

Review

True Grit: A Story of Perseverance Making Two Out of Three the First Non-Animal Testing Strategy (Adopted as OECD Guideline No. 497)

Annette Mehling¹, Susanne N. Kolle², Britta Wareing² and Robert Landsiedel^{2,3,*}

¹ BASF Personal Care and Nutrition GmbH, 40589 Duesseldorf, Germany; annette.mehling@basf.com

² Experimental Toxicology and Ecology, BASF SE, 67056 Ludwigshafen, Germany; susanne.kolle@basf.com (S.N.K.); britta.wareing@basf.com (B.W.)

³ Institute for Pharmacy, Pharmacology and Toxicology, Freie Universität Berlin, 14195 Berlin, Germany

* Correspondence: robert.landsiedel@basf.com

Abstract: In the last two decades, great strides have been made in developing alternative methods to animal testing for regulatory and safety testing. In 2021, a breakthrough in regulatory testing was achieved in that the first test strategies employing non-animal test methods for skin sensitization have been accepted as OECD guideline 497, which falls under the mutual acceptance of data (MAD) by OECD member states. Achieving this goal was a story of hard work and perseverance of the many people involved. This review gives an overview of some of the many aspects and timelines this entailed—just from the perspective of one stakeholder. In the end, the true grit of all involved allowed us to achieve not only a way forward in using test strategies for skin sensitization, but also a new approach to address other complex toxicological effects without the use of animals in the future.



Citation: Mehling, A.; Kolle, S.N.; Wareing, B.; Landsiedel, R. True Grit: A Story of Perseverance Making Two Out of Three the First Non-Animal Testing Strategy (Adopted as OECD Guideline No. 497). *Cosmetics* **2022**, *9*, 22. <https://doi.org/10.3390/cosmetics9010022>

Academic Editors: Cristina Avonto, Amar Gopal Chittiboyina, Elena Gimenez-Arnau and Enzo Berardesca

Received: 9 December 2021

Accepted: 26 January 2022

Published: 5 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: skin sensitization; OECD guideline 497; nonanimal methods; new approach methodologies (NAMs); testing strategies; “two out of three”

1. Introduction

According to the EU Cosmetic Regulation (EU 2009) a “cosmetic product means any substance or mixture intended to be placed in contact with the external parts of the human body (epidermis, hair system, nails, lips and external genital organs) or with the teeth and the mucous membranes of the oral cavity with a view exclusively or mainly to cleaning them, perfuming them, changing their appearance, protecting them, keeping them in good condition or correcting body odours”. Similar definitions can be found in other regulatory frameworks, and these definitions govern products placed on the market in that region. Cosmetic products therefore include not only decorative cosmetics but also products such as shampoos, skin creams, toothpastes, deodorants and sunscreens. Safety is fundamental for any product placed on the market, and for cosmetics the risk of adverse health effects could be high because of the often high and frequent exposure. This is one reason why the Scientific Committee on Consumer Safety published the SCCS Notes of Guidance for Testing of Cosmetic Ingredients and their Safety Evaluation and regularly updates them (currently the 11th revision [1]).

The most prominent adverse reactions to topically applied substances are irritation and sensitization. This makes the identification of potential hazards arising from skin or eye contact an essential endpoint in safety assessments, not only for the consumer setting, but also for the workplace. Consequently, identification of the irritation and sensitization potentials of a substance are core requirements for chemical registrations in many national legislations. With the animal testing ban and concomitant marketing ban imposed by the EU Cosmetics Regulation [2], as of 2013, no animal testing on formulations or their ingredients for the purposes of the regulation are allowed. This, along with the

increased ethical concerns on animal testing [3], makes non-animal approaches to assess these endpoints essential.

For a raw materials supplier, further regulations also exist, the most prominent in Europe probably being EU Regulation (EC) No. 1907/2006 on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) [4] and the European Classification, Labelling and Packaging of substances and mixtures (CLP) [5]. According to the European Chemicals Agency (ECHA), animal testing requirements continue to apply to tests needed to assess the risks of exposure by workers even if a substance is registered exclusively for cosmetic use [6]. REACH and the Cosmetic Regulation are also sometimes conflicting regarding animal testing. Recently, 3206 dossiers were identified in the REACH database on substances with cosmetic use, of which 419 report cosmetics as the only use [7].

As mentioned, these regulations first affect a raw materials manufacturer, as they apply to the raw material, and potential health hazards are not only of concern to consumers but for workers as well. In general, standardized approaches are used in regulatory toxicology to assess a specific toxicological endpoint. Over the years, the Organisation for Economic Co-operation and Development (OECD) has been instrumental in setting international standards for testing. The OECD facilitates mutual acceptance of data (MAD) by OECD member states if results were obtained by methods which were validated, accepted for regulatory use and translated into standardized OECD test guidelines. Results of tests conducted under Good Laboratory Practice (GLP) and according to an OECD test guideline then have international validity and fall under MAD.

Regulatory acceptance is needed to ensure that new methods are really applied. Yet, science often develops at a different pace than does the regulatory landscape. Often, regulations lag behind, and this is a disadvantage for new approach methods (NAM) [8], and this has fashioned a call for closing the gap between regulatory chemicals testing and modern safety science [9]. Since regulations themselves can sometimes be decades old, they have not always been sufficiently adapted to and updated with scientific progress made. Recently, “a big step forward” [10] was taken to close this gap: The first toxicological testing strategies for skin sensitization hazard assessment without the use of animal testing was approved by the OECD. The new OECD Guideline (GL) No. 497, Guideline on Defined Approaches for Skin Sensitisation, is the first OECD GL to include defined approaches (DAs), a new type of (non-test) guideline that uses combined information from non-animal methods to provide toxicological information for hazard and/or potency assessments [11].

The following is a time-lapse view of the events leading to OECD GL 497 from the perspective of just one of the many stakeholders involved and with a focus on the “2 out of 3” DA for skin sensitization hazard assessments. This DA is based on NAMs that had previously been adopted by the OECD, whereas two additional DAs include in silico information, and are somewhat less accurate in terms of hazard identification but can provide some additional information on potency. People who are not involved in the regulatory acceptance process rarely have any appreciation of how long it takes, and how much work goes into it. This review aims at giving the reader an idea of the adventure of bridging the gap and testing the “grit” of everybody involved in various ways.

2. Setting the Scene

The main aim of toxicological testing is to reliably predict adverse effects in humans. For this purpose, tests are performed on animals *in vivo*, or by utilizing cell culture, interactions with biomacromolecule, etc., *in vitro*. However, these are all imperfect proxies for humans—for one, animals are not humans, and cells, although often human origin, are not the whole organism. Often—and fortunately—data on toxic effects in humans are missing and only animal data are available. Hence, new *in vitro* methods are developed using existing animal *in vivo* data as a reference. For skin sensitization, unlike almost all other toxicological effects, a good number of good human data is available. Hence, results from the standard animal tests and non-animal approach can be compared to actual human skin sensitizing potentials of test substances [12,13].

As each *in vitro* method represents only a small part of an organism or pathway, in general, several methods need to be combined to address complex toxicological effects. We hence need to understand the molecular and cellular events leading to a toxic effect to develop and combine the corresponding methods:

Allergic contact dermatitis (ACD) develops in two phases: (1) the sensitization (or induction) phase which develops after the first contact with an allergen, but no visible symptoms occur, and (2) the elicitation (or challenge) phase in which, after renewed contact with the allergen, the actual allergy occurs. Most chemicals of low molecular weight (generally considered to be a molecule size below 500 Da) cannot be recognized by our immune system; in order to trigger skin sensitization (the first of two phases leading to skin allergy), they need to bind to proteins of the skin. The altered skin proteins are the actual antigens leading to delayed type (Type IV, according to Gell [14]) skin sensitization. Usually, the immune system only responds to foreign or altered haptens or proteins if they are also dangerous. These “danger signals” [15] are needed to signal that some stress or injury to the skin has occurred and may be associated with the presence of a foreign protein. Once a protein has been altered in the skin by binding to the test substance, it can be recognized as “foreign” by specialized immune cells in the skin, the antigen presenting cells (e.g., dendritic cells). If these cells recognize a foreign protein in the context of “danger”, they start to change their appearance, they mature, and migrate to the lymph node. There, the mature antigen presenting cells present the altered protein to the T cells and the immune response is initiated.

The process from the first contact of human skin with a substance to the outbreak of an ACD upon a repeated contact is what in modern toxicology is called an adverse outcome pathway (AOP) and the major steps involved are termed “key events”. The AOP for skin sensitization was one of the first AOPs described in human toxicology [16]. Classical toxicology for testing the adverse outcome was, and still is, often done in animals upon test substance exposure; the guinea pig assays (OECD TG 406 [17]) giving information on both the sensitization and induction phases, and the local lymph node assay (LLNA, OECD TG 429 [18]) assessing the sensitization phase only. Testing one test substance in the LLNA requires at least twenty mice and, until 2018, an estimated 33,000 mice were used for skin sensitization testing under REACH [19].

By identifying the key events of an AOP, this offers testing possibilities assessing these using nonanimal tests instead. The well-described skin sensitization AOP (OECD 2014) was the basis to develop the “2 out of 3” testing strategy. Several companies had already developed various *in vitro* methods to test for specific key events of the AOP; selected test methods have been adopted by the OECD since 2015 (described in OECD TGs 442C, 442D and 442E [20–22]), more methods are in the approval process, and the first OECD-adopted defined approach was published in 2021 as OECD GL 497 [11].

How did we get there—prepare yourself for a long story

3. Embarking on the Journey

When we started looking into non-animal alternatives to skin sensitization testing, we were not sure how long it would take to develop something useful and if we would be successful at all. Skin sensitization is a complex process involving different cell types, tissues and processes. Up to then, nothing that complex had been addressed by NAMs. As, however, a good understanding of the underlying biological mechanisms leading to the development of skin sensitization and allergies was already available, starting from scratch was not necessary. Furthermore, other research groups and companies in particular were also looking into this and a number of methods had already been published both in *in vitro* methods [23] and computational methods (*in silico*; [24]). Following a careful review and evaluation, four methods out of all assays available at that time were selected [25]. Luckily, the methods selected were already quite mature and were later included in OECD TG 442C (Key Event-Based Test Guideline for *in chemico* skin sensitisation assays addressing the Adverse Outcome Pathway Key Event on Covalent Binding to Proteins, [22]), OECD

TG 442D (Key Event-Based Test Guideline for in vitro sensitisation assays addressing the AOP key event on keratinocyte activation, [20]), and/or OECD TG 442E (In vitro skin sensitisation assays addressing the key event on activation of dendritic cells on the adverse outcome pathway for skin sensitisation [21]).

Having made the selection of methods deemed to be most developed and relevant, we got the methods up and running, and initiated an in-house validation of the methods [25]. The methods selected addressed known mechanisms involved in the sensitization pathway AOP described by the OECD [16]. The Direct Peptide Reactivity Assay (DPRA; [26]) addressed the protein binding needed for haptenization, the KeratinoSens™ and LuSens assays looked at the signals keratinocytes emit upon contact with potential sensitizers [27], and the dendritic cell activation tests, and the human cell line activation test (h-CLAT) and the modified myeloid U937 skin sensitisation test (mMUSST) assays, detected maturation markers essential to activate T-cell responses [28–30]. The results for 23 test substances from the different methods were then compared to the gold standard animal test method—the local lymph node assay (LLNA), but also with human data where available. The results looked promising [25].

Feeling quite confident, that we had selected the right assays, we embarked on the next leg of our journey—full of drive and optimism. Additional substances were tested, and the decision was made to look into mechanistic AOP-based combinations of these assays to come to an overall conclusion. The strategies (later to be called data interpretation procedures (DIP) in the context of DAs) were based on the known events described above, but proof was needed that this testing strategy using only three assays predicts skin sensitization potentials in humans at least as well as the animal test. This was done in 2011 and published in [31]. In the presented study, the testing strategy using a prediction model based on the outcome of “2 of out of 3” methods addressing the key events driving the classification (two negative results indicating a non-sensitizer, two positives indicating a sensitizer; Figure 1) ultimately provided reliable data with which to predict skin sensitization in humans. This strategy actually demonstrated a better predictivity than the “gold standard” animal test (LLNA; OECD TG 429 [18]). At this point, we were assured that a skin sensitization testing strategy without the use of animals was possible. We therefore submitted the “2 out of 3” strategy to EURL ECVAM for evaluation in 2011 but were put on hold due to the pre-validation and validation studies with NAMs for skin sensitization taking place at that time.

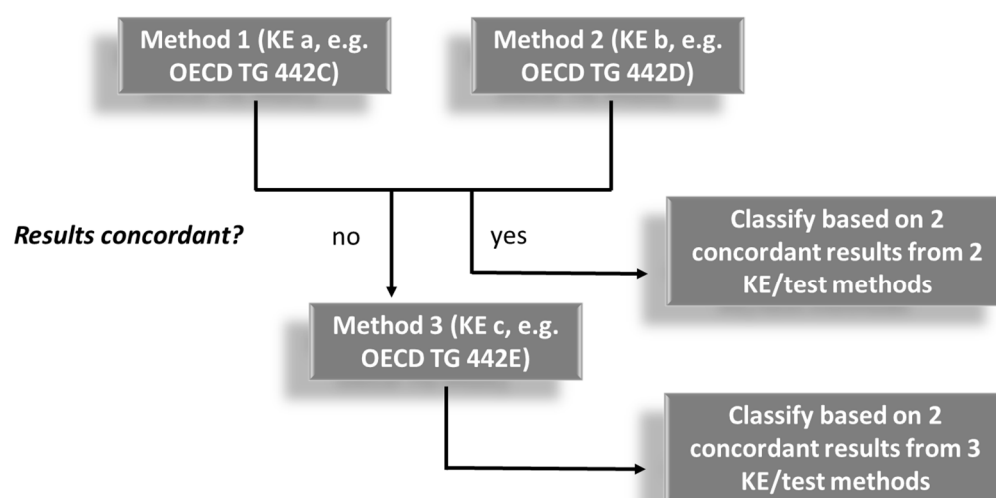


Figure 1. The “2 out of 3” prediction model. Two concordant results of the test methods (DPRA, KeratinoSens or LuSens and h-CLAT or mMUSST) addressing a combination of the first three key events (KE a, b and c) determine the outcome. The sequence of conduct is irrelevant for the prediction.

In parallel, the OECD had been working on defining an “official” AOP “The Adverse Outcome Pathway for Skin Sensitization Initiated by Covalent Binding to Proteins” [16], which was then first published in 2012. It formally defined the key events of the processes leading to skin sensitization and allergy. These included the key events mentioned above, namely, the protein binding, keratinocyte and dendritic cell activation and the resulting T-cell responses (Figure 2). This confirmed that we were moving in the right direction. In the following years, more and other AOP-based DAs were introduced and later published as case studies for the “Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitization” [32].

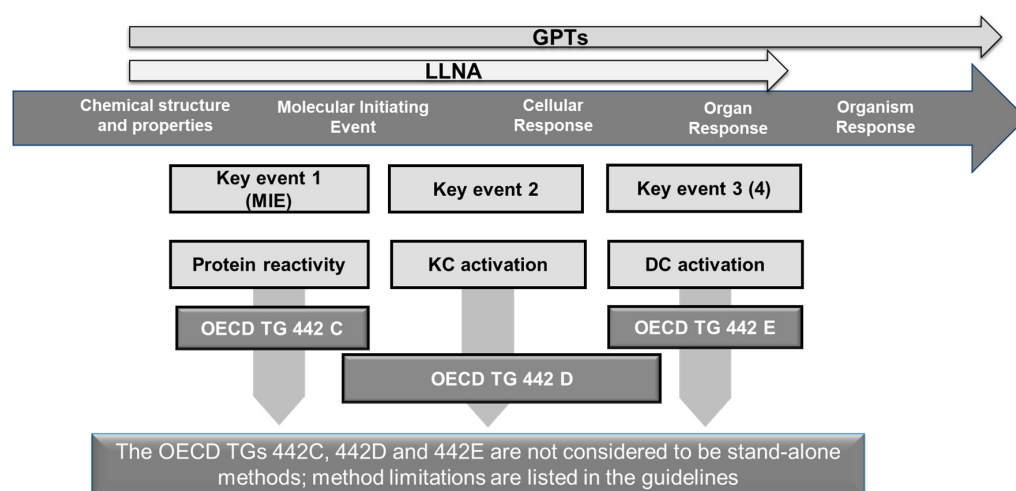


Figure 2. The adverse outcome pathway for skin sensitization and examples of test methods (non-exhaustive) (reproduced from [33] according to [16]).

So, we gritted our teeth and used the delay to further evaluate the methods, and the predictivity of the methods and the “2 out of 3” strategy itself by expanding the data sets. To this accord, we combined forces (and data, Table 1) with other companies and published more data in 2015 [13]. In the meantime, the methods we were using were adopted by the OECD and were included in OECD 442C, D and later E. At that point, it was clear that predicting human skin sensitization without using animals is possible.

Table 1. Predictivity of single in vitro methods, the “2 out of 3” approach and the in vivo LLNA compared to human data. ^a Table reproduced under CC-BY from [13].

Chemical Set and Reference Data	Subset B							
	Human Data ^a				LLNA Data			
Cooper Statistics	Se ^b [%]	Sp ^b [%]	Acc ^b [%]	n	Se ^b [%]	Sp ^b [%]	Acc ^b [%]	n
“2 out of 3” WoE approach	90	90	90	101	81	83	82	103
DPRA	84	84	84	102	77	85	79	105
KeratinoSens TM	82	84	82	102	74	73	74	103
h-CLAT	89	64	82	98	86	68	81	101
LuSens	78	79	79	60	73	70	71	62
(m)MUSST	74	88	78	85	71	83	75	87
LLNA	91	64	82	111	-	-	-	-

^a α -amyl cinnamic aldehyde, α -hexyl cinnamic aldehyde and 6-methyl coumarin could not finally be evaluated as human sensitizers or non-sensitizers and were thus rated as “borderline” ^b Se, sensitivity; Sp, specificity; Acc, accuracy; n, number of substances analyzed.

The next and last step was to convince the regulatory community that this is a reliable approach, and a non-animal testing strategy can be adopted for regulatory use. We did not expect this to take so many years, but we appreciate that methods intended to test for human health hazards require meticulous scrutiny and that the concept of non-animal testing strategies (or DAs), rather than single assays, was new. Along this way of gaining trust and acceptance, several issues had to be clarified, like the applicability to pre- and pro-haptens [34,35], the uncertainties associated with these methods [36,37] and also a careful review of the human data which were used as a reference [38]. Many other papers were also published by the various groups, and more and more information was made available to the scientific and regulatory communities. This included data covering LLNA and human test results, in vitro, in silico and in chemico results, and of course their predictivities (reviewed in, e.g., [12,39]).

4. The Concomitantly Changing Regulatory Landscape

Formal validation to establish and adopt an OECD test guideline was and is often conducted in association with validation centers such as EURL-ECVAM in Europe, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in the US, the Japanese Center for the Validation of Alternative Methods (JaCVAM) in Japan, etc. The formal validation process at the time was a resource consuming process, which was made evident by the workflow found on the EURL ECVAM's Validation Process website [40]. An example of typical steps within the EU in the process leading to the OECD adoption of a test method is depicted in Figure 3 (an example of a workflow only). Even in its simplified depiction, it is probably evident that the whole procedure is/was quite complex. Over the years, some steps have been simplified, responsibilities have changed, etc., thereby speeding up the process.

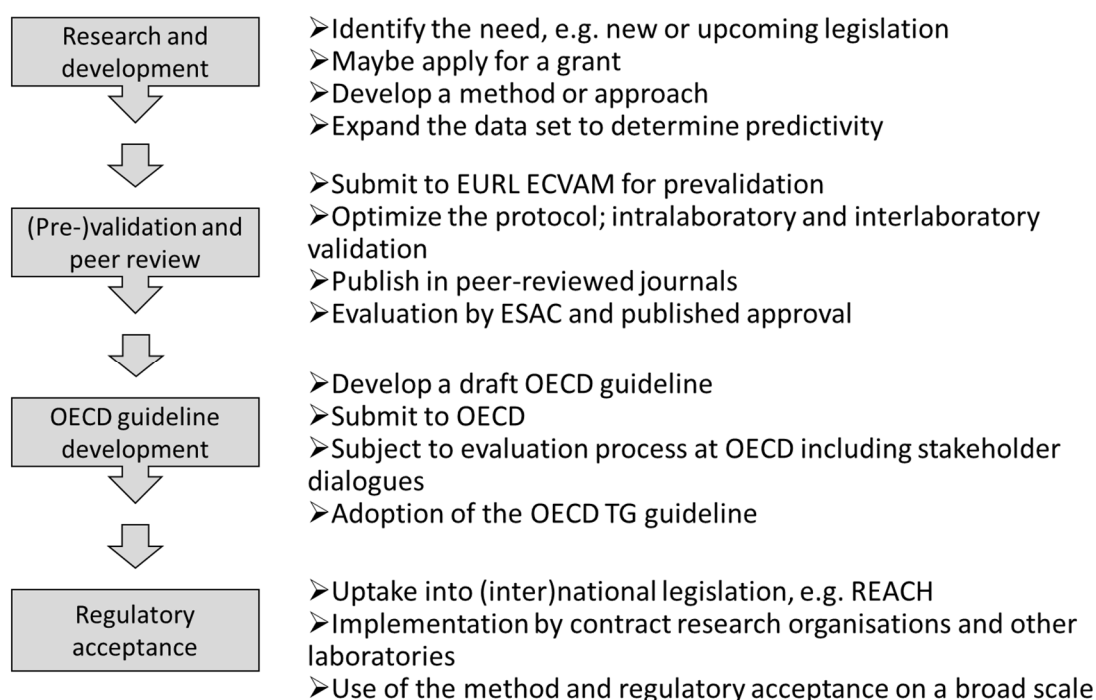


Figure 3. A flow chart of typical steps that comprise the regulatory acceptance procedure of a test method (example only—other steps, etc., possible). Modified from [41]. ESAC: EURL-ECVAM Scientific Advisory Committee.

The OECD test guideline (TG) numbers and structure were originally dedicated to an individual test method and which were often considered to be “stand-alone” methods. Following the publication of the AOP, some OECD TG numbers were then used to denote

the key event and tests addressing that key event are now clustered under the same OECD TG No., e.g., OECD TG 442C (peptide binding); OECD 442D (keratinocyte activation) and OECD TG 442E (dendritic cell activation). Unfortunately, and in contrast to animal tests, *in vitro* tests were not considered to be stand-alone methods and combinations were considered necessary. Yet, combinations and strategies did not fall under the MAD criteria as those back then only covered individual TGs not combinations thereof. In addition, other requirements, such as the definition of proficiency chemicals, etc., were implemented in the nonanimal test methods. More hurdles to be overcome and more time that needed to be invested. A list of currently available NAMs to assess skin sensitization can be found in Table 2. More tests are in the process of being evaluated, and, if deemed suitable, will then be integrated in updated OECD TGs.

Table 2. OECD-adopted NAMs to test skin sensitization potential.

OECD (Test) Guideline Key Event Addressed	Last Updated	AOP-Based Test Method or DA	Outcome
OECD TG 442C Key Event 1 (peptide/protein binding)	2021	DPRA	yes/no information on KE1 as part of an integrated assessment
		ADRA	yes/no information on KE1 as part of an integrated assessment
		kDPRA	quantitative potency information as stand-alone or part of an integrated assessment: Cat ^a 1A or Cat 1B/NS
OECD TG 442D Key Event 2 (Keratinocyte responses)	2018	Keratinos ens TM	yes/no information on KE2 as part of an integrated assessment
		LuSens	yes/no information on KE2 as part of an integrated assessment
OECD TG 442E Key Event 3 (monocyte/dendritic cell responses)	2018	h-CLAT	yes/no information on KE3 as part of an integrated assessment
		U-SENS TM	yes/no information on KE3 as part of an integrated assessment
		IL-8 Luc	yes/no information on KE3 as part of an integrated assessment
OECD GL 497 Combining Key Events 1, 3 and/or 2	2021	2 out of 3	non-sensitizer or non-sensitizer
		ITS v1 and v2	non-sensitizer or sensitizer (Cat 1A or 1B)

^a Category (Cat) 1A, 1B and nonsensitizer (NS) according to Classification, Labelling and Packaging of substances and mixtures (CLP) Regulation (CE) 1272/2008 [5]; AOP: adverse outcome pathway; DA: defined approach; GL: guideline; ITS integrated testing strategy; KE: key event; TG: test guideline.

Over the years, there was also quite a bit of confusion created by the different use of terminologies, with integrated testing strategies (ITS), weight of evidence (WoE), integrated approaches to testing and assessment (IATA), etc., being used interchangeably depending on the authors and context but people understanding something different [8]. The OECD, along with EURL-ECVAM and other stakeholders, then made the effort to not only harmonize terminology but also reporting templates [32] with the aim of facilitating regulatory acceptance by harmonized reporting. The guidance document “Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation” was then published in 2016 [32]. As a very general delineation of the terms, IATA used to describe the full decision tree leading to classification and/or risk assessments. The current OECD working definition of an IATA is as follows: “a structured approach used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment

(potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted testing and therefore optimising and potentially reducing the number of tests that need to be conducted". The term "Defined Approaches" (DAs) was used to describe fixed testing strategies, e.g., ITS (information from tests is gathered in parallel) and sequential testing strategies (STS; information is gathered in a sequential manner), and that are based on defined data integration procedures (DIP, i.e., the fixed prediction model is based on a certain algorithm defined prior to testing). WoE is then used to describe a non-formalized procedure primarily based on expert judgement and the available data and having bearing on hazard and risk assessments. Case studies, including the "2 out of 3" approach, illustrating how to report these testing strategies and the information sources were also published [42,43].

In the meantime, REACH had also come into full force in 2018, and with the acceptance of the non-animal methods by the OECD, animal welfare considerations, etc., testing using animal methods became a "last resort" and justification is needed to conduct them. Finally, in 2018, the USA concluded that the "2 out of 3" was suitable [44]—the finishing line was coming into sight.

5. Heading down to the Finish Line: OECD GL 497

We started developing and validating new non-animal methods and the DA combining these methods in 2009. In 2011 this was submitted to EURL-ECVAM for validation and subsequent submission to the OECD. This was the first testing strategy, now called "defined approach", submitted to OECD; before only single experimental methods, but not a combination of several methods, had been submitted. Hence, it took until 2017 for the DA to be entered in the OECD working plan and it took another four years of discussions and data analysis at the OECD expert group before OECD GL 497 [11] was finally approved [45]. During this extended period of time, further efforts were put into the comprehensive curation and analysis of existing animal and human data within the OECD expert group (Table 3). In addition, this time period was also spent to discuss the new concept of DAs (rather than just single methods), and several fundamental questions with adopting a new concept had to be cleared in these discussions.

Table 3. Predictivity of the LLNA compared to human data ^a (Table reproduced under CC-BY from (Natsch et al., 2021).

	Sensitivity [%]	Specificity [%]	Balanced Accuracy [%]	n
Bauch et al., 2012 [31]	96	81	88	50
Urbisch et al., 2015 [13]	91	64	77.5	111
Hoffmann et al., 2018 and Kleinstreuer et al., 2018 [38,46]	85.2	50.0	67.6	128
OECD LLNA database [11] vs. human data [47]	99	39	69	96
OECD database [11]	94	22	58	56

^a For consistency, values are provided with the same number of decimals as reported in the respective original publications.

Everyone who wants or needs to obtain information about the skin sensitization of a new, yet untested, substance can and should use this testing strategy if the substance is amenable to testing using the OECD test methods. Since it has been adopted by the OECD and all OECD member countries, it will accept results from this testing strategy if it was performed under GLP. Usually, many non-OECD countries adhere to this as well. Some OECD countries adopted DAs before the OECD did; the US EPA issued its Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing [48].

There are, however, technical limitations of the non-animal testing methods; for example, test substances which are insoluble in water do not work well in these methods. In addition, there are test substances for which the validity of the results has not been adequately demonstrated, like mixtures. Some of these limitations are also described in the newer OECD TG documents. It is important to recognize the limitations of the methods—really of any method—since this is true for the animal test alike. The constant development of new tests may prove beneficial to address some of these limitations, e.g., when testing “difficult substances”, the use reconstructed human epidermis (RhE) models (e.g., SENS-IS or IL-18 methodologies) allow better testing of lipophilic substances [23,49,50], or the use of “omics” to identify combinations of genomic markers that may better identify some sensitizing substances (e.g., GARD; [51,52]), or methods that can be used to determine sensitizing potency (e.g., kDPRA, [53,54])—and the future holds more good methods for non-animal sensitization testing.

In the end, true grit of all involved allowed us to achieve not only a testing strategy to test for skin sensitization without animals but also a new approach to address other complex toxicological effects without animals in the future.

6. Looking Forward

Modern toxicology certainly involves *in vitro* testing in human-relevant non-animal models, and hence studying key events that eventually lead to toxic effects observed in animals and humans is here to stay.

AOPs provide a framework to recognize and address these key events leading to toxic effects, and testing strategies—more precisely DAs, as they are called by the OECD—provide a way to utilize results of non-animal testing methods addressing the key events of an AOP to obtain an overall assessment of the toxic effect. The testing strategy (or DA) on skin sensitization is therefore a vanguard of new non-animal testing strategies of the future. Nowadays, the use of AOPs and DAs is inevitably included in designing potential testing strategies (DAs) along with the new methods. Nowadays, the terminology “New approach method (NAMs)”, i.e., *in vitro*, *in chemico* and *in silico* methods has become a buzzword and is found in numerous publications (e.g., [12,55]).

Another way to replace animal testing is called “grouping”. If we can define a group of substances which is sufficiently similar, we may assume the same toxicological profile for all substances in that group—and if we know the toxicity of one representative of this group, the other substances do not have to be tested. This can obviously spare a lot of animal testing. This method has evolved from a quick and dirty data-filling to an elaborated and well-founded tool [56]. The question is how to decide if substances are “sufficiently similar” to be grouped together. Among chemical and physical properties and functionalities, *in vitro* toxicity testing can provide data on similar biological effects of different substances to justify their grouping. This concept has recently been used for nanomaterials (cf. [57]). This concept can be widened to using the available information not only on one similar substance but all the available toxicological information to predict the toxicity of a yet not tested substance. This is what computational toxicology is doing in applying prediction models to toxicological big data [58]. Again, computational methods fortified by biological information obtained from non-animal testing may be the most promising tool [59].

Recently, ECHA estimated the end of animal testing in regulatory toxicology to be more than forty years away (Dr. Hansen speaking at an 18 May virtual forum organised by the Green Chemistry and Commerce Council (GC3), reported by Chemical Watch on 3rd June 2021). This is also what we calculated if we continued replacing one animal method at a time at the current pace. However, trying to find perfect replacements for animal methods, which are imperfect models of humans anyway, may not be the best way. The current risk assessment practice and the regulations of chemicals are based on animal testing and the methodology is designed and fit for animal data. We should look at a system for risk assessment and chemical regulation which is accommodating data obtained from

non-animal methods and existing toxicological data. This will certainly enable replacement of animal testing before four more decades pass by. If we do it right, it will also provide more accurate data and data which are more relevant for human risk assessment. This will also be of importance as the number of countries implementing animal testing bans for cosmetics is increasing globally, and a call for a global ban has even been submitted to the United Nations. Use of animal testing for other sectors is also becoming increasingly difficult. As such, the adoption of non-animal testing strategies to assess human hazards and risks is becoming indispensable. The breakthrough made with OECD test guideline 497 will pave the way to someday making animal testing obsolete.

Author Contributions: Conceptualization, A.M., S.N.K., B.W. and R.L.; Data Compilation, A.M., S.N.K., B.W. and R.L.; Writing—Original Draft Preparation, A.M., S.N.K., B.W. and R.L.; Tables and Figures Preparation, A.M., S.N.K., B.W. and R.L.; Writing—Review & Editing, A.M., S.N.K., B.W. and R.L. All authors have read and agreed to the published version of the manuscript.

Funding: No external funding was obtained for this work.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be found in the publications cited.

Conflicts of Interest: Other than employment, the authors declare no conflict of interest.

References

1. SCCS. SCCS Notes of Guidance for Testing of Cosmetic Ingredients and Their Safety Evaluation. Available online: https://ec.europa.eu/health/sites/default/files/scientific_committees/consumer_safety/docs/sccs_o_250.pdf (accessed on 15 October 2021).
2. EU. Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on Cosmetic Products. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02009R1223-20210823> (accessed on 15 October 2021).
3. Rollin, B.E. The regulation of animal research and the emergence of animal ethics: A conceptual history. *Theor. Med. Bioeth.* **2006**, *27*, 285–304. [[CrossRef](#)] [[PubMed](#)]
4. EU. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), Establishing a European Chemicals Agency, Amending Directive 1999/45/EC and Repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02006R1907-20211001> (accessed on 15 October 2021).
5. EU. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on Classification, Labelling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No 1907/2006. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02008R1272-20211001> (accessed on 15 October 2021).
6. ECHA. Interface between REACH and Cosmetics Regulations. Available online: https://echa.europa.eu/documents/10162/13628/reach_cosmetics_factsheet_en.pdf/2fbcf6bf-cc78-4a2c-83fa-43ca87cfb314 (accessed on 15 October 2021).
7. Knight, J.; Rovida, C.; Kreiling, R.; Zhu, C.; Knudsen, M.; Hartung, T. Continuing Animal Tests on Cosmetic Ingredients for REACH in the EU. *ALTEX* **2021**, *38*, 653–668. [[CrossRef](#)] [[PubMed](#)]
8. Sauer, U.G.; Hill, E.H.; Curren, R.D.; Raabe, H.A.; Kolle, S.N.; Teubner, W.; Mehling, A.; Landsiedel, R. Local tolerance testing under REACH: Accepted non-animal methods are not on equal footing with animal tests. *Altern. Lab. Anim.* **2016**, *44*, 281–299. [[CrossRef](#)] [[PubMed](#)]
9. Fentem, J.; Malcomber, I.; Maxwell, G.; Westmoreland, C. Upholding the EU's Commitment to 'Animal Testing as a Last Resort' Under REACH Requires a Paradigm Shift in How We Assess Chemical Safety to Close the Gap Between Regulatory Testing and Modern Safety Science. *Altern. Lab. Anim.* **2021**, *49*, 122–132. [[CrossRef](#)]
10. Culliney, K. 'A Big Step Forward': BASF and Givaudan Receive OECD Approval on Animal-Free Skin Sensitisation and Allergen Potency Testing Strategy. Available online: <https://www.cosmeticsdesign-europe.com/article/2021/06/30/ANIMAL-FREE-TESTING-FOR-SKIN-SENSITISATION-ALLERGEN-POTENCY-FROM-BASF-AND-GIVAUDAN-RECEIVES-OECD-APPROVAL> (accessed on 15 October 2021).
11. OECD. Guideline No. 497: Defined Approaches on Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/b92879a4-en> (accessed on 15 October 2021).

12. Natsch, A.; Landsiedel, R.; Kolle, S.N. A triangular approach for the validation of new approach methods for skin sensitization. *ALTEX* **2021**, *38*, 669–677. [CrossRef]
13. Urbisch, D.; Mehling, A.; Guth, K.; Ramirez, T.; Honarvar, N.; Kolle, S.; Landsiedel, R.; Jaworska, J.; Kern, P.S.; Gerberick, F.; et al. Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.* **2015**, *71*, 337–351. [CrossRef]
14. Gell, P.G.H.C.R.R.A. *Clinical Aspects of Immunology*; Blackwell: Oxford, UK, 1963.
15. Matzinger, P. The danger model: A renewed sense of self. *Science* **2002**, *296*, 301–305. [CrossRef]
16. OECD. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264221444-en> (accessed on 15 October 2021).
17. OECD. Test No. 406: Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264070660-en> (accessed on 15 October 2021).
18. OECD. Test No. 429: Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264071100-en> (accessed on 15 October 2021).
19. Taylor, K. Ten years of REACH—An animal protection perspective. *Altern. Lab. Anim.* **2018**, *46*, 347–373. [CrossRef]
20. OECD. Test No. 442D: In Vitro Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264229822-en> (accessed on 15 October 2021).
21. OECD. Test No. 442E: In Vitro Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264264359-en> (accessed on 15 October 2021).
22. OECD. Test No. 442C: In Chemico Skin Sensitisation. Available online: <http://www.oecd.org/env/ehs/testing/draft-test-guideline-442C-in-chemico-skin-sensitisation.pdf> (accessed on 15 October 2021).
23. Mehling, A.; Eriksson, T.; Eltze, T.; Kolle, S.; Ramirez, T.; Teubner, W.; van Ravenzwaay, B.; Landsiedel, R. Non-animal test methods for predicting skin sensitization potentials. *Arch. Toxicol.* **2012**, *86*, 1273–1295. [CrossRef]
24. Teubner, W.; Mehling, A.; Schuster, P.X.; Guth, K.; Worth, A.; Burton, J.; van Ravenzwaay, B.; Landsiedel, R. Computer models versus reality: How well do in silico models currently predict the sensitization potential of a substance. *Regul. Toxicol. Pharmacol.* **2013**, *67*, 468–485. [CrossRef]
25. Bauch, C.; Kolle, S.N.; Fabian, E.; Pachel, C.; Ramirez, T.; Wiench, B.; Wruck, C.J.; van Ravenzwaay, B.; Landsiedel, R. Intralaboratory validation of four in vitro assays for the prediction of the skin sensitizing potential of chemicals. *Toxicology* **2011**, *25*, 1162–1168. [CrossRef]
26. Gerberick, G.F.; Vassallo, J.D.; Bailey, R.E.; Chaney, J.G.; Morrall, S.W.; Lepoittevin, J.P. Development of a peptide reactivity assay for screening contact allergens. *Toxicol. Sci.* **2004**, *81*, 332–343. [CrossRef] [PubMed]
27. Emter, R.; Ellis, G.; Natsch, A. Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. *Toxicol. Appl. Pharmacol.* **2010**, *245*, 281–290. [CrossRef] [PubMed]
28. Ashikaga, T.; Yoshida, Y.; Hirota, M.; Yoneyama, K.; Itagaki, H.; Sakaguchi, H.; Miyazawa, M.; Ito, Y.; Suzuki, H.; Toyoda, H. Development of an in vitro skin sensitization test using human cell lines: The human Cell Line Activation Test (h-CLAT). I. Optimization of the h-CLAT protocol. *Toxicology* **2006**, *20*, 767–773. [CrossRef]
29. Python, F.; Goebel, C.; Aeby, P. Assessment of the U937 cell line for the detection of contact allergens. *Toxicol. Appl. Pharmacol.* **2007**, *220*, 113–124. [CrossRef] [PubMed]
30. Sakaguchi, H.; Ashikaga, T.; Miyazawa, M.; Yoshida, Y.; Ito, Y.; Yoneyama, K.; Hirota, M.; Itagaki, H.; Toyoda, H.; Suzuki, H. Development of an in vitro skin sensitization test using human cell lines; human Cell Line Activation Test (h-CLAT). II. An inter-laboratory study of the h-CLAT. *Toxicology* **2006**, *20*, 774–784. [CrossRef]
31. Bauch, C.; Kolle, S.N.; Ramirez, T.; Eltze, T.; Fabian, E.; Mehling, A.; Teubner, W.; van Ravenzwaay, B.; Landsiedel, R. Putting the parts together: Combining in vitro methods to test for skin sensitizing potentials. *Regul. Toxicol. Pharmacol.* **2012**, *63*, 489–504. [CrossRef]
32. OECD. Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation. Available online: <https://www.oecd-ilibrary.org/content/publication/9789264279285-en> (accessed on 15 October 2021).
33. Kolle, S.N.; Teubner, W.; Landsiedel, R. Modern Skin Toxicity Testing Strategies. In *Environment and Skin*; Krutmann, J., Merk, H.F., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 27–40.
34. Urbisch, D.; Becker, M.; Honarvar, N.; Kolle, S.N.; Mehling, A.; Teubner, W.; Wareing, B.; Landsiedel, R. Assessment of Pre- and Pro-haptens Using Nonanimal Test Methods for Skin Sensitization. *Chem. Res. Toxicol.* **2016**, *29*, 901–913. [CrossRef]
35. Patlewicz, G.; Casati, S.; Basketter, D.A.; Asturiol, D.; Roberts, D.W.; Lepoittevin, J.P.; Worth, A.P.; Aschberger, K. Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization? *Regul. Toxicol. Pharmacol.* **2016**, *82*, 147–155. [CrossRef]
36. Gabbert, S.; Mathea, M.; Kolle, S.N.; Landsiedel, R. Accounting for Precision Uncertainty of Toxicity Testing: Methods to Define Borderline Ranges and Implications for Hazard Assessment of Chemicals. *Risk Anal.* **2020**. [CrossRef]
37. Kolle, S.N.; Mathea, M.; Natsch, A.; Landsiedel, R. Assessing Experimental Uncertainty in Defined Approaches: Borderline Ranges for In Chemico and In Vitro Skin Sensitization Methods Determined from Ring Trial Data. *Applied Toxicol.* **2021**, *7*, 102–111. [CrossRef]

38. Kleinstreuer, N.C.; Hoffmann, S.; Alepee, N.; Allen, D.; Ashikaga, T.; Casey, W.; Clouet, E.; Cluzel, M.; Desprez, B.; Gellatly, N.; et al. Non-animal methods to predict skin sensitization (II): An assessment of defined approaches (*). *Crit. Rev. Toxicol.* **2018**, *48*, 359–374. [[CrossRef](#)] [[PubMed](#)]
39. Kolle, S.N.; Natsch, A.; Gerberick, G.F.; Landsiedel, R. A review of substances found positive in 1 of 3 in vitro tests for skin sensitization. *Regul. Toxicol. Pharmacol.* **2019**, *106*, 352–368. [[CrossRef](#)] [[PubMed](#)]
40. EU. EU Reference Laboratory for Alternatives to Animal Testing. Available online: <https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/eurl-ecvams-validation-process/eurl-ecvams-validation-process#figure-1-the-eurl> (accessed on 15 October 2021).
41. EU. Validation and Submission Process. Available online: <https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing/validation> (accessed on 15 October 2021).
42. OECD. Annex I: Case Studies to the Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation. Available online: [https://one.oecd.org/document/ENV/JM/WRPR\(2016\)62/ANN1/en/pdf](https://one.oecd.org/document/ENV/JM/WRPR(2016)62/ANN1/en/pdf) (accessed on 15 October 2021).
43. OECD. Annex II: Information Sources Used within the Case Studies to the Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation. Available online: [https://one.oecd.org/document/ENV/JM/WRPR\(2016\)62/ANN2/en/pdf](https://one.oecd.org/document/ENV/JM/WRPR(2016)62/ANN2/en/pdf) (accessed on 15 October 2021).
44. Strickland, J.; Daniel, A.B.; Allen, D.; Aguila, C.; Ahir, S.; Bancos, S.; Craig, E.; Germolec, D.; Ghosh, C.; Hudson, N.L.; et al. Skin sensitization testing needs and data uses by US regulatory and research agencies. *Arch. Toxicol.* **2019**, *93*, 273–291. [[CrossRef](#)]
45. Kolle, S.N.; Landsiedel, R.; Natsch, A. Replacing the refinement for skin sensitization testing: Considerations to the implementation of adverse outcome pathway (AOP)-based defined approaches (DA) in OECD guidelines. *Regul. Toxicol. Pharmacol.* **2020**, *115*, 104713. [[CrossRef](#)]
46. Hoffmann, S.; Kleinstreuer, N.; Alepee, N.; Allen, D.; Api, A.M.; Ashikaga, T.; Clouet, E.; Cluzel, M.; Desprez, B.; Gellatly, N.; et al. Non-animal methods to predict skin sensitization (I): The Cosmetics Europe database. *Crit. Rev. Toxicol.* **2018**, *48*, 344–358. [[CrossRef](#)]
47. Basketter, D.A.; Alepee, N.; Ashikaga, T.; Barroso, J.; Gilmour, N.; Goebel, C.; Hibatallah, J.; Hoffmann, S.; Kern, P.; Martinozzi-Teissier, S.; et al. Categorization of chemicals according to their relative human skin sensitizing potency. *Dermatitis* **2014**, *25*, 11–21. [[CrossRef](#)]
48. EPA. Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing. Available online: <https://www.regulations.gov/document?D=EPA-HQ-OPP-2016-0093-0090> (accessed on 15 October 2021).
49. Bergal, M.; Puginier, M.; Gerbeix, C.; Groux, H.; Roso, A.; Cottrez, F.; Milius, A. In vitro testing strategy for assessing the skin sensitizing potential of “difficult to test” cosmetic ingredients. *Toxicology* **2020**, *65*, 104781. [[CrossRef](#)]
50. Cottrez, F.; Boitel, E.; Ourlin, J.C.; Peiffer, J.L.; Fabre, I.; Henaoui, I.S.; Mari, B.; Vallauri, A.; Paquet, A.; Barbry, P.; et al. SENS-IS, a 3D reconstituted epidermis based model for quantifying chemical sensitization potency: Reproducibility and predictivity results from an inter-laboratory study. *Toxicology* **2016**, *32*, 248–260. [[CrossRef](#)]
51. Forreryd, A.; Zeller, K.S.; Lindberg, T.; Johansson, H.; Lindstedt, M. From genome-wide arrays to tailor-made biomarker readout—Progress towards routine analysis of skin sensitizing chemicals with GARD. *Toxicology* **2016**, *37*, 178–188. [[CrossRef](#)]
52. Roberts, D.W. Is a combination of assays really needed for non-animal prediction of skin sensitization potential? Performance of the GARD (Genomic Allergen Rapid Detection) assay in comparison with OECD guideline assays alone and in combination. *Regul. Toxicol. Pharmacol.* **2018**, *98*, 155–160. [[CrossRef](#)] [[PubMed](#)]
53. Natsch, A.; Haupt, T.; Wareing, B.; Landsiedel, R.; Kolle, S.N. Predictivity of the kinetic direct peptide reactivity assay (kDPRA) for sensitizer potency assessment and subclassification. *ALTEX* **2020**, *37*, 652–664. [[CrossRef](#)] [[PubMed](#)]
54. Wareing, B.; Kolle, S.N.; Birk, B.; Alepee, N.; Haupt, T.; Kathawala, R.; Kern, P.S.; Nardelli, L.; Raabe, H.; Rucki, M.; et al. The kinetic Direct Peptide Reactivity Assay (kDPRA): Intra- and inter-laboratory reproducibility in a seven-laboratory ring trial. *ALTEX* **2020**, *37*, 639–651. [[CrossRef](#)] [[PubMed](#)]
55. Barthe, M.; Bavoux, C.; Finot, F.; Mouche, I.; Cuceu-Petrenci, C.; Forreryd, A.; Chérouvrier Hansson, A.; Johansson, H.; Lemkine, G.F.; Thénot, J.-P.; et al. Safety Testing of Cosmetic Products: Overview of Established Methods and New Approach Methodologies (NAMs). *Cosmetics* **2021**, *8*, 50. [[CrossRef](#)]
56. Teubner, W.; Landsiedel, R. Read-across for hazard assessment: The ugly duckling is growing up. *Altern. Lab. Anim.* **2015**, *43*, 67–71. [[CrossRef](#)]
57. Janer, G.; Ag-Seleci, D.; Sergent, J.A.; Landsiedel, R.; Wohlleben, W. Creating sets of similar nanoforms with the ECETOC NanoApp: Real-life case studies. *Nanotoxicology* **2021**, *15*, 1016–1034. [[CrossRef](#)]
58. Luechtefeld, T.; Rowlands, C.; Hartung, T. Big-data and machine learning to revamp computational toxicology and its use in risk assessment. *Toxicol. Res.* **2018**, *7*, 732–744. [[CrossRef](#)]
59. Garcia de Lomana, M.; Morger, A.; Norinder, U.; Buesen, R.; Landsiedel, R.; Volkamer, A.; Kirchmair, J.; Mathea, M. ChemBioSim: Enhancing Conformal Prediction of In Vivo Toxicity by Use of Predicted Bioactivities. *J. Chem. Inf. Model* **2021**, *61*, 3255–3272. [[CrossRef](#)]