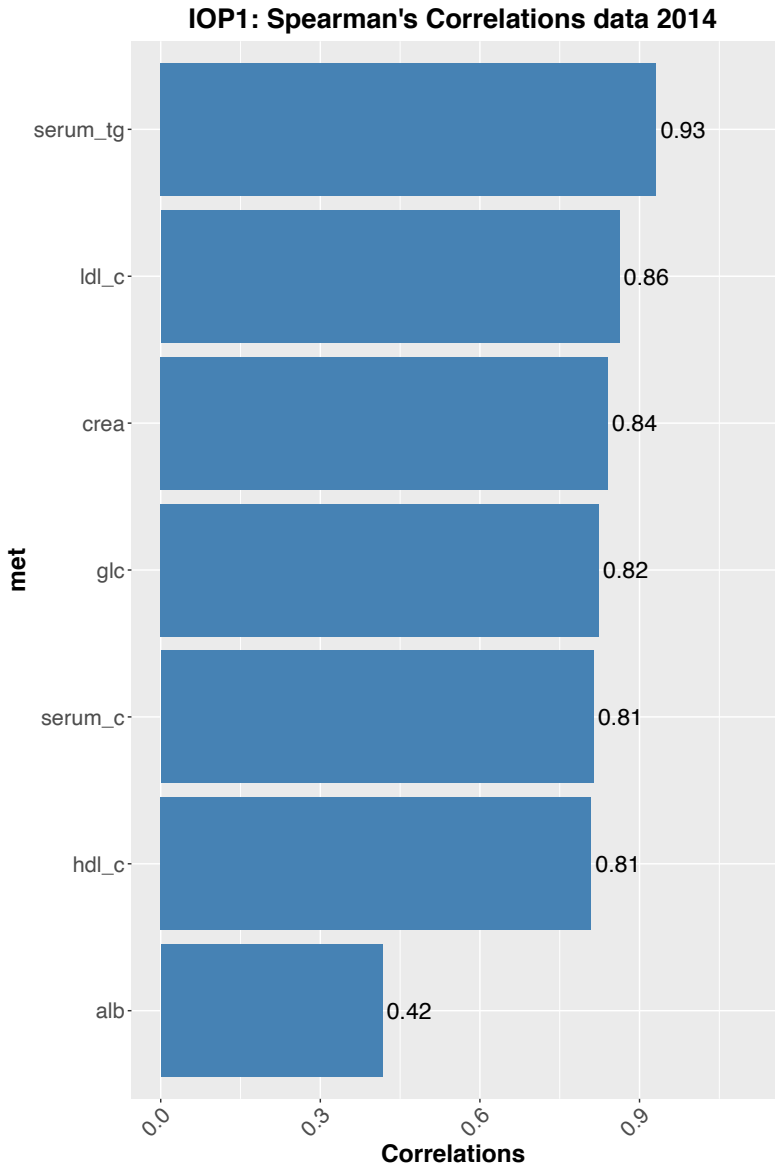


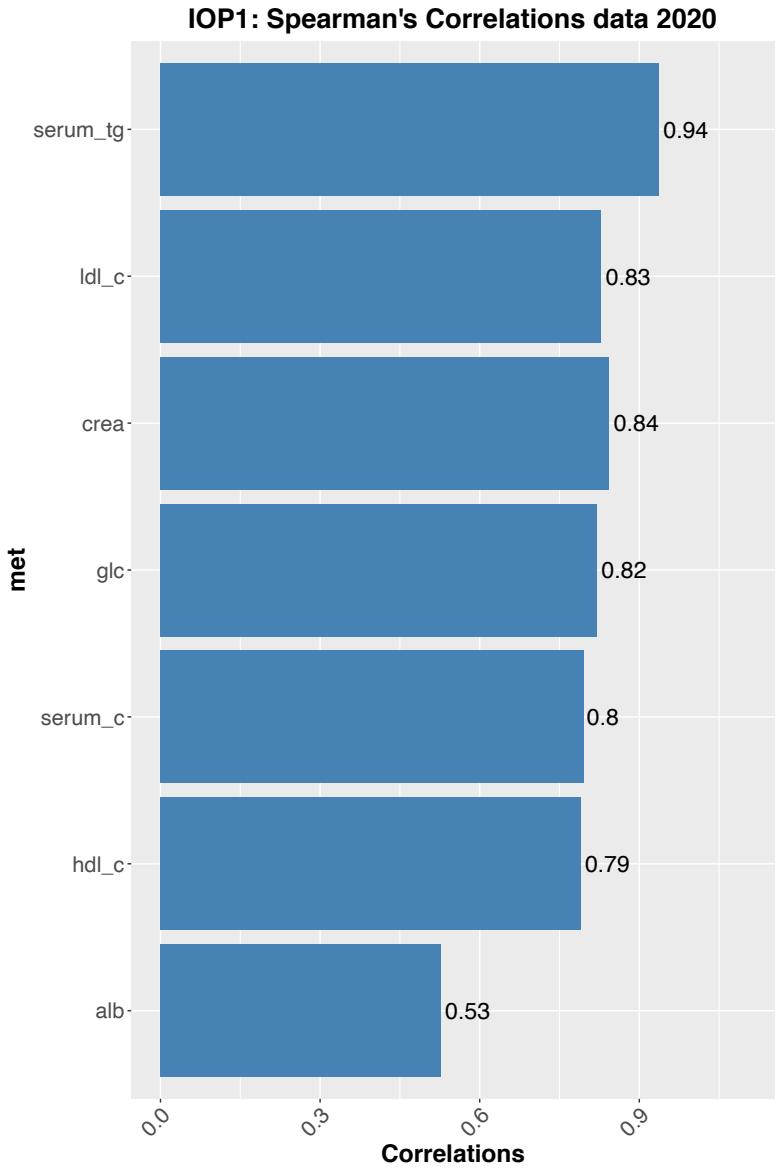
Clinical ldl_c compared to Nightingale 2020 in IOP1



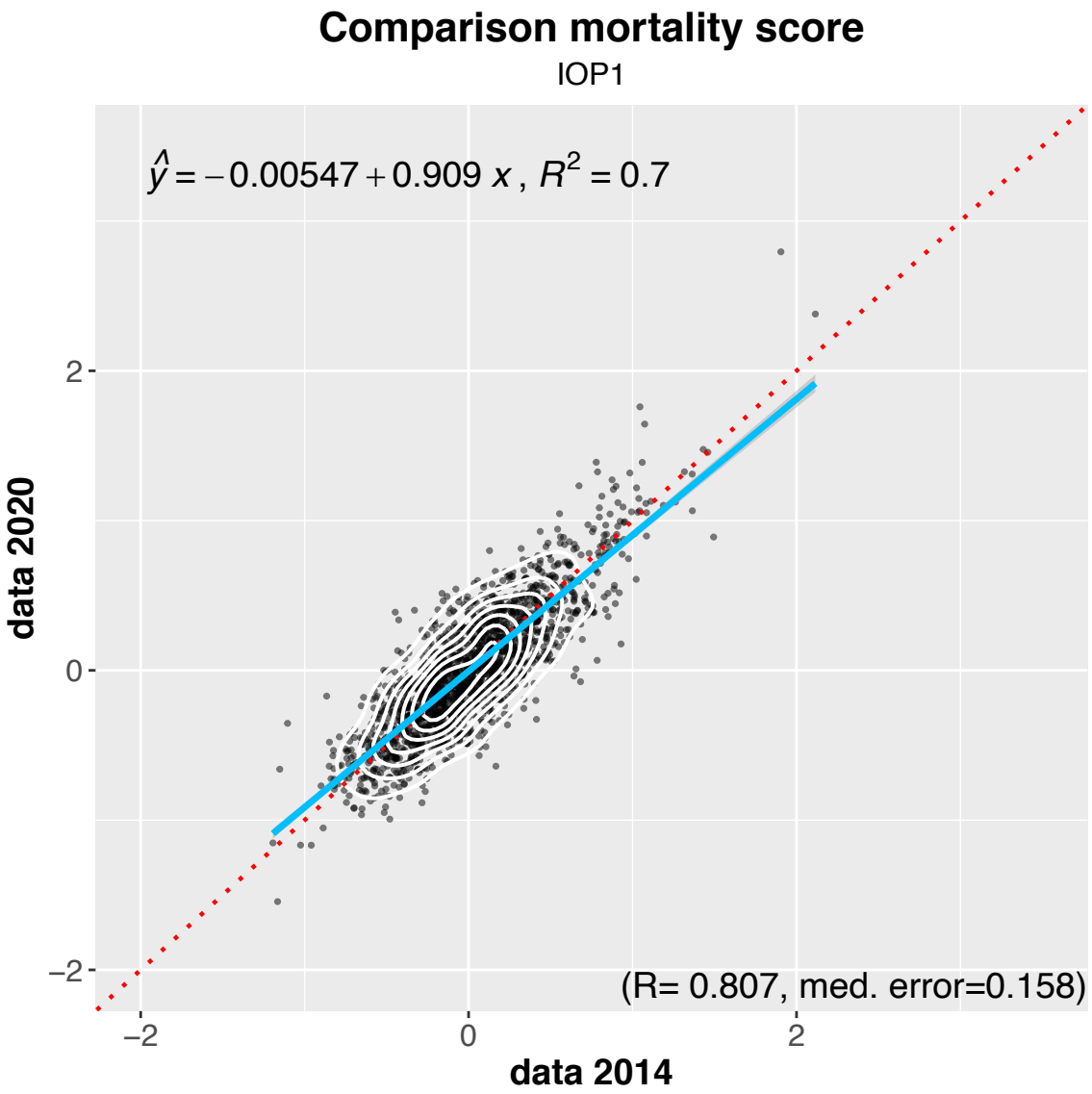
Comparisons with clinical chemistry

Re-calibration does not improve the correlation with clinical variables





Mortality score
IOP1

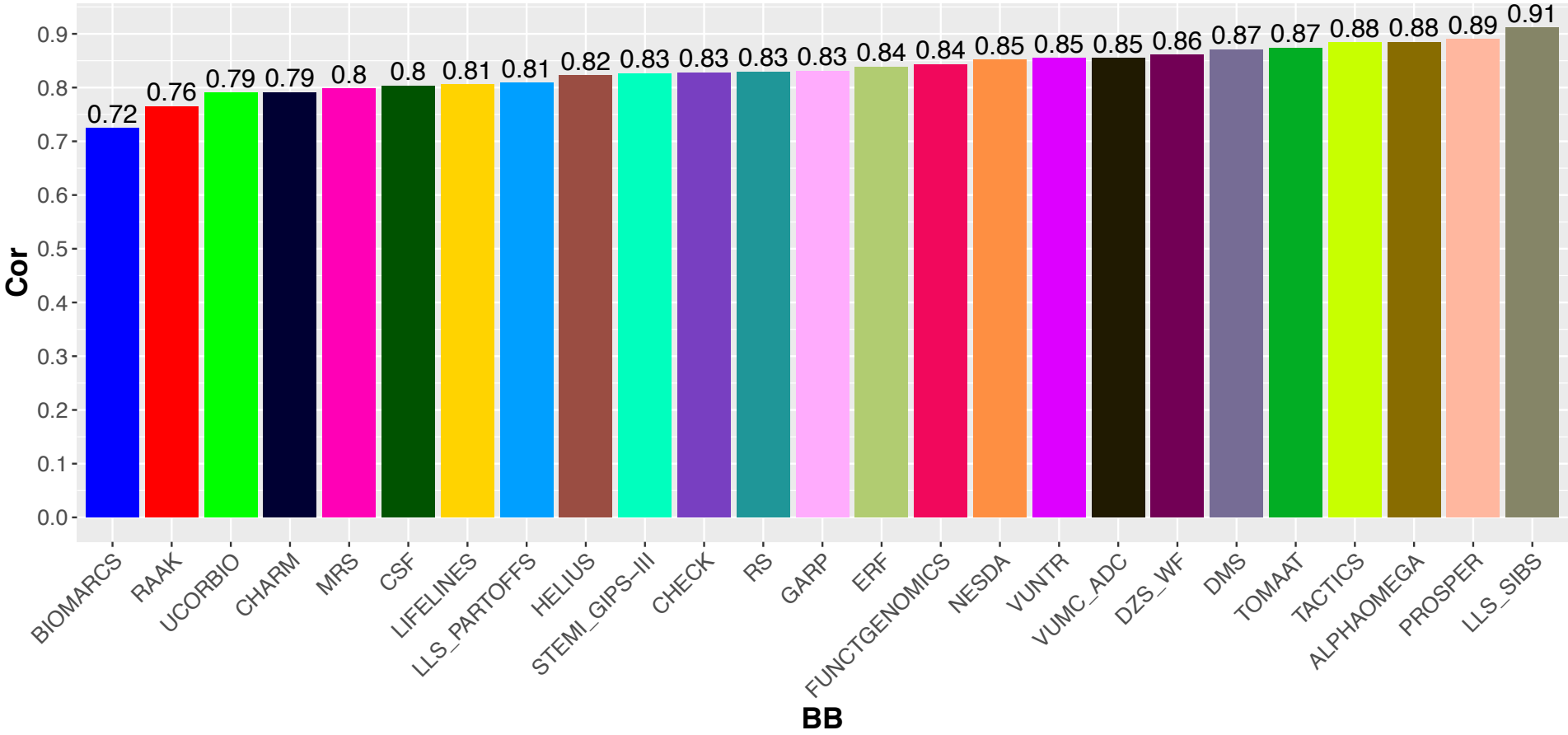


Abbreviation	Beta_value	cor
pufa_fa	-0.25126406	0.9402639
gp	0.28017938	0.9789782
glc	0.14839265	0.9992556
s_hdl_l	-0.14191241	0.8336958
xxl_vldl_l	-0.22368594	0.9569986
alb	-0.11329097	0.8529113
phe	0.12422723	0.9816918
acace	0.07870493	0.7174917
ile	0.20570686	0.9547256
vldl_d	-0.16013759	0.9653296
leu	-0.19574111	0.9936984
val	-0.14369820	0.9029045
his	-0.06922333	0.9240034
lac	0.06226233	0.9450445

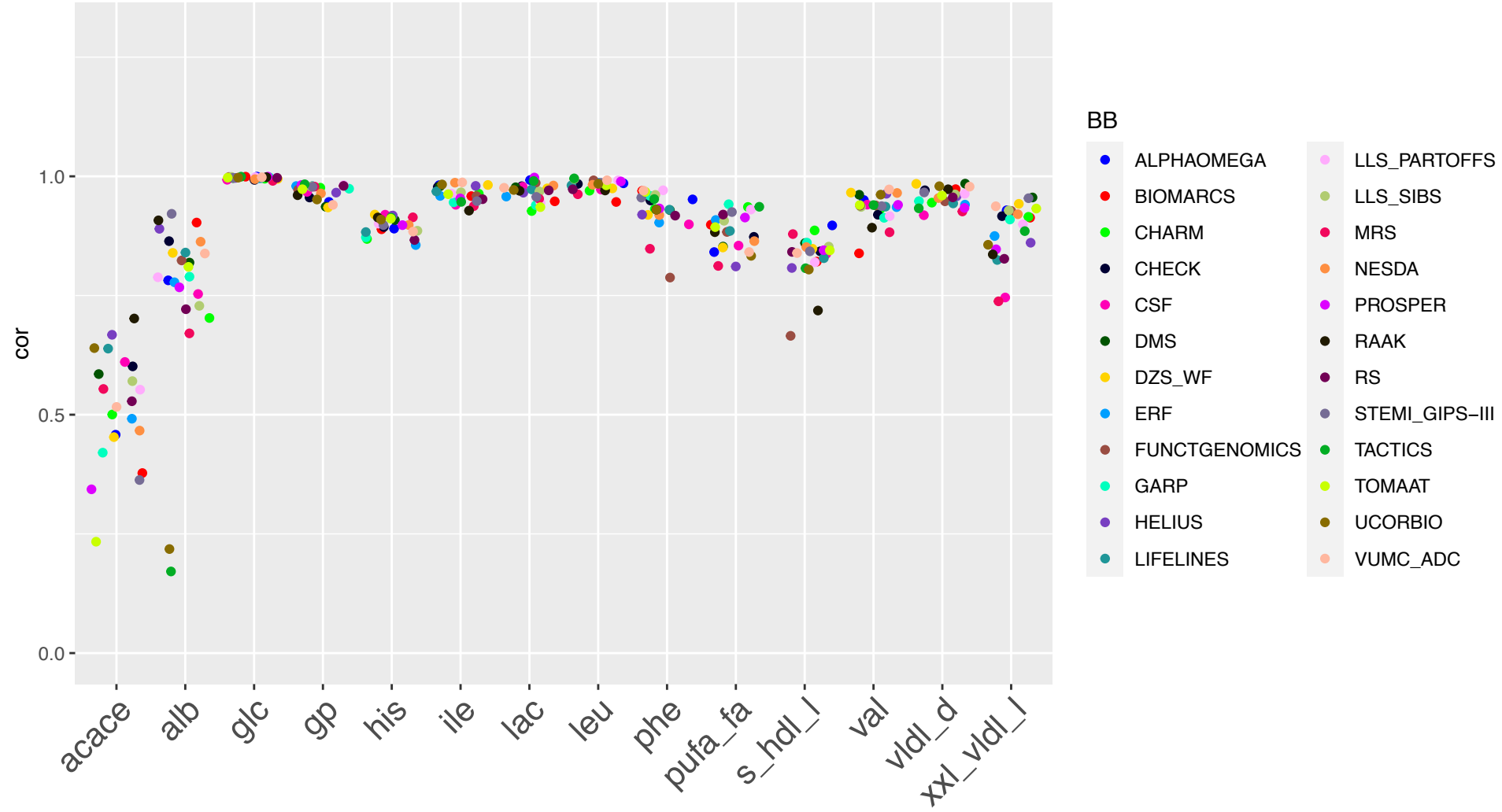
Mortality score

Overview comparison over data 2014 vs 2020

After QC (23,665 samples)



Correlations of the 14 mortality score metabolites vary across cohorts



Inconsistent metabolites: acace, alb, s_hdl_l and xxl_vdl_l

MetaboAge
Can’t be used on data 2020

Discontinued

name	description
HDL2-C	HDL2 cholesterol
HDL3-C	HDL3 cholesterol

Comparable biology

name	description
XL-HDL-C	Cholesterol in very large HDL
L-HDL-C	Cholesterol in large HDL
M-HDL-C	Cholesterol in medium HDL
S-HDL-C	Cholesterol in small HDL

But different signal!

Conclusions

1) Consistent poor correlations in 4 metabolites: acace, ldl_d, s_hdl_c, sm
2) Correlations with data 2020 in 2016 >> 2014
3) Re-calibration does not seem to improve the correlation with clinical variables
4) The mortality score is still working, but validation is warranted
5) We cannot use MetaboAge in 2020

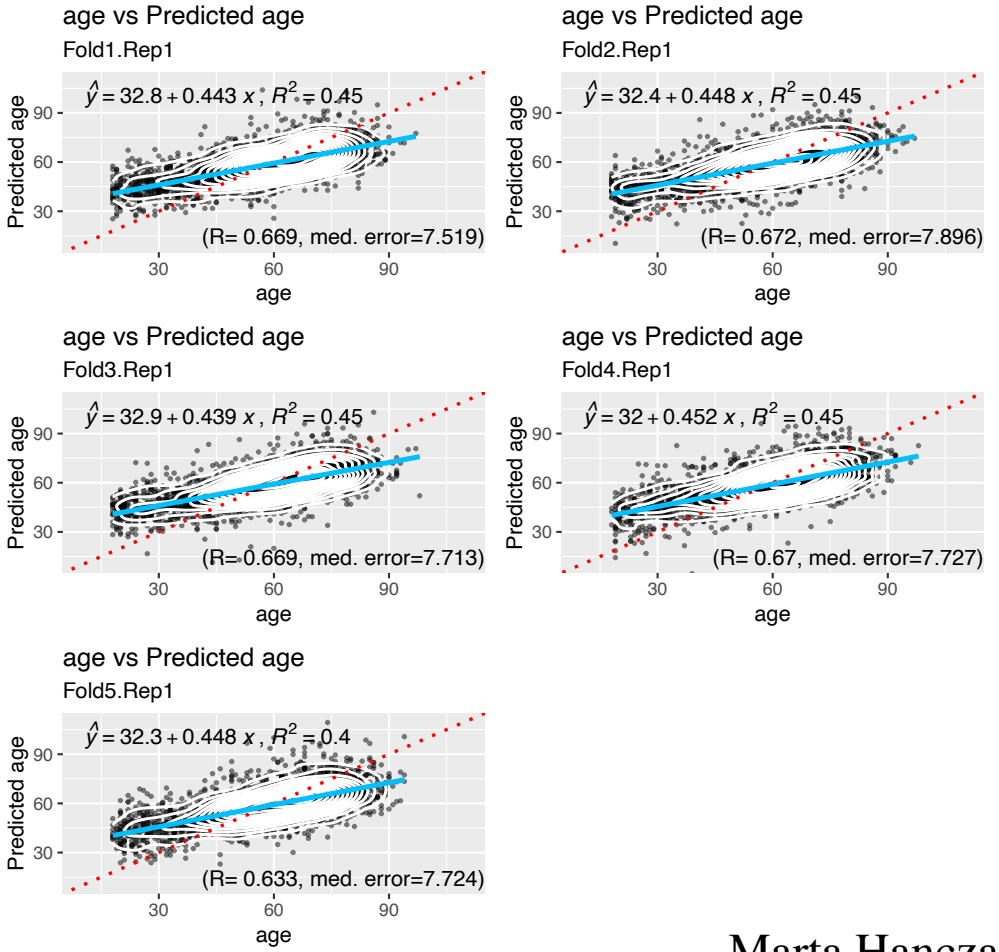
FUTURE WORK: Associations with mortality score

We are going to look if associations with the mortality score still holds in data of 2020!

MetaboAge

We are re-training it (preliminary results)

Linear model scatterplots MET65, mean Rsq= 0.437

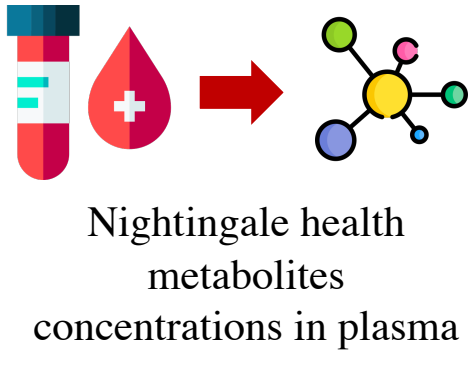


We are training a replacement.

Marta Hanczar ->
Will work on a better version!
Using: stable metabolites, clinically validated, etc.

MiMIR
(Metabolomics-based Models for Imputing Risks)

https://github.com/DanieleBizzarri/MiMIR



MetaboAge [1]

Mortality score [2]

19 surrogate
risk scores [3]

COVID-19 severity [4]

… [5,6]

[1] van den Akker Erik B. *et al, Circ. Genomic Precis. Med.*, vol. 0, no. 0, doi: 10.1161/CIRCGEN.119.002610.
[2] J. Deelen *et al.*, *Nat. Commun.*, vol. 10, no. 1, pp. 1–8, Aug. 2019, doi: 10.1038/s41467-019-11311-9.
[3] D. Bizzarri et al., EBioMedicine, **75**, Jan 2022, doi: 10.1014/j.ebiom.2021.103764.
[4] Nightingale Health UK Biobank Initiative et al., *eLife*, vol. 10, p. e63033, May 2021, doi: 10.7554/eLife.63033.
[5]A. V. Ahola-Olli *et al.*, *Diabetologia*, vol. 62, no. 12, pp. 2298–2309, 2019, doi: 10.1007/s00125-019-05001-w.
[6] P. Würtz *et al.*, *Circulation*, vol. 131, no. 9, pp. 774–785, Mar. 2015, doi: 10.1161/CIRCULATIONAHA.114.013116.

Acknowledgements

LUMC: Molecular Epidemiology

Eline Slagboom

Marian Beekman

Fatih Bogaards

LUMC: Leiden Computational Biology Center and

TU-Delft: The Delft Bioinformatics Lab

Marcel Reinders

Erik Van Den Akker

Jeppe Severens

Onur Karakalar

LUMC: Sequencing Analysis Support Core

Hailiang (Leon) Mei

Davy Cats

BBMRI-nl

Thank you for your attention!

Questions?

Daniele Bizzarri

d.bizzarri@lumc.nl