Organisation for Economic Co-operation and Development

ENV/JM/HA(2020)13

For Official Use

English - Or. English

27 May 2020

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Working Party on Hazard Assessment

Addressing the Mutual Acceptance of Data from Computational Methods within OECD Test Guidelines

Towards a Good Computational Methods Practice

22-24 June 2020, Virtual Meeting

The Working Party on Hazard Assessment is requested to provide feedback on elements of this document including:
- Potential Construct to Address Computational Methods in Test Guidelines under Mutual Acceptance of Data (paragraphs 3-7)
-On the potential elements of a Good Computational Methods Practice document (paragraph 10)

This feedback will aid the secretariat in preparation of a document for consideration by the Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology at their November 2020 meeting.

Patience BROWNE, Policy Analyst
Tel: +(33-1) 45 24 19 11; Email: Patience.BROWNE@oecd.org

JT03462169

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Background

1. In February 2020 the Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology discussed a thought-starter on the Mutual Acceptance of Data (MAD) and computational methods included in OECD Test Guidelines (see attached document ENV/JM(2020)6). The objective of the paper was to support a discussion on ways to continue to maintain the MAD system for adherents accepting results of OECD Guidelines as the Test Guideline Programme considers new and innovative methods that use, to some degree, predictive models. While the language of the Council Act on MAD is sufficiently flexible to encompass new methods, there may be reasons to clarify the language in the future. One of the outcomes of the Joint Meeting discussion was, as a first step, to develop a structured approach towards a new quality assurance document on Good Computational Method Practices (GCMP) in the context of MAD to ensure alignment with the overarching principles of Good Laboratory Practices (GLP), and to complement existing OECD guidance material (e.g. on validation) where needed. Experts from Joint Meeting sub-bodies, in particular groups responsible for GLP, hazard assessment and (Q)SARs[1] would be invited to provide input on the development of this new document.

Terminology

2. It is important that all experts share a common understanding of key terms and this section provides commonly used descriptions:

Computational methods[2]

- "Computational methods" include mathematical operations that are applied to raw data. It is, therefore, a data interpretation procedure. Currently, MAD only references "data", however, results of OECD Test Guidelines that include the data interpretation are covered by MAD
  - For example, a standardised computational model or equation is used to convert raw data (the direct output from in vitro methods such as counts of radioactivity, luminescence, light transmission) to information that can be used for regulatory purposes (e.g. positive/negative; GHS and/or potency categories).
  - Defined Approaches take this a step further by including data interpretation procedures of the data resulting from the combination of more than one information source.
- "Computational methods" may also refer to in silico approaches that predict the toxicological response, such as quantitative structure-activity relationship (QSAR) models. Some Defined Approaches for Skin Sensitisation proposed for inclusion in OECD Guidelines include computational (in silico) methods used with in vitro data. In the future, other approaches may be proposed that do not include any (de novo) laboratory-derived data.

1 (Quantitative) Structure-Activity Relationships [(Q)SARs] are methods for estimating properties of a chemical from its molecular structure.

2 This text is an extract from the Joint Meeting document ENV/JM(2020)6

ED_006472_00014706-00002

Validation

- Within OECD's work on chemicals, the term "validation" has different meanings depending on the context.
- In the GLP context, the word "validation" concerns, among other things, the demonstration that a method or computerised system is suitable throughout its life cycle for its intended purpose.
  - The OECD Principles on Good Laboratory Practices define the responsibilities of a Test Facility or Study Director with respect to validation (§ 1.1.2.b, e and q).
  - The OECD Advisory Document for the Application of GLP Principles to Computerised Systems (Guidance Document 17) provides guidance for test facilities developing a strategy for the validation and operation of computerised systems.
  - The role of the GLP Programme is limited to the verification that OECD Test Guidelines are implemented correctly by test facilities.
- In the Test Guideline context, "scientific validation" is the demonstration of the reliability and relevance of a new or updated test, approach, method, or process established for a specific purpose and international regulatory acceptance.
  - The OECD Guidance Document on the Validation and International Acceptance of New and Updated Test Methods for Hazard Assessment (Guidance Document 34) defines the (scientific) validation principles and provides instructions on the process to validate test methods. The process is meant to be modular and sufficiently flexible to adapt to different situations where the existing body of evidence supporting the validity of a method may vary.
  - Determining the scientific validity of a test method (which may include a prediction model or make use of a computerised system) or defined approach (which includes a data interpretation procedure) is the responsibility of the Test Guidelines Programme. For (Q)SARs, OECD has developed a Guidance Document on the Validation Of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (Guidance Document 69).

Potential Construct to Address Computational Methods in Test Guidelines under MAD

3. Regardless of what guidance exists today or will be developed, in order to provide receiving authorities with the confidence that data generated by future test methods are of high quality, some form of managerial quality control system is necessary. Such a system will need to cover how studies are planned, performed, monitored, recorded, reported and archived.

4. Thus, it is possible to envision that a MAD framework will ensure the results generated from Test Guidelines which include computational methods follow a quality control system consisting of:

a Test Guideline that describes a validated method including any computational component, its execution and how data are analysed and interpreted;

b The Principles on GLP; and

ED_006472_00014706-00003

c Good Computational Method Practices (GCMP).

5. Within this construct, the development of the Test Guideline encompasses the scientific validation of the method(s) (i.e. in vivo, in vitro, computational etc.), the data interpretation procedure, and the study protocol (including instructions for any laboratory and computational components). Then the GLP and GCMP provide the elements of quality assurance.

6. For example, some Defined Approaches for Skin Sensitisation under consideration for inclusion in an OECD Guideline include in vitro data and an in silico model prediction that have been validated and reviewed during Test Guideline development. Anyone submitting data using an in silico model must follow the protocol for running the model provided in the Test Guideline, provide documentation of how they ran the model, and report the prediction following instructions in the Test Report section of the Test Guideline. Typically, this has been done for (Q)SAR models using the (Q)SAR Prediction Reporting Format (QPRF). However, principles for describing computational models (e.g. in silico profilers, expert systems) and reporting the resulting predictions for a broader suite of computational methods would be further elaborated in the Good Computational Method Practices.

7. GLP inspectors would not be responsible for certifying the validity of the computational method or the resulting Defined Approach prediction. Responsibilities for oversight of computational aspects of OECD Guideline methods could be discussed and described in the GCMP.

Drawing Inspiration for a GCMP from GLP Principles

8. In order to determine what elements could be included in the GCMP, it is necessary to consider what elements – if any – of the GLP Principles could apply to computational methods. To that end, a Questionnaire was circulated to the members of the Working Group on GLP.

9. Responses were received from 25 GLP experts representing 20 countries. Questions included in the survey and a summary of responses are included in the table below. Detailed responses can be found in **Annex A**.

Table 1. Summary of responses to survey on computational data circulated to WG GLP Q2 2020

Survey Questions	No	Yes
1a. Does your organisation require (or plan to require) GLP and/or any other quality system for non-clinical studies submitted based, in part, on computational methods?	43%	57%
2a. Must data generated using predictive models be compliant with GLP in order for such data to be accepted under MAD?	39%	61%
3a. From a GLP perspective, does it matter if a computational method is conducted outside a GLP compliant test facility (e.g., at a sponsor's corporate office or in a private home)?	39%	61%
4a. Are there specific elements you can envision now that should be covered by the Good Computational Method Guidance (GCMG)?	46%	54%
5a. Has your GLP Compliance Monitoring Programme already audited data or studies generated by computational methods?	69%	31%
6a. GLP SECTION I/Subsection 2:"Introduction / Definition of Terms" - Do the elements or terms in this Subsection apply?	9%	91%

ED_006472_00014706-00004

7a. GLP SECTION II/Subsection 1: "Test Facility Organisation and Personnel" - Do the elements or terms in this Subsection apply?	4%	96%
8a. GLP SECTION II/Subsection 2: "Quality Assurance Programme" - Do the elements or terms in this Subsection apply?	17%	83%
9a. GLP SECTION II/Subsection 3: "Facilities" - Do the elements or terms in this Subsection apply?	13%	87%
10a. GLP SECTION II/Subsection 4: "Apparatus, Material, and Reagents" - Do the elements or terms in this Subsection apply?	17%	83%
11a. GLP SECTION II/Subsection 5: "Test Systems" - Do the elements or terms in this Subsection apply?	30%	70%
12a. GLP SECTION II/Subsection 6: "Test and Reference Items" - Do the elements or terms in this Subsection apply?	61%	39%
13a. GLP SECTION II/Subsection 7: "Standard Operating Procedures" - Do the elements or terms in this Subsection apply?	9%	91%
14a. GLP SECTION II/Subsection 8: "Performance of the Study" - Do the elements or terms in this Subsection apply?	13%	87%
15a. GLP SECTION II/Subsection 9: "Reporting of Study Results" - Do the elements or terms in this Subsection apply?	9%	91%
16a. GLP SECTION II/Subsection 10: "Storage and Retention of Records and Materials" - Do the elements or terms in this Subsection apply?	13%	87%

Towards a GCMP

10. The GCMP is not envisioned to address aspects of scientific validation, but rather will focus on aspects of quality assurance. In order to determine relevant elements of a GCMP, the secretariat drew upon the feedback of the WG on GLP as well as the Guidance Document on Good In Vitro Methods Practice (GIVIMP). The following areas, at a minimum, will need to be discussed to determine, if there is GLP guidance relevant to computational methods, if the topic is covered in the Test Guideline for computational methods, or if a GCMP can provide complementary information for computational methods. The following sections are proposed, as a start:

a Introduction and terminology (introduce the purpose of the document and key terminology)

b Roles and responsibilities (computational method developers, providers, validation bodies, users, sponsors, etc.; documentation requirements, software version management and control)

c Quality considerations (QA versus QC; requirements for developers and implementation of computational methods; considerations for integrity of the data)

d Conduct of the study (location, validation of data integrity (e.g. checksums), software versioning, site of analyses (e.g. cloud-based or desktop), etc.)

e Reporting of results (reporting of computational method validation, reporting computational data and model predictions for regulatory purposes and MAD, publishing data)

f Storage and retention of records (requirements re: data and records storage and retention, application to data integrity, protection of data and records from deliberate/accidental changes)

ED_006472_00014706-00005

Timelines and Process

11. Following this meeting, a proposal will be prepared for the consideration of the Joint Meeting at their November 2020 meeting, on the general framework of and approach for developing a Good Computational Methods Practices document as described in this paper.

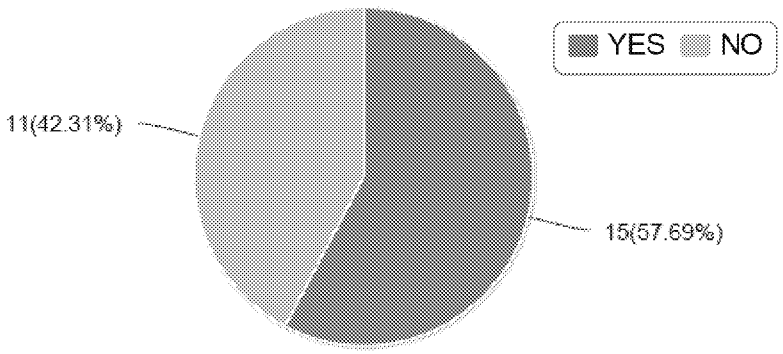
12. The JM will be asked to review the proposal for the GCMP development and consider the value in establishing an ad hoc group to work on a draft GCMP document with representation from the relevant sub-bodies, and a process for nominating experts.

ED_006472_00014706-00006

Annex A. Tally of WP GLP Responses to Questionnaires

Part I General Questions

Q1a Does your organisation require (or plan to require) GLP and/or any other quality system for non-clinical studies submitted based, in part, on computational methods?



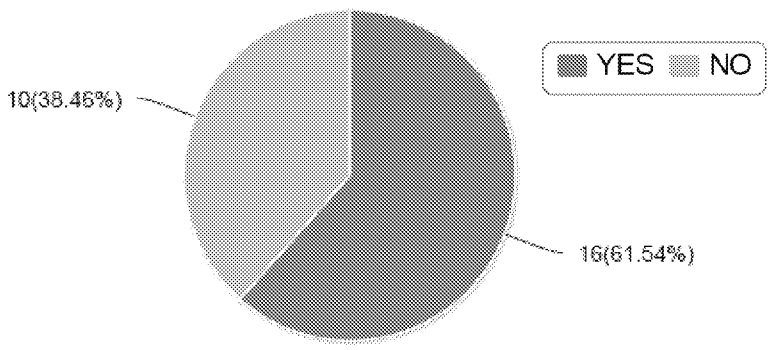
Q1b If your answer to 1a is Yes, please describe (optional)

Australia	The Australian Industrial Chemicals Introduction Scheme (AICIS) will replace the National Industrial Chemicals Notification and Assessment Scheme (NICNAS) on 1 July 2020. AICIS is enacted by the Industrial Chemicals Act 2019. The data required for introduction will vary depending on the type of introduction (exempted, reported and assessed) according to categorisation based on introduction volume and hazard profile of the chemical. These requirements are set out in the Industrial Chemicals General Rules, the Industrial Chemicals Categorisation Guidelines, and several guidance materials published on the NICNAS website. Non-clinical studies derived, in part or wholly, from computational methods are acceptable for chemical introductions in Australia at lower exposure levels for specific human health and environment hazard endpoints. At higher exposure levels, computational methods play a supporting role to in vivo, in vitro and/or in chemico studies for particular human health hazard endpoints. AICIS will require detailed documentation, using QSAR Model Reporting Format (QMRF) and QSAR Prediction Reporting Format (QPRF), for the model used and the predictions derived from computational methods. The detailed documentation will allow us, if needed, to review the generated predictions. Use of QMRF and QPRF would benefit from more guidance to help the user fill in these forms. Scrutiny of the submitted documentation would need to be undertaken to check the validity/reliability of the prediction.
Australia	In accordance with ICH guidelines, GLP is expected for key safety-related studies.
Australia	The APVMA generally requires certain types of data submitted to support the registration of pesticides or veterinary medicines to be conducted in accordance with the OECD principles of GLP. However, some allowances are made in addressing the validity of a study or components of a study that may be outside the scope of GLP. A degree of expert judgement is used in determining the reliability of the study and whether it provides a sound basis for regulatory decision-making.

ED_006472_00014706-00007

Belgium	If required by any legislation to be done under GLP, Sciensano as Belgium GLP CMA has the obligation to inspect any activity that claims GLP in Belgium.
France (ANSM)	GLP compliance is mandated by the European and national regulations for the non clinical safety studies submitted in authorisation files for human health products.
India	All study data needs to be generated in compliance with GLP principles, the same applies to computational methods as well
Israel	GLP is applicable to all non clinical studies. GLP is not required but voluntary in Israel.
Malaysia	Any non-clinical studies to support registration for products fall under NPRA's scope (e.g. new chemical entity, biologics and herbal products with high claims) are require to be conducted following GLP
Slovakia	cloud based archiving, documentation system
Switzerland	Yes: if GLP is required; for all methods that include a laboratory step or if the computational method is a part of a OECD TG.
UK	We would take the view that if an activity is included in a GLP study it must be GLP compliant. How this fits into the study plan and how we inspect it will be dependent on its nature. For example whether the method is a software service and may be treated as a vendor, or in another case the activity/system may be considered as being part of a facility.

Q2a Must data generated using predictive models be compliant with GLP in order for such data to be accepted under MAD?



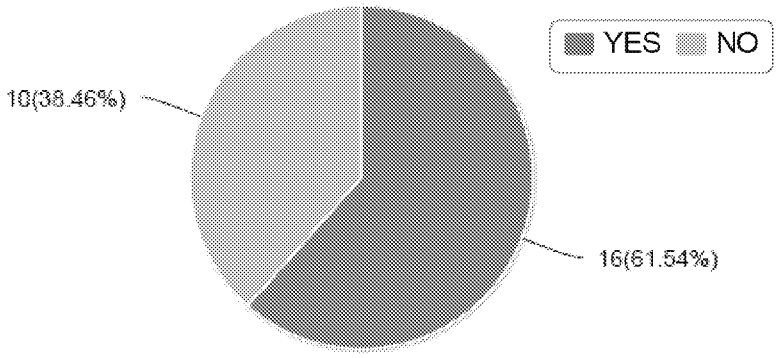
Q2b If your answer to 2a is Yes, please describe (optional)

Australia	Data generated using computational methods must be compliant with principles that are similar or analogous to GLP to facilitate acceptability under MAD system. GLP covers mostly animal-based experimental or field-based systems and does not include quality control criteria for computational methods. In addition to an analogous system, mechanisms to check adherence to the system are also important. We support the development of Good Computational Method Principles (GCMP) and guidance on the practical applicability of GCMP to the MAD system. What is important for us is the transparency of the protocol followed to generate the prediction and whether the prediction is valid based on the applicability domain of the computational method used. There should be a standardised protocol for the use of computational tools and detailed guidance on how to describe the protocol so that regulators can be confident that the predictions are reliable and compliant with a system

	such as a GCMP similar to GLP for experimental or field-based systems. The standardised protocols may differ depending on the computational model. Including computational methods in the MAD system and respective OECD guidance would contribute to a harmonised approach to the international use and acceptance of predictions from computational methods for regulatory purposes.
Australia	Desirable but not essential
India	For data to be accepted under MAD, data reliability and the principle of "Generated once, accepted everywhere" is applicable only if all parts of the study are in compliance with GLP, the same is true with data generated using predictive models
Italy	Data accepted under MAD should be compliant with GLP Principle which assure that the data shared are under a control system with traceability and a chain of responsibility which would avoid any forgery. Computational methods would be used by the Receiving Authorities to assess data included in studies under GLP system. Consequently, they can be seen as instruments used by the assessors. The accuracy and quality of the data generated by computational methods (outputs) are not an issue of the GLP Principles.
New Zealand	Not that it is particularly relevant in NZ, but from first principles, MAD has GLP compliance at its core therefore to be accepted under MAD, GLP compliance must be demonstrated.
Poland	MAD consists of two elements: - GLP, - acceptable test methods. In our point of view in order to accept data generated using predictive models under MAD we have to change the general requirements of GLP. In GLP we have only a definition of "non-clinical health and environment safety study" - ....set of experiments in which a test item is examined under laboratory conditions or in the environment. The definition of study does not include predictive models.
Slovakia	to be sure that are valid
Switzerland	Yes: As part of GLP studies, i.e. methods that include a laboratory step and if the computational method is a part of a OECD TG. No: for other predictive models
The Netherlands	MAD only applies in case an OECD Test Guideline AND the Principles of GLP are being followed [C(81)30(Final), Part 1, Section 1].
UK	If it is key to the conclusion of a GLP study then we feel it will need to be compliant. It may well again depend on the nature of the predictive model application itself whether it is considered as a software/service so treated as a vendor, or a phase of a study that would need to be GLP compliant

ED_006472_00014706-00009

Q3a From a GLP perspective, does it matter if a computational method is conducted outside a GLP compliant test facility (e.g., at a sponsor's corporate office or in a private home)?

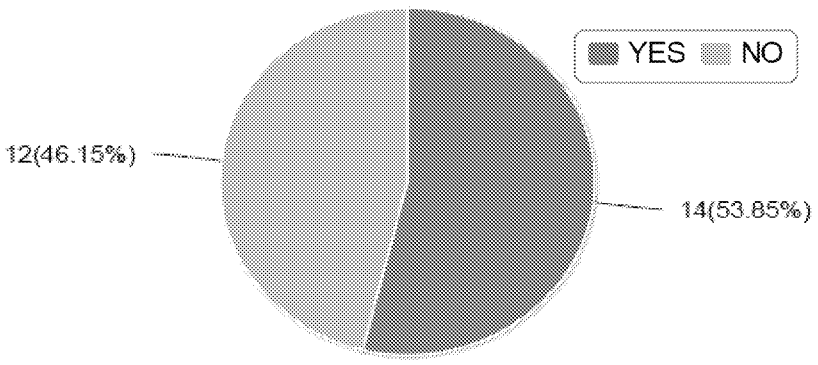


Q3b If your answer to 3a is Yes, please describe (optional)

Australia	The OECD Advisory Document for the Application of GLP Principles to Computerised Systems outlines optimum conditions for running a test facility where computerised systems are located. From a GLP perspective, the physical location of where the predictions were derived using computational methods should not matter as long as there is an up-to-date and highly maintained catalogue of computerised systems and their functionalities. Regardless of whether the entity that performs the computational predictions is under the umbrella of an existing, GLP compliant test facility or is a completely unrelated entity (such as a chemical business), we would require the entity to comply with the GCMP requirements.
Belgium	If the method is claimed to be done under GLP, it should be under responsibility of a SD/PI. Proper documentation should exist on how the relation between Test facility and location of the method is. (e.g. peer review pathology).
India	MAD is applicable only if all parts of a GLP study are conducted in the GLP compliant test facility. So, conducting a computational method outside the GLP compliant test facility may raise concerns regarding the GLP compliance of the study, especially in terms of data reconstructibility.
Italy	The Study Director (SD) is the responsible of the study report. If a computational method is used inside a study, the responsibility of the management of the entire study, in all its parts, is from the Test Facility (TF) where the SD works. Consequently if, for example, the computational method can be treated as a special software, the TF management should apply what asserted by the Advisory document n.17 on Computerised systems. In multi-site studies, agreements among different TFs can be considered. All the parts of a GLP study should be conducted in one TF or among TFs which are included in a GLP Compliance Monitoring Programme.
Slovakia	if there is no contract with strictly specified requirements
Switzerland	In general, we do not think that the location is of great importance assuming that the notebook used to perform the computational method is within the TF's IT infrastructure and compliance to SOPs - can also be used with remote access.
Switzerland	It should be conducted in a GLP test facility.
The Netherlands	A GLP study should be linked to a physical location to allow for inspection. A computational method used in a GLP study should be considered as part of a GLP test

	facility.
UK	We feel that this will depend on what is being conducted or supplied (e.g. a service). This may dictate if a method is a supplied software service that can be validated by the user to ensure fit for purpose, or treated as a distinct part of a study
US (EPA)	In my opinion a computational method should be conducted in a GLP compliant test facility in order to be considered GLP compliant.

Q4a Are there specific elements you can envision now that should be covered by the GCMG?

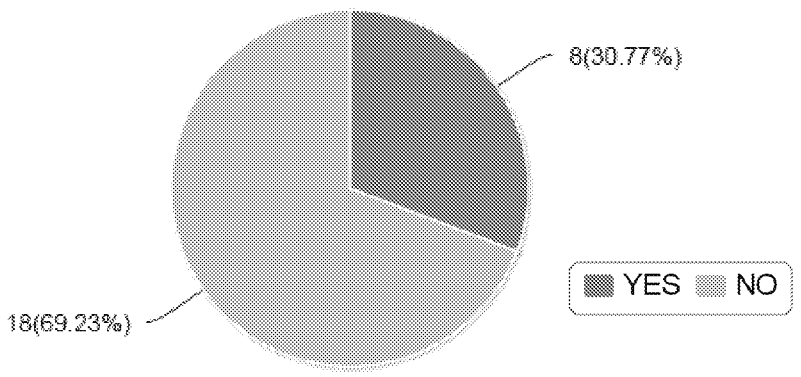


Q4b If your answer to 3a is Yes, please describe (optional)

Australia	Elements of the guidance to GCMP should include how predictions correspond to internationally agreed principles such as the OECD Principles of (Q)SAR Model Validation and its associated guidance, the OECD Guidance Document on (Q)SAR Model Validation. Considerations covered in these documents include (but are not limited to): clear definition of model performance and predictions; availability of model training sets and evidence of overfitting for statistical models; reliability of the predictions to analogues; applicability domain; and reproducibility of the predictions. In the MAD system, the data must be generated in a test facility that has been inspected by a national GLP monitoring compliance program, and that the program must have undergone successful evaluation by the OECD. Options for implementing a GCMP monitoring compliance program may need to be covered, including the application of principles outlined in national GLP monitoring programs. Compliance program could also include monitoring whether appropriate and updated tools are used and the predictions made were reliable. The GCMP could be based on the OECD Principles for validation of computational models. Following are some questions to be considered: - Should we consider an analogous GCMP monitoring compliance program? - How would the guidance material for GCMP ensure GCMP compliance is achieved? - Should the GCMP compliance program only apply to testing facilities that are GLP compliant or, if the extension of MAD to the generation of data through computational methods also applies to such data generated by non-traditional test facilities (such as chemical businesses that are currently not GLP-compliant), should a GCMP compliance monitoring program also include such non-traditional testing facilities?
Australia	The proposed framework to ensure computational methods follow a quality control system consisting of the test guideline, GLP principles, and good computational method guidance seems to be appropriate elements to be covered by the CGMG.

Belgium	Validation of software, Archiving of libraries, Data migration, data mitigation, ...
Colombia	statistical elements
France (ANSM)	The most important challenge would be the archiving of the software's versions and of the libraries if any in the date of their use in the study, so that it could be reconstructed or reloaded even years after..
India	1. The entire study (including Good computational methods) should be conducted in compliance with GLP principles. 2. The key personnel conducting the study should be identified and assigned roles and responsibilities which should be similar to the personnel conducting other type of studies. 3.Role of QA unit in auditing such studies should be clearly defined, based on risk-based approach, QA audit planner and identification of critical phases by the QA would be important steps that might need standardization within the TF. 4. Method validation and Validation of software used in the studies should be ensured by the TF
Israel	Audit trail Computational method validation
Poland	GCMS should cover, inter allia, detail information about test system and test item.
Switzerland	establish harmonized definitions of Terms (e.g. qualification, validation) - use Good In vitro Practice Approach as example.
Switzerland	Elements to be included: - Quality assurance and quality control checks - Validation of computerized system, calculation and scientific method - Integrity/stability of the method - User (training, functions, responsibility) - Documentation (Retention of calculations) - Data integrity - Test facility (GLP)
The Netherlands	Access control, integrity of used databases, controlled use of algorithms, validation of applications, audit trail, reporting of results and conclusions, QA inspections and audits.
UK	If viewed as a computer system then vendor approval, validation and Document 17 become important. If a method is viewed as a GLP study then the principles will apply (for example could a method replace an animal study?)

Q5a Has your GLP Compliance Monitoring Programme already audited data or studies generated by computational methods?



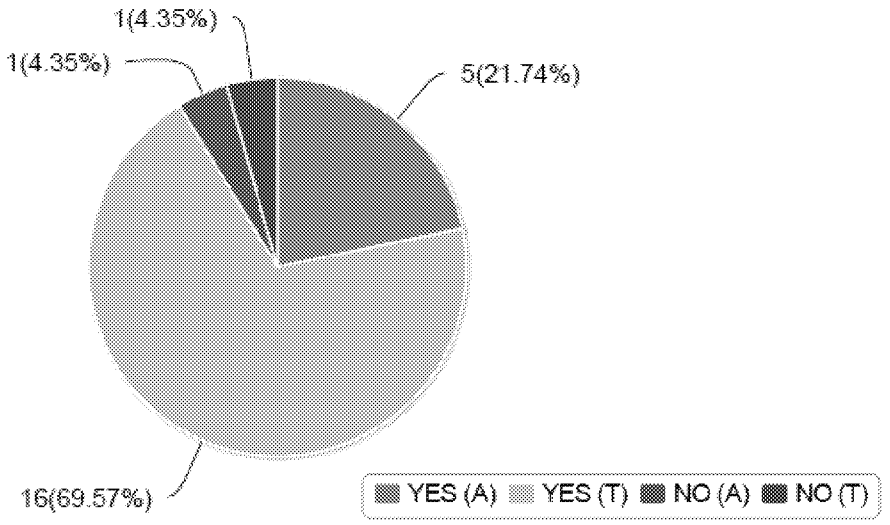
Q5b If your answer to 5a is Yes, please describe how your audit was conducted and whether it was conducted within a GLP context

France (ANSM)	QSAR studies declared in compliance with GLP were detected during a routine GLP inspection. The issue was then presented to the OECD GLP WG and after the discussions, decision was taken to remove the GLP compliance of the studies, waiting for a new consensus (if any) on that issue.
India	The audit was conducted in compliance with the OECD Principles of GLP, with special emphasis on: 1. Roles and responsibilities of key personnel involved in the study-Study director, study personnel, QA, test facility management and sponsor. 2. QA identification of critical phases 3. Validation of Software used and the source of validation-whether done at the test facility or by a vendor or sponsor himself 4. Characterization of test item/reference item. 5. Characterization of test systems, if applicable 6. Raw data capture and the processes followed for processing of raw data 7. Study Report and archiving
Japan (PMDA)	We audit Computalysed system (statistics, chromatography, etc...) based on the Document No.17.
Switzerland	yes - within GLP context e g. TK analysis was a study phase Audit included inspection of computerized system, personal records incl. knowledge of SOPs,
The Netherlands	A test facility specialised in pharmacokinetic and PK/PD data analysis claims GLP for this work. Usually these analyses are carried out as part of a multisite nonclinical safety study. Data (test item concentration in samples) are received in electronic format and analysed with validated computer applications such as WinNonlin. The test facility is inspected for compliance with the relevant requirements of the GLP principles.
UK	Based on the definitions in the background document the UK has inspected statistics companies who analyse parts of GLP studies. Inspections use the same approach and apply the relevant principles to the environment.
US (EPA)	I was not involved with the GLP data audit of computational methods. Frances Liem and one of our inspectors Elmer Griffin were involved and Frances shared with the WG US EPA's experience. Over the past fifteen years or so, the US EPA received six (Q)SAR/modeling studies and conducted study audits. The studies were not conducted according to GLP because the test facilities did not believe it was feasible to do so since there was no test system.

ED_006472_00014706-00013

Part II GLP Principles that could apply to Computational Methods

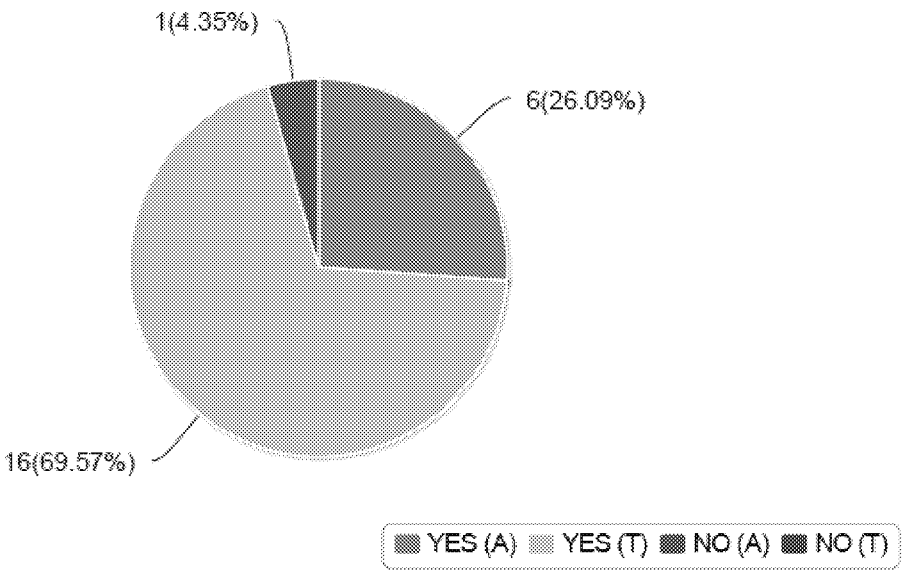
Q6a GLP SECTION I/Subsection 2: "Introduction / Definition of Terms" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q6b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Italy	It is not necessary based on our limited experience in complex computational methods.
Switzerland	Additional terms should be defined.

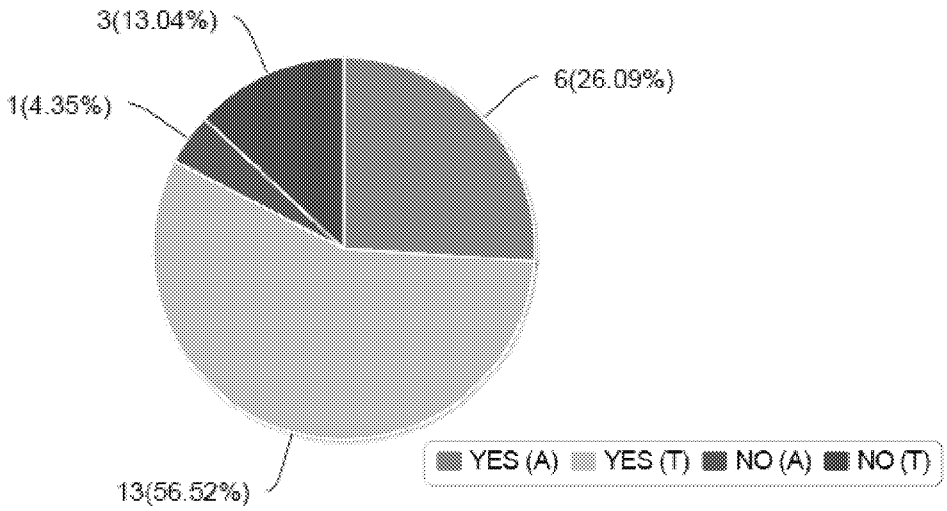
Q7a GLP SECTION II/Subsection 1: "Test Facility Organisation and Personnel" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q7b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Italy	It is not necessary based on our limited experience in complex computational methods.

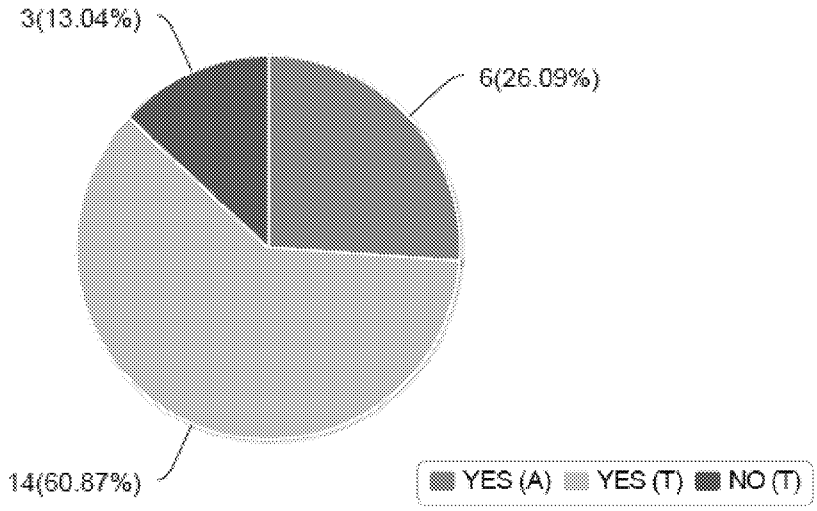
Q8a GLP SECTION II/Subsection 2: "Quality Assurance Programme" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q8b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Italy	It is not necessary based on our limited experience in complex computational methods.
Slovak	QA needs to know what is required to be checked and documented
Switzerland	Procedures for QA need to be described.

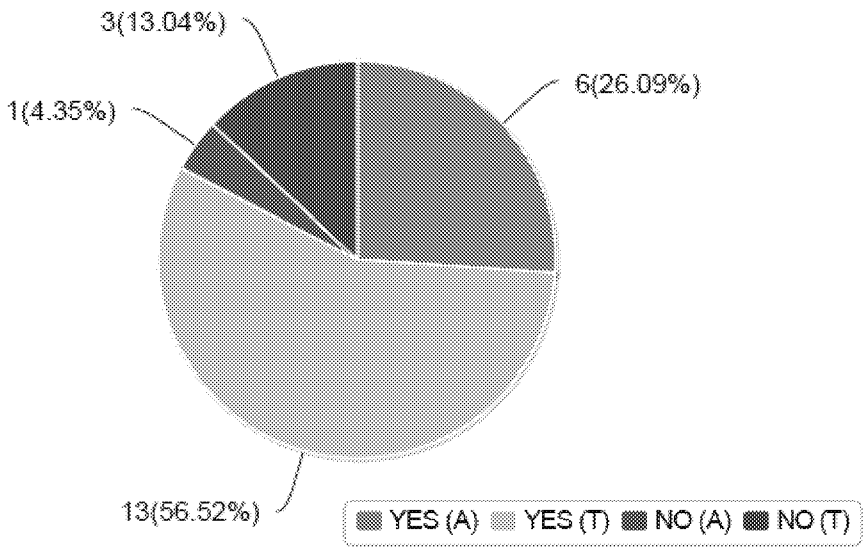
Q9a GLP SECTION II/Subsection 3: "Facilities" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q9b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Belgium	A chapter could be added on what the requirements can be to produce the severs/computers handling the data or the method. (some is already available in OECD 17, might be included in the new cloud document)

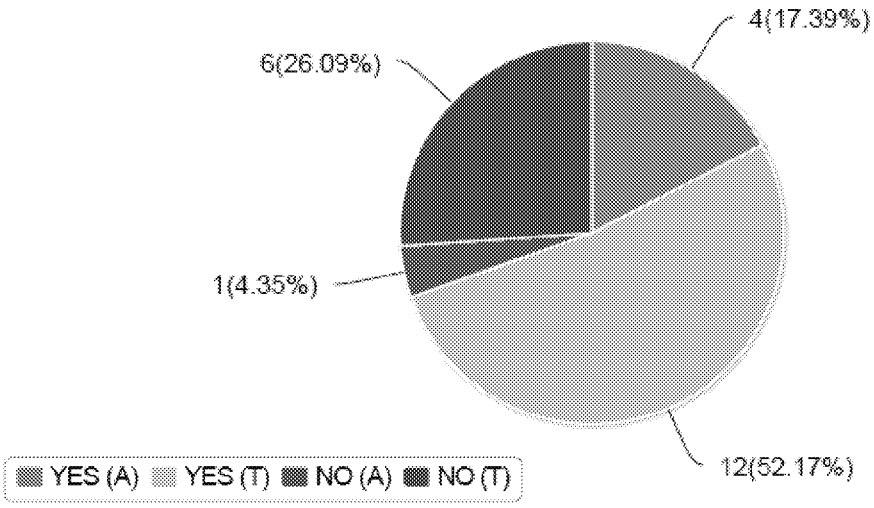
Q10a GLP SECTION II/Subsection 4: "Apparatus, Material, and Reagents" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q10b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Italy	It is not necessary based on our limited experience in complex computational methods.
Slovakia	new technologies and requirements are not described anywhere
US (EPA)	Since the GLP principles are specific to laboratory or field testing and not for *in silico* predictive methods, the development of "Good Computational Practices (GCP)" analogous to "Good Laboratory Practices (GLP)" will be highly appreciated. Also, see General Comments.

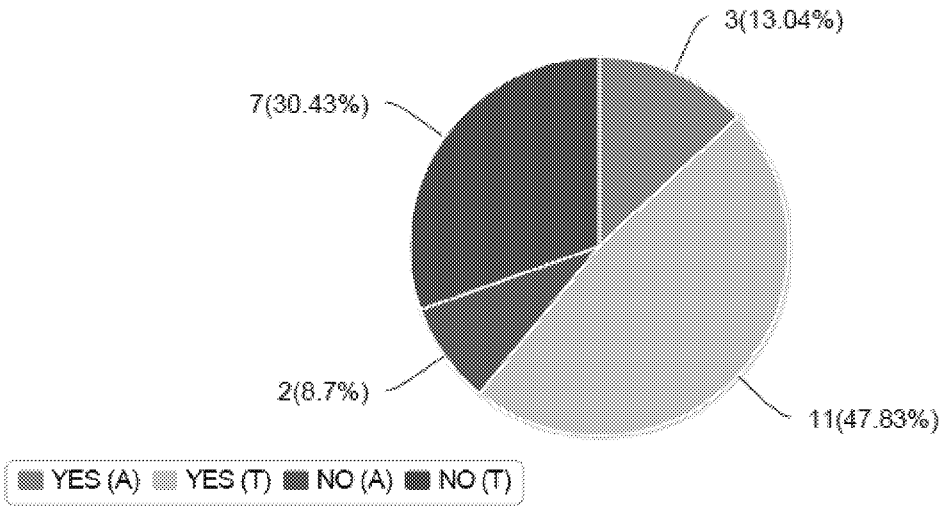
Q11a GLP SECTION II/Subsection 5: "Test Systems" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q11b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Austria	A software based modelling or prediction system should be a test system on its own. It is neither a physical nor a biological test system.
Italy	It is not necessary based on our limited experience in complex computational methods
New Zealand	There is no test system (to which a test item is applied)?
Slovakia	new technologies and requirements are not described anywhere
Switzerland	Computational models can be treated as special test systems; however, the current regulations do not address this situation.
The Netherlands	Computational methods probably cannot be considered as test system according to the current definition of test systems.
US (EPA)	Since the GLP principles are specific to laboratory or field testing and not for *in silico* predictive methods, the development of "Good Computational Practices (GCP)" analogous to "Good Laboratory Practices (GLP)" will be highly appreciated. Also, see General Comments.

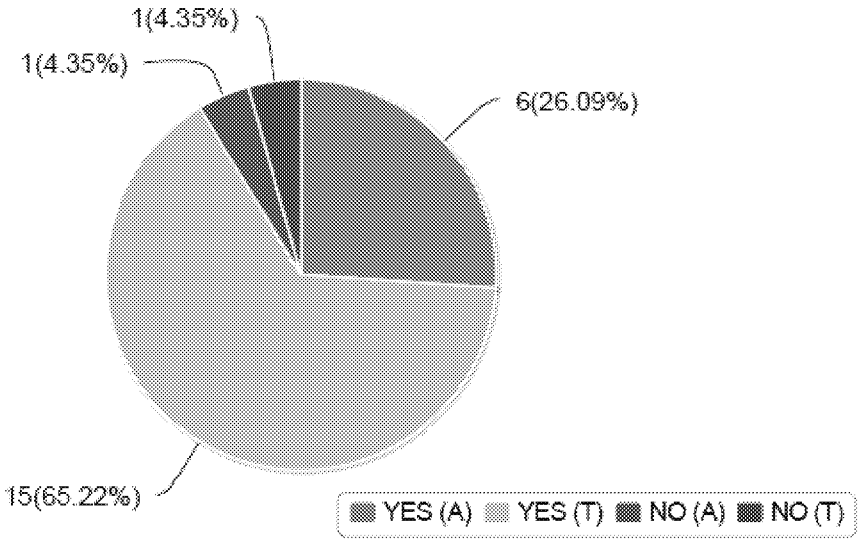
Q12a GLP SECTION II/Subsection 6: "Test and Reference Items" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q12b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Australia	I am not sure about this one. If the computational methods is using data derived from an in vitro study then the test item would apply to the in vitro part only. If, however, it is an in silico method which does not include an laboratory derived data then I don't see how Section 6 would apply....what is the test item?
Austria	I would not expect that a computer based systems needs test items.
Belgium	There is no physical test item in these methods.
France (ANSM)	It is difficult to characterise a non real test item (but only a chemical formula). The chapter 6 of the GLP Principles is not entirely applicable.
Italy	It is not necessary based on our limited experience in complex computational methods
Slovakia	new technologies and requirements are not described anywhere
Switzerland	Handling of structures as virtual test items is not covered currently.
US (EPA)	Since the GLP principles are specific to laboratory or field testing and not for in silico predictive methods, the development of "Good Computational Practices (GCP)" analogous to "Good Laboratory Practices (GLP)" will be highly appreciated. Also, see General Comments.

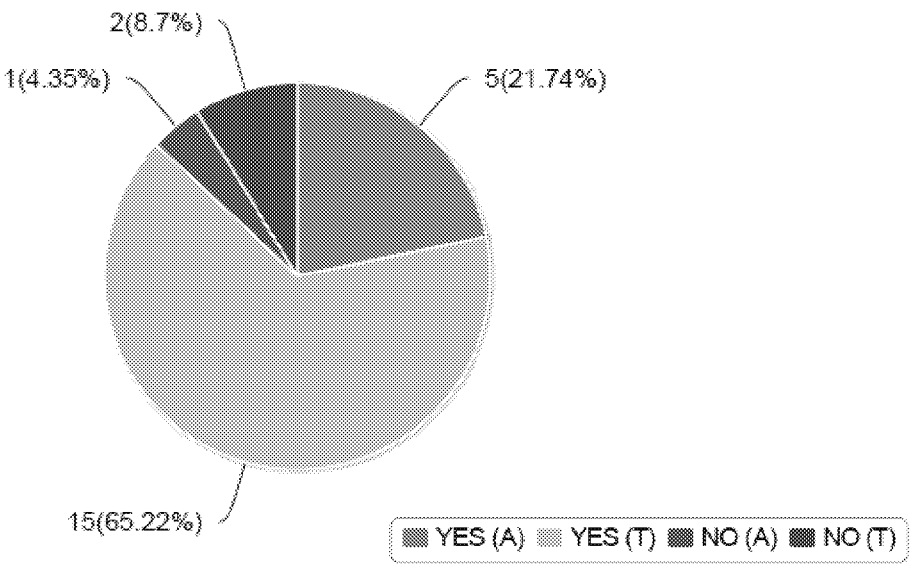
Q13a GLP SECTION II/Subsection 7: "Standard Operating Procedures" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q13b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Slovakia	new technologies and requirements are not described anywhere

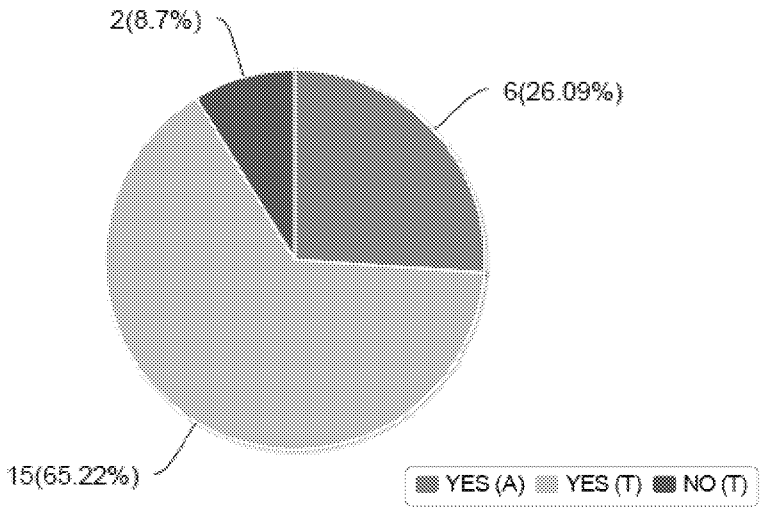
Q14a GLP SECTION II/Subsection 8: "Performance of the Study" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q14b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element? If based on actual experience, choose "A" if based in theory choose "T".

France (ANSM)	The definition of a study should remove the notion of experiments conducted in laboratory to stick to computational methods.
Slovakia	new technologies and requirements are not described anywhere
Switzerland	Current regulations cover raw data generation and calculations within a test facility. Computational methods based on cloud applications (SaaS) or calculations by the software supplier are not yet covered (raw data, meta data and results of calculations).

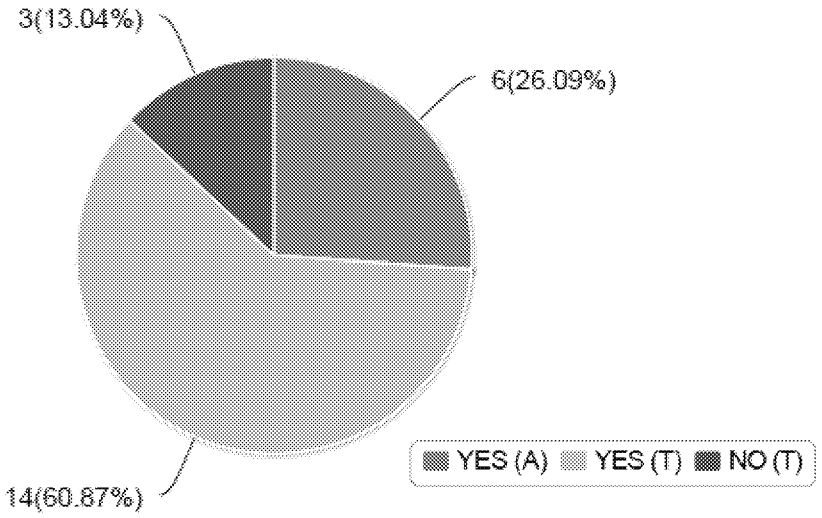
Q15a GLP SECTION II/Subsection 9: "Reporting of Study Results" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q15b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element? If based on actual experience, choose "A" if based in theory choose "T".

Mexico	It would be no appropriate in case that the GLP test facility does not use the computational Method for the report.
Slovakia	new technologies and requirements are not described anywhere

Q16a GLP SECTION II/Subsection 10: "Storage and Retention of Records and Materials" - Do the elements or terms in this Subsection apply? If based on actual experience, choose "A" if based in theory choose "T".



Q16b If you answered "NO" for any of the elements, please explain why and whether it is essential that the GCMG address this element?

Belgium	Specific arching requirement should be written in order to allow for full reconstructability of the method at any time. Algorithm, libraries, ...
Slovakia	new technologies and requirements are not described anywhere
Switzerland	Current regulations cover data and calculations within a test facility. Computational methods based on cloud applications (SaaS) or calculations by the software supplier are not yet covered.

Q17 General Comments

Australia	With all of the Yes (in theory) responses above, for some Section only relevant aspects of the Principles may apply. It should be possible to adapt the GLP Principles to computational methods. Alternatively, the proposed GCMP document could incorporate these. The questions posed by NICNAS in Question 4 above are ones that need to be considered. What is the role of the GLP compliance monitoring authority in this situation? Is the GCMP document analogous to a Test Guideline or more than this?
Australia	See Australia's response submitted for NICNAS as the answers are the same
Australia	See responses submitted by Australia for NICNAS
Austria	Most of GLP principles can be relied on when producing a guidance document for computational methods.
Belgium	In my opinion, this is very interesting and valuable initiative. I think the GLP principles are broad and open enough to be applied to computational methods, on the conditions of some additions and adaptations. Best regards Martijn

ED_006472_00014706-00021

France (ANSM)	My personal opinion is that most of the GLP requirements are applicable to computational method studies, but some of the requirements may be adapted to stick to the specificities of them. This opinion was not shared by everyone, then the need to issue a specific standard.
Israel	Examples for computational methods: NGS- Next generation sequencing Pathology- Computational slide evaluation
Italy	A quality system assuring the accuracy of the data generation process should be considered for predictive methods. Generally, such models would be used by the Assessors and would not be part of a study report.
Malaysia (NPRA)	GLP principles are followed wherever possible.
Malaysia (NPRA)	So far, we don't have any test facility that utilizes computational method in GLP studies.
New Zealand	We do not want another 'standard' or GXP that sits alongside the GLP Principles for safety studies done by computational methods; it is not necessary. It would be overtly bureaucratic to have to write GCMP into MAD and every other regulatory instrument. They are still safety studies and thus GLP can and should be applied. What we want is an application document that sits 'under' the GLP Principles - how the Principles are applied to computational method studies. CMAs can apply this document should they come across a compliant test facility conducting these sorts of studies - they are then covered by existing compliance statements, without having to issue new statements calling up another GXP (GCMP).
Poland	Generally we think that GLP is not required now for computational methods. But if the RA decide that computational methods should be cover by GLP we need to change GLP requirements and from GLP inspectors perspective we need a guidance how to control such studies.
Slovakia	IT specialist should be involved to GLP more than nowadays
Switzerland	The Broad Definition of Computational Methods in the background document leads to rather ambiguous answers.
Switzerland	The new guidance on Good Computational Method Practice should be established in collaboration with WNT and WG on GLP.
UK	Our general comments are that discussion may be needed on when a computational method is either treated as software (so can be treated as any other vendor), or could be considered a "GLP study" maybe for example when an in-silico method replaces an animal study - in this case would regulators still need assurance that data is reliable, and still remain in a regulated environment so still be required to be a GLP study.
US (EPA)	The US EPA regulations allow studies, including Q(SAR), to be submitted with a compliance statement describing the differences between the study and our GLPs. The compliance statement must say the study was either conducted in accordance with our GLP regulations or describe in detail all the differences between the practices used in the study and those required by US EPA regulations or that the person was not a sponsor of the study, did not conduct the study, and does not know whether the study was conducted in accordance with US EPA regulations. Please note the US position from the February 4-6, 2020 JM agenda item #9: Thought-starter on the expansion of MAD to computational methods. Chemical safety evaluation has benefited from the expansion in innovative toxicological methods and the revolution in data science. Regulators acknowledge that modern toxicology is no longer restricted to laboratory experiments, and that MAD is not

limited to the experimental raw data. Indeed, many in vitro methods adopted as OECD TGs (e.g. skin irritation, skin corrosion, eye irritation) include prediction models to translate raw data into results interpretable in a given context (i.e. subcategories defined according to the UN Globally Harmonized System for Classification and Labelling of substances). Adoption of innovative toxicological approaches will increasingly require interpretation of raw data to be meaningful in a given regulatory context. In order for the results to be covered by MAD, OECD TGs need to be precise regarding data interpretation to limit possible diverging interpretations. New in vitro, in chemico, and in silico methods are proposed for testing chemicals as stand-alone methods and to be used in combination to predict increasingly complex endpoints. To date, OECD TGs describe procedures for evaluating chemical effects using a single method. Test chemicals are added to the test system and effects are observed. However, the OECD now has proposals for Guidelines using methods in combination. Methods (i.e. information sources) can be combined in different ways, and thus introduce potential variability in the approaches for evaluating chemical effects and the interpretation of the resulting data. To avoid this potential variability, the OECD launched work to define the information sources and data interpretation procedures for methods used in combination to predict chemical effects on a specified endpoint (i.e. Defined Approaches). Because Defined Approaches (DAs) fix the information sources, how information sources are combined, and the interpretation of resulting data, any two parties using the same DA will come to the same conclusion. The current DAs that are being considered for inclusion in OECD TGs include in vitro and in silico methods used in specified combinations, and data interpretation procedures are relatively simple additive or rules-based models. Other DAs that have been reviewed as Case Studies are entirely in silico and use complex computational data interpretation models. The consideration of increasingly computational approaches for evaluating chemical safety has led to a need to clarify what types of "data" are covered under MAD. In some cases, the information sources (e.g. in silico predictive models) or the translation of raw data using a complex data interpretation procedure to come to a result (e.g. omics approaches) may not easily conform to MAD or principles of GLP, as originally conceived for animal experimental data generated in a laboratory or the field. Revisions to the guidance and instruments that support MAD may be needed to assure Member Countries continue to benefit from international harmonization of chemicals safety testing. Specific considerations regarding computational methods: The 1981 Council Decision on MAD refers to acceptance of "data" and does not explicitly specify the diversity of possible types of data covered (e.g., in silico, in vitro, etc.) or tests used to generate the data (e.g. traditional animal tests, alternative methods). "Computational Methods" may refer to mathematical operations that are applied to raw data resulting from in vitro methods. In most cases, an equation or model is used to convert the raw data in order to make assessments on the safety of test chemicals (e.g. data interpretation procedure). For example, the direct output from in vitro methods (e.g. counts of radioactivity, luminescence, light transmission) are not considered, but rather a standardized computational model is used to convert raw data to something that can be easily used for regulatory purposes (e.g. positive/negative; potency categories). DAs take this a step further by including data interpretation procedures for the data resulting from the combination of more than one information source. Currently, MAD only references "data", however, results of OECD TGs that include the data interpretation are covered by MAD. "Computational Methods" may also refer to in silico approaches that predict the toxicological response, such as quantitative structure-activity relationship (QSAR) models. Methods proposed in OECD Guidelines for DAs for skin sensitization include computational (in silico) methods to be used with in vitro data. In the future, other approaches may be proposed that do not include any (de novo) laboratory-derived data. However, the principles of GLP are specific to laboratory-generated data and are not relevant for in silico predictive methods. Another quality system may therefore be needed to assure computational data are high quality, reproducible, and accepted for regulatory decision under MAD. Recent Developments in Countries and OECD International regulatory authorities are exploring opportunities to reduce or ban animal testing and expand the use of non-animal methods. For example, the European Union Directive

2010/63/EU restricts the manufacture or marketing of cosmetic products that have undergone animal tests, nor can companies rely on in vivo data for cosmetics products imported from outside the European Union. In December 2016, a report produced by the Netherlands National Committee for the Protection of Animals called for eliminating animal testing for chemical safety, food ingredients, pesticides and medicines by 2025. In September 2019, the US Environmental Protection Agency issued a directive to reduce animal testing by 30% by 2025 and to completely eliminate animal testing by 2035. An OECD project to develop a developmental neurotoxicity testing battery to address regulatory needs is underway. The project is supported by complementary activities from the Danish EPA, EFSA, and the US EPA. Neurodevelopment may be affected by a variety of complex processes and accurate prediction of adverse outcomes will involve compilation of in vivo and in vitro data from a variety of regulatory agencies, development of in vitro assays measuring a suite of molecular targets, and sophisticated computational approaches to integrate data in a predictive model. Developing alternative methods for predicting complex endpoints will also rely on more intelligent test systems. Organotypic 2D and 3D cell culture systems are capable of expressing physiological biomarkers of organ systems function and are robust human tissue mimetics. In 2016, the US National Institutes of Health established three Tissue Chip Testing Centers to test and validate microphysiological tissue chips. The validation effort aims to adhere to OECD standards for method validation, guidance for non-guideline methods, and guidance documents published by other agencies regarding validation of alternative methods for regulatory application. A similar effort has been undertaken with the 2017 European Union ORCHID (Organ-on-Chip development) project involving seven European research institutions. It will likely be several years before organotypic cell culture models and tissue chip technologies are proposed for inclusion in OECD Test Guidelines, but guidelines proposed in the interim are expected to include many of the toxicological and computational advances of the past decade. In order for the OECD Test Guidelines Programme and Mutual Acceptance of Data to remain relevant, there need to be instruments that anticipate uptake of these new technologies for chemical safety testing. Short-, Long- and Near-Term Instruments and Guidance that Could be Adapted or Developed Short-Term: Existing language in Part 1.1 of the 1981 Council Decision could be amended as follows: "…(in chemico, in silico, in vitro, in vivo, etc.) data generated in the standardized evaluation or testing of chemicals in an OECD Member country in accordance with OECD Test Guidelines and OECD Principles of Good Laboratory/Computational Practice shall be accepted in other Member countries for purposes of assessment and other uses relating to the protection of man and the environment". The same validation principles that apply to laboratory methods also apply to alternative approaches. Computational methods, such as in silico models, would still be required to adhere to validation standards discussed below which are intended to assure quality of computational data (equivalent to the aims of GLP for laboratory data). Long-Term: The 1981 Council Decision could be updated to explicitly include language regarding quality principles for computational data, analogous to Good Laboratory Practices for lab-derived data (e.g. "Good Computational Practices"). Development of quality assurance standards for computational data. Current OECD guidance on validation of in silico models for regulatory purposes provides a set of principles (GD No 49 and 69), but no formalized process for acceptance of predictions. There are also existing OECD formats for documentation than can be used to report in silico models and predictions. In addition, relevant sections (e.g. Documentation) of existing guidance on Application of GLP Principles to Computerized Systems (GD No 17) could be further developed to cover computational models. I agree there needs to be a comprehensive discussion and phased evaluation of this proposal. OECD has identified reasonable processes for evaluating methods; however, care should be exercised regarding the extent to which relatively well vetted tools based on mammalian processes can be extrapolated to non-mammalian species. A proposed path forward focuses on developing "good computational practices" which is equivalent to GLP. The proposed "good computational practices" is definitely needed. Approaches to standardize will take time to develop for Mutual Acceptance of Data and will also have to evolve. It seems to be critical

and, yet, difficult to define the scope and details of MAD from computational approaches. Since the technology evolves along with the time, and algorithms usually appear like inaccessible/hard-to-understand concepts, setting up a way to govern the computational approaches can be challenging. Traditional programming uses input and rules to get the output; machine learning, on the other hand, uses the input and output to figure out the rules. Either approach still carries uncertainties. However, machine learning seems capable of taking care of more complex, big data, which was impossible in the past. Another important consideration is data quality and sources (related to the input)…the context of "use" is the key. Other points to be considered include, what level of certainty can be acceptable and what performance criteria can be defined. Several FDA scientists are involved in a consortium-driven project led by Leadscope to develop frameworks for systematic integration of in silico and empirical (in vitro, in vivo, clinical) data to predict complex endpoints of regulatory interest. The project is funded by an NIH/NIEHS grant and has already resulted in a couple of published papers. The project was highlighted at FDA's Predictive Tox Roadmap workshop. It differs from the CiPA models in that the integration is only semi-automatic, and includes the application of human expert knowledge, but it still fits the description of DAs outlined in the MAD document. Acceptance of computational predictions are context dependent. Computational predictions are not always acceptable substitutes for the biological endpoint. For example, computational prediction of bacterial mutagenicity is acceptable for drug impurities per ICH M7(R1) but not for the parent active ingredient. Regarding near-term proposals: Expanding MAD to computational methods by having the plan to amend the existing language in Part 1.1 of the 1981 Council Decision with the inclusion of in silico data along with in chemico, in vitro, and in vivo data, etc. is appropriate. Since the GLP principles are specific to laboratory or field testing and not for in silico predictive methods, the development of "Good Computational Practices (GCP)" analogous to "Good Laboratory Practices (GLP)" will be highly appreciated. US EPA will consider the current OECD guidance on validation of in silico models for regulatory purposes, QMRF and QSAR Prediction Reporting Formats (QPRF) as a first step of developing Good Computational Practices (CGP). However, recommendations and agreements on issues related to validation, transparency, confidence and reproducibility for in silico models need to be considered, e.g., unbiased test set selection, protection against overtraining, confidence estimates for individual predictions, acceptable level of confidence in prediction, etc. For long-term recommendations: Updating the 1981 Council Decision to include language regarding principles for computational data (e.g. "Good Computational Practice") analogous to Good Laboratory Practices for lab-derived data would be acceptable. I agree the development of an assessment framework for in silico models predictions would be helpful.