Proposed Framework to Evaluate the Quality and Reliability of Targeted Metabolomics Assays – a position statement and working proposal from the UK Consortium on Metabolic Phenotyping (MAP/UK)

Sarir Sarmad[1], Mark Viant[2], Warwick Dunn[3,4], Royston Goodacre[3], Ian Wilson[5], Katie Chapell[6], Jules Griffin[7], Valerie O'Donnell[8], Brendon Naicker[8], Matthew Lewis[6], Toru Suzuki[1,9]; on behalf of the UK Consortium on Metabolic Phenotyping (MAP/UK)

[1] Department of Cardiovascular Sciences, University of Leicester and NIHR Leicester Biomedical Research Centre, Leicester, UK.
[2] Phenome Centre Birmingham, University of Birmingham, Birmingham, UK
[3] Department of Biochemistry and Systems Biology, Institute of Systems, Molecular, and Integrative Biology, University of Liverpool, UK.
[4] Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom.
[5] Computational & Systems Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College, London, UK.
[6] The National Phenome Centre, Imperial College, London, UK.
[7] School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK.
[8] Systems Immunity Research Institute, School of Medicine, Cardiff University, Cardiff, UK.
[9] The Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

*** Address for Correspondence:**

Prof Toru Suzuki
University of Leicester and NIHR Leicester Cardiovascular Biomedical Research Centre
Glenfield Hospital, Leicester, LE3 9QP, UK
Email: ts263@le.ac.uk
Tel: (0044) 116 204 4741

27 **Abstract**

28 Targeted metabolite assays that measure tens or hundreds of pre-selected metabolites,
29 typically using liquid chromatography mass spectrometry (LC-MS), are increasingly being
30 developed and applied to metabolic phenotyping studies. These are used both as standalone
31 phenotyping methods and for the validation of putative metabolic biomarkers obtained from
32 untargeted metabolomics studies. However, there are no widely accepted standards in the
33 scientific community for ensuring reliability of the development and validation of targeted
34 metabolite assays (i.e. what we refer to here as targeted metabolomics). Most current practice
35 attempts to adopt, with modification, the strict guidance provided by drug regulatory
36 authorities for analytical methods designed largely for measuring drugs and other xenobiotic
37 analytes.

38 Here, the regulatory guidance provided by the European Medicines Agency, U.S.
39 Food and Drug Administration, and International Council for Harmonisation of Technical
40 Requirements for Pharmaceuticals for Human Use are summarised. A less onerous 'tiered'
41 approach to evaluate the reliability of a wide range of metabolomics analyses is proposed,
42 addressing the need for community-accepted, harmonised guidelines for tiers other than full
43 validation. This 'fit-for-purpose' tiered approach comprises 4 levels – discovery, screening,
44 qualification and validation – and is discussed in the context of a range of targeted and
45 untargeted metabolomics assays. Issues arising with targeted multiplexed metabolomics
46 assays, and how these might be addressed, are considered. Furthermore, guidance is provided
47 to assist the community with selecting the appropriate tier of reliability for a series of well-
48 defined applications of metabolomics.

49 **Keywords:** Metabolic phenotyping, metabolomics, LC-MS, multiplexed assays, validation,
50 qualification, screening, discovery, regulatory, tiered framework.

Introduction

Metabolomics – or metabolic phenotyping - is a multidisciplinary field of research that investigates the metabolome, the terminal downstream products of the genome consisting of a repertoire of low molecular weight biomolecules involved in cellular metabolism and other biochemical processes (i.e. metabolites) in cells, tissues and bodily fluids [1,2]. Metabolomics facilitates the characterization of a system from genomic to metabol(om)ic activity and its interaction with its environment, and reveals dynamic insight into multiple metabolic pathways and networks that are the consequences of cellular activity, to understand molecular pathophysiology [3]. In addition, metabolomics aims to identify biomolecules (metabolite biomarkers) that modulate phenotype in physiological and/or disease status, reflective of biological processes as well as dysregulated pathways [4,5,6]. The analytical approaches applied in metabolomics research are generally categorised as either untargeted, targeted, or a hybrid approach (otherwise defined as semi-targeted approach) that combines some aspects of both types of analyses [7].

The techniques that are most widely used for untargeted analysis include liquid chromatography high-resolution mass spectrometry (LC-MS), gas chromatography mass spectrometry (GC-MS) and ^1H nuclear magnetic resonance (NMR) spectroscopy, while liquid chromatography-triple quad-tandem mass spectrometry (LC-MS/MS) remains one of the traditional techniques for targeted analysis of limited numbers of analytes, the other being GC-MS due to the fragmentation of the metabolite during electron ionisation [8,9]. Untargeted metabolomics is a discovery-based approach where the objective is to analyse as many detectable metabolites without biological bias, including unknowns, to determine which, if any, are significantly perturbed in the diseased phenotype, followed by post-hoc identification of those putative metabolic biomarkers [10]. The major disadvantage of untargeted approaches is that relative responses and not actual concentrations are reported.

Targeted approaches on the other hand, involve the (multiplexed) analysis of known metabolites, and such methods often focus on a subset of metabolites representative of key pathways, or of metabolites determined to be important from prior untargeted metabolomics [11]. The major disadvantage of targeted approaches is their limited coverage of the metabolome [12].

Advances in metabolomics have led to new clinical and toxicological diagnostic biomarkers [13, 14, 15], which can contribute to stratified medicine and safety assessment of drugs [16, 17]. Metabolomics is also central to the screening of inborn errors of metabolism [18].

However, there are several challenges in the translation of metabolomics research to clinical and toxicological applications under regulatory control. Issues include analytical reproducibility, accuracy, precision, metabolite identification/quantification, study design, sample handling, lack of harmonised reporting frameworks for published data and metadata, insufficient open-access data to enable data-mining by other researchers [19], lack of harmonisation in bio-banking, batch-to-batch variation, and between-methods bias [20]. Assessing the reliability of bioanalytical methods for metabolomics is challenging when compared to validation of other types of bioanalytical methods. Data from the metabolomics field are variable, and heterogeneity among data formats, data analysis pipelines, algorithms and applied statistical methods should be addressed. There is a need to define the scope and extent of assessing the reliability of these methods, and how the standards applied and methods for reporting should be set in order to ensure appropriate data quality for use in regulatory processes [21]. To eliminate some of these problems, communication between the research and regulated clinical and toxicological communities needs to be more fully developed, and the establishment of a system to assess and cross-correlate metabolic profiles obtained by different laboratories and instruments is needed [19]. The new Metabolomics Reporting Framework for regulatory toxicology, developed by multiple stakeholders from research laboratories, industry

101 and government regulatory agencies and coordinated by the Organisation for Economic Co-
102 operation and Development (OECD) provides evidence on how progress can be made to
103 achieve harmonised reporting of methods, data, metadata and findings, and thereby advance
104 the application of metabolomics within regulatory settings [22]. There are also a plethora of
105 publications that provide comprehensive guidelines for assessing the quality of untargeted
106 metabolomics assays [23, 24, 25, 26, 27]. Whilst these guidelines provide the foundation for
107 metabolomics system suitability and quality assurance/quality control (QA/QC) proficiency, a
108 community-initiated approach towards harmonised guidelines that ultimately achieve
109 acceptance via their consensus use for evaluating the reliability of targeted metabolomics
110 within research, clinical and toxicological settings is still required.

111 Our scientific collaboration, the UK Consortium on Metabolic Phenotyping (MAP/UK,
112 https://mapuk.org), is a partnership of eight specialised research laboratories and two Phenome
113 Centres, which has been funded by the Medical Research Council to improve UK-wide
114 metabolic phenotyping expertise and capabilities. The MAP/UK partnership brings together a
115 critical mass of methodological, analytical, and computational platforms to develop, optimise,
116 transfer, harmonise, and validate efficient, high-quality metabolomics research and training
117 methods, specifically tailored to the growing need for biomedical studies that require robust
118 metabolic phenotyping. The overall aim of the MAP/UK partnership is to investigate new
119 biomarkers within metabolic signatures of disease, novel targeted quantitative metabolomic
120 and hybrid approaches, and developing untargeted metabolomics to meet gaps in molecular
121 coverage of key disease-related pathways, alongside a variety of other factors, including
122 dietary, lifestyle/environmental, gut microbial and genetics. As a collective of scientists with
123 the aim of harmonisation of metabolic phenotyping, existing regulatory guidelines have been
124 reviewed to extract commonalities from these guidelines that can be adopted to 'fit-for-
125 purpose' and tiered approaches for untargeted and targeted metabolomics.

126 The aim of this manuscript is to propose harmonised guidelines for evaluating the
127 reliability of targeted (multiplexed) mass spectrometry-based metabolomics assays taking into
128 consideration intra-laboratory precision, accuracy, reproducibility, and cross-laboratory
129 harmonisation of methods and data acquired on different instrumental platforms. First, existing
130 guidelines for bioanalytical method validation, including an existing 4-tiered framework
131 applied in drug discovery, are reviewed. Then, after introducing the applications of clinical and
132 toxicological metabolomics in regulatory settings, a new 'fit-for-purpose' 4-tiered (discovery,
133 screening, qualification and validation) framework for assessing analytical reliability that is
134 suitable for targeted and hybrid untargeted metabolomics assays is proposed.

135 In addition, a checklist on the bioanalytical process has been provided to facilitate better
136 understanding and emphasising the importance of harmonisation at each step.

137 **Checklist for bioanalytical assay process:**

138 **1- Pre-analytical:**

139 • Hypothesis/study design/ sample size

140 • Data acquisition of demographics for groups/individuals including clinical, diet,
141 medications and life-style data

142 • Sample type (plasma/serum/urine/feces), collection method, preservation, and
143 timing

144 • Sample storage

145 **2- Analytical:**

146 • Sample preparation and purification

147 • Authentic reference materials (external standards), quality control (QC) samples
148 and suitable internal standards

149 • Maintaining assay reliability and quality by selecting the right tier based on
150 number of metabolites and assay purpose (consult Table I).
151 • Select validation parameters and acceptance criteria for targeted assays (tier 1
152 and 2), by consulting Table II. Note that Tier 1 parameters are the same as
153 suggested by regulatory guidelines (FDA/EMA/ICH2019) for validation, and
154 Tier 2 (qualification) has a wider range of acceptance criteria.
155 • Select appropriate instrumentation such as liquid chromatography high-
156 resolution mass spectrometry (LC-MS), liquid chromatography-triple quad-
157 tandem mass spectrometry (LC-MS/MS), and considerations regarding
158 instrument calibration, settings, analytical batches, and quality assurance
159 (QA)/performance.

160 **The concept of regulatory bioanalytical validation**

161 Validation is defined as a process that provides proof of assay integrity within given
162 specifications with the parameters of an assay used for quantification being statistically reliable
163 between assays over time. Multiple guidelines exist that describe the regulation of bioanalytical
164 assays such as those from the U.S. Food and Drug Administration (FDA) [28], the European
165 Medicines Agency (EMA) [29], the International Council for Harmonisation of Technical
166 Requirements for Pharmaceuticals for Human Use (ICH) [30], Japanese Ministry of Health,
167 Labour and Welfare (MHLW) [31], Chinese (State) Food and Drug Administration (CFDA,
168 currently the National Medical Products Administration, NMPA) [32], Australian Therapeutic
169 Goods Administration (TGA) [33], and Brazilian National Health Surveillance Agency (Anvisa)
170 [34, 35, 36]. The regional differences along with differences in terminology, parameters and
171 acceptance criteria can cause confusion amongst bioanalysts and/or pharmaceutical companies
172 given the globalisation of the pharmaceutical sector.

173 Whilst these regulatory guidelines are comprehensive, they are largely developed for
174 the measurement of drugs and other xenobiotic analytes. Endogenous biomarkers are often
175 measured in metabolomics which requires different considerations of matrix use. For
176 example, with endogenous metabolites, the issue of evaluation of (L) LOQ (lower limit of
177 quantification) encountered due to matrix effect. Evaluating these limits using standard
178 solutions in neat solvent, and/or matrix deprived of specific classes of metabolites (such as
179 stripped plasma) are not an ideal solution as what has been depleted is not defined.
180 Furthermore, measurement of specificity/selectivity for endogenous metabolites is much
181 more challenging due to presence of multiple isoforms.

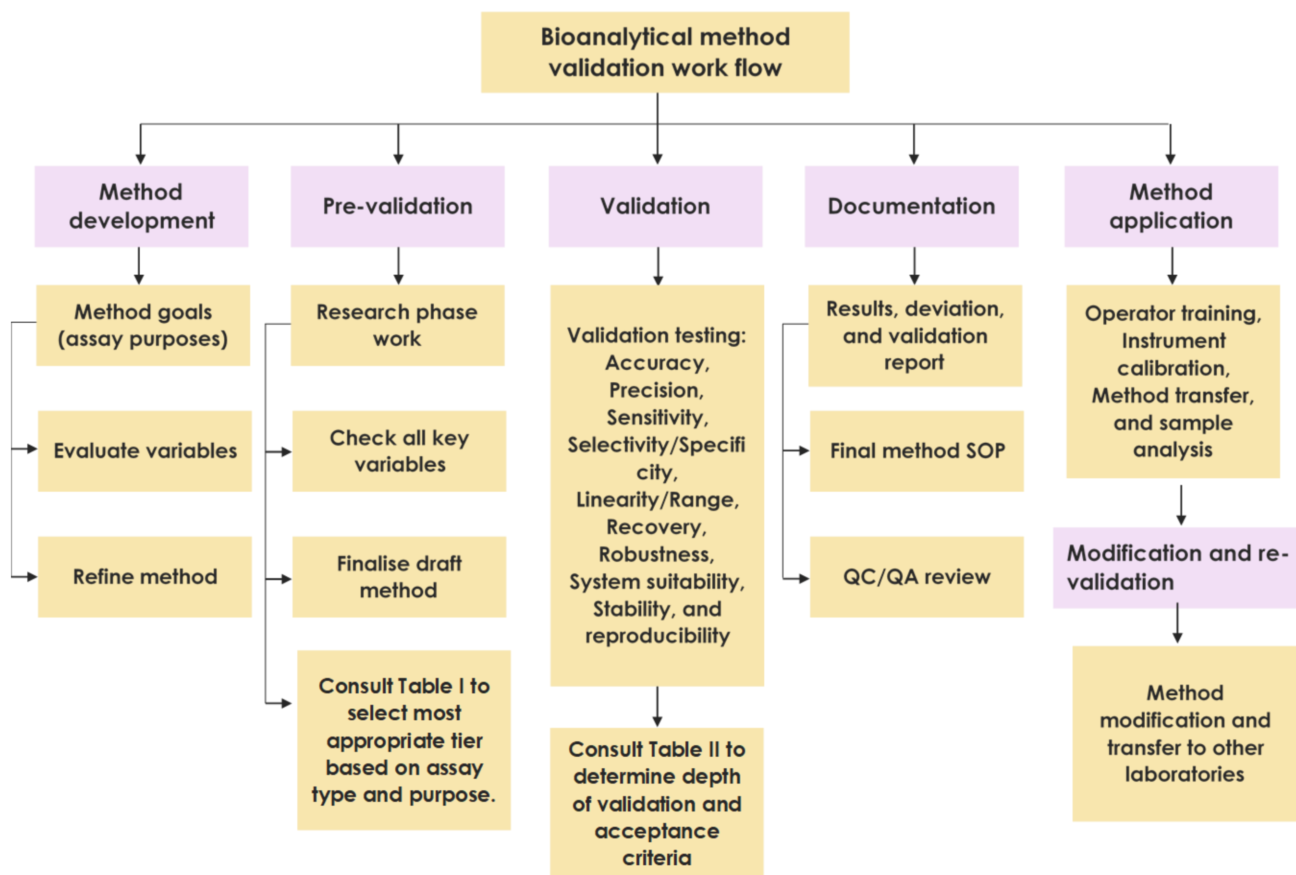
182 The two most practised bioanalytical guidelines from the EMA and FDA are similar
183 but not identical. The scientific basis for the evaluation of parameters is the same across both
184 guidelines. However, there are also differences in terminology, recommended validation
185 parameters, acceptance criteria and methodology.

186 Standards setting and harmonisation was advanced by the ICH, which is an
187 international organisation with the mission to achieve greater harmonisation worldwide to
188 ensure that safe, effective, and high-quality medicines are developed and registered in the
189 most resource-efficient manner. The ICH consolidated best practices from the FDA and EMA
190 guidelines in 2019 into a harmonised M10 bioanalytical method validation draft in order to
191 clarify any areas of uncertainty between the two guidelines. A comparison between the FDA
192 and EMA guidelines and the consolidated ICH M10 draft guideline are summarised in Table
193 1a (to be deposited on Figshare as supplemental material to this paper).

194 An analytical assay starts with a definition of its purpose (i.e. intended application),
195 and to define what is 'fit-for-purpose' and is then followed by method development and
196 optimisation, then subsequently by assay validation (dependent upon the tier, as introduced
197 above) and documentation before it can finally be applied for the intended purpose. Prior to

198 initiating a validation study, a well-planned validation protocol should be written and
199 reviewed for scientific soundness and completeness. The protocol should describe the
200 procedure in detail and should include pre-defined acceptance criteria and pre-defined
201 statistical methods, and should be approved by all participants in the analytical pipeline.

202 There are numerous validation parameters (accuracy, precision, calibration curve,
203 lower limit of quantitation, selectivity/specificity, carryover, analyte stability, recovery,
204 dilution integrity, system suitability test, matrix effect/factor, parallelism, incurred sample
205 re- analysis, quality control, robustness/ruggedness, hook/prozone effect, and minimum
206 required dilution) to incorporate into the validation process. One should justify the required
207 level of validation to be 'fit-for-purpose' based on the differing applications of a particular
208 method. Theoretically, there are no limits to the extent of validation and verification
209 procedures. However, in practice, there are both time and economic constraints on what can
210 be achieved. Therefore, it is crucial to have optimised guidelines that are generally accepted,
211 harmonised and cost-effective [47]. The validation workflow has been summarized in a visual
212 format (Fig. 1).



213 Figure 1- Validation workflow steps and positioning of the suggested framework (Table I
214 and Table II) to select the most appropriate tier and degree of validation.

215 Before introducing our proposed framework to assist bioanalyst in selecting
216 appropriate tier of validation for a series of well-defined applications of metabolomics, we
217 give a brief introduction to the tiered approach within regulatory perspective in the next
218 section.

219 **Tiered approach within regulatory perspectives**

220 The concept of defensible scientific flexibility has been a debate within the

221 bioanalytical community in the pharmaceutical industry. The Crystal City III workshop
222 proposed the concept of 'fit-to-purpose' in 2006 as an alternative for the full validation
223 workflow already described by the FDA regulatory documents to address uncertainties in the
224 bioanalytical community on what level of data scrutiny is required to generate quality data
225 whilst optimising resources to meet study objectives with adequate level of data quality and
226 reliability [37]. Furthermore, the European Bioanalysis Forum (EBF) proposed the consolidation
227 of tiered approaches to include three levels (or tiers) of quality standards for metabolite
228 quantification for screening, qualified, and validated assays [38]. Consequently, the MHLW
229 and FDA adapted 'flexible adjustments and modifications' of their bioanalytical method
230 validation guidelines to meet the intended use of the assay, and this perspective was extended
231 to tiered approaches for metabolite quantification [39, 40, 41].

232 The Crystal City workshop VI in 2015 [42] defined a less rigorous level of validation
233 than the FDA guidelines for drug metabolite quantification at early stages of development.
234 The Global Bioanalytical Consortium (GBC) assigned Team A2 with the objective of
235 providing a framework to rationalise the level of bioanalytical methods for drug
236 characterization and proposed a clear path for implementation and use of tiered approaches
237 [39]. Furthermore, two globally recognised teams within the GBC (S1 and L1) provided
238 acceptance standards for validation methods for small and large pharmaceutical molecules,
239 respectively [43]. However, different terminologies have been used as part of the 'fit-to-
240 purpose' concept, such as tiered assays, scientific validation, qualified assays or partial
241 validation. Thus, it has been a source of confusion for academia and the
242 biotechnology/pharmaceutical industry due to lack of clear guidance [39]. More recently, these
243 alternative validation assay workflows in the bioanalytical industry have been categorised
244 into four tiered levels of method performance and evaluation based on the final purpose of
245 the derived analytical data ranging from the most to least stringent: level 1) validation,
246 intended for regulatory studies; 2) qualification; 3) research; and 4) the least stringent defined

247 as ‘screening’[39, 44, 45]. These four tiered levels are described in more detail below, and whilst
248 these concepts have been designed for drug development and submission to regulatory
249 authorities, they provide a framework that could be adapted for a range of assays used in
250 metabolomics studies.

251 - Level 1) **validated bioanalytical assays** are designed for intended pharmaceutical
252 products and thus require the highest level of confidence in analytical results as suitable
253 for regulated good laboratory practice (GLP), pre-clinical/clinical, pharmacokinetic
254 and/or toxicological studies, and identification of active metabolites in safety testing
255 (MIST). These mandate that assay precision, accuracy, selectivity, sensitivity, and
256 stability of the analytes should be determined throughout the bioanalytical measurement
257 process. FDA recommended evaluations should be performed [38].
258 - Level 2) **qualified bioanalytical assays** do not need to demonstrate that the
259 measurement methods are as robust as validated assays. This tier is suitable for non-
260 regulated studies in the drug development process, with additional assessment of tissue
261 concentrations or other matrices during preclinical or late discovery phases, and in
262 decision-making for context of use (COU) statements. Single method performance with
263 a statistically appropriate number of quality controls (QC) samples (n=5) at each level
264 and a suitable calibration range, precision and accuracy should be performed.
265 - Level 3) **research-grade bioanalytical assays** are suitable for mid- to late-discovery
266 phases of drug development projects for decision-making evaluations and/or
267 verification of additional biomarkers or metabolites for non-GLP regulated studies.
268 They use limited characterization with calibration standards prepared using a
269 comparator reference material such as an *in situ* (in solution) standard with the
270 concentration estimated by radioactivity measurement, NMR or ultraviolet (UV)
271 absorption as representative methods. The method provides semi-quantitative analyte

272 concentrations within wider accuracy and precision limits than for the two higher tiers
273 [39]. This approach enables the partial characterisation of an analytical method that may
274 eventually move to a qualified or validated assay. It should provide sufficient scientific
275 rigor to ensure that it is fit-for-purpose and that there is confidence in the data. Method
276 evaluation should be conducted prior to sample analysis, with the precision and
277 accuracy needed to achieve the more relaxed criteria of 20% relative standard deviation
278 (RSD) and 30% reduction of error (RE) at the LLOQ (Lowest Limit of Quantitation).

279 - Level 4) **screening bioanalytical assays** apply a generic method (not specific to the
280 analyte) to provide adequate results for the analyte of interest and are suitable for early
281 discovery and qualitative (present/absent) analysis. Screening assays undergo limited
282 characterization based on relative instrument analyte response where reference material
283 is not available. The assay provides relative analyte measurements (i.e. response and
284 not concentration) only but may still be suitable for decision-making processes. An
285 abbreviated set of QCs with large margins of variability of 30% RSD and 40% RE is
286 advisable. As such, screening bioanalytical assays are most similar to untargeted
287 metabolomics assays.

288 Apart from the four-tiered levels approach in the bioanalytical industry, there is a
289 general concept of ‘full’ and ‘partial’ validation. Full validation is necessary when developing
290 and implementing a bioanalytical method for the first time such as when analytes are added to
291 a panel for bioanalytical quantification. In targeted metabolomics, full validation of a method
292 by the accredited clinical laboratory is required when the result from that assay (e.g.
293 concentration of a biomarker in terms of molarity for liquids or µg/mg for tissue) is used for
294 making a clinical decision. Partial validation is required in the case of bioanalytical method
295 transfers between laboratories or the method parameters such as instrument and/or software

296 platform changes, such as changes in species within matrix (e.g. human plasma to murine
297 plasma) or within a species (e.g. human plasma to human serum/urine). Partial validation can
298 range from as little as one intra-assay accuracy and precision determination to nearly full
299 validation [46] depending on the degree of change required being undertaken.

300 The sections above have introduced concepts and terminologies within bioanalytical
301 validation as well as highlighting the need for the standardisation of guidelines for the
302 validation of endogenous metabolite analysis with the aim of maximising the cross-
303 comparability of generated data. In the next section, a flexible and practical framework to
304 assist bioanalysts to select the appropriate tier of reliability for multiplexed metabolic
305 biomarker assays, each with a defined use, is proposed.

306 **Framework for assessing the reliability of metabolomics bioanalytical methods**

307 A fundamental question is how stringently regulatory bodies view these guidelines as
308 being hard rules, or whether they could be adopted as 'fit-for-purpose' for targeted
309 metabolomics assays, and used within a 'tiered' framework. The intended use (or application)
310 of metabolomics drives which level of reliability assessment should be used, not the type of
311 assay. Selecting the most appropriate tier for measuring multiple metabolic biomarkers
312 simultaneously for targeted metabolomics assays is challenging if the intended data use is not
313 carefully defined. Hence, the first step in selecting an appropriate tier is to define the intended
314 use of the data and which type of assay is needed. and then the most appropriate reliability tier
315 can be further defined. Considering that there are a range of applications for metabolomics and
316 new advances in LC-MS techniques for multiplexed measurement of metabolites, there is a clear
317 need to propose a new framework that describes which reliability tier is most 'fit-for-purpose'
318 for different applications. Evaluation of being 'fit-for-purpose' involves questions such as: 1)
319 what is the context of use for the assay (i.e. what will the data be used for); 2) should it be a
320 quantitative, semi-quantitative or relatively quantitative assessment; and 3) what level of

321 uncertainty can be tolerated in the assessment.

322 Consolidating the concept of 'fit-for-purpose' assists bioanalysts in decision-making on 323 whether to qualify or validate a biomarker assay, and which parameters to choose in addition to 324 the number of appropriate replicates [48]. The end-result of a 'fit-for-purpose' validation of an 325 assay using relative quantification is a resource-effective and -efficient demonstration of the 326 bioanalytical method's performance that is tailored to meet the objective of the application. This 327 ultimately provides reliable study data to make important decisions. The decisions may involve 328 further assay development and progression to a fully validated method. The following 329 framework is proposed as a guideline for the metabolomics community to assess the reliability 330 of targeted metabolomics assays for different types of applications (i.e. from biomarker 331 discovery by a research laboratory, transfer of a method to a different laboratory, through to the 332 use of biomarker within a clinical setting). The proposed framework is summarised in Table I 333 (Tiers 1-4) to assist bioanalysts in selecting the most appropriate tier based on their purpose and 334 assay type. Tiers 1 and 2 (targeted metabolomics) are the main focus of this manuscript, and all 335 related parameters for safeguarding scientific rigor for robust validation and bioanalytical 336 quantification for these two tiers (termed validation and qualification) are summarised in Table 337 II. These tiers differ in depth, robustness of parameters, and the number of replicates performed 338 for each parameter (See Table II).

339 Tier 1 - Validation

340 Diagnosis of disease/toxicity phenotype using traditional targeted metabolite analysis 341 with absolute quantification of typically one to a few (less than 10) metabolites. Tier 1 validation 342 is required for compliance with regulatory agencies for clinical diagnostics. This requires an 343 authentic standard (external standard) for each metabolite. The proposed procedure is in 344 alignment with current FDA and ICH M10 bioanalytical method validation guidelines, and is 345 applicable to quantitative analytical assays such as chromatographic, liquid chromatography-

346 mass spectrometry (LC-MS and/or LC-MS/MS), and ligand binding assays (LBA) (see Table
347 II).

348 Tier 2 - Qualification

349 Diagnosis of disease/toxicity phenotype using a multiplexed targeted metabolomics assay
350 with absolute quantification of more than 10 metabolites. This requires an authentic external
351 standard for each metabolite. The criteria for qualifying a method are less strict than for tier 1
352 validation of a method (see Table II).

353 Tier 3 - Screening

354 Screening for a disease/toxicity phenotype using a multiplexed targeted or hybrid
355 metabolomics assay with relative or semi-quantification of a panel of hundreds of metabolites.
356 This does not require an authentic external standard for each metabolite. The criteria to meet in
357 a screening method are less strict than for tier 2 qualification of a method.

358 Tier 4 - Discovery

359 Discovery of putative metabolic biomarkers using untargeted or hybrid metabolomics
360 with relative quantification in a research laboratory. Untargeted methods have the least strict
361 criteria. Tiers 3 and 4 are not within the scope of this manuscript as they do not require absolute
362 quantification. Furthermore, the use of system suitability tests, intra-study QC samples,
363 phenotyping QCs (healthy vs. disease), inter-laboratory QC samples, and dilution series of
364 pooled QCs have been previously discussed [7, 49] and provide a dimension of semi-quantitative
365 nature to these untargeted assays.

366 **Table I. Four-tiered framework for assessing the reliability of metabolomics assays**

Tiers of framework to evaluate reliability	Purpose (example)	Assay type	Assay quantification
1- Validation	Diagnosis of disease/toxicity phenotype	Targeted metabolite analysis of 1 to < 10 metabolites	Absolute quantification with authentic standard(s)
2- Qualification	Diagnosis of disease/toxicity phenotype	Multiplexed targeted metabolomics analysis of > 10 metabolites	Absolute quantification with authentic standards
3- Screening	Screening for a disease/toxicity phenotype	Multiplexed targeted metabolomics analysis of panel of hundreds of metabolites	Relative or semi-quantitative; does not require an authentic standard for each metabolite
4- Discovery	Discovery of putative metabolic biomarkers	Untargeted metabolomics	Relative quantification

367 **Table II. Biomarkers - validation *vs*. qualification**

Parameters	Tier 1- Validation	Acceptance criteria	Tier 2- qualification	Acceptance criteria
Calibrators/linearity	• 5 independent calibration lines, minimum of 6 non-zero calibrators covering the range of incurred samples	• R^2 >0.98, closer to 1 is better • Setting LLOQ as lowest acceptable standard	• 3 independent calibration lines, minimum of 8 non-zero calibrators covering the range of incurred samples	• R^2 >0.98, closer to 1 is better • Setting LLOQ as lowest acceptable standard
Assay range - lower/upper limit of quantification (LLOQ/ULOQ)	• Over 6 runs	• R^2 >0.98	• Over 3 runs	• R^2 >0.98
Calibration Quality Control (QC) levels	• Prepare LLOQ, low, medium and high QCs in 5 replicates	• RSD<15%, except for LLOQ (RSD<20%)	• Prepare LLOQ, low, medium and high QCs in 5 replicates	• RSD<20%, except for LLOQ (RSD<25%)
Intra-study QC (pooled QC) levels	• After every 6 unknown samples with the minimum number of 6 per assay	• At least 67% (e.g. at least four out of six) of the QCs concentration results should be within CV<15 %	• After every 6 unknown samples with the minimum number of 6 per assay	• At least 67% (e.g. at least four out of six) of the QCs concentration results should be within CV<20 %
Precision (within-day/intra-precision)	• Over 1 Run, 5 replicates, 4 levels (LLOQ, low, medium and high)	• Should not exceed 15% of the coefficient of variation (CV% or RSD%) except for the LLOQ, where it should not exceed 20% of the CV	• Over 1 Run, 5 replicates , 3 levels (low, medium and high)	• RSD<20-25%
Precision (between-day/inter-precision)	• Over 6 runs, 5 replicates, 4 levels (LLOQ, low, medium and high)	• RSD <20%, at LLOQ RSD<25%	• Over 3 runs, 5 replicates, 3 levels (low, medium and high)	• RSD<30%
Accuracy (within-day/intra-accuracy)	• Over 1 Run, 5 replicates, 4 levels (LLOQ, low, medium and high)	• Within 15% of nominal value, except for LLOQ within 20%	• Over 1 Run, 5 replicates, 3 levels (low, medium and high)	• Within 20-25% of the nominal value
Accuracy (between-day/inter-accuracy)	• Over 6 runs, 5 replicates, 4 levels (LLOQ, low, medium and high)	• Within 20-25% of the nominal value	• Over 3 runs, 5 replicates, 3 levels (low, medium and high)	• Within 25-30% of the nominal value
Selectivity/specificity/ matrix effect	• Yes	• Absence of interfering compound accepted where the	• N/A	• N/A

Parameters	Tier 1- Validation	Acceptance criteria	Tier 2- qualification	Acceptance criteria
		response is less than 20% of LLOQ and/or less than 5% for IS		
Carry over	• Yes	• Absence of interfering compound accepted where the response is less than 20% of LLOQ and/or less than 5% for IS	• Yes	• Absence of interfering compound accepted where the response is less than 20% of LLOQ and/or less than 5% for IS
Parallelism	• Yes, depending on availability of sample with high endogenous analyte from 6 individual sources of blank matrix	• Precision between samples in a dilution series should not exceed 30%	• Perform 1 or 2 tests depending on availability of sample with high level of endogenous analyte	• The precision between samples in a dilution series should be 30%-40%
Dilutional Linearity/integrity	• Yes	• Spike blank matrix to concentration above ULOQ and dilute it down with blank matrix (5 determinations per dilution) • Accuracy: ± 15% of nominal concentrations • Precision: ± 15% CV • R^2 >0.98	• If applicable	• Spike blank matrix to concentration above ULOQ and dilute it down with blank matrix (1 determinations per dilution) • R^2 >0.98
Prozone (hook) effect	• Yes, as applicable	• The calculated concentration for each dilution should be within ±20% of the nominal concentration after correction for dilution and the precision of the final concentrations across all the dilutions should not exceed 20%	• N/A	• N/A
Stability - room temperature	• Yes	• The accuracy (% nominal) at each level should be ± 15%	• Recommended	• The accuracy (% nominal) at each level should be ± 25%
Stability - 4◦C	• Yes	• Same as above	• Recommended	• Same as above
Stability - freeze/thaw	• Yes	• Same as above	• Recommended	• Same as above
Stability - long-term (-20◦C and/or -80◦C)	• Yes	• Same as above	• N/A	• N/A

368

369 **Bioanalytical considerations for generation of quality data in targeted and untargeted**
370 **or hybrid metabolomics assays**

371 The importance of good laboratory practice at different stages (e.g. sample collection,
372 storage integrity) should be considered for bioanalysis. Sample, analyte and data integrity as
373 well as basic laboratory record keeping are essential. Implementing a laboratory information
374 management system (LIMS) is recommended. Routine calibration of laboratory instruments,
375 pipettes and balances with well-written standard operating procedures (SOPs), as well as
376 selection of suitable blank matrices, internal standards, system suitability test and intra-study
377 QCs are essential. Intra-study QCs should be placed in the analytical run in such a way that the
378 precision of the whole run is ensured by taking into account that study samples should always
379 be bracketed by QCs [7]. Phenotyping QCs (e.g. healthy *vs*. diseased) are recommended. A QC
380 is typically produced by pooling a small aliquot of all study samples, and these are analysed
381 throughout the analytical run. For untargeted metabolomics, a dilution series of the intra-study
382 QC is highly recommended to help differentiate features of biological origin from LC-MS
383 chemical background [50]. Application of isotopically-labelled standards can provide a
384 generalised measure of precision across the study. Furthermore, use of isotopically labelled
385 internal standards helps to compensate for matrix-induced ionisation effects, thereby enhancing
386 the accuracy of the assay when quantification/semi-quantification is applied [25]. Choice of
387 suitable surrogate matrices are recommended to improve sensitivity and selectivity of
388 biomarkers quantification [51, 52, 53, 54]. Blank matrices with the minimum level of endogenous
389 analyte should be used wherever possible. This approach is suitable for multianalyte assays
390 (spiked with appropriate concentration of each analyte), but matrix effects and stability
391 should be investigated for each analyte. In the absence of blank matrices or surrogate
392 matrices, standard addition approaches which take into account the native concentration of
393 the targeted analyte(s) can be used for recovery and matrix effect checks; and the use of QCs

394 or standards prepared only in solvent and/or buffer considered for accuracy and
395 repeatability/reproducibility tests represents the approach that makes the least assumptions.
396 Artificial blank matrices may be used. A solution of 4% fatty acid-free bovine serum albumin
397 (BSA) in saline buffer that represents the same concentrations of salts and electrolytes in
398 human plasma is a good example of blank matrix for human plasma (artificial surrogate
399 matrix). Normalisation strategies to correct for differences in sample amount should be
400 considered. For example, urinary creatinine is often used to adjust the concentration of
401 urinary biomarkers.

402 All targeted assays should have a clearly defined limit of detection (LOD) and limit of
403 quantitation (LOQ). A clearly discernible peak must be visible above clearly visible baseline
404 noise and should be comprised of a specified number of data points (often 6 or above is used).
405 As a general rule, LOQ of S:N (signal to noise) at least 5-10 is used by research laboratories,
406 with an LOD of around 3:1. This approach is fully in line with guidelines from international
407 bodies [28, 55, 56, 57, 58, 59, 60, 61, 62, 63].

408 For targeted assays, all peaks should be checked to ensure they reach the specified S:N
409 ratio, as well as the required number of data points. However, for large scale metabolomics,
410 manual checking is not feasible for all peaks, but if certain metabolites or features are judged
411 to be discriminatory (e.g. predictive of sample type), then those should be prioritised for manual
412 post-processing checks to ensure that the differences are real and the data is of good quality.

413 Discussion

414 Validation is defined as the process of proving that any procedure, process, equipment,
415 material, activity or system performs as expected within defined acceptance criteria under a
416 given set of conditions, and that the performance characteristics of the procedure meet the
417 requirements for the intended analytical applications [64, 65]. Although implementing fail/pass

418 criteria advised by bioanalytical method validation guidelines have provided a useful degree of
419 standardisation and consistency between regulated laboratories, new advances in technology,
420 multiplexing, and metabolomics studies require tiered and/or 'fit-for-purpose' approaches [66]
421 for pragmatic/practical use.

422 One of the challenges in targeted metabolomics is that obtaining the suitable internal
423 standards are often difficult. On the other hand, one of the advantages of targeted biomarker
424 assays is that the biology of the biomarker has often already been understood, so the anticipated
425 levels, turnover rate, the intra- and inter-subject variability is known, thus enabling the analyst
426 to develop the right assays with appropriate level of validation to generate quality data.
427 However, for newly discovered biomarkers for which little is known, assay development
428 should start with a focus on parallelism, selectivity and sensitivity. Then, at a later stage, the
429 assay could be fine-tuned to the required acceptance criteria [67].

430 Pre-determined or fixed acceptance criteria are established and appropriate for
431 validated assays (Tier 1); however, for qualified, research, and screening methods (Tiers 2-4),
432 it may be appropriate to define these after the method performance experiments have been
433 conducted to fine-tune the assay to the required acceptance criteria. Minimally, it is expected
434 that *a priori* acceptance criteria can be relaxed for the higher tiers if such method
435 performance still supports the intended use of the data and ultimately supports the necessary
436 decisions that will be made [39].

437 Recently, regulatory bodies have begun to address the requirements needed to achieve
438 robust and reliable data in biomarker assays applying omics data. To our knowledge, the
439 Omics subgroup report [21] and C-Path report [68] are the only documents published by the
440 regulatory agencies on assessment of biomarkers assays. The Omics subgroup report [21] on
441 behalf of the EMA and Heads of Medicines Agencies (HMA) published in 2017 a checklist
442 to introduce considerations for successful qualification of novel methodologies such as

443 biomarker quantification, clinical outcome assessment, imaging methods and big data
444 approaches.

445 This checklist entails brief recommendations for context-of-use (CoU), selection of
446 endpoints, statistical analysis plan, demonstration of clinical utility, standard of
447 truth/surrogate standard of truth, suitability of the analytical platform, as well as a link to
448 ICH E16 and ICH E18 guidelines that focus on pharmacogenomics biomarkers, and sampling
449 and management of genomic data (EMA/750178/2017 document). Furthermore, the FDA in
450 conjunction with the Path Institute (C-Path) published a document entailing broad scientific
451 insight to biomarker assay challenges, and a complete description of necessary approaches
452 that can be applied to biomarkers qualification [68].

453 Targeted metabolomic studies often require the quantification (e.g. absolute, semi-
454 and/or relative) of multiple analytes (e.g. multiplexing) in order to exploit putative
455 biomarkers identified via untargeted metabolomics methods, and validate derived hypotheses.
456 The gap between targeted and untargeted metabolomics is very narrow and often
457 overlapping. For example, in assays for the quantification of hundreds of polar or lipophilic
458 metabolites, authentic external standards and internal standards may not be available for all
459 analytes. Many of these assays also satisfy the criteria for the accuracy and precision of
460 metabolite measurements as defined by the FDA. However, they should be reported as
461 estimated rather than absolute concentrations mainly due to lack of standard and/or internal
462 standard availability.

463 LC-MS multiplexing allows for the measurement of numerous analytes in the same
464 analytical run, thus providing significantly more information about molecular biomarker
465 signatures than measurements of single analytes. As the number of analytes increases,
466 favourable accuracy and precision values are often more difficult to obtain. As noted by
467 regulatory guidelines, all quantified analytes in the same assay need to meet the same

468 acceptance criteria. If one of the analytes fails to meet acceptance criteria, the whole
469 analytical run fails. However, in multiplexing assays, re-analysis of the whole panel of
470 analytes should not be necessary if most of the analytes are within the pre-defined quality
471 specifications.

472 Furthermore, acceptance criteria should be widened [69], in which the variation at the
473 LLOQ is increased from 20% to 30%-40%. One should bear in mind that increasing the
474 number of replicates at the LLOQ will result in lower variation (RSD%). The degree of
475 analytical variability that can be tolerated depends on biological variation. Higher variation is
476 often expected for large biomolecules compared to metabolites. Incurred sample reanalysis
477 (ISR) of macromolecules as recommended by the FDA is within 30% of the average of
478 original and reanalysed values compared to 20% for small molecules [70]. In the proposed
479 framework, acceptance criteria for Tier 2 is more relaxed as size and number of replicates are
480 lowered. However, increased calibration points for Tier 2 when the number of metabolites
481 are increased are recommended. Furthermore, biomarkers should be simultaneously
482 evaluated in both absolute and semi/relative quantification manners for multiplexed assays [69].
483 For instance, identification or presence of a particular compound (e.g. qualitative evaluation)
484 alongside quantification of related metabolites or a precursor could provide better insight into
485 metabolic phenotyping.

486 Validation beyond the intended use of the data means significant re-work, loss of time
487 and increased cost in the blind pursuit of absolute requirements. For metabolomics at its current
488 state of development, what is required is the definition of a simple, pragmatic and easy- to-follow
489 framework that reflects realistic and practical needs that allow for the most efficient practices.
490 For instance, an assay that does not pass the criteria for full validation but, nevertheless, fulfils
491 the essential requirements for linearity, accuracy, precision, LLOQ and carryover criteria may
492 be devised. In that case, guidance should focus on minimum requirements. Specifications of

493 merit might include: linearity with an LLOQ set as first calibrant, accuracy, precision and
494 carryover.

495 Overall, the guidelines for assays developed for drugs that have been devised by
496 regulatory authorities to ensure safety and efficacy in humans represent a 'gold standard' that
497 may not be required for many types of targeted and untargeted metabolomics applications.
498 This is not to suggest that metabolic phenotyping methods should not be developed to the
499 standards necessary to provide reliable and scientifically valid data but to suggest that the
500 use of tiered approaches linked to the type of investigation is at (i.e. discovery, hypothesis
501 validation, biomarker/panel, and/or qualification stages), should drive the level of validation
502 performed. A number of intricate analytical factors (e.g. pre-analytical factors) defining core
503 assay expectations, and setting acceptable assay performance criteria, should be taken into
504 account for assessing the reliability and quality of metabolomics assays. Our MAP/UK
505 consensus framework provides a bench guide for the two major categories of validation and
506 qualification of targeted metabolomics analysis that have been described in Table II.

507 **Conclusions**

508 Metabolomics has the potential to lead advances in the discovery of clinically and
509 toxicologically relevant biomarkers, yet the lack of harmonisation at different levels of
510 processes throughout the whole metabolomics pipeline from study design, sample handling,
511 biobanking, metabolite quantification and data analysis remain issues that need to be addressed.
512 Metrological tracability and future development of certified matrix reference materials similar
513 to National Institute of Standards and Technology reference standards (NIST SRM 1950), and
514 standard calibration mixtures should be established and harmonized within both the research
515 and regulatory communities.

516 The MAP/UK consortium proposes the pragmatic development of a 'fit-for- purpose' 4-

517 tiered framework for assessing the reliability of metabolomics assays via a decision-making
518 process and adaptation of existing drug regulatory guidance. The required level of analytical
519 rigour and/or qualification that bioanalytical methods need to achieve scientifically valid
520 studies in metabolomics has been considered. This framework is intended to guide bioanalysts
521 and to facilitate improved communication between the research and regulatory communities,
522 and to enable the establishment of appropriately qualified targeted metabolomics assays to meet
523 the needs of multiple applications of this technology in the regulatory sciences. Ultimately, this
524 community-initiated framework can accelerate the application of metabolomics in regulatory
525 applications and achieve acceptance via its consensus use.

526 **Acknowledgements**

527 Funding – Medical Research Council (grant reference: MR/S010483/1) to the MAPUK project.
528 Wellcome Trust for LIPID MAPS (203014/Z/16/Z) to VO'D.

References

1- Roberts, Lee D., Amanda L. Souza, Robert E. Gerszten, and Clary B. Clish. Targeted Metabolomics. *Curr. Protoc. Mol. Biol*. **98**, no. 1 (2012): 30.2.1-30.2.24. https://doi.org/https://doi.org/10.1002/0471142727.mb3002s98.

2- Goodacre, Royston, Seetharaman Vaidyanathan, Warwick B. Dunn, George G. Harrigan, and Douglas B. Kell. Metabolomics by Numbers: Acquiring and Understanding Global Metabolite Data. *Trends Biotechnol*. **22**, no. 5 (2004/05// 2004): 245-52. https://doi.org/10.1016/j.tibtech.2004.03.007.

3- Wei, Ru, Guodong Li, and Albert B. Seymour. High-Throughput and Multiplexed Lc/Ms/Mrm Method for Targeted Metabolomics. Anal. Chem. **82**, no. 13 (2010/07// 2010): 5527-33. https://doi.org/10.1021/ac100331b.

4- Dunn, Warwick B., *et al*. Molecular Phenotyping of a Uk Population: Defining the Human Serum Metabolome. *Metabol.* : Official journal of the *Metabol. Soc.* **11** (2015): 9-26. https://doi.org/10.1007/s11306-014-0707-1.

5- Guijas, Carlos, J. Rafael Montenegro-Burke, Benedikt Warth, Mary E. Spilker, and Gary Siuzdak. Metabolomics Activity Screening for Identifying Metabolites That Modulate Phenotype. *Nat. Biotechnol*. **36**, no. 4 (2018/04/01 2018): 316-20. https://doi.org/10.1038/nbt.4101.

6- Rinschen, M. M., J. Ivanisevic, M. Giera, and G. Siuzdak. Identification of Bioactive Metabolites Using Activity Metabolomics*. Nat Rev Mol Cell Biol*. **20**, no. 6 (Jun 2019): 353-67. https://doi.org/10.1038/s41580-019-0108-4.

7- Viant, Mark R., *et al.* Use Cases, Best Practice and Reporting Standards for Metabolomics in Regulatory Toxicology. *Nat. Commun.* **10**, no. 1 (2019/07/10 2019): 3041. https://doi.org/10.1038/s41467-019-10900-y.

8- Want, Elizabeth J., Benjamin F. Cravatt, and Gary Siuzdak. The Expanding Role of Mass Spectrometry in Metabolite Profiling and Characterization. *Chem. BioChem*. **6**, no. 11 (2005): 1941-51.

9- Dunn, W. B., *et al.* Procedures for Large-Scale Metabolic Profiling of Serum and Plasma Using Gas Chromatography and Liquid Chromatography Coupled to Mass Spectrometry. *Nat Protoc.* **6**, no. 7 (Jun 30 2011): 1060-83. https://doi.org/10.1038/nprot.2011.335.

10- Gertsman, Ilya, and Bruce A. Barshop. Promises and Pitfalls of Untargeted Metabolomics. *J. Inherited Metab. Dis.* **41**, no. 3 (2018): 355-66. https://doi.org/https://doi.org/10.1007/s10545-017-0130-7.

11- Rattray, Nicholas J. W., *et al*. Metabolic Dysregulation in Vitamin e and Carnitine Shuttle Energy Mechanisms Associate with Human Frailty. *Nat. Commun*. **10**, no. 1 (2019/11/05 2019): 5027. https://doi.org/10.1038/s41467-019-12716-2.

12- Ribbenstedt, Anton, Haizea Ziarrusta, and Jonathan P. Benskin. Development, Characterization and Comparisons of Targeted and Non-Targeted Metabolomics Methods. *PLoS One.* **13**, no. 11 (2018): e0207082. https://doi.org/10.1371/journal.pone.0207082.

13- Tabassum, R., *et al.* Genetic Architecture of Human Plasma Lipidome and Its Link to Cardiovascular Disease. *Nat. Commun*. **10**, no. 1 (Sep 24 2019): 4329. https://doi.org/10.1038/s41467-019-11954-8.

14- Goutman, S. A., *et al.* Untargeted Metabolomics Yields Insight into Als Disease Mechanisms. *J Neurol Neurosurg. Psychiatry* **91**, no. 12 (Dec 2020): 1329-38. https://doi.org/10.1136/jnnp-2020-323611.

15- Palmer, Jessica A, *et al.* A Targeted Metabolomics-Based Assay Using Human Induced Pluripotent

Stem Cell-Derived Cardiomyocytes Identifies Structural and Functional Cardiotoxicity Potential. *Toxicol. Sci.* **174**, no. 2 (2020): 218-40. https://doi.org/10.1093/toxsci/kfaa015.

16- Frédérich, Michel, Bernard Pirotte, Marianne Fillet, and Pascal de Tullio. Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* **59**, no. 19 (2016/10/13 2016): 8649-66.
https://doi.org/10.1021/acs.jmedchem.5b01335.

17- Beger, R. D., *et al.* Metabolomics Enables Precision Medicine: A White Paper, Community Perspective. *Metabol.* **12**, no. 10 (2016): 149. https://doi.org/10.1007/s11306-016-1094-6.

18- Mandal, Rupasri, Danuta Chamot, and David S. Wishart. The Role of the Human Metabolome Database in Inborn Errors of Metabolism. *J. Inherited Metab. Dis*. **41**, no. 3 (2018/05/01 2018): 329-36. https://doi.org/10.1007/s10545-018-0137-8.

19- Burla, B., *et al.* Ms-Based Lipidomics of Human Blood Plasma: A Community-Initiated Position Paper to Develop Accepted Guidelines. *J Lipid Res.* **59**, no. 10 (Oct 2018): 2001-17. https://doi.org/10.1194/jlr.S087163.

20- Thompson, J. W., *et al.* International Ring Trial of a High Resolution Targeted Metabolomics and Lipidomics Platform for Serum and Plasma Analysis. *Anal. Chem.* **91**, no. 22 (Nov 19 2019): 14407-16. https://doi.org/10.1021/acs.analchem.9b02908.

21- Renate König, Alison Cave, Mark Goldammer, and Didier Meulendijks. Bioanalytical Omics Subgroup Report. Heads of Medicines Agencies (HMA) and European Medicine Agency (EMA) (n.d.). https://www.ema.europa.eu/en/documents/report/bioanalytical- omics-subgroup-report_en.pdf.

22- Harrill, J. A., *et al.* Progress towards an OECD Reporting Framework for Transcriptomics and Metabolomics in Regulatory Toxicology. *Regul Toxicol Pharmacol.* **125** (Oct 2021): 105020. https://doi.org/10.1016/j.yrtph.2021.105020.

23- Dunn, W. B., D. I. Broadhurst, H. J. Atherton, R. Goodacre, and J. L. Griffin. Systems Level Studies of Mammalian Metabolomes: The Roles of Mass Spectrometry and Nuclear Magnetic Resonance Spectroscopy. *Chem Soc Rev*. **40**, no. 1 (Jan 2011): 387-426. https://doi.org/10.1039/b906712b.

24- Dunn, Warwick B., *et al.* Quality Assurance and Quality Control Processes: Summary of a Metabolomics Community Questionnaire. *Metabolom.* **13**, no. 5 (2017/03/06 2017): 50. https://doi.org/10.1007/s11306-017-1188-9.

25- Broadhurst, David, *et al.* Guidelines and Considerations for the Use of System Suitability and Quality Control Samples in Mass Spectrometry Assays Applied in Untargeted Clinical Metabolomic Studies. *Metabolom.* : Official J of the *Metabolom. Soci.* **14**, no. 6 (2018 2018): 72. https://doi.org/10.1007/s11306-018-1367-3.

26- Hyötyläinen, Tuulia, Linda Ahonen, Päivi Pöhö, and Matej Orešič. Lipidomics in Biomedical Research-Practical Considerations. *Biochimica et Biophysica Acta*. (BBA) – *Mol. and Cell Biol of Lipids* **1862**, no. 8 (2017/08/01/ 2017): 800-03. https://doi.org/https://doi.org/10.1016/j.bbalip.2017.04.002.

27- Beger, R. D., *et al.* Towards Quality Assurance and Quality Control in Untargeted Metabolomics Studies. *Metabolom*. **15**, no. 1 (Jan 3 2019): 4. https://doi.org/10.1007/s11306-018-1460-7.

28- U.S. Department of Health and Human, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), and Center for Veterinary Medicine (CVM).
Bioanalytical Method Validation, Guidance for Industry. (2018). https://www.fda.gov/media/70858/download.

29- European Medicine. Guideline on Bioanalytical Method Validation. Committee for Medicinal Products for Human Use (CHMP), (2011). https://www.ema.europa.eu/en/documents/scientific-

guideline/guideline-bioanalytical- method-validation_en.pdf.

30- International Commitee of Harmonisation/European Medicine. Ich Guideline M10 on Bioanalytical Method Validation. (2019). https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-guideline- m10-bioanalytical-method-validation-step-2b_en.pdf.

31- Ministry of Health and Welfare, Japan. Draft Guideline on Bioanalytical Method Validation in Pharmaceutical Development (2013). https://www.nihs.go.jp/drug/BMV/BMV_draft_130415_E.pdf.

32- China Food and Drug Administration (CFDA), Good Manufacturing Practice for Drugs (2010 Revision), MOH Decree No. 79, https://www.gmpsop.com/RegulatoryReference/CFDA/China_SFDA_Good_Manufacturi ng_Practice_for_Drugs.pdf. Accessed March 2022.

33- Austrilian government, Therapeutic Goods Administration, Finished product (medicine) analytical procedure validations for Complementary Medicines, March 2006, https://www.tga.gov.au/sites/default/files/cm-analytical-procedure-finished.pdf. Accessed March 2022.

34- ANVISA Guide for validation of analytical and bioanalytical methods. Resolution RE no. 899, 23 May 2003 Brazilian Sanitary Surveillance Agency, BrazilGoogle Scholar.

35- Anvisa. Guide for Validation of Analytical and Bioanalytical Methods. . ANVISA Brazil (2012).

36- Huynh-Ba, Kim, and Alexandra Beumer Sassi. Anvisa: An Introduction to a New Regulatory Agency with Many Challenges. *AAPS Open* **4**, no. 1 (2018/12/12 2018): 9. https://doi.org/10.1186/s41120-018-0029-x.

37- Viswanathan, C. T., *et al.* Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays. *Pharm Res.* **24**, no. 10 (Oct 2007): 1962-73. https://doi.org/10.1007/s11095-007-9291-7.

38- Timmerman, P., *et al.* Best Practices in a Tiered Approach to Metabolite Quantification: Views and Recommendations of the European Bioanalysis Forum. *Bioanal.* **2**, no. 7 (Jul 2010): 1185-94. https://doi.org/10.4155/bio.10.90.

39- Lowes, S., *et al.* Tiered Approaches to Chromatographic Bioanalytical Method Performance Evaluation: Recommendation for Best Practices and Harmonization from the Global Bioanalysis Consortium Harmonization Team. *The AAPS Journal* **17** (2015): 17- 23.

40- Nakamura, T. 7th Japan Bioanalysis Forum Symposium: Regulated Bioanalysis, to a New Stage. *Bioanal.* **8**, no. 20 (Oct 2016): 2097-102. https://doi.org/10.4155/bio-2016-4995.

41- Kadian, N., *et al.* Comparative Assessment of Bioanalytical Method Validation Guidelines for Pharmaceutical Industry. *J Pharm Biomed Anal.* **126** (Jul 15 2016): 83-97. https://doi.org/10.1016/j.jpba.2016.03.052.

42- Booth, B., *et al.* Workshop Report: Crystal City V--Quantitative Bioanalytical Method Validation and Implementation: The 2013 Revised Fda Guidance. *The AAPS journal* **17**, no. 2 (2015/03// 2015): 277-88. https://doi.org/10.1208/s12248-014-9696-2.

43- GBC, Global Bioanalysis Consortium on Harmonisation of bioanalytical guidance, accessed Sept 2020, https://sites.google.com/site/globalbioanalysisconsortium/

44- Timmerman, P., *et al.* Tiered Approach into Practice: Scientific Validation for Chromatography-Based Assays in Early Development - a Recommendation from the European Bioanalysis Forum. *Bioanal.* **7**, no. 18 (Sep 2015): 2387-98. https://doi.org/10.4155/bio.15.168.

45- Lavezzari, G., and A. W. Womack. Industry Perspectives on Biomarker Qualification. *Clin Pharmacol Ther*. **99**, no. 2 (Feb 2016): 208-13. https://doi.org/10.1002/cpt.264.

46- Shah, Vinod P., *et al.* Bioanalytical Method Validatiož a Revisit with a Decade of Progress. *Pharm. Res.* **17** (2004): 1551-57.

47- Theodorsson, Elvar. Validation and Verification of Measurement Methods in Clinical Chemistry.

Bioanal. **4**, no. 3 (2012): 305-20. https://doi.org/10.4155/bio.11.311. https://www.future-science.com/doi/abs/10.4155/bio.11.311.

48- Safavi A., Explotatory Biomarker Testing-to Qualify or Validate the Assay? Bioagilityx Insight, 2019, accessed Feb. 2021, 2021, https://www.bioagilytix.com/blog/2019/08/28/webinar-sneak-peek-exploratory- biomarker-testing-to-qualify-or-validate-the-assay/.

49- Lewis M, *et al.* An Open Platform for Large Scale LC-MS-Based Metabolomics . *ChemRxiv.* Cambridge: Cambridge Open Engage; 2022; This content is a preprint and has not been peer-reviewed.

50- Saigusa D, *et al.* (2016) Establishment of Protocols for Global Metabolomics by LC-MS for Biomarker Discovery. *PLoS ONE* **11**(8): e0160555. https://doi.org/10.1371/journal.pone.0160555.

51- Van de Merbel, Nico C. Quantitative Determination of Endogenous Compounds in Biological Samples Using Chromatographic Techniques. *TrAC Trends Anal. Chem.* **27**, no. 10 (2008/11/01/ 2008): 924-33. https://doi.org/https://doi.org/10.1016/j.trac.2008.09.002. https://www.sciencedirect.com/science/article/pii/S016599360800191X.

52- Thakare, R., Y. S. Chhonker, N. Gautam, J. A. Alamoudi, and Y. Alnouti. Quantitative Analysis of Endogenous Compounds. *J Pharm Biomed Anal.* **128** (Sep 5 2016): 426-37. https://doi.org/10.1016/j.jpba.2016.06.017.

53- Tsikas, Dimitrios. "Bioanalytical Method Validation of Endogenous Substances According to Guidelines by the Fda and Other Organizations: Basic Need to Specify Concentration Ranges." Journal of Chromatography B 1093-1094 (2018/09/01/ 2018): 80-81. https://doi.org/https://doi.org/10.1016/j.jchromb.2018.07.005. https://www.sciencedirect.com/science/article/pii/S1570023218308419.

54- Wakamatsu, A., *et al*. Proposed Selection Strategy of Surrogate Matrix to Quantify Endogenous Substances by Japan Bioanalysis Forum Dg2015-15*. Bioanal*. **10**, no. 17 (2018/09/01 2018): 1349-60. https://doi.org/10.4155/bio-2018-0105.

55- Research, C. f. D. E. a., and C. f. B. E. a. Research. 2015. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/analytical-procedures-and-methods- validation-drugs-and-biologics. https://www.fda.gov/regulatory-information/search-fda- guidance-documents/analytical-procedures-and-methods-validation-drugs-and- biologics.

56- Araujo, P. 2009. Key aspects of analytical method validation and linearity evaluation*. J Chromatogr B Analyt Technol Biomed Life Sci.* **877**: 2224-2234.

57- Commission, E. 2002. 2002/657/EC: Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (Text with EEA relevance) (notified under document number C (2002) 3044) https://op.europa.eu/en/publication-detail/-/publication/ed928116-a955-4a84-b10a-cf7a82bad858/language-en.

58- Desimoni, E., Brunetti, B. 2015. Signal to noise ratio; Limit of detection; Standard error of the regression; Chromatographic/Voltammetric/Spectroscopic signals. *Pharm. Anal. ACTA*. **6** (04).

59- Heyden, Y. V. B., R. (2009) The limit of detection, LCGC Europe, https://www.researchgate.net/profile/Stefano_Polesello/post/How_to_calculate_LOD_and_LOQ_of_analyte_by_hplc/attachment/59d6356379197b8077992f7e/AS:384058174525440@1468578012222/download/The+Limit+of+Detection-LCGC.doc.

60- Wenzl, T., Haedrich, J, Schaechtele, A, Robouch, P, Stroka, J. 2016. Guidance Document on the Estimation of LOD and LOQ for Measurements in the Field of Contaminants in Feed and Food. EUR 28099; Publications Office of the European Union: Luxembourg, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC102946/eur%2028099%2

0en_lod%20loq%20guidance%20document.pdf.
61- http://www.pharmacopeia.cn/v29240/usp29nf24s0_c1225.html.
62- https://www.ema.europa.eu/en/ich-q2-r1-validation-analytical-procedures-text-methodology.
63- https://www.who.int/medicines/areas/quality_safety/quality_assurance/28092018Guideline_Validation_AnalyticalMethodValidation-Appendix4_QAS16-671.pdf.
64- Watson, D.G. *Pharmaceutical Analysis*. Second Edition. Oxford: Elsevier churchill livingstone, 2005.
65- Bliesner, D.M. *Validating Chromatographic Methods: A Practical Guide*, Hoboken: Wiely Interscience, 2006.
66- Lee, J. W., *et al.* Fit-for-Purpose Method Development and Validation for Successful Biomarker Measurement. *Pharm Res.* **23**, no. 2 (Feb 2006): 312-28. https://doi.org/10.1007/s11095-005-9045-3.
67- Goodman, J., *et al.* Update to the European Bioanalysis Forum Recommendation on Biomarkers Assays; Bringing Context of Use into Practice. *Bioanal.* **12**, no. 20 (Oct 2020): 1427-37. https://doi.org/10.4155/bio-2020-0243.
68- EMA. Essential Considerations for Successful Qualification of Novel Methodologies. European Medicine Agency (2017). https://www.ema.europa.eu/en/documents/other/essential-considerations-successful- qualification-novel-methodologies_en.pdf.
69- Biomarker assay collaborative evidentiary consideration writing group. Points to Consider Document: Scientific and Regulatory Considerations for the Analytical Validation of Assays Used in the Qualification of Biomarkers in Biological Matrices. Critical Path Institute (2019). https://c-path.org/wp- content/uploads/2019/06/evidconsid-whitepaper-analyticalsection2019.pdf.
70- Christians, Uwe, *et al.* Mass Spectrometry-Based Multiplexing for the Analysis of Biomarkers in Drug Development and Clinical Diagnostics — How Much Is Too Much? *Microchem. J.* **105** (2012/11/01/ 2012): 32-38.
71- Thway, Theingi M., et al. Assessment of Incurred Sample Reanalysis for Macromolecules to Evaluate Bioanalytical Method Robustness: Effects from Imprecision. The AAPS J. 13, no. 2 (2011): 291-98. https://doi.org/10.1208/s12248-011-9271-z. https://pubmed.ncbi.nlm.nih.gov/21461973.