NMR-based Quality Control for Biobanks

Spraul M, Cannet C., Fang F., Leiminger.R and Schäfer H. Bruker BioSpin GmbH, Rheinstetten, Germany

Introduction

Quality control, standardization and coverage of relevant metadata are major issues for successful biobanking. Keeping in mind the rapidly increasing number of biobanks worldwide, the solution to solve such issues becomes a high priority. NMR is a technology that has rapidly grown into one of the 2 major tools in mixture analysis. It benefits from its outstanding reproducibility and transferability. Such spectra generated by different groups worldwide, working under standardized NMR conditions, can generate spectra that can go into common statistical analysis. This is a huge advantage for large clinical studies or epidemiology. The NMR (Nuclear Magnetic Resonance) based IVDr platform (In vitro Diagnostic Research) has been developed at 600 MHz with completely standardized hardware and standard operation procedures for the most common body fluids, and is already widely used in clinical and translational research. It also forms the basis of the International Phenome Center Network (IPCN, https://phenomenetwork.org) NMR-based investigations. The intrinsic reproducibility and transferability of NMR is demonstrated in figures 1 and 2.

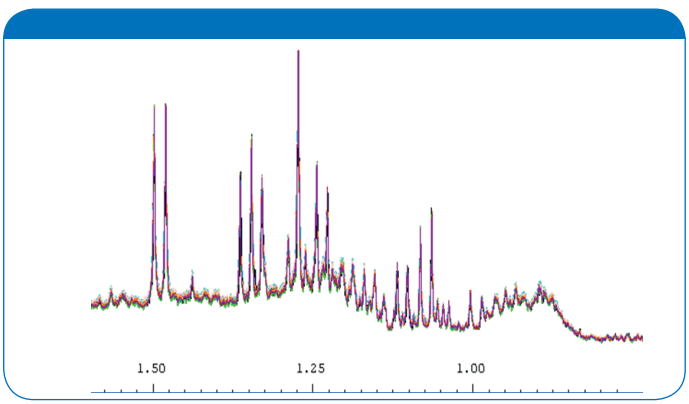


Figure 1: Reproducibility and transferability in urine Left part shows the overlay of 30 spectra generated from 30 aliquots of one urine sample, prepared by 6 people (5 aliquots each) and analyzed on 3 identical NMR platforms

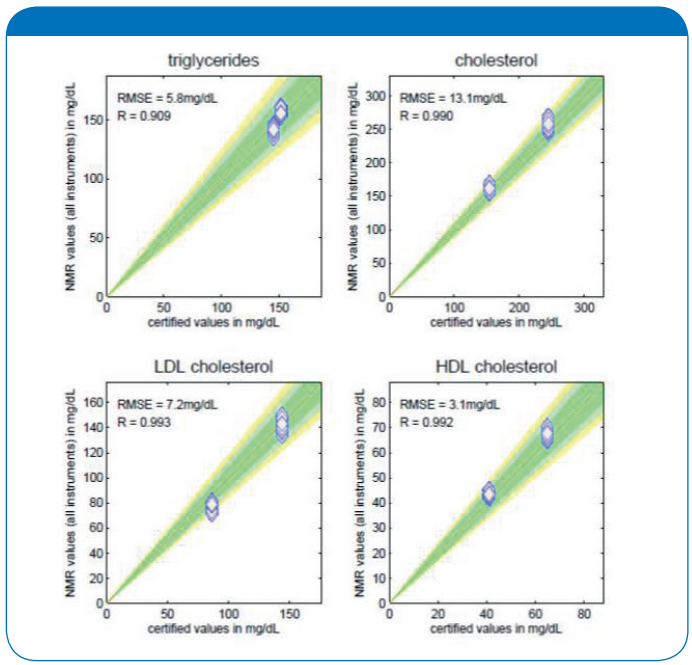


Figure 2: Reproducibility and transferability in plasma/serum Result of an 11 instrument ringtest on 2 NIST standard serum samples (1951c). In green, the error allowed by the NCEP (national cardiovascular education program) is shown, 11 results of each sample are shown as blue squares

NMR in Biobanks

1. Advantages beyond reproducibility and transferability

- Minimum sample preparation (buffer addition and mixing)
- Complete push button operation, including report generation in high throughput mode
- Robust and minimum maintenance (sample does not get in contact with detection)
- Detection system with high dynamic range (> 2*105)
- Delivers large number of parameters from each sample
- Targeted and non-targeted analysis in one experiment
- Retrospective analysis (re-analyze spectra with quantification or statistical models)

2. Quality control before input of samples into biobanks

Identifying quality issues can reduce cost or allow a warning remark on input to avoid wrong interpretation or outliers in future analysis. Such statistical outliers in clinical trials or epidemiological studies can be explained based on quality issues.

- Preanalytic conditions (under development: time at room temperature, freezing cycles)
- Control of status description (e.g.fasting, declared drugs, alcohol, food…)
- Unreported drugs
- Accidental mislabeling/exchange of samples
- Differentiate serum/plasma
- Validate type of plasma (e.g. EDTA, Citrate)
- Dilution of plasma/serum
- Contaminations (e.g. disinfectants, cleaning agents, contrast agents, skin creams, ..)

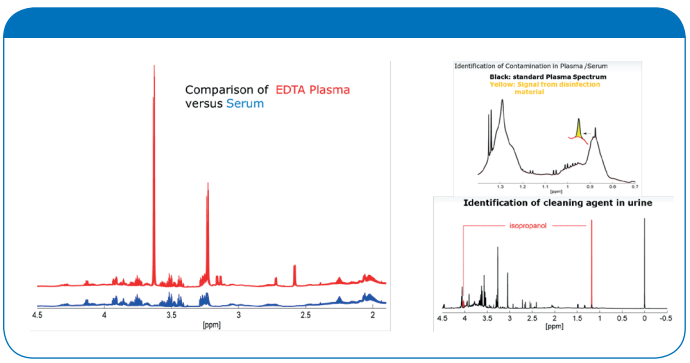


Figure 3: Visualization of 3 NMR based QC functions in plasma/serum and urine

3. Added biobank value of NMR beyond QC

After running NMR-based QC, spectra are available for storage into the biobank. Once spectra are generated under the platform SOPs, the analysis tools for urine and plasma/serum can be executed and additional information can be entered into the biobank sample description. Such information can be enriched retrospectively, when updated or new analysis tools become available.

- Added value by NMR generating additional information for each sample:
- Quantification of 150 metabolites and disease markers in newborn urine
- Classification against healthy newborn urine model
- Quantification of 150 metabolites/disease-markers in urine of children and adults
- Quantification of 115 lipoprotein parameters including subclasses
- Quantification of 20 small molecules in plasma/serum

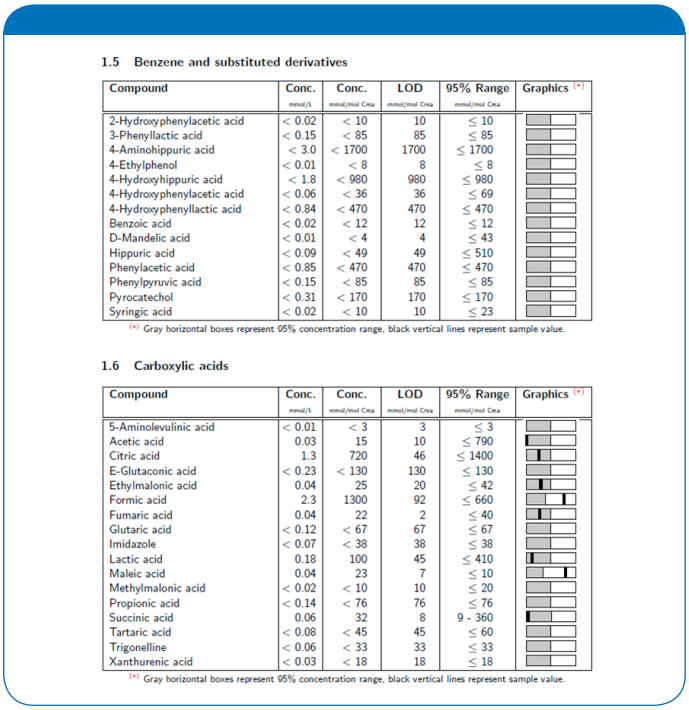
Figures 4 and 5 show excerpts of urine and plasma/serum quantification, for all parameters generated reference ranges are given and the actual sample is shown by a black bar. In the newborn samples univariate and multivariate classification is performed and deviating spectral regions are indicated.

4 Amino acids and derivatives

Compound	Conc. mmol/L	Conc. mmol/mol Crea	LOD mmol/mol Crea	95% Range mmol/mol Crea	Graphics (*)
1-Methylhistidine	< 0.01	< 15	15	≤ 15	
2-Furoylglycine	< 0.04	< 39	39	≤ 40	
3-Aminoisobutyric acid	< 0.08	< 85	85	≤ 85	
3-Methylcrotonylglycine	< 0.01	< 8	8	≤ 8	
4-Aminobutyric acid	< 0.02	< 20	20	≤ 20	
5-Aminopentanoic acid	< 0.09	< 94	94	≤ 94	
Alanine	0.24	250	10	11 - 72	
Arginine	< 0.72	< 750	750	≤ 750	
Argininosuccinic acid	< 0.03	< 29	29	≤ 29	
Betaine	0.44	460	110	≤ 110	
Citrulline	< 0.66	< 690	690	≤ 690	
Creatine	0.46	480	50	≤ 280	
Cystine	< 0.01	< 0	0	≤ 1	
DL-Alloisoleucine	< 0.05	< 48	48	≤ 48	
DL-Tyrosine	< 0.04	< 44	44	≤ 44	
Glutamic acid	< 0.44	< 460	460	≤ 460	
Glutamine	< 0.42	< 440	440	≤ 440	
Glycine	0.72	750	34	38 - 440	
Guanidinoacetic acid	2.2	2300	100	≤ 140	
Isobutyrylglycine	< 0.01	< 7	7	≤ 7	
L-Carnosine	< 0.12	< 130	130	≤ 130	
L-Homocystine	< 0.88	< 910	910	≤ 910	
L-Isoleucine	< 0.01	< 16	16	≤ 16	
L-Pyroglutamic acid	0.11	120	32	≤ 67	
L-Tryptophan	< 0.09	< 97	97	≤ 97	
Leucine	0.03	28	22	≤ 22	
Methionine	< 0.05	< 54	54	≤ 54	
N,N-Dimethylglycine	0.06	64	5	≤ 15	
N-Acetylaspartic acid	< 0.10	< 99	99	≤ 99	
N-Acetylglutamate	< 0.04	< 42	42	≤ 42	
N-Acetylphenylalanine	< 0.13	< 130	130	≤ 130	
N-Acetyltyrosine	< 0.36	< 380	380	≤ 380	
N-Isovaleroylglycine	0.00	4	2	≤ 5	
Phenylalanine	< 0.19	< 200	200	≤ 200	
Proline betaine	0.03	30	20	≤ 280	
Propionylglycine	< 0.01	< 12	12	≤ 12	
Sarcosine	0.01	8	5	≤ 7	
Taurine	< 0.14	< 140	140	≤ 170	
Tiglylglycine	< 0.02	< 19	19	≤ 19	
Valine	0.03	27	13	≤ 13	

(*) Gray horizontal boxes represent 95% concentration range, black vertical lines represent sample value.

Figure 4: Excerpts of a neonate urine NMR report under full automation



1.5 Benzene and substituted derivatives

Compound	Conc. mmol/L	Conc. mmol/mol Crea	LOD mmol/mol Crea	95% Range mmol/mol Crea	Graphics (*)
2-Hydroxyphenylacetic acid	< 0.02	< 10	10	≤ 10	
3-Phenyllactic acid	< 0.15	< 85	85	≤ 85	
4-Aminohippuric acid	< 3.0	< 1700	1700	≤ 1700	
4-Ethylphenol	< 0.01	< 8	8	≤ 8	
4-Hydroxyhippuric acid	< 1.8	< 980	980	≤ 980	
4-Hydroxyphenylacetic acid	< 0.06	< 36	36	≤ 69	
4-Hydroxyphenyllactic acid	< 0.84	< 470	470	≤ 470	
Benzoic acid	< 0.02	< 12	12	≤ 12	
D-Mandelic acid	< 0.01	< 4	4	≤ 43	
Hippuric acid	< 0.09	< 49	49	≤ 510	
Phenylacetic acid	< 0.85	< 470	470	≤ 470	
Phenylpyruvic acid	< 0.15	< 85	85	≤ 85	
Pyrocatechol	< 0.31	< 170	170	≤ 170	
Syringic acid	< 0.02	< 10	10	≤ 23	

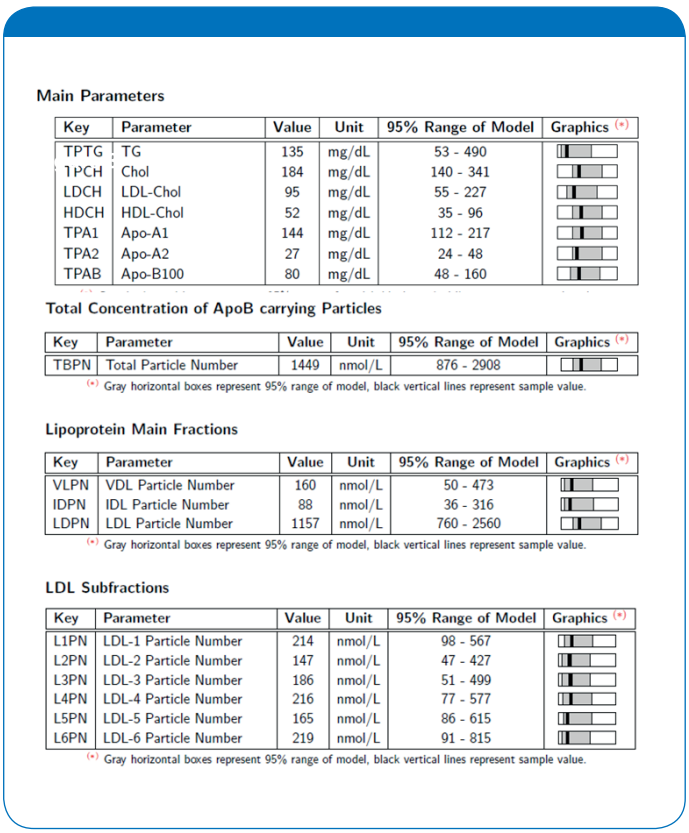
(*) Gray horizontal boxes represent 95% concentration range, black vertical lines represent sample value.

1.6 Carboxylic acids

Compound	Conc. mmol/L	Conc. mmol/mol Crea	LOD mmol/mol Crea	95% Range mmol/mol Crea	Graphics (*)
5-Aminolevulinic acid	< 0.01	< 3	3	≤ 3	
Acetic acid	0.03	15	10	≤ 790	
Citric acid	1.3	720	46	≤ 1400	
E-Glutaconic acid	< 0.23	< 130	130	≤ 130	
Ethylmalonic acid	0.04	25	20	≤ 42	
Formic acid	2.3	1300	92	≤ 660	
Fumaric acid	0.04	22	2	≤ 40	
Glutaric acid	< 0.12	< 67	67	≤ 67	
Imidazole	< 0.07	< 38	38	≤ 38	
Lactic acid	0.18	100	45	≤ 410	
Maleic acid	0.04	23	7	≤ 10	
Methylmalonic acid	< 0.02	< 10	10	≤ 20	
Propionic acid	< 0.14	< 76	76	≤ 76	
Succinic acid	0.06	32	8	9 - 360	
Tartaric acid	< 0.08	< 45	45	≤ 60	
Trigonelline	< 0.06	< 33	33	≤ 33	
Xanthurenic acid	< 0.03	< 18	18	≤ 18	

(*) Gray horizontal boxes represent 95% concentration range, black vertical lines represent sample value.

Figure 5: Excerpts of a neonate urine NMR report under full automation



Main Parameters

Key	Parameter	Value	Unit	95% Range of Model	Graphics (*)
TPTG	TG	135	mg/dL	53 - 490	
TPCH	Chol	184	mg/dL	140 - 341	
LDCH	LDL-Chol	95	mg/dL	55 - 227	
HDCH	HDL-Chol	52	mg/dL	35 - 96	
TPA1	Apo-A1	144	mg/dL	112 - 217	
TPA2	Apo-A2	27	mg/dL	24 - 48	
TPAB	Apo-B100	80	mg/dL	48 - 160	

Total Concentration of ApoB carrying Particles

Key	Parameter	Value	Unit	95% Range of Model	Graphics (*)
TBPN	Total Particle Number	1449	nmol/L	876 - 2908	

(*) Gray horizontal boxes represent 95% range of model, black vertical lines represent sample value.

Lipoprotein Main Fractions

Key	Parameter	Value	Unit	95% Range of Model	Graphics (*)
VLPN	VDL Particle Number	160	nmol/L	50 - 473	
IDPN	IDL Particle Number	88	nmol/L	36 - 316	
LDPN	LDL Particle Number	1157	nmol/L	760 - 2560	

(*) Gray horizontal boxes represent 95% range of model, black vertical lines represent sample value.

LDL Subfractions

Key	Parameter	Value	Unit	95% Range of Model	Graphics (*)
L1PN	LDL-1 Particle Number	214	nmol/L	98 - 567	
L2PN	LDL-2 Particle Number	147	nmol/L	47 - 427	
L3PN	LDL-3 Particle Number	186	nmol/L	51 - 499	
L4PN	LDL-4 Particle Number	216	nmol/L	77 - 577	
L5PN	LDL-5 Particle Number	165	nmol/L	86 - 615	
L6PN	LDL-6 Particle Number	219	nmol/L	91 - 815	

(*) Gray horizontal boxes represent 95% range of model, black vertical lines represent sample value.

Figure 6: Parts of a lipoprotein subclass analysis report from plasma or serum

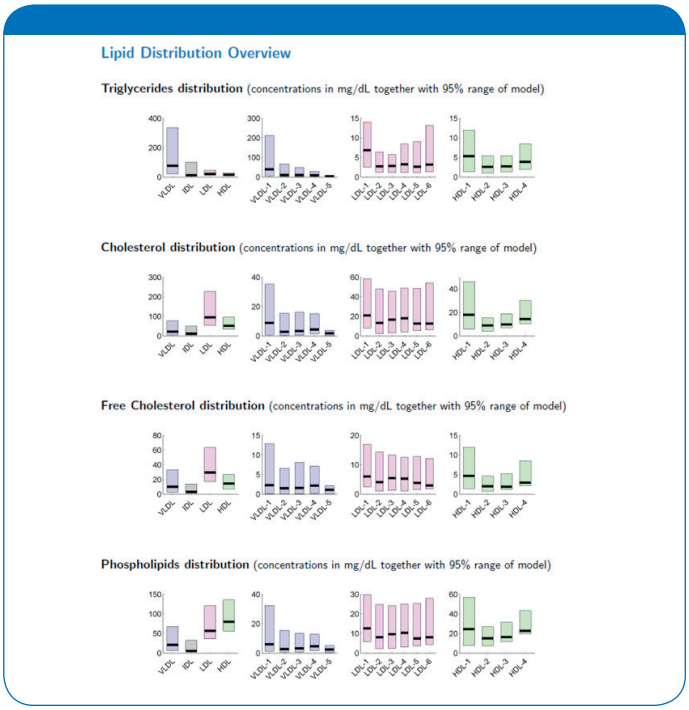


Figure 7: Parts of a lipoprotein subclass analysis report from plasma or serum

4. NMR spectra instead of new aliquots for clinical trials or research

Having multiple biobanks offering standardized NMR spectra from their aliquots, it is possible to reduce the number of new samples to be collected for clinical trials and clinical/ translational research as well as epidemiology studies. This allows worldwide integration of spectral data into statistical analysis to develop new assays or run clinical trials on new drug candidates. Any assay such can such be tested on their specificity and sensitivity against worldwide samples. Benefits of this approach are

- Reduce cost and time of clinical trials and research projects
- Less new samples, making use of standardized NMR spectra in biobanks
- Metadata selected spectrum collection to fit trials or research projects
- Integrated use of spectra from biobanks worldwide
- Use of analysis results (e.g. quantification) stored with NMR spectra

Conclusion

NMR analysis of body fluids can be used efficiently for biobanks to enhance quality control in a fully automatic process, testing multiple criteria out of one screening experiment. At the same time the offering of biobanks can be expanded to NMR spectra (produced during QC) and analysis results thereof with hundreds of parameters. An NMR platform under complete standardization allows to integrate spectra and results from different biobanks worldwide (under identical SOPs) to reduce the number of samples needed for research or clinical trials, saving cost and time.

Bruker BioSpin

info@bruker.com
www.bruker.com