Biomarker Qualification at the European Medicines Agency: A Review of Biomarker Qualification Procedures From 2008 to 2020

Elisabeth Bakker[1,‡], Natalie M. Hendrikse[2,‡], Falk Ehmann[2], Daniëlla S. van der Meer[1,3], Jordi Llinares Garcia[2,†], Thorsten Vetter[2], Viktoriia Starokozhko[1,3] and Peter G.M. Mol[1,3,4,*]

Regulatory qualification of biomarkers facilitates their harmonized use across drug developers, enabling more personalized medicine. This study reviews various aspects of the European Medicines Agency's (EMA's) biomarker qualification procedure, including frequency and outcome, common challenges, and biomarker characteristics. Our findings provide insights into the EMA's biomarker qualification process and will thereby support future applications. All biomarker-related "Qualification of Novel Methodologies for Medicine Development" procedures that started from 2008 to 2020 were included. Procedural data were extracted from relevant documents and analyzed descriptively. In total, 86 biomarker qualification procedures were identified, of which 13 resulted in qualified biomarkers. Whereas initially many biomarker qualification procedures were linked to a single company and specific drug development program, a shift was observed to qualification efforts by consortia. Most biomarkers were proposed (*n* = 45) and qualified (*n* = 9) for use in patient selection, stratification, and/or enrichment, followed by efficacy biomarkers (37 proposed, 4 qualified). Overall, many issues were raised during qualification procedures, mostly related to biomarker properties and assay validation (in 79% and 77% of all procedures, respectively). Issues related to the proposed context of use and rationale were least common yet were still raised in 54% of all procedures. While few qualified biomarkers are currently available, procedures focus increasingly on biomarkers for general use instead of those linked to specific drug compounds. The issues raised during qualification procedures illustrate the thorough discussions taking place between applicants and regulators—highlighting aspects that need careful consideration and underlining the importance of an appropriate validation strategy.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ To our knowledge, the biomarker qualification procedure has been reviewed using qualified biomarkers only and no analysis of issues raised in qualification procedures has been performed.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This study reviews the utilization of the European Medicines Agency (EMA) biomarker qualification procedure, regarding frequency and outcome (confidential qualification advice or published qualification opinion), but also other characteristics including type of applicant, disease area, and common issues raised during qualification.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ This study provides a broad overview of characteristics of both qualified and nonqualified biomarkers evaluated by the EMA. This information is not publicly available and cannot be accessed elsewhere.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ The broad use of qualified biomarkers may expedite drug development, especially in areas where a high unmet medical need exists, and can support the broad implementation of precision medicine in research and clinical practice. This study may stimulate future applicants by providing considerations for these procedures, which may be useful when preparing a biomarker qualification.

[1]Department of Clinical Pharmacy and Pharmacology, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands; [2]European Medicines Agency, Amsterdam, The Netherlands; [3]College ter Beoordeling van Geneesmiddelen–Medicines Evaluation Board, Utrecht, The Netherlands; [4]Scientific Advice Working Party, European Medicines Agency, Amsterdam, The Netherlands. *Correspondence: Peter G.M. Mol (p.g.m.mol@umcg.nl)

† Deceased July 17, 2021.
‡ These authors contributed equally.

Linked article: This article is linked to Biomarker Qualification at the European Medicines Agency: A Look Under the Hood by Berman, S. and Siegel, J., *Clin Pharmacol Ther.*, **112**, 28–30 (2022).

Received August 10, 2021; accepted February 2, 2022. doi:10.1002/cpt.2554

Over the past two decades, the discovery, validation and qualification of biomarkers for new medicine development and assistance in regulatory decision making has become increasingly important.[1] The Biomarkers, Endpoints and Other Tools (BEST) resource describes a biomarker as "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions."[2] From a precision medicine perspective, biomarkers can be useful to identify patients who are likely to benefit from a certain therapeutic agent or exclude patients that are likely to be harmed by that therapeutic agent, i.e., to predict an individual's response (*predictive biomarkers*). Furthermore, biomarkers can be used to track disease progression or predict whether an individual is at increased risk of experiencing a certain clinical event (*prognostic biomarkers*). Therefore, using biomarkers to select the right treatment for the right patient can increase the chance of treatment success in clinical development and clinical practice.[3] Successful implementation of new biomarkers depends on their discovery as well as a thorough validation process. Endorsement of validated biomarkers by regulators in the form of qualification facilitates their harmonized use across drug developers, expedites drug development, and may be an important step in adoption of precision medicine in clinical practice.[1,4,5]

One of the core recommendations of the European Medicines Agency (EMA) Regulatory Science Strategy to 2025 is to support developments in precision medicine, biomarkers, and 'omics. Underlying actions that are suggested to achieve this goal include "enhancing early engagement with novel biomarker developers to facilitate regulatory qualification" and "critically review the EMA's biomarker validation process, including duration and opportunities to discuss validation strategies in advance, in order to encourage greater uptake and use."[6] The EMA introduced the "Qualification of Novel Methodologies for Medicine Development" in 2008.[7] This procedure provides support to "the qualification of innovative development methods for a specific intended use in the context of research and development into pharmaceuticals," such as biomarkers, and is provided by the EMA's Committee for Medicinal Products for Human Use (CHMP) based on recommendations by the Scientific Advice Working Party (SAWP).[7] The first biomarkers to obtain such regulatory approval were qualified through a joint pilot procedure by the EMA and the US Food and Drug Administration (FDA) that took place from 2007 to 2008.[8] A qualification procedure can result in a qualification advice (QA) or qualification opinion (QO). Although applicants are asked to indicate the anticipated scope (QA/QO) when submitting their request, the outcome is determined during the course of the procedure and based on the level of evidence as assessed by the qualification team.[9] Targeting early stages of the qualification exercise, the confidential QA aims to facilitate biomarker validation by providing a platform for discussion and reaching consensus between the EMA and the applicant on scientific rationale, proposed context of use (CoU), preliminary data, and evidence generation strategy.[10] Multiple QAs may precede a QO, which is only issued when the evidence is deemed adequate to support the biomarker's targeted CoU. Before it is finally adopted, a draft QO is first published on the EMA website for 2 months of public consultation, to confirm validity with the scientific community.[11] In case a QA is pursued, a public Letter of Support may be proposed by EMA when the novel methodology is shown to be promising based on preliminary data, which aims to encourage data sharing and facilitate further studies towards qualification.[9]

Many current publicly and privately sponsored biomarker projects, such as Innovative Medicines Initiative (IMI) consortia, including the Biomarker Enterprise to Attack Diabetic Kidney Disease (BEAT-DKD) consortium, the HORIZON program, and the Biomarker Commercialization (BIC) consortium, aim for regulatory qualification and broad use of newly discovered biomarkers.[12–14] It is currently unclear, however, how much the EMA qualification procedure is used and how often it leads to qualified biomarkers. In addition, available guidance for applicants is mainly focused on the procedural aspects[9] and, although some essential considerations for successful qualification are provided,[15] evidentiary requirements are mainly driven by the proposed biomarker-specific CoU and cannot be generalized. In this collaborative effort between academic researchers and the EMA, we analyzed the use of this procedure in terms of frequency and outcome (QA or QO), but also with respect to other characteristics, such as type of applicant, disease area, and issues that are commonly raised. The findings presented may provide insights into the biomarker qualification process and could support future applications.

METHODS

Selection of procedures

For this study, all finalized "Qualification of Novel Methodologies for Medicine Development" procedures that started between January 1, 2008 and December 31, 2020 (i.e., since the start of the procedure), were obtained from an internal EMA database (**Figure 1**). For all procedures, the lists of issues and final advice letters were screened for their eligibility to be included in the study. Joint procedures with the FDA were excluded in line with the provisions of the confidentiality arrangement between the FDA and the EMA. Procedures on nonbiomarker methodologies, such as patient-reported outcomes, registries and other data platforms, clinical trial methodologies, and statistical or modeling methods were also excluded (**Figure 1**). The selection process was performed by six researchers (E.B., N.M.H., D.S.v.d.M., T.V., V.S., and P.G.M.M.) and in case of disagreements, these were discussed until a consensus was reached. The unit of analysis in this study is the qualification procedure, where a single procedure may comprise several (candidate) biomarkers, either for use as a single biomarker or as a panel of biomarkers. Related procedures, such as follow-ups, were treated independently.

Data extraction, characterization, and interpretation

The following characteristics of the *procedure* were extracted: the outcome of the procedure (QO or QA), whether the procedure was a follow-up procedure, the type of applicant (consortium or company), and whether the biomarker procedure was linked to a clinical development program of a specific drug of that company. The duration of the procedure was also extracted, which was defined as the number of months from the start of the procedure until adoption of the FAL by the CHMP. For QOs, a distinction was made between the duration until adoption of the draft opinion for public consultation and adoption of the final opinion.

The following characteristics of the *biomarkers* were extracted: the CoU proposed by the applicant, the biomarker type, and disease area for which the biomarker is intended. Based on the CoU claimed by the applicant, biomarkers were assigned to the following broad CoU categories: (i)

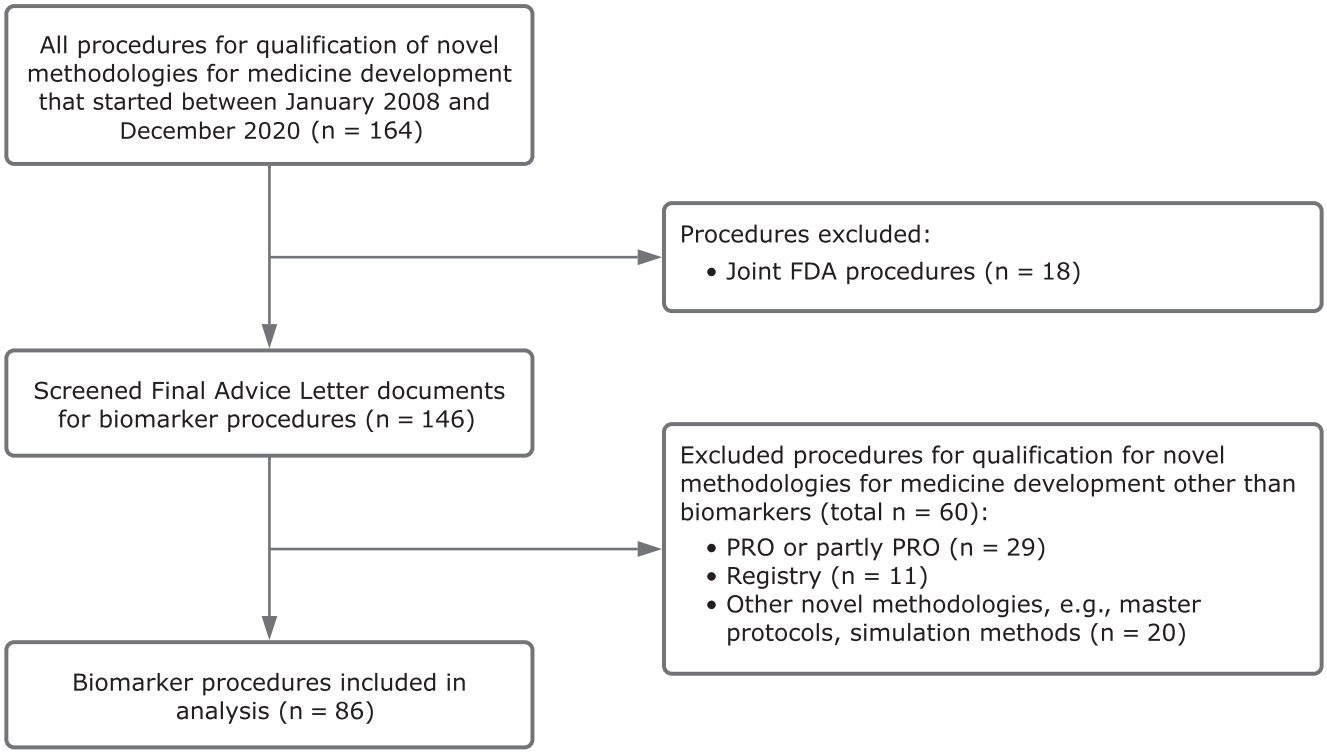


Figure 1 Selection of biomarker-related qualification procedures at the EMA (European Medicines Agency). All procedures for *qualification of novel methodologies* that started between January 2008 and December 2020 were screened for eligibility according to the flowchart. FDA, US Food and Drug Administration; PRO, patient-reported oucome.

patient selection, stratification, and/or enrichment, (ii) efficacy, and (iii) safety. In addition, every biomarker was assigned to a category based on its function, according to definitions that were adapted from the BEST Resource, as outlined in **Table 1**.[2] A single biomarker procedure could be assigned to more than one intended CoU, multiple biomarker categories, biomarker types, and/or disease categories.

Lastly, we characterized frequently arising *issues* in the biomarker qualification procedures, i.e., topics of controversy that required further discussion between regulator and applicant. These issues were identified in an iterative process that was performed as follows: two researchers from the University Medical Center Groningen (E.B. and V.S.) independently identified issues in a sample set of ten randomly selected procedures, which were compared, and related issues were grouped. One EMA researcher (N.M.H.) identified issues in another sample set of eight procedures, which were discussed and grouped in collaboration with two additional EMA researchers (F.E. and T.V.). Then, a pilot of ten procedures was performed by three researchers (E.B., N.M.H., and V.S.), in which the issue groups from the two separate exercises were merged and final categories (all coauthors) were defined that covered all issues identified in the ten pilot procedures. Eventually, eight different categories of qualification issues were identified: biomarker properties, general study design, assay validity, evidence, data analysis/statistics, anchoring, rationale, and CoU (**Table 2**). Subsequently, all remaining procedures were reviewed, and issues were identified and assigned to one of these eight predefined categories by two independent researchers (E.B. and N.M.H.). Differential classifications were resolved by consensus. A final quality check of the assigned issue categories was performed by four other researchers (F.E., T.V., V.S., and P.G.M.M.), who each assessed the assigned issue categories for a sample of two procedures each. For all procedures, issues were extracted from the lists of issues and final advice letters, more specifically from the numbered issue section and the CHMP answers, respectively. All extracted data were analyzed descriptively.

RESULTS

Biomarker-related qualification procedures from 2008 to 2020

Between 2008 and 2020, the EMA handled a total of 164 qualification procedures, of which 86 were biomarker related and included in the analysis (see **Figure 1**). Of the 18 excluded joint FDA procedures, 16 were biomarker related, of which one resulted in a QO. The number of biomarker-related procedures fluctuated between 2 and 12 per year and included 13 QOs (**Figure 2a**). Out of the total of 86, 16 were follow-ups to previous procedures concerning the same candidate biomarker(s) and/or CoU, resulting in seven QOs. The maximum number of iterative interactions was three (initial procedure and two follow-ups), which occurred three times. The median duration of QAs was 4 months (minimum (min.) 1, maximum (max.) 11). The median duration of QOs was 11.5 months in total (min. 4, max. 31); 6 months (min. 3, max. 25) from the start of the procedure until adoption of the draft opinion, and 6 months (min. 2, max. 18) from the start of the public consultation until adoption of the final opinion. Whereas up to 2014 when most biomarker qualification procedures were linked to a single company and a specific drug development program, most procedures in later years targeted a general use of biomarkers and were increasingly submitted by consortia (**Figure 2b**).

Context of use and disease areas

Most biomarker procedures (n = 45), including nine QOs, were assigned to the *patient selection*, *stratification*, *and/or*

Table 1 CoU categories, associated biomarker categories, and main diseases and disease areas

CoU category	QA	QO	Biomarker category	QA	QO	Definition	Main diseases/disease areas	QA	QO
Patient selection, stratification, and/or enrichment	36	9	Diagnostic/ Stratification	23	6	Confirms or detects the presence of a condition or disease of interest or identifies individuals with a subtype of the disease (diagnostic) or can be used to divide the population into subgroups (stratification).	Autism spectrum disorder	10	
							Alzheimer's disease	3	4
							Crohn's disease	1	
							Drug-induced injury	2	
							Kidney	1	
							Liver	1	
							NASH/NAFLD	4	
							Parkinson's disease	2	1
							Other	1	1
			Prognostic	19	8	Indicates the likelihood of a clinical event, disease recurrence, or disease progression in patients with a confirmed disease or medical condition of interest.	Alzheimer's disease	4	4
							Autism spectrum disorder	8	
							Parkinson's disease	2	1
							Diabetes mellitus type 1	1	1
							NASH/NAFLD	1	
							Oncology	1	
							Other	2	2
			Predictive	11	3	Identifies individuals that are more likely to experience a favorable or unfavorable effect from exposure to a certain treatment.	Alzheimer's disease	2	3
							Drug-induced injury	2	
							Kidney	1	
							Liver	1	
							Oncology	1	
							Schizophrenia	1	
							Other	2	

(Continued)

Table 1 (Continued)

CoU category	QA	QO	Biomarker category	QA	QO	Definition	Main diseases/disease areas	QA	QO
Efficacy (pharmacodynamic/response)	33	4	(Co)primary endpoint	10	1	A precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address the primary research question.	Alzheimer's disease	4	
							Autism spectrum disorder	2	
							Crohn's disease	1	
							Multiple sclerosis	1	
							Oncology	1	
							Other	1	1
			Surrogate endpoint	9	0	An endpoint that is used in clinical trials as a substitute for a direct measure of how a patient feels, functions, or survives.	Duchenne muscular dystrophy	2	
							Oncology	4	
							Other	3	
			(Key) secondary endpoint	2	0	A precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address a secondary research question.	Parkinson's disease	1	
							X-linked retinitis pigmentosa	1	
			Type of endpoint not clearly predefined	2	2	Applicant did not indicate one specific type of endpoint for qualification, e.g., "key endpoints" or "either primary or secondary endpoint."	Duchenne muscular dystrophy		1
							Multiple sclerosis		1
							Parkinson's disease	1	
							Schizophrenia	1	
			PD/response, no endpoint	10	1	Changes in response to exposure to a medicinal product.	Alzheimer's disease	4	
							Autism spectrum disorder	1	
							Parkinson's disease	3	
							Multiple sclerosis	1	
							Other	3	1
Safety	11	1	Safety/Monitoring	11	1	Assesses the status of a medical condition or disease by means of sequential measurements and may indicate the likelihood, presence, or extent of toxicity as a result of exposure to a medicinal product.	Autism spectrum disorder	1	
							Drug-induced injury	9	1
							Kidney	2	1
							Muscle	2	
							Vascular	2	
							CNS	1	
							Liver	1	
							Skin	1	
							Other	1	

All procedures were assigned one or several CoU categories and biomarker categories based on information from the List of Issues and Final Advice Letter. The biomarker subtypes were adapted from the BEST tool.[2] The main diseases and disease areas for which the biomarkers were proposed are presented per category. In the pharmacodynamic/response category one procedure covered MS (multiple sclerosis), PD (Parkinson's disease), and several other diseases. This procedure was counted for MS, PD, and "other."

CNS, central nervous system; CoU, context of use; NASH/NAFLD, Duchenne muscular dystrophy steatohepatitis / Duchenne muscular dystrophy fatty liver disease; PD, pharmacodynamic; QA, qualification advice; QO, qualification opinion.

enrichment CoU category. This was followed by 37 procedures with an *efficacy* and 12 procedures with a *safety* CoU, including four QOs and one QO, respectively (**Table 1**). In the largest CoU category, most biomarkers were aimed at diagnosis/stratification for the purpose of enrichment of patient populations for clinical trials. In the *efficacy* category, although all biomarkers were pharmacodynamics/response markers, only a selection of those were aimed to be qualified as surrogate (n = 9), (co) primary (n = 11) or "key" secondary endpoints (n = 2). In four applications, the applicant did not indicate one specific type of endpoint to be qualified (**Table 1**).

The most common diseases and disease areas per CoU category are listed in **Table 1**. Alzheimer's disease is the most common disease across all CoU categories (17 out of 86 procedures), as well as the main disease in CoU categories *patient selection*, *stratification*, *and/or enrichment*, and *efficacy*. Nearly all procedures (n = 15) related to Alzheimer's disease occurred before 2015 and only two additional procedures started in 2017 (**Figure 3**). Most biomarkers that were intended as safety markers were related to drug-induced organ injury (e.g., kidney injury, liver injury, and vascular injury). As opposed to biomarkers in Alzheimer's disease, safety markers for drug-induced injury were submitted throughout the studied period. Qualifications of biomarkers intended for Parkinson's disease did not take place until 2015, after which at least one procedure was started per year. All five biomarker qualifications for nonalcoholic fatty liver disease / nonalcoholic steatohepatitis biomarkers started in the same year (2019).

The majority of biomarkers were soluble biomarkers (n = 30), imaging biomarkers (n = 29), or performance scores (n = 29) (**Figure 4**). Between the different CoU categories, some differences were noted: Whereas imaging and soluble biomarkers are most frequent among the *patient selection*, *stratification*, *and/or enrichment* biomarkers (n = 19 and n = 16, respectively), performance scores are most frequent among the *efficacy* biomarkers (n = 18). In relation to the remaining biomarker types, seven clinical scores, two functional biomarkers, one histology biomarker, and one genetic biomarker were identified.

Issues that were raised during qualification procedures

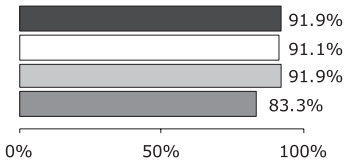
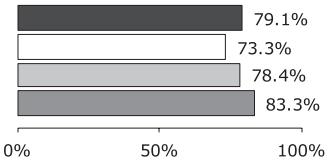
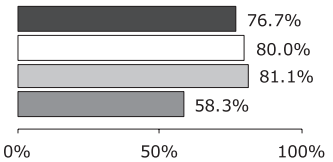
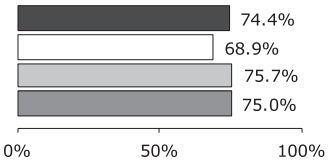
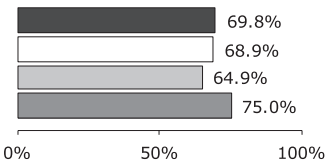
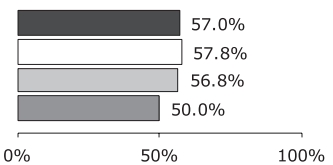
In all procedures, the issues raised by the CHMP during qualification covered several of the previously defined categories and all identified issues could be assigned to one of those categories. The most commonly raised issues related to biomarker properties (92%), general study design (79%), and assay validity (77%) (**Table 2**). The first category often involved discussions around chosen cutoff values and the minimal clinically relevant change. Reoccurring issues with respect to study design typically revolved around the study population, as illustrated in **Table 2**, as well as study set up in terms of options for generating sufficient data for biomarker discovery and validation. Such a prespecified "learn and confirm" concept was often missing, both in early-stage projects and advanced qualification efforts, and was categorized as related to evidence, since the evidence generation plan was not deemed acceptable. Issues related to assay validity were also frequently raised and referred to sensitivity and specificity of the assay or test used to measure the biomarker, as well as interlab and intralab reproducibility and standardization (**Table 2**). Issues around the proposed CoU and rationale were raised in 54% of all procedures (**Table 2**). These categories involved important discussions especially in early-stage projects, as they are key to determining the strategy for biomarker validation and qualification. No difference in distribution of issues was observed between biomarkers that were linked to a specific drug compound and those that were not. Between CoU categories, the distributions of issues were rather similar, with procedures in the *safety* category deviating most from the average and/or other COU categories in nearly all of the issue categories. In particular, fewer issues were raised about assay validity for safety biomarkers compared with the other COU categories (**Table 2**).

DISCUSSION

In this study we identified only 13 qualified biomarkers out of a total of 86 biomarker-related qualification procedures that started between January 2008 and December 2020 at the EMA. Sixteen of these were follow-ups to previous procedures, of which seven resulted in QOs. Whereas most biomarker procedures were intended to support patient selection (n = 45) or measure treatment outcome (n = 37), fewer were aimed at monitoring drug safety (n = 12). In terms of disease areas, the biomarker qualification procedures were dominated by biomarkers developed for Alzheimer's disease (AD), autism spectrum disorder, and disease-independent drug-induced organ injury, and these applications seemed to cluster in time. Perhaps not surprisingly, most procedures concerned soluble (n = 30) and imaging biomarkers (n = 29), which are often relatively easy to obtain concomitantly with regular clinical trial procedures. More surprising, however, was the high number of procedures related to performance scores (n = 29) and low number of procedures related to genetic biomarkers (n = 1). Finally, various issues were raised in the biomarker qualification procedures, which were mostly related to biomarker properties, study design, assay validity, and evidence. Issues related to the rationale behind the suggested biomarker, as well as the suggested CoU were raised in >50% of the procedures, highlighting the importance of these aspects for determining the biomarker validation strategy.

In general, the use of the EMA qualification of novel methodologies procedure is still modest, also when it comes to biomarkers. Out of all 164 qualification procedures that started between 2008 and 2020, approximately half are related to biomarkers (n = 86)—almost seven per year on average. Interestingly, the number of biomarker qualification procedures has not significantly increased since 2015 (**Figure 2**), despite the growing interest for biomarker identification and development following the rise of precision medicine.[1] However, the number of procedures in 2019 and 2020 were relatively high compared with the previous 3 years, which might indicate a positive trend for the years to come. As of December 2021, 19 Letters of Support for the included procedures have been published on the EMA website, and it may be expected that applicants return for future qualification. A possible explanation for the low number of qualified biomarkers is that biomarker validity can also be assessed on a case-by-case basis as part of a marketing authorization application (MAA) evaluation, and regulatory qualification of biomarkers is not a requirement for their use

Table 2 Issues identified in the qualification procedures

Category		Description	Example
Biomarker properties		Issues related to the properties of the biomarker *e.g., issues related to sensitivity, specificity, baseline/ reference measurements, thresholds of detection, cutoff values, clinically relevant change*	"… a full evaluation of the prognostic properties including a potential assessment of its clinical usefulness/utility would also require the generation of ROC analyses with the calculation of sensitivity and specificity and the derived parameters of PPV and NPV. While this is partially foreseen for the long-term outcome evaluations, the evaluations for the "intermediate endpoint" evaluations do currently not foresee such."
General study design		Issues related to the design of the studies that are part of the qualification exercise (either planned or performed) *e.g., study type and study population, study period*	"Inclusion of healthy controls would be acceptable in proof-of-concept and phase II trials, but their inclusion in confirmatory trials is questionable. Their inclusion in the autopsy study, as a means of assuring recruitment of at least some regions negative for β-amyloid deposition, may be helpful."
Assay validity		Issues related to the proposed test, assay, or methodology to measure the biomarker *e.g., validation data, (interlab/intralab) reproducibility, sensitivity and specificity of the assay, sampling techniques, image processing, protocols*	"Validation data should be provided supporting the appropriateness of the analytical methods: accuracy, linearity (calibration range), precision, selectivity (including consideration of sample interference) and standard/sample solution stability should be demonstrated … it should be confirmed that the method is reproducible (cross-validated) between laboratories."
Evidence		Issues related to the evidence (data) provided to support the claims with respect to the biomarker itself (not the assay) *e.g., data is not robust, no or inappropriate external validation (confirmatory studies, cross-validation), contradictory data, etc.*	"Regarding use as primary endpoint for pivotal trials in this setting, although promising, more robust data gained with additional patients and longer follow-up could be beneficial."
Data analysis/statistics		Issues related to analysis of data that are used to support the claims *e.g., statistical methods, handling of missing data or confounders, etc.*	"The use of the Tobit regression should be explained. It is not clear what makes the Tobit regression the method of choice. The technical procedure in and the properties of the Tobit regression and the parametrization of the factors included should be discussed. In addition, the model fit of this analysis should be discussed. Alternative analyses should also be employed."
Anchoring		Issues related to the additional value of the biomarker *e.g., compared with a gold standard or clinically relevant outcome*	"The Applicant is invited to further discuss the added clinical usefulness of [the biomarker] when compared with [the current gold standard], which would justify its qualification."

(Continued)

Table 2 (Continued)

Category	Description	Example
Rationale	Issues related to the choice for the specific biomarker *e.g., based on the biological rationale, availability, or accessibility (invasiveness), etc.*	"Biomarkers for diagnostic, predictive, and negative predictive use may have different abilities in detecting these endpoints. Please discuss the scientific rationale for the biomarkers selected in relation to these different functions."
Context of use	Issues related to the definition of the context of use *e.g., proposed context of use not specific enough in terms of intended population, clinical setting, therapeutic area, etc.*	"On the basis of available clinical experience, the Applicant is invited to elaborate on the intended clinical use of IS [ingestible sensor] with particular reference to the clinical setting (GP, specialist, private) and the therapeutic area."

The bar charts in the left column show the percentage of all procedures (dark gray), and the percentage of procedures in the CoU categories *patient selection, stratification, and/or enrichment* (white), *efficacy* (light gray), and *safety* (mid gray) that contain issues in the respective categories.
CoU, context of use; GP, general practitioner; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic.

in drug development. This may be reflected in the development context of the reviewed biomarkers throughout the years, with a shift away from qualification procedures that were linked to a specific drug compound over time (**Figure 2b**). Applicants, especially companies, aiming at qualifying a biomarker in the context of a clinical development program for a specific drug compound, may opt for a Scientific Advice (SA) procedure instead. In contrast to the qualification procedures that aim to publish the evidence supporting a qualified biomarker, these product-related SA procedures remain confidential. Although SA may include questions around the acceptability of the use of established biomarkers or plans for novel biomarker validation, biomarker validity itself will not be assessed until an applicant submits its product for MAA evaluation. The aspect of confidentiality may be a determining factor in a company's strategy for regulatory interactions, including those related to biomarkers.[16] The increasing number of procedures initiated by consortia can only be encouraged, as these collaborations typically have the financial and intellectual resources (e.g., access to extensive clinical data) to move the broad implementation of precision medicine forward by qualifying biomarkers that will become publicly available.[16,17] In comparison, as of February 2020, 49 biomarkers had been submitted to the FDA Biomarker Qualification Program (BQP), of which eight resulted in qualified biomarkers.[16] Overall, the BQP is similar to the EMA biomarker qualification procedure in the sense that biomarkers can be qualified for a specific CoU and that qualification data become publicly available. Out of the 49 BQP procedures, 16 were conducted in collaboration with the EMA but were not included in this study due to confidentiality arrangements. As only one of them led to a QO, exclusion of these procedures did not substantially influence the findings of our study. Notably, the FDA has provided draft guidance on its evidentiary requirements for biomarker qualification.[1,18]

The majority of qualification procedures concerned biomarkers to be used for patient selection, stratification, and/or enrichment ($n = 45$, **Table 1**), which also included the highest number of QOs (n = 9). From an industry perspective, this may reflect the need for increased efficiency in clinical trials, which can be achieved by selecting a study population in which establishing efficacy of a drug is more likely. From a public health perspective, on the other hand, these numbers highlight the increasing ability to implement precision medicine, i.e., identifying the right medicine for the right patient. Another factor to consider is that different levels of evidence are required for qualification in the three CoU categories, which has also been widely discussed by the Biomarkers Consortium during workshops on defining an evidentiary criteria framework.[19] The higher burden of evidence might explain the lower number of qualifications for efficacy biomarkers ($n = 37$, **Table 1**) and the lower proportion that received a QO ($n = 4$). The four qualified biomarkers to establish efficacy include an activity index for use in pediatric ulcerative colitis, an ingestible sensor system for medication adherence, gait measurement through a wearable device in Duchenne muscular dystrophy, and a multiple sclerosis clinical outcome assessment.[20–23]

When it comes to disease areas, the large number of qualification procedures for biomarkers to be used in Alzheimer's disease is remarkable but might be explained by the unmet medical need

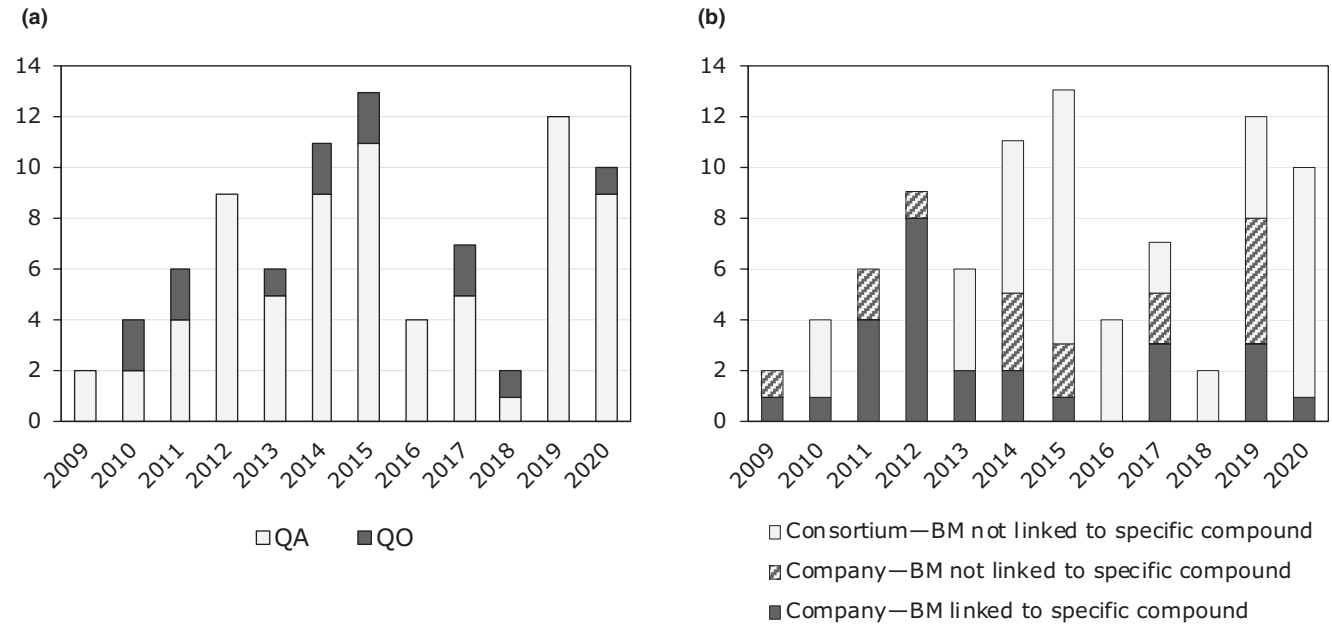


Figure 2 Biomarker qualification procedures from 2008 to 2020. (**a**) All biomarker-related qualification procedures are grouped according to the year in which the procedure was initiated: Stacked bars show the number of qualification advices (light gray) and qualification opinions (dark gray) for each year. (**b**) The type of applicant (company vs. consortium) and development context of the biomarker to be qualified were extracted for all 86 procedures. A distinction was made between biomarkers linked to a specific drug and clinical development program and those that were not. BM, biomarker; QA, qualification advice; QO, qualification opinion.

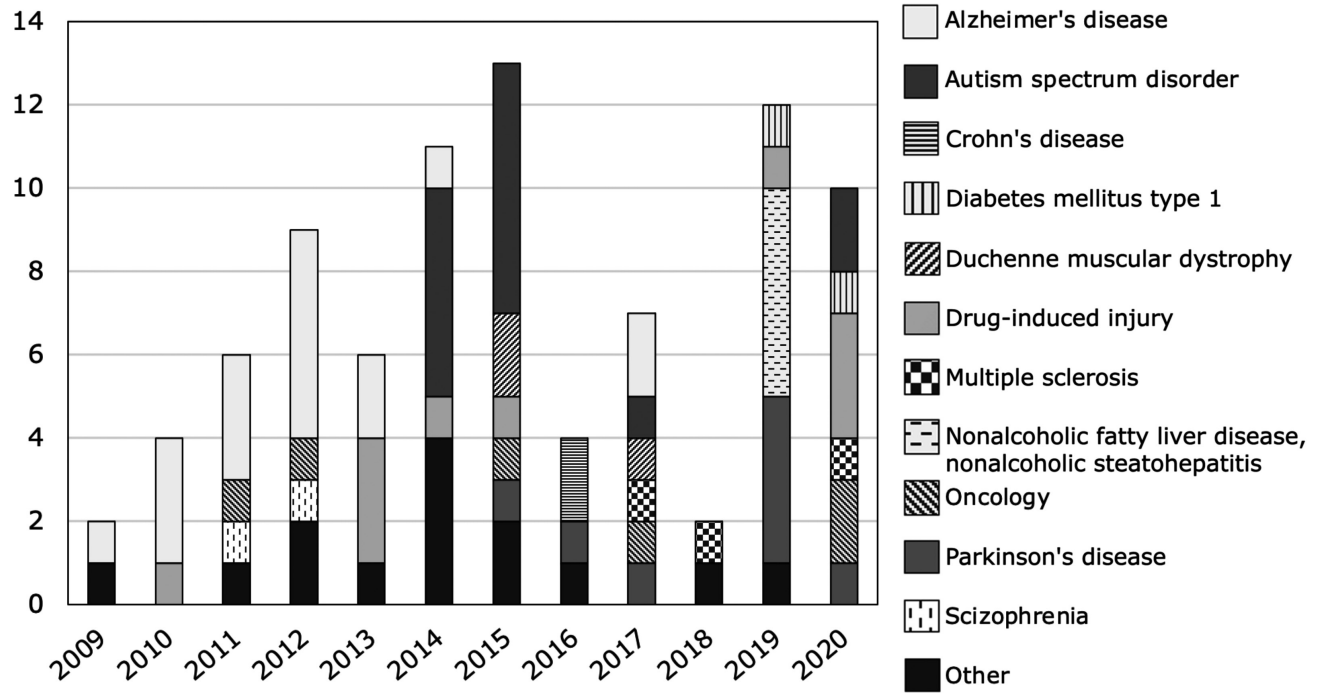


Figure 3 Main diseases or disease areas over time from 2008 to 2020. Number of biomarker procedures per disease or disease areas for which they were proposed over time. Diseases or disease areas for which only one procedure was started are grouped in the category "other." In 2017 one procedure covered MS (multiple sclerosis) and two other diseases, and was therefore assigned to both categories MS and "other." In 2020 one procedure covered MS, PD (Parkinson's disease), and several other diseases. This procedure was counted for MS, PD, and "other."

that exists in this field. Interestingly, most qualification procedures that involve biomarkers for use in Alzheimer's disease were initiated before 2015, with only two procedures starting between 2016 and 2020 (Figure 3). This finding is likely related to developments at the time, i.e., scientific hypotheses that have been driving research questions and drug development efforts in the field.[24]

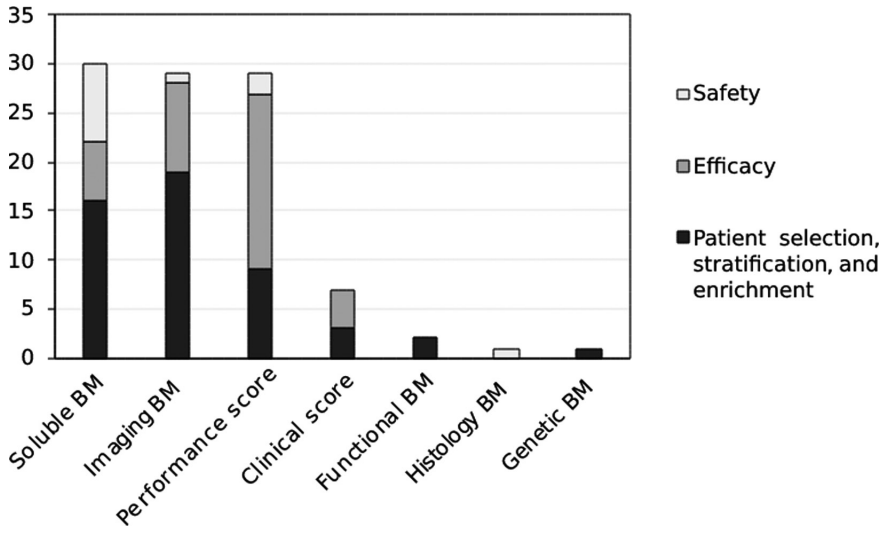


Figure 4 Different types of biomarkers in each CoU category. The stacked bars show the number of procedures in the three CoU categories for each type. BM, biomarker; CoU, context of use.

Although it is known that numerous successful biomarker-guided cancer treatments already exist on the market, only six biomarker qualification procedures concerned biomarkers in the area of oncology.[25] The oncology field is generally considered a pioneer in terms of biomarker use compared with other disease areas, both in clinical trials and clinical practice, as many forms of cancer are associated with known and validated driver mutations.[26] Since these are often highly specific, it is likely that such mutations are topics of discussion in SA procedures and become endorsed as acceptable biomarkers during the MAA review,[27] which is in line with the low number of genetic biomarkers observed in this study. The most common types of biomarkers were imaging biomarkers, soluble biomarkers, and performance scores (**Figure 3**). Whereas imaging and soluble biomarkers are most frequent among the patient selection, stratification, and/or enrichment biomarkers, performance scores are most frequent among the efficacy biomarkers. This may be partly explained by the most common disease areas for which biomarkers are developed in these two CoU categories. For example, for Alzheimer's disease, Parkinson's disease, and Duchenne muscular dystrophy that were dominating in the efficacy CoU category, performance scores are likely to be the most appropriate means to measure efficacy. In disease areas, such as nonalcoholic fatty liver disease and nonalcoholic steatohepatitis, there is a need for less invasive diagnostic tools, e.g., imaging tools in this case, to replace liver biopsy.[28]

An analysis of the issues raised by the CHMP during qualification procedures revealed that a typical procedure contains a high number of issues across multiple categories, underlining that thorough discussions between regulators and applicants take place. In a few cases (n = 4), additional rounds of discussion occurred, which may contribute to the observed variation in procedure length. Perhaps unexpectedly, the fewest issues—approximately 50% of the procedures—were raised on rationale and CoU. Both are crucial aspects in determining the validation strategy toward biomarker qualification and should be thoroughly discussed and properly defined at an early stage, which was the case for many QAs. By nature, early-stage procedures and later-stage procedures result in different issues, which is reflected in the comments from the CHMP. Whereas comments in early-stage QAs mostly concern validation plans, the comments in later stage QAs and QOs concern the data already obtained. It should be noted that final advice letters are typically modeled around specific questions from the applicant to the CHMP, which may also affect the types of issues that were identified from the CHMP answers. For example, one applicant asked: "Does the EMA agree that the Context of Use clearly describes how TKV [total kidney volume] will be used by applicants as a prognostic biomarker to enrich clinical trial population in clinical trials at all stages of ADPKD [autosomal dominant polycystic kidney disease] drug development, including proof of concept, dose-ranging, and confirmatory trials?" which resulted in a discussion and possible issues around the CoU.[29]

In summary, only a limited number of qualified biomarkers is currently available. Despite high heterogeneity among the included procedures, some trends could be identified. Whereas initially most procedures were brought forward by developers aiming to qualify a biomarker linked to a specific drug compound, a shift has taken place in recent years to projects funded by public/private partnerships that aim to develop biomarkers for a CoU that is not linked to a specific compound. The majority of proposed and qualified biomarkers were intended for use in patient selection, stratification, and/or enrichment, followed by biomarkers for efficacy. Disease areas for which biomarkers were proposed to be qualified seem to follow time trends that correspond to scientific developments and unmet medical needs in the field. Moreover, biomarker types differed between CoU categories, which may also be partly explained by the predominant disease areas within these categories. The issues raised in the procedures illustrate that the biomarker qualification procedure is a complex, interactive process, which is also reflected in the longer median timelines of qualification procedures as compared with

the guidance document and the fact that more than half of the QOs resulted from follow-ups to initial procedures.[9] The issue analysis also highlights several aspects that need careful consideration by applicants. These insights in the biomarker qualification procedure at EMA should be used to guide and stimulate future applicants for biomarker qualification, to increase the number of qualified biomarkers and, thereby, facilitate the implementation of precision medicine in research and clinical practice. It would be of interest to further investigate the impact of regulatory biomarker qualification on medicines development, in particular uptake and use in clinical trials and MAAs.

Strengths and limitations

A strength of this study is that it includes not only the publicly available QOs, but also QAs, and provides a comprehensive overview of the biomarker qualification procedures from the regulatory perspective. However, we only considered biomarkers that were proposed for qualification at the EMA, which do not encompass all novel biomarkers that are used in clinical studies. A limitation of the study is that the analysis is based on the procedures and not on the biomarkers itself. The reason for this decision was the great variability among the procedures (e.g., some procedures concerned early-stage projects that presented 20 possible biomarker candidates). For the same reason, follow-up procedures were treated as individual procedures, which underestimates the proportion of QOs, but overestimates the number of different biomarkers proposed. Given that 7 out of 16 follow-up procedures led to a QO, it is possible that several biomarkers that have only been discussed in QAs thus far will eventually proceed to a QO. Due to the overall limited number of biomarker-related procedures and low proportion of QOs, a trend analysis over the years was not informative. Finally, the results from the issue analysis should be interpreted with some caution, despite being performed by two researchers independently and checked by four other researchers. Nevertheless, the results highlight some important considerations for biomarker qualification.

ACKNOWLEDGMENTS

It was a great shock to learn that Jordi Llinares Garcia passed away suddenly at the end of our research project. We will dearly miss his insight in regulatory science matters and his friendly but decided input.

FUNDING

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement no. 115974 (Biomarker Enterprise to Attack DKD; BEAT-DKD). This Joint Undertaking receives support from the European Union's Horizon 2020 Research and Innovation Programme and European Federation of Pharmaceutical Industries and Associations (EFPIA) with JDRF. The funding sources did not have a role in the design, conduct, or analysis of the study.

CONFLICT OF INTEREST

E.B. received funding from the Biomarker Enterprise to Attack DKD (BEAT-DKD). All other authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

E.B., N.M.H., F.E., J.L.G., T.V., V.S., and P.G.M.M. wrote the manuscript. E.B., N.M.H., F.E., T.V., V.S., and P.G.M.M. designed the research. E.B., N.M.H., and D.S.v.d.M. performed the research. E.B. and N.M.H. analyzed the data.

DISCLAIMER

This article expresses the opinion of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the Dutch Medicines Evaluation Board, the European Medicines Agency, or one of its committees or working parties.

1. Amur, S., LaVange, L., Zineh, I., Buckman-Garner, S. & Woodcock, J. Biomarker qualification: toward a multiple stakeholder framework for biomarker development, regulatory acceptance, and utilization. *Clin. Pharmacol. Ther.* **98**, 34–46 (2015).
2. FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet] (*Food and Drug Administration (US) and National Institutes of Health (US), Bethesda (MD), Silver Spring, MD, 2016). <https://www.ncbi.nlm.nih.gov/books/NBK326791/>. Accessed February 23, 2022.
3. Galante, D. & Buyse, M. Integrating personalized medicine in the health care. In *Handbook of Biomarkers and Precision Medicine 1st edition* (eds. Carini, C. *et al*.) (Chapman and Hall/CRC, Boca Raton, 2019). < https://www.taylorfrancis.com/chapters/edit/10.1201/9780429202872-3/integrating-personalized-medicine-health-care-dominic-galante-marylou-buyse>. Accessed February 23, 2022.
4. Lathia, C.D. *et al*. The value, qualification, and regulatory use of surrogate endpoints in drug development. *Clin. Pharmacol. Ther.* **86**, 32–43 (2009).
5. Landeck, L., Kneip, C., Reischl, J. & Asadullah, K. Biomarkers and personalized medicine: current status and further perspectives with special focus on dermatology. *Exp. Dermatol.* **25**, 333–339 (2016).
6. European Medicines Agency. EMA Regulatory Science to 2025: strategic reflection <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf> (2020). Accessed June 16, 2021.
7. Manolis, E., Vamvakas, S. & Isaac, M. New pathway for qualification of novel methodologies in the European medicines agency. *Proteomics Clin. Appl.* **5**, 248–255 (2011).
8. European Medicines Agency. Final Conclusions on the Pilot Joint EMEA/FDA VXDS Experience on Qualification of Nephrotoxicity Biomarkers <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/final-conclusions-pilot-joint-european-medicines-agency/food-drug-administration-vxds-experience-qualification-nephrotoxicity-biomarkers_en.pdf> (2009). Accessed June 16, 2021.
9. European Medicines Agency. Qualification of novel methodologies for drug development: guidance to applicants <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-novel-methodologies-drug-development-guidance-applicants_en.pdf> (2014). Accessed June 16, 2021.
10. European Medicines Agency. Qualification of novel methodologies for medicine development <https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance/qualification-novel-methodologies-medicine-development-0> Accessed June 16, 2021.
11. European Medicines Agency. Opinions and letters of support on the qualification of novel methodologies for medicine development <https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance/novel-methodologies-biomarkers/opinions-letters-support-qualification-novel-methodologies-medicine-development>. Accessed June 16, 2021.
12. Goldman, M., Seigneuret, N. & Eichler, H.-G. The Innovative Medicines Initiative: an engine for regulatory science. *Nat. Rev. Drug Discov.* **14**, 1–2 (2015).

13. Biomarker Enterprise to Attack Diabetic Kidney Disease (BEAt-DKD). About our project <https://www.beat-dkd.eu/project/> Accessed September 10, 2021.
14. Biomarker Commercialization. From BIC to BIC BRIDGE <https://biomarker.nu/about> Accessed September 10, 2021.
15. European Medicines Agency. Essential considerations for successful qualification of novel methodologies <https://www.ema.europa.eu/en/documents/other/essential-considerations-successful-qualification-novel-methodologies_en.pdf> (2017). Accessed December 21, 2021.
16. Gromova, M., Vaggelas, A., Dallmann, G. & Seimetz, D. Biomarkers: opportunities and challenges for drug development in the current regulatory landscape. *Biomark. Insights* **15**, 1177271920974652 (2020).
17. Bakker, E. *et al*. Perspectives on a way forward to implementation of precision medicine in patients with diabetic kidney disease; results of a stakeholder consensus-building meeting. *Front. Pharmacol.* **12, 662642** (2021).
18. Center for Drug Evaluation and Research. Draft guidance for industry and FDA Staff: biomarker qualification — evidentiary framework. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/biomarker-qualification-evidentiary-framework> (2018). Accessed December 23, 2021.
19. Wagner, J. Foundation for the National Institutes of Health. A framework for defining evidentiary criteria for biomarker qualification <https://fnih.org/sites/default/files/final/pdf/4-Wagner-Framework%20Overview.pdf> (2018). Accessed June 16, 2021.
20. European Medicines Agency. Qualification opinion: paediatric ulcerative colitis activity index (PUCAI) <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-paediatric-ulcerative-colitis-activity-index-pucai_en.pdf> (2016). Accessed June 16, 2021.
21. European Medicines Agency. Qualification opinion on ingestible sensor system for medication adherence as biomarker for measuring patient adherence to medication in clinical trials <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-ingestible-sensor-system-medication-adherence-biomarker-measuring-patient_en.pdf> (2016). Accessed June 16, 2021.
22. European Medicines Agency. Qualification opinion on stride velocity 95th centile as a secondary endpoint in Duchenne muscular dystrophy measured by a valid and suitable wearable device* <https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-stride-velocity-95th-centile-secondary-endpoint-duchenne-muscular-dystrophy_en.pdf> (2019). Accessed June 16, 2021.
23. European Medicines Agency. Qualification opinion on multiple sclerosis clinical outcome assessment (MSCOA) <https://www.ema.europa.eu/en/documents/other/qualification-opinion-multiple-sclerosis-clinical-outcome-assessment-mscoa_en.pdf> (2020).
24. Kametani, F. & Hasegawa, M. Reconsideration of amyloid hypothesis and tau hypothesis in Alzheimer's disease. *Front. Neurosci.* **12**, 25 (2018).
25. Jørgensen, J.T. A paradigm shift in biomarker guided oncology drug development. *Ann. Transl. Med.* **7**, 148 (2019).
26. Duffy, M.J. Clinical use of tumor biomarkers: an overview. *Klin. Biochem. Metab.* **25**, 157–161 (2017).
27. Pignatti, F. *et al*. Cancer drug development and the evolving regulatory framework for companion diagnostics in the European union. *Clin. Cancer Res.* **20**, 1458–1468 (2014).
28. Piazzolla, V.A. & Mangia, A. Noninvasive Diagnosis of NAFLD and NASH. *Cells* **9**, 1005 (2020). https://dx.doi.org/10.3390%2Fcells9041005
29. European Medicines Agency Qualification opinion: total kidney volume (TKV) as a prognostic biomarker for use in clinical trials evaluating patients with autosomal dominant polycystic kidney disease (ADPKD) <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-total-kidney-volume-tkv-prognostic-biomarker-use-clinical-trials-evaluating_en.pdf> (2015). Accessed June 16, 2021.