UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

A proposed framework to evaluate the quality and reliability of targeted metabolomics assays from the UK Consortium on Metabolic Phenotyping (MAP/UK)

UK Consortium on Metabolic Phenotyping (MAP/UK)

DOI:
10.1038/s41596-022-00801-8

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
UK Consortium on Metabolic Phenotyping (MAP/UK) 2023, 'A proposed framework to evaluate the quality and reliability of targeted metabolomics assays from the UK Consortium on Metabolic Phenotyping (MAP/UK)', *Nature protocols*. https://doi.org/10.1038/s41596-022-00801-8

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1038/s41596-022-00801-8

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 02. May. 2024

A Proposed Framework to Evaluate the Quality and Reliability of Targeted Metabolomics Assays from the UK Consortium on Metabolic Phenotyping (MAP/UK)

Sarir Sarmad[1], Mark R.Viant[2], Warwick B. Dunn[3,4], Royston Goodacre[3], Ian D. Wilson[5], Katie E. Chapell[6], Jules L. Griffin[7], Valerie B. O'Donnell[8], Brendon Naicker[8], Matthew R. 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

EDITORIAL SUMMARY: In this Perspective, authors from the UK Consortium on Metabolic Phenotyping propose a 'fit-for-purpose' 4-tiered framework to evaluate the reliability of targeted metabolomics analyses, addressing the need for community-accepted, harmonised guidelines for tiers other than full validation.

TWEET: A new Perspective from the UK Consortium on Metabolic Phenotyping proposes a 'fit-for-purpose' 4-tiered Framework to Evaluate the Quality and Reliability of Targeted Metabolomics Assays

TEASER: 4-tier Framework to Evaluate Metabolomics Assays

RELATED LINKS:
https://pubs.acs.org/doi/abs/10.1021/acs.analchem.9b02908
https://chemrxiv.org/engage/chemrxiv/article-details/61ebd6fa0716a8529e3823dc
https://www.sciencedirect.com/science/article/pii/S0002870321000107

Abstract

Targeted metabolite assays that measure tens or hundreds of pre-selected metabolites, typically using liquid chromatography mass spectrometry (LC-MS), are increasingly being developed and applied to metabolic phenotyping studies. These are used both as standalone phenotyping methods and for the validation of putative metabolic biomarkers obtained from untargeted metabolomics studies. However, there are no widely accepted standards in the scientific community for ensuring reliability of the development and validation of targeted metabolite assays (referred to here as targeted metabolomics). Most current practices attempt to adopt, with modifications, the strict guidance provided by drug regulatory authorities for analytical methods designed largely for measuring drugs and other xenobiotic analytes.Here, the regulatory guidance provided by the European Medicines Agency, U.S. Food and Drug Administration, and International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use are summarised. In this Perspective, we have adapted these guidelines and propose a less onerous 'tiered' approach to evaluate the reliability of a wide

range of metabolomics analyses, addressing the need for community-accepted, harmonised guidelines for tiers other than full validation. This ‘fit-for-purpose’ tiered approach comprises 4 levels – discovery, screening, qualification and validation – and is discussed in the context of a range of targeted and untargeted metabolomics assays. Issues arising with targeted multiplexed metabolomics assays, and how these might be addressed, are considered. Furthermore, guidance is provided to assist the community with selecting the appropriate degree of reliability for a series of well- defined applications of metabolomics.

Keywords: Metabolic phenotyping, metabolomics, LC-MS, multiplexed assays, validation, 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 (known as metabolites) involved in cellular metabolism and other biochemical processes in cells, tissues and bodily fluids [1,2]. Metabolomics facilitates the characterization of a system from genomic to metabol(om)ic activity and interaction with the environment, and reveals dynamic insights 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 (sometimes defined as a semi-targeted approach) that combines some aspects of both types of analyses [7]. Untargeted metabolomics is

84 a discovery-based approach where the objective is to analyse as many detectable metabolites
85 without biological bias, including unknowns, to determine which, if any, are significantly
86 perturbed in the diseased phenotype, followed by post-hoc identification of those putative
87 metabolic biomarkers [8]. Targeted approaches on the other hand, involve the (multiplexed)
88 analysis of known metabolites, and such methods often focus on quantification of a subset of
89 metabolites representative of key pathways, or of metabolites determined to be important from
90 prior untargeted metabolomics [9]. Targeted metabolomics is hypothesis driven, with the
91 significant advantage of quantifying known metabolites with greater sensitivity and selectivity
92 [1], while untargeted metabolomics is hypothesis generating, with the advantage of increased
93 metabolite coverage and potential of biomarker discovery [8]. The major disadvantage of
94 untargeted approaches is that relative responses and not actual concentrations are reported,
95 while the major disadvantage of targeted approaches is their limited coverage of the
96 metabolome [10]. The techniques that are most widely used for untargeted analysis include liquid
97 chromatography high-resolution mass spectrometry (LC-MS), gas chromatography mass
98 spectrometry (GC-MS) and ^{1}H nuclear magnetic resonance (NMR) spectroscopy, while liquid
99 chromatography-triple quad-tandem mass spectrometry (LC-MS/MS) remains one of the
100 traditional techniques for targeted analysis of a limited numbers of analytes, with another
101 approach being GC-MS which involves fragmentation of the metabolite during electron
102 ionisation [11,12]. One of the challenges in targeted metabolomics is that obtaining suitable
103 internal standards is often difficult. On the other hand, one of the advantages of targeted
104 biomarker assays is that the biology of the biomarker has often already been understood, so the
105 anticipated levels, turnover rate, the intra- and inter-subject variability is known, thus enabling
106 the analyst to develop the right assays with appropriate level of validation to generate quality
107 data. However, for newly discovered biomarkers for which little is known, assay development
108 should start with a focus on parallelism, selectivity and sensitivity. Then, at a later stage, the
109 assay could be fine-tuned to the required acceptance criteria [13]..

110 Advances in metabolomics have led to new clinical and toxicological diagnostic
111 biomarkers [14, 15, 16], which can contribute to stratified medicine and safety assessment of
112 drugs [17, 18]. Metabolomics is also central to the screening of innate errors of metabolism [19].
113 However, there are several challenges in the translation of metabolomics research to clinical
114 and toxicological applications under regulatory control. Issues include analytical
115 reproducibility, accuracy, precision, metabolite identification/quantification, study design,
116 sample handling, lack of harmonised reporting frameworks for published data and metadata,
117 insufficient open-access data to enable data-mining by other researchers [20], lack of
118 harmonisation in bio-banking, batch-to-batch variation, and between-methods bias [21].
119 Assessing the reliability of bioanalytical methods for metabolomics is challenging when
120 compared to the validation of other types of bioanalytical methods. Data from the
121 metabolomics field are variable, and heterogeneity among data formats, data analysis
122 pipelines, algorithms and applied statistical methods need to be addressed. There is a need to
123 define the extent to which assessing the reliability of these methods is required, and the scope
124 of such assessments, as well as how the standards applied and methods for reporting should
125 be chosen in order to ensure appropriate data quality for use in regulatory processes [22]. To
126 eliminate some of these problems, communication between the research and regulated
127 clinical and toxicological communities needs to be more fully developed, and the
128 establishment of a system to assess and cross-correlate metabolic profiles obtained by different
129 laboratories and instruments is needed [20]. The new Metabolomics Reporting Framework for
130 regulatory toxicology, developed by multiple stakeholders from research laboratories, industry
131 and government regulatory agencies and coordinated by the Organisation for Economic Co-
132 operation and Development (OECD) provides evidence on how progress can be made to
133 achieve harmonised reporting of methods, data, metadata and findings, and thereby advance
134 the application of metabolomics within regulatory settings [23]. There are a plethora of
135 publications that provide comprehensive guidelines for assessing the quality of untargeted

136 metabolomics assays [24, 25, 26, 27, 28]. Whilst these guidelines provide the foundation for
137 metabolomics system suitability and quality assurance/quality control (QA/QC) proficiency, a
138 community-initiated approach towards harmonised guidelines that ultimately achieves
139 acceptance via their consensus use for evaluating the reliability of targeted metabolomics
140 within research, clinical and toxicological settings is still required.

141 Our scientific collaboration, the UK Consortium on Metabolic Phenotyping (MAP/UK,
142 https://mapuk.org), is a partnership of eight specialised research laboratories and two Phenome
143 Centres, which has been funded by the Medical Research Council to improve UK-wide
144 metabolic phenotyping expertise and capabilities. The MAP/UK partnership brings together a
145 critical mass of methodological, analytical, and computational platforms to develop, optimise,
146 transfer, harmonise, and validate efficient, high-quality metabolomics research and training
147 methods, specifically tailored to the growing need for biomedical studies that require robust
148 metabolic phenotyping. The overall aim of the MAP/UK partnership is to investigate new
149 biomarkers within metabolic signatures of disease, novel targeted quantitative metabolomic
150 and hybrid approaches, and developing untargeted metabolomics to meet gaps in molecular
151 coverage of key disease-related pathways, alongside a variety of other factors, including diet,
152 lifestyle/environment, microbiome and genetics. As a collective of scientists with the aim of
153 harmonisation of metabolic phenotyping, existing regulatory guidelines have been reviewed to
154 extract commonalities from these guidelines that can be adopted to 'fit-for- purpose' and tiered
155 approaches for untargeted and targeted metabolomics.

156 The aim of this manuscript is to propose harmonised guidelines for evaluating the
157 reliability of targeted (multiplexed) mass spectrometry-based metabolomics assays taking into
158 consideration intra-laboratory precision, accuracy, reproducibility, and cross-laboratory
159 harmonisation of methods and data acquired on different instrumental platforms. First, existing
160 guidelines for bioanalytical method validation, including an existing 4-tiered framework

161 applied in drug discovery, are reviewed. Then, after introducing the applications of clinical and
162 toxicological metabolomics in regulatory settings, a new 'fit-for-purpose' 4-tiered (discovery,
163 screening, qualification and validation) framework for assessing analytical reliability that is
164 suitable for targeted and hybrid untargeted metabolomics assays is proposed. In addition, a
165 checklist for the bioanalytical process has been provided to facilitate better understanding and
166 emphasise the importance of harmonisation at each step, as described in Box 1.

167 **START BOX 1**

168 **Checklist for bioanalytical assay process:**

169 **1- Pre-analytical:**

170 • Hypothesis/study design/ sample size
171 • Data acquisition of demographics for groups/individuals including clinical, diet,
172 medications and life-style data
173 • Sample type (plasma/serum/urine/feces), collection method, preservation, and
174 timing
175 • Sample storage

176 **2- Analytical:**

177 • Sample preparation and purification
178 • Authentic reference materials (external standards), quality control (QC) samples
179 and suitable internal standards
180 • Maintaining assay reliability and quality by selecting the right tier based on
181 number of metabolites and assay purpose (consult Table 1).
182 • Select validation parameters and acceptance criteria for targeted assays (tier 1
183 and 2), by consulting Table 2. Note that Tier 1 parameters are the same as

184 suggested by regulatory guidelines (FDA/EMA/ICH2019) for validation, and
185 Tier 2 (qualification) has a wider range of acceptance criteria.
186 • Select appropriate instrumentation such as liquid chromatography high-
187 resolution mass spectrometry (LC-MS), liquid chromatography-triple quad-
188 tandem mass spectrometry (LC-MS/MS), and considerations regarding
189 instrument calibration, settings, analytical batches, and quality assurance
190 (QA)/performance.

END BOX 1

191

192 **The concept of regulatory bioanalytical validation**

193 An analytical assay starts with a definition of its purpose (i.e. intended application),
194 defining what is 'fit-for-purpose', followed by method development and optimisation, then
195 subsequently by assay validation (dependent upon the tier, as introduced above) and
196 documentation before it can finally be applied for the intended purpose. Validation is
197 defined as a process that provides proof of assay integrity within given specifications with
198 the parameters of an assay used for quantification being statistically reliable between assays
199 over time. Prior to initiating a validation study, a well-planned validation protocol should be
200 written and reviewed for scientific soundness and completeness. The protocol should
201 describe the procedure in detail and should include pre-defined acceptance criteria and pre-
202 defined statistical methods, and should be approved by all participants in the analytical
203 pipeline.

204 There are numerous validation parameters (accuracy, precision, calibration curve,
205 lower limit of quantitation, selectivity/specificity, carryover, analyte stability, recovery,
206 dilution integrity, system suitability test, matrix effect/factor, parallelism, incurred sample
207 re-analysis, quality control, robustness/ruggedness, hook/prozone effect, and minimum

208 required dilution) to incorporate into the validation process (please see Supplementary
209 Table 1 for comparison of validation parameters by multiple guidelines, and Supplementary
210 Table 2 for definition of validation parameters). The validation workflow has been
211 summarized in a visual format (Fig. 1). This workflow is a modification of general
212 validation workflow in combination with two extra steps based on our proposed framework
213 to advise analysts for choice of appropriate tier of the assay, and depth of required
214 validation. One should justify the required level of validation to be 'fit-for-purpose' based
215 on the differing applications of a particular method. Theoretically, there are no limits to the
216 extent of validation and verification procedures. However, in practice, there are both time
217 and economic constraints on what can be achieved. Therefore, it is crucial to have optimised
218 guidelines that are generally accepted, harmonised and cost-effective [29].

219 Multiple guidelines exist that describe the regulation of bioanalytical assays such as
220 those from the U.S. Food and Drug Administration (FDA) [30], the European Medicines
221 Agency (EMA) [31], the International Council for Harmonisation of Technical Requirements
222 for Pharmaceuticals for Human Use (ICH) [32], the Japanese Ministry of Health, Labour and
223 Welfare (MHLW) [33], Chinese (State) Food and Drug Administration (CFDA, currently the
224 National Medical Products Administration, NMPA) [34], Australian Therapeutic Goods
225 Administration (TGA) [35], and Brazilian National Health Surveillance Agency (Anvisa) [36, 37,
226 38].

227 The two most well-used bioanalytical guidelines from the EMA and FDA are similar
228 but not identical. The scientific basis for the evaluation of parameters is the same across both
229 guidelines. However, there are also differences in terminology, recommended validation
230 parameters, acceptance criteria and methodology, which can cause confusion amongst
231 bioanalysts and/or pharmaceutical companies given the globalisation of the pharmaceutical
232 sector. Standards setting and harmonisation was advanced by the ICH, which is an

233 international organisation with the mission to achieve greater harmonisation worldwide to
234 ensure that safe, effective, and high-quality medicines are developed and registered in the
235 most resource-efficient manner. The ICH consolidated best practices from the FDA and EMA
236 guidelines in 2019 into a harmonised M10 bioanalytical method validation draft in order to
237 clarify any areas of uncertainty between the two guidelines. A comparison between the FDA
238 and EMA guidelines and the consolidated ICH M10 draft guideline are summarised in
239 Supplementary Table 1.

240 Whilst these regulatory guidelines are comprehensive, they are largely developed for
241 the measurement of drugs and other xenobiotic analytes. Endogenous biomarkers are often
242 measured in metabolomics which requires different considerations including matrix effect.
243 Matrix effect is referred to a phenomena usually encountered in LC-MS/MS where ionisation
244 efficiency of target analytes are altered in the presence of co-eluting compounds in the same
245 matrix. It could cause either ion suppression or enhancement. Quantitation of matrix effect is
246 termed matrix factor (MF), and should be determined within the lowest limit of
247 quantification (LLOQ), and upper limit of quantification (ULOQ) of a matrix-matched
248 calibration curve.

249 Lower limit of quantification (LLOQ) represents sensitivity of the assay and
250 determines the lowest concentration of analyte in a sample which can be quantified reliably
251 with an acceptable accuracy and precision. LLOQ should be considered as the lowest point in
252 the calibration curve where signal-to-noise (S:N) ratio should be at least 5:1. Evaluating
253 these limits using standard solutions in neat solvent, and/or matrix deprived of specific
254 classes of metabolites (such as stripped plasma) are not an ideal solution as what has been
255 depleted is not defined. Furthermore, measurement of specificity/selectivity for endogenous
256 metabolites is much more challenging due to presence of multiple isoforms. Recently,
257 regulatory bodies have begun to address the requirements needed to achieve robust and

258 reliable data in biomarker assays applying omics data. To our knowledge, the Omics
259 subgroup report [22] and C-Path report [39] are the only documents published by the regulatory
260 agencies on assessment of biomarkers assays. The Omics subgroup report [22], on behalf of the
261 EMA and Heads of Medicines Agencies (HMA), published in 2017 a checklist to introduce
262 considerations for successful qualification of novel methodologies such as biomarker
263 quantification, clinical outcome assessment, imaging methods and big data approaches. This
264 checklist entails brief recommendations for context-of-use (CoU), selection of endpoints,
265 statistical analysis plan, demonstration of clinical utility, standard of truth/surrogate standard
266 of truth, suitability of the analytical platform, as well as a link to ICH E16 and ICH E18
267 guidelines that focus on pharmacogenomics biomarkers, and sampling and management of
268 genomic data (EMA/750178/2017 document). Furthermore, the FDA in conjunction with the
269 Path Institute (C-Path) published a document entailing broad scientific insight to biomarker
270 assay challenges, and a complete description of necessary approaches that can be applied to
271 biomarker qualification [39]. Before introducing our proposed framework to assist bioanalysts
272 in selecting the appropriate tier of validation for a series of well-defined applications of
273 metabolomics, we provide a brief introduction to the existing tiered regulatory guidance for
274 the targeted measurement of single drugs.

275

276 Figure 1- Validation workflow steps and positioning of the suggested tiered framework
277 (Table 1 and Table 2) within the general workflow to select the most appropriate tier and
278 degree of validation.

279 **Existing tiered regulatory guidance for bioanalysis** A fundamental question is
280 how stringently regulatory bodies view these guidelines as being hard rules, or whether they
281 could be adopted as 'fit-for-purpose' for targeted metabolomics assays, and used within a
282 'tiered' framework. The concept of defensible scientific flexibility has been a debate within
283 the bioanalytical community in the pharmaceutical industry. The Crystal City III workshop
284 proposed the concept of 'fit-to-purpose' in 2006 as an alternative for the full validation
285 workflow already described by the FDA regulatory documents in order to address
286 uncertainties from the bioanalytical community as to what level of data scrutiny is required
287 to generate quality data whilst optimising resources to meet study objectives with an
288 adequate level of data quality and reliability [40]. Furthermore, the European Bioanalysis
289 Forum (EBF) proposed the consolidation of tiered approaches to include three levels (or tiers)
290 of quality standards for metabolite quantification for screening, qualified and validated
291 assays [41]. Consequently, the MHLW and FDA permitted 'adjustments and modifications' of
292 their bioanalytical method validation guidelines to fit the intended use of the assay, and this
293 perspective was extended to tiered approaches for metabolite quantification [42, 43, 44].

294 The Crystal City VI workshop in 2015 [45] defined a less rigorous level of validation
295 than the FDA guidelines for drug metabolite quantification at early stages of development.
296 The Global Bioanalytical Consortium (GBC) assigned Team A2 with the objective of
297 providing a framework to rationalise the level of bioanalytical methods for drug
298 characterization and proposed a clear path for implementation and use of tiered approaches
299 [42]. Furthermore, two globally recognised teams within the GBC (S1 and L1) provided
300 acceptance standards for validation methods for small and large pharmaceutical molecules,
301 respectively [46]. However, different terminologies have been used as part of the 'fit-to-
302 purpose' concept, such as tiered assays, scientific validation, qualified assays or partial
303 validation. Thus, it has been a source of confusion for academia and the
304 biotechnology/pharmaceutical industry due to a lack of clear guidance [42]. More recently,

305 these alternative validation assay workflows in the bioanalytical industry have been
306 categorised into four tiered levels of method performance and evaluation based on the final
307 purpose of the derived analytical data ranging from the most to least stringent, as follows:
308 level 1) validation, intended for regulatory studies; level 2) qualification; level 3) research;
309 and level 4) the least stringent defined as 'screening' [42, 47, 48]. These four tiered levels are
310 described in more detail below, and whilst these concepts have been designed for drug
311 development and submission to regulatory authorities, they provide a framework that could
312 be adapted for a range of assays used in metabolomics studies.

313 • Level 1) **validated bioanalytical assays** are designed for intended pharmaceutical
314 products and thus require the highest level of confidence in analytical results as suitable
315 for regulated good laboratory practice (GLP), pre-clinical/clinical, pharmacokinetic
316 and/or toxicological studies, and identification of active metabolites in safety testing
317 (MIST). These mandate that assay precision, accuracy, selectivity, sensitivity, and
318 stability of the analytes should be determined throughout the bioanalytical measurement
319 process. FDA recommended evaluations should be performed [41].

320 • Level 2) **qualified bioanalytical assays** do not need to demonstrate that the
321 measurement methods are as robust as validated assays. This tier is suitable for non-
322 regulated studies in the drug development process, with additional assessment of tissue
323 concentrations or other matrices during preclinical or late discovery phases, and in
324 decision-making for context of use (COU) statements. Single method performance with
325 a statistically appropriate number of quality controls (QC) samples ($n \geq 5$) at each level
326 and a suitable calibration range, precision and accuracy should be performed.

327 • Level 3) **research-grade bioanalytical assays** are suitable for mid- to late-discovery
328 phases of drug development projects for decision-making evaluations and/or
329 verification of additional biomarkers or metabolites for non-GLP regulated studies.

330 They use limited characterization with calibration standards prepared using a
331 comparator reference material such as an *in situ* (in solution) standard with the
332 concentration estimated by radioactivity measurement, NMR or ultraviolet (UV)
333 absorption as representative methods. The method provides semi-quantitative analyte
334 concentrations within wider accuracy and precision limits than for the two higher tiers
335 [42]. This approach enables the partial characterisation of an analytical method that may
336 eventually move to a qualified or validated assay. It should provide sufficient scientific
337 rigor to ensure that it is fit-for-purpose and that there is confidence in the data. Method
338 evaluation should be conducted prior to sample analysis, with the precision and
339 accuracy needed to achieve the more relaxed criteria of 20% relative standard deviation
340 (RSD) and 30% reduction of error (RE) at the LLOQ (Lowest Limit of Quantitation).

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

350 Apart from the four-tiered levels approach in the bioanalytical industry, there is a
351 general concept of 'full' and 'partial' validation. Full validation is necessary when developing
352 and implementing a bioanalytical method for the first time such as when analytes are added to
353 a panel for bioanalytical quantification. In targeted metabolomics, full validation of a method

354 by the accredited clinical laboratory is required when the result from that assay (e.g.
355 concentration of a biomarker in terms of molarity for liquids or µg/mg for tissue) is used for
356 making a clinical decision. Partial validation is required in the case of bioanalytical method
357 transfers between laboratories or when the method parameters such as instrument and/or
358 software platform change, such as changes in species (e.g. human plasma to murine plasma)
359 or matrix (e.g. human plasma to human serum/urine). Partial validation can range from as little
360 as one intra-assay accuracy and precision determination to nearly full validation [49] depending
361 on the degree of change required being undertaken.

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

368 **Proposed tiered framework for assessing the reliability of metabolomics bioanalytical**
369 **methods**

370 Considering that there are a range of applications for metabolomics and new advances
371 in LC-MS techniques for multiplexed measurement of metabolites, there is a clear need to
372 propose a harmonisedframework that describes which reliability tier is most 'fit-for-purpose'
373 for different applications. Evaluation of being 'fit-for-purpose' involves questions such as: 1)
374 what is the context of use for the assay (i.e. what will the data be used for); 2) should it be a
375 quantitative, semi-quantitative or relatively quantitative assessment; and 3) what level of
376 uncertainty can be tolerated in the assessment. Consolidating the concept of 'fit-for-purpose'
377 assists bioanalysts in decision-making on whether to qualify or validate a biomarker assay, and
378 which parameters to choose in addition to the number of appropriate replicates [50]. The end-result

379 of a 'fit-for-purpose' validation of an assay using relative quantification is a resource-effective
380 and -efficient demonstration of the bioanalytical method's performance that is tailored to meet
381 the objective of the application. This ultimately provides reliable study data to make important
382 decisions. The decisions may involve further assay development and progression to a fully
383 validated method.

384 The intended use (or application) of a metabolomics assay determines which level of
385 reliability assessment should be used, not the type of assay. Selecting the most appropriate tier
386 for measuring multiple metabolic biomarkers simultaneously for targeted metabolomics assays
387 is challenging if the intended data use is not carefully defined. Hence, the first step in selecting
388 an appropriate tier is to define the intended use of the data and which type of assay is needed,
389 then the most appropriate reliability tier can be further defined.

390 The following framework is proposed as a guideline for the metabolomics community
391 to assess the reliability of both targeted and untargeted metabolomics assays for different types of
392 applications (i.e. from biomarker discovery by a research laboratory, transfer of a method to a
393 different laboratory, through to the use of biomarkers within a clinical setting). The proposed
394 framework is summarised in Table 1 (Tiers 1-4) to assist bioanalysts in selecting the most
395 appropriate tier based on their purpose and assay type. Tiers 1 and 2 (targeted metabolomics)
396 are the main focus of this manuscript, and all related parameters for safeguarding scientific rigor
397 for robust validation and bioanalytical quantification for these two tiers (termed validation and
398 qualification) are summarised in Table 2. These tiers differ in depth, robustness of parameters,
399 and the number of replicates performed for each parameter (See Table 2).

400 Tier 1 - Validation

401 Diagnosis of disease/toxicity phenotype using traditional targeted metabolite analysis
402 with absolute quantification of typically one to a few (less than 10) metabolites. Tier 1 validation
403 is required for compliance with regulatory agencies for clinical diagnostics. This requires an

404 authentic standard (external standard) for each metabolite. The proposed procedure is in
405 alignment with current FDA and ICH M10 bioanalytical method validation guidelines, and is
406 applicable to quantitative analytical assays such as chromatographic, liquid chromatography-
407 mass spectrometry (LC-MS and/or LC-MS/MS), and ligand binding assays (LBA) (see Table
408 2).

409 Tier 2 - Qualification

410 Diagnosis of disease/toxicity phenotype using a multiplexed targeted metabolomics assay
411 with absolute quantification of more than 10 metabolites. This requires an authentic external
412 standard for each metabolite. The criteria for qualifying a method are less strict than for tier 1
413 validation of a method (see Table 2).

414 Tier 3 - Screening

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

419 Tier 4 - Discovery

420 Discovery of putative metabolic biomarkers using untargeted or hybrid metabolomics
421 with relative quantification in a research laboratory. Untargeted methods have the least strict
422 criteria. Furthermore, the use of system suitability tests, intra-study QC samples, phenotyping
423 QCs (healthy vs. disease), inter-laboratory QC samples, and dilution series of pooled QCs have
424 been previously discussed [7, 51] and provide a dimension of semi-quantitative nature to these
425 untargeted assays.

426

427 **Table 1. 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

428 **Table 2. Parameters for validation (Tier 1) *vs*. qualification (Tier 2) of a metabolomics assay**

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 QC 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	• Perform the test	• Absence of interfering compound accepted where the	• Not applicable.	• Not applicable.

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		
Carryover	• Perform the test	• Absence of interfering compound accepted where the response is less than 20% of LLOQ and/or less than 5% for IS	• Perform the test	• Absence of interfering compound accepted where the response is less than 20% of LLOQ and/or less than 5% for IS
Parallelism	• Perform the test, 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	• Precision between samples in a dilution series should be 30-40%
Dilutional Linearity/integrity	• Perform the test	• 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	• Perform the test if applicable	• Spike blank matrix to concentration above ULOQ and dilute it down with blank matrix (1 determination per dilution) • R^2 >0.98
Prozone (hook) effect	• Perform the test, 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%	• Not applicable.	• Not applicable.
Stability - room temperature	• Perform the test	• The accuracy (% nominal) at each level should be ± 15%	• Recommended	• The accuracy (% nominal) at each level should be ± 25%
Stability - 4◦C	• Perform the test	• The accuracy (% nominal) at each level should be ± 15%	• Recommended	• The accuracy (% nominal) at each level should be ± 25%
Stability - freeze/thaw	• Perform the test	• The accuracy (% nominal) at each level should be ± 15%	• Recommended	• The accuracy (% nominal) at each level should be ± 25%
Stability - long-term (-20◦C and/or -80◦C)	• Perform the test	• The accuracy (% nominal) at each level should be ± 15%	• Not applicable.	• Not applicable.

429 Table 2- Required validation parameters to be tested and their acceptance criteria for validation (Tier 1), and qualification (Tier 2) of metabolomics assays. Comparison of

430 evaluation of validation parameters by three regulatory bodies FDA, EMA, and ICH is summarised as Supplementary Table 1. Furthermore, Definition of validation
431 parameters and their assessment methods has been provided as Supplementary Table 2. Abbreviations: LLOQ, lower limit of quantification; ULOQ, upper limit of
432 quantification; IS, internal standard.
433

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

436 Targeted metabolomic studies often require the quantification (e.g. absolute, semi-
437 and/or relative) of multiple analytes (e.g. multiplexing) in order to exploit putative biomarkers
438 identified via untargeted metabolomics methods, and validate derived hypotheses. The gap
439 between targeted and untargeted metabolomics is very narrow and often overlapping. For
440 example, in assays for the quantification of hundreds of polar or lipophilic metabolites,
441 authentic external standards and internal standards may not be available for all analytes. Many
442 of these assays also satisfy the criteria for the accuracy and precision of metabolite
443 measurements as defined by the FDA. However, they should be reported as semi-quantitative
444 concentration rather than absolute concentrations mainly due to lack of standard and/or
445 internal standard availability.

446 LC-MS multiplexing allows for the measurement of numerous analytes in the same
447 analytical run, thus providing significantly more information about molecular biomarker
448 signatures than measurements of single analytes. As the number of analytes increases,
449 favourable accuracy and precision values are often more difficult to obtain. As noted by
450 regulatory guidelines, all quantified analytes in the same assay need to meet the same
451 acceptance criteria. If one of the analytes fails to meet acceptance criteria, the whole
452 analytical run fails. However, in multiplexing assays, re-analysis of the whole panel of
453 analytes should not be necessary if most of the analytes are within the pre-defined quality
454 specifications.

455 Additionally, acceptance criteria should be widened [52], in which the variation at the
456 LLOQ is increased from 20% to 30%-40%. One should bear in mind that increasing the
457 number of replicates at the LLOQ will result in lower variation (RSD%). The degree of
458 analytical variability that can be tolerated depends on biological variation. Higher variation
459 is often expected for large biomolecules compared to metabolites. Incurred sample reanalysis

460 (ISR) of macromolecules as recommended by the FDA is within 30% of the average of
461 original and reanalysed values compared to 20% for small molecules [53]. In the proposed
462 framework, acceptance criteria for Tier 2 is more relaxed as size and number of replicates are
463 lowered. However, increased calibration points for Tier 2 when the number of metabolites
464 are increased are recommended. Furthermore, biomarkers should be simultaneously
465 evaluated in both absolute and semi/relative quantification manners for multiplexed assays
466 [52]. For instance, identification or presence of a particular compound (e.g. qualitative
467 evaluation) alongside quantification of related metabolites or a precursor could provide better
468 insight into metabolic phenotyping.

469 The importance of good laboratory practice at different stages (e.g. sample collection,
470 storage integrity) should be considered for bioanalysis. Sample, analyte and data integrity as
471 well as basic laboratory record keeping are essential. Implementing a laboratory information
472 management system (LIMS) is recommended. Routine calibration of laboratory instruments,
473 pipettes and balances with well-written standard operating procedures (SOPs), as well as
474 selection of suitable blank matrices, internal standards, system suitability test and intra-study
475 QCs are essential. Intra-study QCs should be placed in the analytical run in such a way that the
476 precision of the whole run is ensured by taking into account that study samples should always
477 be bracketed by QCs [7]. Phenotyping QCs (e.g. healthy *vs*. diseased) are recommended. A QC
478 is typically produced by pooling a small aliquot of all study samples, and these are analysed
479 throughout the analytical run. For untargeted metabolomics, a dilution series of the intra-study
480 QC is highly recommended to help differentiate features of biological origin from LC-MS
481 chemical background [12]. Application of isotopically-labelled standards can provide
482 a generalised measure of precision across the study. Furthermore, use of isotopically labelled
483 internal standards helps to compensate for matrix-induced ionisation effects, thereby
484 enhancing the accuracy of the assay when quantification/semi-quantification is applied [26].
485 Choice of suitable surrogate matrices are recommended to improve sensitivity and selectivity

486 of biomarker quantification [54, 55, 56, 57]. Blank matrices with the minimum level of endogenous
487 analyte should be used wherever possible. This approach is suitable for multianalyte assays
488 (spiked with appropriate concentration of each analyte), but matrix effects and stability should
489 be investigated for each analyte. In the absence of blank matrices or surrogate matrices,
490 standard addition approaches which take into account the native concentration of the targeted
491 analyte(s) can be used for recovery and matrix effect checks; and the use of QCs or standards
492 prepared only in solvent and/or buffer considered for accuracy and
493 repeatability/reproducibility tests represents the approach that makes the least assumptions.
494 Artificial blank matrices may be used. A solution of 4% fatty acid-free bovine serum albumin
495 (BSA) in saline buffer that represents the same concentrations of salts and electrolytes in
496 human plasma is an example of blank matrix for human plasma (artificial surrogate matrix).
497 Normalisation strategies to correct for differences in sample amount should be considered.
498 For example, urinary creatinine is often used to adjust the concentration of urinary
499 biomarkers.
500 All targeted assays should have a clearly defined limit of detection (LOD) and limit
501 of quantitation (LOQ). A clearly discernible peak must be visible above clearly visible
502 baseline noise and should be comprised of a specified number of data points (often 6 or above
503 is used). As a general rule, LOQ of S:N (signal-to-noise ratio) of at least 5:1 is used by
504 research laboratories, with an LOD of around 3:1. This approach is fully in line with guidelines
505 from international bodies [30, 58, 59], [60, 61, 62, 63, 64, 65, 66]. For targeted assays, all peaks should be
506 checked to ensure they reach the specified S:N ratio as well as the required number of data
507 points. However, for large scale metabolomics, manual checking is not feasible for all peaks,
508 but if certain metabolites or features are judged to be discriminatory (e.g. predictive of sample
509 type), then those should be prioritised for manual post-processing checks to ensure that the
510 differences are real and the data is of good quality.

511 **Discussion**

512 Validation is defined as the process of proving that any procedure, process, equipment,
513 material, activity or system performs as expected within defined acceptance criteria under a
514 given set of conditions, and that the performance characteristics of the procedure meet the
515 requirements for the intended analytical applications [67, 68]. Although implementing fail/pass
516 criteria advised by bioanalytical method validation guidelines have provided a useful degree of
517 standardisation and consistency between regulated laboratories, new advances in technology,
518 multiplexing, and metabolomics studies require tiered and/or 'fit-for-purpose' approaches [69]
519 for pragmatic/practical use.

520 Pre-determined or fixed acceptance criteria are established and appropriate for
521 validated assays (Tier 1); however, for qualified, research, and screening methods (Tiers 2-4),
522 it may be appropriate to define these after the method performance experiments have been
523 conducted to fine-tune the assay to the required acceptance criteria. Minimally, it is expected
524 that *a priori* acceptance criteria can be relaxed for the higher tiers if such method
525 performance still supports the intended use of the data and ultimately supports the necessary
526 decisions that will be made [42].

527 Validation beyond the intended use of the data means significant re-work, loss of time
528 and increased cost in the blind pursuit of absolute requirements. For metabolomics at its current
529 state of development, what is required is the definition of a simple, pragmatic and easy- to-follow
530 framework that reflects realistic and practical needs that allow for the most efficient practices.
531 For instance, an assay that does not pass the criteria for full validation but, nevertheless, fulfils
532 the essential requirements for linearity, accuracy, precision, LLOQ and carryover criteria may
533 be devised. In that case, guidance should focus on minimum requirements. Specifications of
534 merit might include: linearity with an LLOQ set as first calibrant, accuracy, precision and
535 carryover.

536 Overall, the guidelines for assays developed for drugs that have been devised by
537 regulatory authorities to ensure safety and efficacy in humans represent a 'gold standard' that
538 may not be required for many types of targeted and untargeted metabolomics applications.
539 This is not to suggest that metabolic phenotyping methods should not be developed to the
540 standards necessary to provide reliable and scientifically valid data but to suggest that the
541 use of tiered approaches linked to the type of investigation (i.e. discovery, hypothesis
542 validation, biomarker/panel, and/or qualification stages) should drive the level of validation
543 performed. A number of intricate analytical factors (e.g. pre-analytical factors) defining core
544 assay expectations, and setting acceptable assay performance criteria, should be taken into
545 account for assessing the reliability and quality of metabolomics assays. Our MAP/UK
546 consensus framework provides a bench guide for the two major categories of validation and
547 qualification of targeted metabolomics analysis that have been described in Table 2.

548 **Conclusions**

549 Metabolomics has the potential to lead advances in the discovery of clinically and
550 toxicologically relevant biomarkers, yet the lack of harmonisation at different levels throughout
551 the whole metabolomics pipeline from study design, sample handling, biobanking, metabolite
552 quantification to data analysis remains an issue that needs to be addressed. Metrological
553 tracability and future development of certified matrix reference materials similar to National
554 Institute of Standards and Technology reference standards (NIST SRM 1950) [70], and standard
555 calibration mixtures should be established and harmonized within both the research and
556 regulatory communities. The MAP/UK consortium proposes the pragmatic development of a
557 'fit-for- purpose' 4-tiered framework for assessing the reliability of metabolomics assays via a
558 decision-making process and adaptation of existing drug regulatory guidance. The required
559 level of analytical rigour and/or qualification that bioanalytical methods need to show in order
560 to achieve scientifically valid studies in metabolomics has been considered. This framework is

561 intended to guide bioanalysts and to facilitate improved communication between the research
562 and regulatory communities, in order to enable the establishment of appropriately qualified
563 targeted metabolomics assays to meet the needs of multiple applications of this technology in
564 the regulatory sciences. Ultimately, we hope that such a community-initiated framework can
565 accelerate the application of metabolomics in regulatory applications and achieve acceptance
566 via its consensus use.

567 **Acknowledgements**

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

570

571 **Author Contributions**

572 The necessity of developing a pragmatic and 'fit-for- purpose' 4-tiered framework for assessing
573 the reliability of metabolomics assays via a decision-making process and adaptation of existing
574 drug regulatory guidance, were discussed by MAP/UK consortium members. The Initial draft
575 was written by dr. Sarir Sarmad, under supervision and guidance of Professor Toru Suzuki.
576 Further direction and input were given by dr. Matthew Lewis, and Professor Ian Wilson. We
577 received further contribution and scientific advice from Professor Mark Viant, Professor
578 Royston Goodacre, and Professor Warwick Dunn. Professor Mark Viant and Sarir shaped the
579 framework. We received further advice and constructive comments from Katie Chapell,
580 Professor Jules Griffin, Professor Valerie O'Donnell, and Brendon Naicker. Dr. Sarir Sarmad
581 and Professor Suzuki made substantial contribution in writing the collective ideas of our
582 expertise, and further formatting and revision of the paper.

583

584 **Competing Financial Interests**

585 The authos declare no competing interests as defined by Nature Research, or other interests that
586 might be perceived to influence the interpretation of the article.

587

588 **Data availability**

589

590

591 **Supplementary Information**

592 Supplementary Table 1. Comparison of the FDA, EMA and ICH M10 draft guidelines.

593

594 Supplementary Table 2. Definition, methodology, and acceptance criteria of validation
595 parameters.

References

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

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

3- Wei, Ru, Guodong L., and Seymour. A.B. High-Throughput and Multiplexed LC/MS/MRM Method for Targeted Metabolomics. Anal. Chem. **82**, no. 13 (2010): 5527-33. https://doi.org/10.1021/ac100331b.

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

5- Guijas, C., Montenegro-Burke, J.R., Warth, B., Spilker, M.E. and Siuzdak, G. Metabolomics Activity Screening for Identifying Metabolites That Modulate Phenotype. *Nat. Biotechnol*. **36**, no. 4 (2018): 316-20. https://doi.org/10.1038/nbt.4101.

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

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

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

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

10- Ribbenstedt, A., Ziarrusta, H., and Benskin, J.P. 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.

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

12- 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 (2011): 1060-83. https://doi.org/10.1038/nprot.2011.335.

13- Goodman, J., *et al.* Update to the European Bioanalysis Forum Recommendation on BiomarkersAssays; Bringing Context of Use into Practice. *Bioanal.* **12**, no. 20 (2020): 1427-37. https://doi.org/10.4155/bio-2020-0243.

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

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

16- Palmer, J. 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.

17- Michel, F., Pirotte, B., Fillet, M., and Tullio, P. Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* **59**, no. 19 (2016): 8649-66. https://doi.org/10.1021/acs.jmedchem.5b01335.

18- 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.

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

20- 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 (2018): 2001-2017. https://doi.org/10.1194/jlr.S087163.

21- 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 (2019): 14407-16. https://doi.org/10.1021/acs.analchem.9b02908.

22- König, R., Cave, A., Goldammer, M. and Meulendijks, D. 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.

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

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

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

26- Broadhurst, D., *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): 72. https://doi.org/10.1007/s11306-018-1367-3.

27- Hyötyläinen, T., Ahonen, L., Pöhö, P. and Orešič, M. Lipidomics in Biomedical Research-Practical Considerations. Biochimica et Biophysica Acta. (BBA) – *Mol. and Cell Biol of Lipids* **1862**, no. 8 (2017): 800-03. https:// doi.org/10.1016/j.bbalip.2017.04.002.

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

29- Theodorsson, E. 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.futurescience.com/doi/abs/10.4155/bio.11.311.

30- 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.

31- 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.

32- 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.

33- 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.

34- 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_Manufacturing_Practice_for_Drugs.pdf. Accessed March 2022.

35- 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.

36- ANVISA Guide for validation of analytical and bioanalytical methods. Resolution RE no.899, 23 May 2003 Brazilian Sanitary Surveillance Agency, BrazilGoogle Scholar.
Anvisa. Guide for Validation of Analytical and Bioanalytical Methods. ANVISA Brazil (2012).

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

38- 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.

39- 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.

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

41- 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 (2010): 1185-94. https://doi.org/10.4155/bio.10.90.

42- 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.

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

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

45- 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.

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

47- 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 (2015): 2387-98. https://doi.org/10.4155/bio.15.168.

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

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

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

51- Lewis M, *et al.* An Open Platform for Large Scale LC-MS-Based Metabolomics.*ChemRxiv.* Cambridge Open Engage; 2022; DOI:10.26434/chemrxiv-2022-nq9k0https://chemrxiv.org/engage/chemrxiv/article-details/61ebd6fa0716a8529e3823dc.

52- Uwe, C. *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): 32-38. https://doi.org/10.1016/j.microc.2012.02.011.

53- Thway, T.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.

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

55- Thakare, R., *et al.* Quantitative analysis of endogenous compounds. *J Pharm Biomed Anal.* **128** (2016): 426-37. https://doi.org/10.1016/j.jpba.2016.06.017.

56- Tsikas, D. Bioanalytical methodvalidation of endogenous substances according to guidelines by the FDA and other organizations: basic need to specify concentration ranges. *J Chromatogr B Analyt Technol Biomed Life Sci***1093-1094** (2018): 80-81. https://doi.org/ j.jchromb.2018.07.005.

https://www.sciencedirect.com/science/article/pii/S1570023218308419.

57- 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): 1349-60. https://doi.org/10.4155/bio-2018-0105.

58- European Medicine Agency, ICH guideline Q2(R2) on validation of analytical procedures, EMA/CHMP/ICH/82072/2006. https://www.ema.europa.eu/en/ich-q2-r1-validation-analytical-procedures-text- methodology. Accessed Sept 2020.

59- WHO technical series report No 937. Analytical method validation (2018) https://pharmaguidances.com/analytical-method-validation-as-per-who-technical-report-series-no-937/. Accessed Sept 2021.

60- Araujo, P. Key aspects of analytical method validation and linearity evaluation. *J Chromatogr B Analyt Technol Biomed Life Sci.* **877**, no. 23 (2009): 2224-2234.

61- Boque, R. and Heyden, Y.V.The limit of detection, *LCGC Europe*, **22**, no.2 (2009): 82-85. https://www.chromatographyonline.com/view/limit-detection.

62- Desimoni, E., Brunetti, B. Signal to noise ratio; Limit of detection; Standard error of the regression; Chromatographic/Voltammetric/Spectroscopic signals. *Pharm. Anal.ACTA.* **6**, no. 4 (2015): 355. doi:10.4172/2153-2435.1000355.

63- U.S. Pharmacopeia, document 1225, Validation of compendial methods, *Pharmacoepial Forum*, **31**, no. 2 (2005): 549. http://www.pharmacopeia.cn/v29240/usp29nf24s0_c1225.html.

64- 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.

65- European, Commission, Centre Joint Research, P. Robouch, J. Stroka, J. Haedrich, A. Schaechtele, and T. Wenzl. Guidance Document on the Estimation of LOD and LOQ for Measurements in the Field of Contaminants in Feed and Food. Publications Office, 2016. doi:doi/10.2787/8931.

66- Commission, E. 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.

67- Watson, D.G. Pharmaceutical Analysis. 5thEdition. Oxford: Elsevier churchill livingstone, 2020. Paperback ISBN: 9780702078071.

68- Bliesner, D.M. Validating Chromatographic Methods: A Practical Guide, Hoboken: Wiely Interscience, 2006. ISBN: 978-0-471-74147-3.

69- Lee, J. W., *et al.* Fit-for-Purpose Method Development and Validation for Successful Biomarker Measurement. *Pharm Res.* **23**, no. 2 (2006): 312-28. https://doi.org/10.1007/s11095-005-9045-3.

70- Yamil, S. *et al.* Metabolite Profiling of a NIST Standard Reference Material for Human Plasma (SRM 1950): GC-MS, LC-MS, NMR, and Clinical Laboratory Analyses, Libraries, and Web-Based Resources. *Anal. Chem.*

85, no. 24 (2013): 11725-31. https://doi.org/10.1021/ac402503m.