In silico Prediction of Skin Sensitization: Quo vadis?

Giang Huong Ta¹, Ching-Feng Weng²* and Max K. Leong¹*

¹Department of Chemistry, National Dong Hwa University, Shoufeng, Taiwan, ²Department of Basic Medical Science, Institute of Respiratory Disease, Xiamen Medical College, Xiamen, China

Skin direct contact with chemical or physical substances is predisposed to allergic contact dermatitis (ACD), producing various allergic reactions, namely rash, blister, or itchy, in the contacted skin area. ACD can be triggered by various extremely complicated adverse outcome pathways (AOPs) remains to be causal for biosafety warrant. As such, commercial products such as ointments or cosmetics can fulfill the topically safe requirements in animal and non-animal models including allergy. Europe, nevertheless, has banned animal tests for the safety evaluations of cosmetic ingredients since 2013, followed by other countries. A variety of non-animal in vitro tests addressing different key events of the AOP, the direct peptide reactivity assay (DPRA), KeratinoSens™, LuSens and human cell line activation test h-CLAT and U-SENS™ have been developed and were adopted in OECD test guideline to identify the skin sensitzers. Other methods, such as the SENS-IS are not yet fully validated and regulatorily accepted. A broad spectrum of in silico models, alternatively, to predict skin sensitization have emerged based on various animal and non-animal data using assorted modeling schemes. In this article, we extensively summarize a number of skin sensitization predictive models that can be used in the biopharmaceutics and cosmeceuticals industries as well as their future perspectives, and the underlined challenges are also discussed.

Keywords: skin sensitization, in silico models, human test methods, non-animal test methods, animal test methods, in chemico test methods

INTRODUCTION

Skin is a protective barrier against the external environment to guard internal organs, bones, and muscles. The skin is an organ of the integumentary system made of multiple layers containing epidermis (surface layer), and dermis (deeper layer) (Rehfeld et al., 2017). The topical application and transepidermal delivery of natural or synthesized chemicals are striking approaches for the discovery and development of drugs and medicines by physicians and pharmacologists (Alkilani et al., 2015); and for maintaining healthy skin in general by dermatologists (Kraft and Lynde, 2005; Spada et al., 2018). Upon skin contacts with chemicals or substances that could be non-allergy or allergen caused the hypersensitivity to the subject (Lee and Thomson, 1999). There are four types of skin hypersensitivity, fundamentally, based on the immunologic mechanism that mediates the disease, namely type I (immediate/IgE-related), in which the cutaneous skin test reaction reaches the peak at 2 h; type II (antibody and complement related cytotoxicity); type III (antigen-antibody complex mediated); and type IV or delayed type hypersensitivity (DTH) response, that can occur within 48–72 h (Lee and Thomson, 1999; Posadas and Pichler, 2007). Notably, skin sensitization or allergic contact dermatitis (ACD) is a type IV DTH or type IV allergy (Ouyang et al., 2014). ACD can...
substantially affect the life quality of patients with uncomfortable symptoms of skin rash, blister, and/or swollen that could persist for a lifetime in some cases (Strickland et al., 2016). In illness observation, ACD has affected more than 20% of North America’s and Western Europe’s population based on the data collected from all age groups, and the contact allergy tends to be more prevalent in younger children as the comparison with adults (Thyssen et al., 2007).

The adverse outcome pathways (AOPs) of skin sensitization are the sequential events from the initial skin exposure to chemicals, followed by triggering the downstream cascade pathways, which include induction and elicitation phases. The chemical sensitization pathway (CSP) is initialized by the adduct formation, viz. a covalent bond between skin proteins and chemicals to subsequently form a full antigen (Enoch et al., 2011). Moreover, skin sensitizers have the same physicochemical properties as haptons, which are electrophilic per se (Roberts and Lepoittevin, 1998) or lead to the formation of free radicals (Gáfvert et al., 1994). In some cases, skin sensitizers can be named as prehapten, which initially are not electrophilic or radicals but can be activated through air exposure, photoactivation, bacterial degradation on the skin surface (Karlberg et al., 1992; Sköld et al., 2002) and skin sensitizers also can be termed as prohapten, which can be triggered also through the metabolic pathway (Nilsson et al., 2005; Gerberick et al., 2008; van Eijl et al., 2012). As such, those skin sensitizers may act as electrophiles, whereas the skin protein functions as a nucleophile in the process of addition formation. More specifically, those nucleophilic amino acids such as cysteine (thiols), histidine, lysine (primary amines), methionine, and tyrosine within the skin protein can interact or react with electrophilic hapten (Ahlfors et al., 2003; Gerberick et al., 2004; Schwöbel et al., 2011). This interaction with cysteine and/or lysine leads to the formation of covalent bonds and production of the hapten–protein complex consequently processed by both epidermal and dermal dendritic cells (DCs), which constitute the skin immune system (Ochoa et al., 2008; Clausen and Stoitzner, 2015). Subsequently, the DC presents the part of this protein complex (antigen) on major histocompatibility complex (MHC) and activates naïve T lymphocyte in the lymph node (Martin et al., 2010; Huppert et al., 2018; Johnson et al., 2020). In addition, this can induce the differentiation and proliferation of T cells that, in turn, will propagate the inflammatory response throughout the whole body (Gefen et al., 2015). After the initial exposure, the secondary exposure to the same allergen will initiate the elicitation phase, in which the activated T cells are triggered to secrete specific cytokines to attract inflammatory cells entering into the epidermis of infected parts, causing rash, itchy, and burning on the exposed skin surface. The detail of induction and elicitation phases has been illustrated elsewhere (OECD, 2012). Markedly, the response in the elicitation phase of the immune system is faster than that in the induction phase (OECD, 2012).

The complications of whitening-cosmetics have been documented in recent years since the cosmetic ingredients are not only a major concern in the beauty industry but also a critical factor in human health. Some ingredients in brightening cosmetics such as hydroquinone, corticosteroids, and mercury can cause severe complications. For instance, chronic application of hydroquinone-contained cosmetics can result in exogenous ochronosis or “fish odor syndrome,” the accumulation of mercury can lead to increased pigmentation, nail discoloration, and ACD; and the aggregation of corticosteroids will produce “steroid addiction syndrome” or induce acne on the anterior chest (O lumide et al., 2008; Ladizinski et al., 2011; Mahé, 2014). Cosmeceuticals is a burgeoning industry, in which cosmetic products can exert therapeutic effects (Martin and Glaser, 2011). Vitamin, hydroxy acids, growth factors, peptides, and botanicals, for example, are considered as the cosmeceutical ingredients (Martin and Glaser, 2011). Some skincare products also include ingredients with pharmaceutical properties as exemplified by Oz. Or. Oil 30, which cannot only soften the skin but also show the potential in antibacteria and ameliorating dermal wound healing (Serio et al., 2017). Sargafuran, which is extracted from marine brown alga, is a promising compound to be used in skincare cosmetics for preventing acne because of its antibacterial properties (Kamei et al., 2009). In addition, Food and Drug Administration (FDA) has already approved some antibiotics such as quinupristin-dalfopristin, linezolid, and daptomycin for the treatment of skin-structure infections (Schweiger and Weinberg, 2004). According to FDA regulation, a product can be both a cosmetic and a drug if it can meet the definitions of both cosmetics and drugs. However, the category such as “cosmeceuticals” has not been recognized by FDA. In contrast, cosmeceuticals are a subclass of cosmetics in Europe and Japan, and considered as a subclass of drugs in the UK (Pandey et al., 2021). However, the criteria for classifying compounds or raw materials from the same plant are various in different countries. For instance, the raw materials of C. limon have been considered as the natural ingredients to potentially threat human health by the European Food Safety Authority (EFSA). However, the oil and extracts from this species is classified as safe products by FDA (Klimek-Szczykutowicz et al., 2020), suggesting that these criteria are not universally applicable.

Skin sensitization is an increasingly important issue that can be manifested by the number of publications about skin
sensitization as illustrated in Figure 1. It can be observed that the number of published literature has gradually increased in recent years, especially the dramatic increase after 2000. The consumption and interest in the cosmetic market have progressively increased by 5% every year and it is expected to reach 31.75 billion US dollars by 2023 (Kumar, 2005; Orbis-Research, 2018). The potential benefits and demand are still high and the information about toxicity, physicochemical, and bioactivity properties of the cosmetics ingredients need to be promoted (Kumar, 2005; Panico et al., 2019), the growth of the global cosmetics market is updated annually at http://www.statista.com/statistics/297070/growth-rate-of-the-global-cosmetics-market/.

### SKIN SENSITIZATION ASSAY

Various tests have been devised to evaluate the potential of the human skin sensitization of new substances and they can be basically classified into human tests, animal tests, and non-animal tests as enlisted in Table 1. The more detailed test information will be discussed as follows.

#### Human Tests

Human tests for skin sensitization include human repeat insult patch test (HRIFT) and human maximization test (HMT) (Kligman, 1966; Kligman and Epstein, 1975; Marzulli and Maibach, 1974). In both tests, the human skin reaction is recorded after the secondary contact between a tested substance and human skin. Generally, the response of the tested substance is classified into 5 levels according to the incidence of the positive response from test subjects in the HMT system: weak (0–2/25), mild (3–7/25), moderate (8–13/25), strong (14–20/25), or extreme sensitizer (21–25/25) (Kligman, 1966). The HRIFT classification system is instituted according to the grades of skin reactions: 1) erythema; 2) erythema and induration; 3) vesiculation; and 4) bulla formation and only the substances of grade 1 are qualified as non-sensitizers (Marzulli and Maibach, 1974). To date, the classification systems of skin sensitization in human tests are not consistent and often depend on the subjective judgment of experts (Gerberick et al., 2001; Roberts, 2018). In the classification and labeling of chemicals of the globally harmonized system (GHS), chemicals are classified as subcategory 1A or 1B if their HRIFT or HMT values are ≤500 µg/cm² or HRIFT or HMT values are >500 µg/cm², respectively. Both subcategories are experimentally considered as skin sensitizers. Non sensitizers are not classified in this classification system (ICCVAM, 2011; United Nation, 2013). Nevertheless, the other classification system has also been proposed, in which chemicals are classified into six skin sensitization categories based on their no observed effect level (NOEL) values of HRIFT as enlisted in Table 2 (Basketter et al., 2014). Nowadays, human tests are only implemented to confirm skin sensitization of test chemicals under specific conditions and no maximum concentrations are allowed to apply due to ethical issues (Kimber et al., 2001).

#### Animal Tests

Besides human tests, there are various animal tests have been conducted to evaluate the potential of the human skin sensitization for new substance, namely local lymph node assay (LLNA), which depends on the nature of AOP key events as listed in Table 1 (OECD, 2012), Guinea pig maximization test (GPMT), and Buehler tests. Of various animal assay systems, LLNA (OECD, 2010a; 2010b), which is based on the extent of induced proliferative responses in draining lymph nodes after the topical exposure of chemicals to mice (stimulation index, SI), is the preferred animal test model and has been adopted by various regulatory agencies (Cockshott et al., 2006; Gerberick et al., 2007a). The LLNA system is designated to measure the substance concentration when the lymphocyte proliferation of the lymph node is three-fold higher than that of the vehicle-treated controls, viz. SI ≥ 3, and is defined as the LLNA EC3 value. The risk potential of skin sensitizers is categorized into various classes according to the measured LLNA EC3 values as summarized in Table 3 (ICCVAM, 2011). The LLNA EC3 value has been converted from the percentage to µg/cm² since 2001 to develop the correlation between LLNA EC3 and NOEL value from human tests, namely HRIFT and HMT. It has been found that the EC3 value can be used to quantitatively estimate the skin sensitization potency in human since EC3 values can be highly correlated with NOELs (Gerberick et al., 2001) that also has been confirmed by Api et al. (Api et al., 2015).

The GPMT method is another popular animal model, in which intradermal injection and/or epidermal application were employed in induction periods to expose guinea pig skin with the test substances. The animals are repeatedly exposed to test substances with a challenge dose after 10–14 days. The skin reaction to the challenge exposure in the test animals is determined by comparing it with the untreated control animals (OECD, 1992). The skin sensitization potential of

---

**Table 1**: Animal and non-animal tests to evaluate the potential of the human skin sensitization of new substance depending on key events in AOP (OECD, 2012).

| Key event | Function and focus | Method | References |
|-----------|--------------------|--------|------------|
| 1         | The molecular interaction with skin proteins through cysteine and/or lysine residue | DPRA, kDPRA, ADRA | OECD (2015a), OECD (2020) |
| 2         | The inflammatory response through keratinocyte | KeratinoSens™ | OECD (2017), OECD (2018a) |
| 3         | The activation of dendritic cells | h-CLAT | OECD (2016b), OECD (2018c) |
| 4         | The proliferation of T cells | U-SENS™ | OECD (2010a), OECD (2010b) |

OECD: Organization for Economic Co-operation and Development (France, 1961).
The skin sensitization categories based on the NOEL value of HRIPT.

| Category | Characteristics | HRIPT NOEL value |
|----------|-----------------|------------------|
| 1        | High intrinsic skin sensitization potency | Less than 25 μg/cm² |
| 2        | Less sensitizing than category 1, the contact with moderate concentration can trigger 1–10% positive induction of subjects | Between 25 and 500 μg/cm² |
| 3        | Substances known as contact allergens produce sensitization in 0.01–0.1% of those exposed | Between 500 and 2,500 μg/cm² |
| 4        | Chemicals in this category require prolonged exposure to higher dose level to produce sensitization and are rarely regarded as important clinical allergens | More than 2,500 μg/cm² |
| 5        | Very low intrinsic ability to cause skin sensitization. Even in the highly selected patient groups, the incidence should not exceed 1% | The NOEL values are variable or absent, because of the inaccuracy of determination of a threshold |
| 6        | Free from skin sensitization activity | |

The skin sensitization potency based on LLNA EC3 values.

| Potency category | Threshold (%) |
|------------------|---------------|
| Extreme          | EC3 < 0.1     |
| Strong           | 0.1 ≤ EC3 < 1 |
| Moderate         | 1 ≤ EC3 < 10  |
| Weak             | 10 ≤ EC3 ≤ 100 |

The skin sensitization potential based on GPMT.

| Induction concentration (%) | GPMT incidence (%) |
|----------------------------|--------------------|
|                             | 30 to < 60 | ≥ 60 |
| <0.1                        | Strong    | Extreme |
| ≥0.1 to <1                 | Moderate  | Strong  |
| ≥1 to <10                   | Weak      | Moderate|
| ≥10 to ≤100                 | Weak      | Weak    |

Table 4 shows the skin sensitization potential based on GPMT.

chemicals can be classified into various extents, namely extreme, strong, moderate, or weak levels, depending on the induction concentration and the incidence of subjects as listed in Table 4 (ICCVAM, 2011). The Buehler method is another test to use guinea pig skin as a module. The only difference between GPMT and Buehler test is the way of sample preparation in that the test substance is mixed with Freund’s complete adjuvant (FCA) in the GPMT test, whereas that step is absent in the non-adjuvant Buehler method (OECD, 1992). Moreover, LLNA and GPMT can be carried out in a synergistic fashion to evaluate skin sensitization. More specifically, there is no need to carry out the GPMT or Buehler test for further validation once the test substance is defined as skin sensitization positive by LLNA. Nevertheless, a substance is subjected to further evaluation by GPMT or Buehler test in case it is qualified as skin sensitization negative by LLNA (OECD, 1992).

The LLNA and GPMT skin sensitization models cannot completely serve as a surrogate to predict skin sensitization potential in human since they can only accurately predict 70% of human tests in addition to the fact that LLNA and GMPT do not always reach the same agreement. It has been shown that the LLNA model is more effective in predicting human skin sensitization better than GPMT in case of discordance between LLNA and GMPT assays (Dean et al., 2001). As such, animal tests cannot completely replace their human counterparts because of their limitations. Surprisingly, it has been observed that one-third of strong sensitizers in the human test were predicted to be weak sensitizers by LLNA despite the fact that the LLNA test is commonly recognized as the gold standard for the human skin sensitization test (ICCVAM, 2011; Strickland et al., 2017; Roberts and Api, 2018), indicating the discrepancy in both assay systems. In addition, LLNA predictions can correlate with human tests well as long as those sensitizers lie within the applicability domain of the LLNA model. The inconsistent predictions, nevertheless, are not due to randomness. More specifically, the under-estimation of human skin sensitization potency by the LLNA model can be principally attributed to the fact that the test chemicals contain electrophilic aromatic Schiff bases or impurities (Roberts and Api, 2018). Conversely, the LLNA model is prone to over-estimating the potency when compared with the human test if the test chemicals under LLNA conditions will undergo autoxidation or have cutaneous pharmacological potentials other than skin sensitization (Roberts and Api, 2018). Furthermore, chemicals, namely pre- or pro-haptens, can be falsely predicted (Roberts and Api, 2018). Moreover, animal tests for skin sensitization that have been adopted for a long time still comprised some controversial issues concerning their effectiveness and ethical problems (Rollin, 2003). In 2017, Predictive Toxicology Roadmap was established by FDA. In this project, they evaluate new methods and technologies that can expand the predictive capabilities of toxicology and reduce the use of animal testing. With the same goals, the in vitro testing methods have been evaluated by a Consortium comprising the Institute for In vitro Sciences, Inc (IVS) the Consumer Healthcare Products Association (CHPA), and the PETA International Science Consortium (PETA-ISC) to substitute rabbit vaginal irritation (RVI) test (Costin et al., 2020). There is a growing tendency, nevertheless, to use non-animal tests as an alternative approach to assess skin sensitization (Doke and Dhawale, 2015).

Non-animal Assays

Animal testing approaches for cosmetic products have been banned due to animal rights and welfare by the 7th amendment to the EU Cosmetics Directive in Europe since 2013 (European Union, 2003; European Commission, 2013a). Notably, some non-animal testing methods have been developed...
to resolve this challenge and approved by the European Union Reference Laboratory for alternatives to animal testing (EUR-L ECVAM) (European Commission, 2013a; European Commission, 2013b). Nonetheless, non-animal test data may still have limitations in predicting skin sensitization indeed. For instance, those methods accepted by the Organization for Economic Co-operation and Development (OECD) only focus on one AOP key event or activation of some specific genes, such as depletion of cysteine and/or lysine-containing protein in Direct Peptide Reactivity Assay (DPRA) (OECD, 2015a), CD86 and CD54 overexpression in human cell line activation test (h-CLAT) (OECD, 2018c), induction of nuclear factor-erythroid-2 related factor 2 (Nrf2)-Kelch-like ECH-associated protein 1 (Keap1)-antioxidant/ electrophile response element (ARE) pathway in KeratinoSens™ (OECD, 2015b), CD86 overexpression for U-SENS™ test (OECD, 2016b), and the expressions of anti-inflammatory, and cell migration genes in SENS-IS test (Cottrez et al., 2016).

DPRA is a non-animal model focused on the hapten and protein interaction due to skin exposure to chemical substances. This method was first proposed by Gerberick and has been further accepted by OECD since 2015 (Gerberick et al., 2004; Gerberick et al., 2007b; OECD, 2015a), in which the synthesized peptides such as cysteine (Ac-RFAACAA-COOH) and lysine (Ac-RFAAKKAA-COOH) (R: Arginine, F: Phenylalanine, A: Alanine, C: Cysteine K: Lysine) are incubated with test substances, followed by measuring the absorption peaks at 220 nm to determine the concentration of cysteine and lysine after the reaction. Generally, skin sensitization can be divided into 4 classes, namely negative, positive with low, moderate, or high reactivity level (OECD, 2015a). Nevertheless, DPRA can be limited by solubility and complex mixture (Gerberick, 2016). In addition, the accuracy of measurement results would be hampered by the fact that chemicals could be co-eluted with the peptide (Natsch et al., 2007; Natsch and Gfeller, 2008). Until now, the DPRA has been improved by another version such as amino acid derivative reactivity assay (ADRA) to prevent the co-elution of test chemicals and nucleophilic agents (Fujita et al., 2019; Waniibuchi et al., 2019; Imamura et al., 2021). This method has been accepted by OECD (OECD, 2020). Another modified DPRA called kinetic direct peptide reactivity assay (kDPRA), in which several concentrations of tested compounds are incubated with synthetic peptide for different incubation times, has been accepted by OECD (OECD, 2020). The matrix of depletion values and incubation times and concentrations are constructed to measure the rate constant (log k\text{max}). The test compounds are further classified into GHS classification scales by their log k\text{max} values. (Natsch et al., 2020). The reproducibility between intra- and inter-laboratories for this method achieved 96 and 88%, respectively (Wareing et al., 2020).

The KeratinoSens™ method takes a different approach by focusing on a second AOP key event (Table 1), namely the inflammatory responses and gene expression associated with specific cell signaling pathways such as ARE-dependent pathways. Keap1 binds to the transcription factor Nrf2 in the un-induced state that helps ubiquitin bind to Nrf2 by Cul2-mediated ubiquitinylation, which, in turn, can degrade Nrf2 into the proteasome (de Freitas Silva et al., 2018). However, Keap1 cannot bind to Nrf2 protein once the covalent bond is formed between Keap1 and small molecules such as sensitizers, leading to the accumulation of Nrf2 protein in the nucleus. The released Nrf2 protein binds to ARE sequence in the promoter regions of detoxification, antioxidant, and anti-inflammatory genes, triggering the expression of target genes (de Freitas Silva et al., 2018). The mechanism of Nrf2-ARE pathway activation was illustrated in Figure 2 of de Fritas Silva et al. (de Freitas Silva et al., 2018). According to the human keratinocytes HaCaT cells are stably transfected with the selected plasmid, which contains the ARE sequence, SV40 promoter, and luciferase gene (luc2) in the KeratinoSens™ test (Emter et al., 2010; Steinberg, 2013). The test chemicals are designated as sensitizers in the KeratinoSens™ test provided that they can produce the induction of luciferase activity above 1.5 folds with respect to the negative control or non-sensitizers otherwise (OECD, 2015b). The LuSens assay is another method accepted by OECD and is developed based on the same concept as KeratinoSens™ (Ramirez et al., 2014; OECD, 2018b).

The activation of DCs is the AOP key event investigated in the h-CLAT method (Table 1). In the induction phase of skin sensitization, the co-expression of CD86 and CD54 on the Langerhans cells is used as the indicator of the antigen-presenting process (Nuriya et al., 1996; Reiser and Schneebaeger, 1996; Tuschl et al., 2000). Thus, the expressions of CD86 and CD54 on THP-1 cells, which are a human monocytic leukemia cell line, are measured in the event when THP-1 cells are exposed to sensitizers in the h-CLAT method (Sakaguchi et al., 2006; Sakaguchi et al., 2009). The upregulation of these markers indicates the occurrence of DCs activation and the skin sensitization activity caused by the test chemical. Of note, chemicals are further classified as sensitizers or non-sensitizers in this test (OECD, 2018c). In addition, the U-SENS™ test method for skin sensitization testing is based on the expression of CD86 cell surface marker on the U937 cells, which are a human histiocytic lymphoma cell line (Piroird et al., 2015). Brieﬂy, a compound is considered to be a sensitizer when CD86 expression in the U937 cell line is 1.5 fold higher than the untreated control and non-sensitizer otherwise. It is of interest to note that this method has been submitted to OECD and the drafted proposal has been publicized on the OCED website (OECD, 2016b; OECD, 2018c). This method recently has been used to evaluate the role of nanomaterials in skin sensitization (Bezerra et al., 2021).

SENS-IS is another non-animal method to measure the skin sensitivity of chemicals using the commercially reconstituted human skin (EpiSkin) (Netzlaff et al., 2005; Cottrez et al., 2015), in which the expression levels of Redox and SENS-IS genes are measured. The former includes 17 genes that contained an ARE in their promoter (Cottrez et al., 2016), which are related to the target genes modulated by the Nrf2-Keap1-ARE signaling pathway, whereas the latter includes 21 genes, which are linked to the activities of DCs and associated with inflammation, danger signals, and cell migration. Those genes measured in the SENS-IS group can be triggered by sensitizers but not under the control of the Nrf2-Keap1-ARE pathway (Cottrez et al., 2015; Cottrez et al., 2016). There are four chemical concentrations, namely 50, 10, 1,
and 0.1%, applied onto the artificial skin, followed by collecting the mRNA from the EpiSkin cells and analyzing the gene expression using reverse transcriptase polymerase chain reaction (RT-PCR). When the number of expressed genes is more than 7 and less than 20 in both groups, the test chemicals are defined as sensitizers and subsequently categorized as weak, moderate, strong, or extreme sensitivity depending on the chemical concentrations that, in fact, is similar to the classification system adopted by LLNA. The test chemical concentration will be lowered when there are 20 genes expressed that are termed overexpression. Moreover, a chemical is considered as negative in case of failures in all tested concentrations (Cottrez et al., 2016). Until now, this method is still under the validation process, which has been announced at https://tsar.jrc.ec.europa.eu/.

Various research groups have published their assay data using those above-mentioned methods and the results are summarized in Table 5, which provides affluence of data source for building in silico models. Until now, there are still many researchers endeavoring to improve the accuracy of in vitro assay for assessing the skin sensitization potential such as finding new biomarkers for predicting skin sensitization (Hirotta and Moro, 2006) or developing a novel assay like Genomic Allergen Rapid Detection (GARD™) to define the skin sensitization activity by only one assay (Roberts, 2018). GARD™ depends on the changes of the gene expression when myeloid cells are exposed to the chemicals (Johansson et al., 2013; Johansson et al., 2019; Johansson et al., 2011; Roberts, 2018; Stevenson et al., 2019). This method was validated by numerous laboratories with an inter-laboratory reproducibility of 92.0% in 2019 (Johansson et al., 2019), and has been used in the peer-review process for EURL ECVAM validation, which has been announced at https://tsar.jrc.ec.europa.eu/. Furthermore, the conformal prediction has been implemented into GARD™ protocol with an accuracy of 88%. In 2021, Masinja et al., nevertheless, used GARD™ to predict the skin sensitization potential of agrochemical active ingredients in total 7/12 GARD™ results concurred with mammalian data, suggesting that GARD™ still needs to be improved to validate the skin sensitization potential of not only cosmetics, but also the other active ingredients (Masinja et al., 2021). Another way is to modify or to improve the current non-animal methods to further increase the accuracy. For instance, the spectro-DPRA method using 5, 5-dithiobis-2-nitrobenzoic acid, or fluorescamine™ as the detection reagent was designed to investigate the unreacted peptide in 2014. It was demonstrated that the accuracy of this method could increase up to 91.5 and 94.9% when compared with LLNA and human data, respectively (Cho et al., 2014; Cho et al., 2019).

Most of the non-animal tests such as U-SENS™, h-CLAT, and KeratinoSens™ are qualitative per se, in which compounds are divided into skin sensitization positive and negative, viz. a binary classification fashion (OECD, 2015b; 2016b; 2018c), whereas DPRA and SENS-IS are basically quantitative, in which the levels of skin sensitization potential are determined (OECD, 2015a). Additionally, the non-animal tests such as DPRA, h-CLAT, and KeratinoSens™ are routinely used as the preliminary screening by Europe, whereas others such as U-SENS™ and SENS-IS can be implemented to further characterize the nature of skin sensitization (OECD, 2016b, 2019). The non-animal models for skin sensitization have been adopted for a long time and the first non-animal DPRA model has been accepted by OECD since 2015. However, not all chemicals such as insoluble chemicals, pro-haptens, and chemicals co-eluting with the model peptide can be assessed by DPRA that may severely limit their applications. These chemicals, nevertheless, can be evaluated by in silico models in the preliminary phase (Urbisch et al., 2016). As such, in silico models are expected to be a useful method for predicting skin sensitization of novel chemicals in this aspect.

### IN SILICO MODELS

#### Data Source

Europe has banned animal tests to verify the safety of cosmetic products such as toxicity in repeated dose systems, skin sensitization, reproductive toxicity, carcinogenicity, and toxicokinetics since 2013 (European Commission, 2013a). Alternatively, various non-animal tests, namely DPRA, KeratinoSens™, and h-CLAT, have been derived and accepted by OCED (vide supra). In addition, various skin sensitization data have been published according to the collection of animal and non-animal data, as well as chemical structure information are listed in Table 5. Some online skin sensitization data sources, which have collected the data and structural alerts, can be used to build predictive models and are listed in Table 6. Of various skin sensitization databases, SkinSensDB, which has collected the animal and non-animal tests and contains 710 unique...
chemicals with 2,078, 467, 1,323, and 1,060 assay values for peptide reactivity (DPRA), keratinocyte activation (KeratinoSen™), dendritic cell activation (h-CLAT), and T-cell activation (LLNA-EC3), respectively (Tung et al., 2019; Wang et al., 2017), is freely accessible. The non-confidential substance data, which have been submitted to European chemicals agency (ECHA) under the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) regulation, are publicly available and free of charge. This dataset contains many types of assays and study categories. Another available source is Cosmetic ingredient database (CosIng). This European Commission database includes the information about cosmetic substances and ingredients (Table 6). Another source is Chemical Evaluation and Risk Estimation System (CERES), which is developed by FDA’s Center for Food Safety and Nutrition. This database contains toxicity data including the skin sensitization hazard and potency (Ghosh et al., 2020). Various predictive models and packages to predict skin sensitization have been published (Wilm et al., 2018). Some models provide structure alerts based on the analysis of chemical characteristics that are responsible for skin sensitization (Sushko et al., 2011). ToxAlerts was established in 2012 serving as a valuable data source for model development to predict the chemical toxicity. Initially, 600 structural alerts for carcinogenicity, mutagenicity, skin sensitization, acute aquatic toxicity, and potential idiosyncratic drug toxicity were issued (Sushko et al., 2012), and the number has increased to more than 3,000 structural alerts to date. The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is a permanent committee of the National Toxicology Program (NTP) Division, which is responsible for evaluating the toxic potential, developing and validating the toxicology methods, collecting the data to strengthen the scientific base for risk assessment. ICCVAM has established a database for skin sensitization with the collection of 1,060 chemicals for the LLNA test and 208 chemicals for the GPMT and Buehler tests (ICCVAM, 2011). Vitic is a commercial toxicity database and information management system developed by Lhasa, consisting of more than 38,000 skin sensitization data for more than 10,000 structures. eChemPortal, which has been developed by OECD, is a free public source and provides the chemical characteristics of physical-chemical properties, environmental fate and behavior, ecotoxicity, and toxicity. The chemical information can be searched using chemical names and numbers or GHS classifications.

### Commercial Package

Numerous commercially available packages and/or models have been published and are listed in Table 7. Computer automated structure evaluation (case) Ultra program is a commercial package to issue the structure alerts, in which, principally, molecular structures are divided into various subunits and those ones responsible for specific activities are identified and termed biophore (Klopman, 1992). Moreover, different subunits can give rise to different biophore activities. A subunit is termed synergistic or biophobic if it could increase or decrease the activity, respectively (Graham et al., 1996; Chakravarti et al., 2012). As such, this package can predict the effective level of skin sensitization for a given compound and has been validated by predicting various adverse effects of drugs, namely carcinogenicity, hepatotoxicity, cardiotoxicity, renal toxicity, and reproductive toxicity. Saiakhov et al. have carried out a pilot study using case Ultra to analyze other adverse effects, including skin sensitization (Saiakhov et al., 2013).

A non-sensitizer might be converted into a sensitizer through a biodegradation metabolism pathway (Jaworska et al., 2002). CATABOL (http://oasis-lmc.org/products/models/environmental-fate-and-ecotoxicity/catabol-301c.aspx) is an online package that can simulate the metabolic pathways of chemicals by predicting the abiotic molecular transformation and enzyme-mediated reactions such as reductive, hydrolytic, oxidative, redox, conjugative reactions, reactions with skin protein, as well as predicting the chemical transformation through spontaneous reactions, enzyme-catalyzed metabolism reactions, and reactions with protein nucleophiles (Jaworska et al., 2002). The tissue metabolism simulator (TIMES) model based on the prediction from CATABOL consists of simulators: 1) using the microbial metabolism simulator to generate the metabolic maps from the training samples; 2) evaluation of skin sensitization potential in light of the metabolic maps and structural alerts (Dimitrov et al., 2005). The TIMES model for skin sensitization (TIMES-SS) package is commercially available and the information about skin metabolism associated with skin

#### Online skin sensitization databases

| database          | Website                                                                 |
|-------------------|-------------------------------------------------------------------------|
| SkinSensDB        | http://cw tung.kmu.edu.tw/skinsensdb                                     |
| ToxAlert          | http://to chem.eu/alerts                                                 |
| National Toxicology Program | https://ntp.niehs.nih.gov/whatwestudy/niceatm/test-method-evaluations/immunotoxicity/llna/index.html |
| vitic             | https://www.hasalmitted.org/products/vitic.htm                           |
| eChemPortal       | https://www.chemportal.org/echemportal/index.action                     |
| REACH             | https://uclcl6.echa.europa.eu/reach-study-results                        |
| CosIng            | https://ec.europa.eu/growth/sectors/cosmetics/cosing_en                 |
| CERES             | https://www.accessdata.fda.gov/scripts/fdatrack/view/track_project.cfm?program%3f&%id%3f%2FCFSANOFAS-Chemical-Evaluation-and-Risk-Estimation-System |

---

**Table 6**: Online skin sensitization databases.
sensitization is available online (http://oasis-lmc.org/products/models/metabolism-simulators/skin-metabolism.aspx). The training samples were excerpted from LLNA, GMPT, and human datasets (Dimitrov et al., 2005; Mekenyan et al., 2012). Ivanova et al. expanded the development of the kinetic component into the TIMES-SS model in 2020. In this model, they tried to implement the kinetic of biotic transformations to predict the skin sensitization potential. The initial predictions were consistent with the experimental data for those tested compounds (Ivanova et al., 2020).

The skin sensitization activity of a chemical will also depend on the transformation ability from prohapten into hapten in that the sensitizers themselves are not electrophilic per se. Nevertheless, they can undergo enzymatic or oxidative processes to become electrophilic that, in turn, can facilitate the interaction with skin protein, producing antigens consequently (Aptula et al., 2005). Unlike the other packages, Computer Aided Discovery and Redesign-Skin Sensitization (CADRE-SS) is focused on such biological transformation and is comprised of three modules to analyze the reaction in each step: I) skin permeability; II) haptenation and hapten-activation mechanisms, and III) conjugation with protein. The interaction potential between chemicals and skin protein is analyzed by module II using the Smiles Arbitrary Target Specification (SMARTS) pattern structure, and compounds are subjected to further analysis by module III once the chemicals are identified as potential hapten. The key event in this process is the adduct formation between the chemical and the Keap1 protein, which contains highly reactive cysteine and lysine amino acids (Kostal and Voutchkova-Kostal, 2016).

SMARTS patterns have been mined by ToxTree (Enoch et al., 2008) to identify the potential of skin sensitization. A series of SMARTS patterns based on the previously identified mechanisms of action have been identified, namely aromatic nucleophilic substitution (SNAr), Schiff base formation (SB), Michael-type addition (MA), aliphatic nucleophilic substitution (SN2), and acylation (Ac) (Aptula et al., 2005), in which the covalent bond can be formed between skin protein and sensitizer (Enoch et al., 2008; Enoch et al., 2011). Totally, 104 structural alerts were issued in 2011 (Enoch et al., 2011) and the most updated version is commercially available at https://www.daylight.com/products/toolkit.html through Daylight Toolkit.

Deductive estimation of risk from existing knowledge (DEREK) is an expert knowledge system-based commercial predictive package, in which the structure alerts are proposed to predict the binding potential between electrophilic chemicals and skin protein. Initially, only 40 structure-activity rules for skin sensitization were issued in 1994 (Barratt et al., 1994), and that number increased up to 70 in 2006 (Langton et al., 2006) from the GPMT input data. The modified version of Derek Nexus was released in 2017 using the LLNA EC3 value from over 650 compounds in the Lhasa EC3 dataset (https://www.lhasalimited.org/products/skin-sensitization-assessment-using-derek-nexus.htm) instead of GPMT, which was used in the previous versions. This version features the qualitative prediction for mammalian skin sensitization and the quantitative EC3 prediction for skin sensitizers (Canipa et al., 2017). The skin sensitization structure alerts in Derek Nexus increased from 73 to 90 between 2014 and 2018 and the performance was validated against a dataset over 2,500 chemicals with LLNA and/or GPMT data (Chilton et al., 2018).

### Models Based on Animal Tests

Computer Assisted Evaluation of Industrial Chemical Substances According to Regulations (CAESAR) was developed according to the QSAR validation principles issued by OECD. This model was

### TABLE 7 | The commercial package and models based on animal, non-animal tests, and mixed test type.

| Commercial package | Model |
|--------------------|-------|
| Klopman (1992) | Case |
| Jaworska et al. (2002) | CATABOL (http://oasis-lmc.org/products/models/environmental-fate-and-ecotoxicity/catabol-301c.aspx) |
| Dimitrov et al. (2005) | TIMES-SS (http://oasis-lmc.org/products/models/metabolism-simulators/skin-metabolism.aspx) |
| Kostal and Voutchkova-Kostal (2016) | CADRE-SS |
| Enoch et al. (2008) | ToxTree |
| Barratt et al. (1994), Langton et al. (2006) | Derek Nexus |
| Wilm et al. (2018) | CAESAR (http://www.caesar-project.eu/index.php?page=result&section=endpoints&ne=2) |
| Benfenati et al. (2013) | VEGA (https://www.vegahub.eu/) |
| Toropova and Toropov (2017) | CORAL (http://www.insilico.eu/coral) |
| Otsubo et al. (2017) | binary classifier based on KeratinoSens™ and h-CLAT |
| Asturio et al. (2016) | qualitative skin sensitization predictive model using DT by combining DPRA, KeratinoSens™, and h-CLAT |
| Roberts and Patlewicz (2018) | Build the model based on DT by combining DPRA, h-CLAT |
| Tung et al. (2018); Tung et al. (2019) | SkinSensPred (https://cwtung.kmu.edu.tw/skinsensdb/predict) |
| Borba et al. (2020a), Braga et al. (2017) | Pred-skin (http://predskin.labmol.com.br/) |
| Ohtake et al. (2017) | model based on in silico Derek Nexus, in chemico DPRA, and in vitro h-CLAT |
| Zang et al. (2017) | DPRA, h-CLAT, KeratinoSens™ results and six physicochemical properties of compounds |
| Wilm et al. (2020) | Skin Doctor CP (https://nerdd.zbh.uni-hamburg.de/skinDoctorII/) |
| Borba et al. (2020a) | https://stoptox.mml.unc.edu/ |
built by the EU (Cassano et al., 2010; Wilm et al., 2018) and is freely available (http://www.caesar-project.eu). CAESAR can be used to develop QSAR models for five endpoints, namely skin sensitization, carcinogenicity, mutagenicity, bioconcentration factor, and developmental toxicity. The CAESAR model for skin sensitization was derived based on 209 compounds excerpted from a previous study (Gerberick et al., 2005) to classify compounds into sensitizer or non-sensitizer (http://www.caesar-project.eu/index.php?page=result&section=endpoint&n=2).

Afterward, Virtual models for property Evaluation of chemicals within a Global Architecture (VEGA) derived from CAESAR model can predict skin sensitization based on the LLNA data. This binary classifier, which is freely accessible, can be downloaded at http://www.vega-qsar.eu (Benfenati et al., 2013) and the latest version can be found at https://www.vegahub.eu/. Fitzpatrick et al. compared the performance of VEGA, TIME-SS, and Derek Nexus in skin sensitization by applying 1,249 substances from the eChemportal skin sensitization dataset (http://www.chemportal.org/echemportal/index.action) and 515 substances from the Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEA™) LLNA database (https://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/immunotoxicity/index.html). The results showed that the accuracy of any expert models was about 65%, especially with the substances that were within the application domain of TIME-SS, the accuracy could be reached to 79 and 82% for both datasets (Fitzpatrick et al., 2018). This comparison, in fact, is consistent with the observation made by Teubner et al., in which it has been demonstrated that TIME-SS executed better than the others such as VEGA and DEREK (Teubner et al., 2013).

Another model developed by Istituto di Ricerche Farmacologiche Mario Negri (IRCCS) and the Joint Research Center (JRC) is available at https://www.vegahub.eu/. The model is built based on decision trees using 8 descriptors, which are listed at https://www.vegahub.eu/vegahub-dwn/qmr/qmrf/qmrfskin_JRC.pdf. The endpoint of this model is skin sensitization on mice (LLNA). When applied to external validation, this model can obtain the accuracy, specificity and sensitivity of 71, 82 and 65%, respectively. All 75 in silico models imbedded in VEGA have been implemented into the OECD QSAR Toolbox.

The Correlation and Logic (CORAL) package has been used as a tool for QSAR analyses (Toropov et al., 2013). Afterward, this software was used to develop a tool to predict the skin sensitization (Toropova and Toropov, 2017). Various QSAR models were built based on 204 compounds with the local lymph node assay results using the Monte Carlo technique. The hybrid descriptors calculated via the representation of the molecular structure by SMILES with molecular graph were used to generate the models. The model is available at http://www.insilico.eu/coral.

Very recently, the first ternary predictive model has been developed by Wilm et al. termed Skin Doctor CP (available at: https://nerdd.zbh.uni-hamburg.de/skinDoctorII/) based on the LLNA database. The most distinguishing characteristic of this model is that compounds are initially categorized into sensitizers or non-sensitizers by the first classifier and the predicted sensitizers are further grouped into weak to moderate sensitizers and strong to extreme sensitizers by a second classifier. The model showed the accuracies of 0.90 and 0.73 as well as the efficiencies of 0.42 and 0.90 at the significance levels of 0.10 and 0.30, respectively. However, this ternary classifier did not achieve good performance since the validity values were 0.70, 0.58, and 0.63 for non-sensitizers, weak to moderate sensitizers, and strong to extreme sensitizers, respectively, at the significance level of 0.30 (Wilm et al., 2020).

### Models Based on Non-animal Tests

Unlike the other predictive models, which rely on a single data type, some predictors make decisions based on multiple data types. For instance, Otsubo et al. have built a binary classifier based on KeratinoSens™ and h-CLAT, and chemicals are designated as skin sensitizers if they have positive results by either one of the assays and non-sensitizers otherwise. The predictions produced the sensitivity values of 93.4 and 94.4% as compared with the LLNA and human data, respectively (Otsubo et al., 2017). After collecting data from three non-animal assays, namely DPRA, KeratinoSens™, and h-CLAT, this study was further extended to build a majority voting system. Compounds were defined as sensitizers when at least two positive responses were obtained from those three assays. The accuracy obtained from this model was 90% when compared with human data, whereas that resulted from LLNA alone was merely about 80% as compared with the human data (Urbisch et al., 2015). According to this, multiple data type models can execute better than their single-data-type counterparts.

Asturiol et al. took a different approach to develop a qualitative skin sensitization predictive model using decision tree (DT) (Asturiol et al., 2016). The model was derived by combining 3 non-animal test data types, namely DPRA, KeratinoSens™, and h-CLAT. The accuracy of the model was defined by comparing with the LLNA classification (sensitizer/non-sensitizer). The model showed 93% accuracy, 98% sensitivity, and 85% specificity for 269 chemicals (Asturiol et al., 2016). A different approach was taken to build various DTs based on non-animal test results, in which compounds were classified as sensitizers when DPRA gave rise to positive results. Further evaluation by h-CLAT was carried out once compounds were considered as negative by DPRA. Compounds were classified as skin sensitizers if h-CLAT showed positive results, whereas compounds were labeled as non-sensitizers otherwise (Roberts and Patlewicz, 2018). Additionally, various models were developed according to binary combinations of those three non-animal tests, namely DPRA, h-CLAT, and KeratinoSens™. It was found that the combination of DPRA and h-CLAT performed best in distinguishing sensitizers from non-sensitizers. More importantly, it was observed that all of the models based on combinations of non-animal tests usually performed better than their counterparts based on a single test that, actually, is consistent with the previous observation (vide supra) (Urbisch et al., 2016; Otsubo et al., 2017), suggesting that predictive models based on the single non-animal test are not sufficient to comprehensively render the skin sensitization complicated.
process (Adler et al., 2011; Hartung et al., 2011; Wilm et al., 2018; Madden et al., 2020).

### Models Based on Mixed Test Types

Most of the published packages or models are binary classification systems, viz. sensitizer vs. non-sensitizer, based on one or more than one non-animal tests. Integrated approaches to testing and assessment (IATA) has taken a different approach by combining various animal tests, non-animal tests, and in silico models to predict the latency of skin sensitization (OECD, 2016a). IATA includes the models, which are flexible and non-normalized judgment based, e.g. grouping and read-across or more structured, rule based approaches such as Integrated Testing Strategy (ITS) (OECD, 2016a). ITS can combine DPRA, KeratinoSens™, and h-CLAT (Jaworska et al., 2015; Urbisch et al., 2015), DPRA, SENS-IS and/or h-CLAT (Clouet et al., 2017), two of 3 non-animal tests, namely DPRA, KeratinoSens™, and h-CLAT, to generate the predictive model (Otsubo et al., 2017) based on in chemico, in vitro, and in silico data (Jaworska et al., 2011). It has been found that the models based on this strategy showed better performance. For instance, this approach was implemented to develop various models based on the combinations of 2 or 3 non-animal datasets, namely DPRA, KeratinoSens™, and h-CLAT and the built model with the selection of 3 non-animal datasets displayed the highest sensitivity and yet the lowest specificity as compared with its counterparts with the combination of only 2 of 3 non-animal datasets (Otsubo et al., 2017). In 2017, Douglas Connect Integrated Testing Strategy (DC ITS) SkinSens was launched to access the integrated testing strategy developed by Jaworska (Jaworska et al., 2013). The latest updated version for DC ITS SkinSens is SaferSkin™ (https://saferworldbydesign.com/saferskin/).

SkinSensPred, which is a skin sensitization predictive function, was developed in 2019 based on SkinSensDB, is freely accessible at https://cw.tung.kmu.edu.tw/skinsensdb/predict (Tung et al., 2018; Tung et al., 2019). This multitask learning model is based on three AOP key events and human skin sensitization test using protein binding (DPRA), keratinocyte activation, dendritic cell activation to binarily classify results in the human test. This model can analyze the application domain (AD) and structure alerts (SA) to predict the human sensitization potential of a chemical. When applied to novel chemicals within the defined AD, this model could reach an accuracy of 84.3% (Tung et al., 2019). In addition, a majority voting model (2 out of 3) (Urbisch et al., 2015) and a DT model (Roberts and Patlewicz, 2018) can be implemented as the read-across predictive methods.

In 2018, Del Bufalo et al. developed an alternative integrated testing for skin sensitization using the combination of 3 in vitro methods (DPRA, Keratinosens™, U-SENS™), two in silico tools (TIMES-SS, TOXTREE) and physicochemical parameters (volatility, pH). These data were run in 5 different classification models (Boosting, Naive Bayes, support vector machine (SVM), Sparse Partial Least Squares Discriminant Analysis, and Expert Scoring). The validation results were used in the stacking meta-model to evaluate the skin sensitization potential (Del Bufalo et al., 2018). The predictions achieved the accuracies of 93 and 91% for the training set and test set, respectively when compared with LLNA hazard data. A larger data set was used to validate this model in 2020 (Toureix et al., 2020). The Defined Approach (version 5) was used (Toureix et al., 2020) to evaluate the skin sensitization potential of 219 compounds.

Pred-skin, which is accessible at http://predskin.labmol.com.br/, is a consensus Naïve Bayes model that employs multiple QSAR models based on various human, LLNA, and non-animal data to predict skin sensitization. This model exhibited good performance in predicting human skin sensitization with sensitivity (94%) and specificity (84%). When applied to 11 new potential sensitizers, which were not included in the dataset, Pred-skin exerts an efficient approach to identify nine sensitizers (Borba et al., 2020a; Braga et al., 2017).

Zang et al. have published an in silico model, which was derived by combining the non-animal data, namely DPRA, h-CLAT, and KeratinoSens™, and six physicochemical properties, namely octanol/water partition coefficient, water solubility, vapor pressure, melting point, boiling point, and molecular weight, as the descriptors to predict LLNA and human outcomes (Zang et al., 2017). Compounds were classified into sensitizers or nonsensitizers in this model, and the sensitizers were further divided into 1A (strong) or 1B (weak) sensitizer subcategories based on the GHS. The model achieved the accuracy of 88% for the prediction of LLNA outcomes and 81% prediction for human test outcomes (Zang et al., 2017).

Ohtake et al. have published a predictive model based on highly heterogeneous data, namely in silico Derek Nexus, in chemico DPRA, and in vitro h-CLAT, in which the results of DPRA and h-CLAT were scaled between 0 and 3, and the outcomes from Derek Nexus were reduced between 0 and 1, and the final total score was generated by summing those scores. A compound is defined as a strong sensitizer when its total score is larger than 7, and a weak sensitizer when its total score is between 2 and 6 (Ohtake et al., 2018). The unique characteristic of this model is the fact that the skin sensitizers are further divided into the strong and weak ones in this model despite the fact that this multiple classification system is not the same as the animal or human test classification. However, only nine isocyanates were included and the prediction results indicated that this model underestimated the skin sensitization potential when compared with LLNA data.

In 2020, Silva et al. used the different combinations of in vitro (human information), in chemico (DPRA), and/or in silico (the formation descriptor calculated by the TIMES-SS) data to build the models, which can predict the skin sensitization potential. The results showed that the combination of in vitro, in chemico, and in silico achieved the best prediction results. Moreover, the models reached an accuracy of 100% in differentiating sensitizers from non-sensitizers. When the same model was, it exhibited the accuracies of 98.8 and 97.5% accuracy when applied to the compounds based on GHS classification (3-level scales) and human data (6-level scales), respectively (Silva et al., 2020).

### Machine Learning-Based Models

Currently, a variety of simulation approach DT, artificial neural network (ANN), support vector machine (SVM), AdaBoost, the iterative least squares linear discriminant (TILSQ), logistic
strategy was illustrated by Sato et al. in 2012. All 593 compounds through multiple steps in this study. Compounds in the gray compounds were allotted to negative, positive, and gray zones.

The SH test is designed to measure changes in cell surface thiols on hapten-treated cells and used to develop the first version of ANN-based iSENS to predict skin sensitization with the combination of h-CLAT data (Suzuki et al., 2009; Hirota et al., 2013). The second version was released afterward using the combination of the ARE assay data and n-octanol–water partition coefficient (log $p$) value (Natsch and Emter, 2008; Tsujita-Inoue et al., 2014). Further extended versions of the ANN model were based on various combinations of h-CLAT, DPRA, KeratinoSens™, and SH test (Hirota et al., 2015). It has been observed that the performance of an ANN model actually depended on the combination of data types. For instance, the ANN model based on the combination of h-CLAT and DPRA showed a better correlation with LLNA than other combinations such as DPRA and ARE assay or the SH test and ARE assay. The predictive models based on three descriptors such as the selection of h-CLAT, DPRA, and ARE assay or h-CLAT, SH test, and ARE assay produced higher correlation coefficients, viz. $r$ values, and smaller prediction errors than their two-data-type counterparts (Hirota et al., 2015).

More recently, Macmillan and Chilton have combined Derek Nexus and non-animal KeratinoSens™, h-CLAT, DPRA, and U-SENS™ tests to develop a DT model. The derived DT model showed great performance with 73 and 76% accuracy of LLNA and human data, respectively, depending on the GHS classification (Macmillan and Chilton, 2019). A variety of machine learning-based schemes, namely ANN, SVM, AdaBoost, and TILSQ, were employed to build skin sensitization predictive models based on linear and non-linear discriminant analyses of 291 samples. It was found that SVM and AdaBoost models based on 32 descriptors to encode the 2-D and 3-D structural characteristics showed the highest performance with 100% accuracy of negative and positive (Sato et al., 2009). This investigation was further extended by including more samples (593 compounds) and adopting a novel KY scheme (Sato et al., 2012). Unlike any binary classification models, compounds were allotted to negative, positive, and gray zones through multiple steps in this study. Compounds in the gray zone, which was a confusing area, were repeatedly deposited into the positive and negative zones until no compound was left, the strategy was illustrated by Sato et al. in 2012. All 593 compounds were classified impeccably in 3 steps (Sato et al., 2012).

Strickland et al. have adopted the LR and SVM schemes to develop predictive models based on non-animal tests, namely DPRA, h-CLAT, and KeratinoSens™ using 6 physicochemical properties, namely log $p$, water solubility, vapor pressure, melting point, boiling point, and molecular weight. It was found that log $p$ was the most pivotal factor in determining skin sensitization among various physicochemical properties that was further assured by Gleeson et al. (Gleeson and Gleeson, 2020). Of various combinations of non-animal test data, models that included the combination of DPRA and h-CLAT produced the highest accuracy (Strickland et al., 2017). Pre- and pro-hapten sensitizers, which need to go through chemical transformation through air exposure (Karlberg et al., 1992; Sköld et al., 2002) or metabolism pathway (Nilsson et al., 2005; van Eijl et al., 2012) prior to the sensitization process, hinder the accuracy of current in vitro assays. Accordingly, a novel tri-culture assay system, which includes MUTZ-3-derived Langerhans cells, HaCaT keratinocytes, and primary dermal fibroblasts, and then measures the secretion levels of cytokines after these cells are exposed to test compounds, viz. sensitizer or non-sensitizer, has been proposed. Numerous SVM models were developed based on the stimulation indices (SI) of 27 human cytokines, namely IL-1$\beta$, IL-1ra, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, IL-15, IL-17, eotaxin, basic FGF, G-CSF, GM-CSF, IFN-γ, IP-10, MCAF, MIP-1α, MIP-β, PDGF-BB, RANTES, TNF-α, and VEGF to identify the most significant cytokines associated with skin sensitization. It was observed that the SVM model based on the top three ranking biomarkers, namely IL-8, MIP-1β, and GM-CSF, in tri-culture assay showed the highest performance with the prediction accuracy of 91%, and the detection of pre- and pro-hapten was improved accordingly (Lee et al., 2018).

Matsumura et al. performed a study using QSAR-like deep neural network (DNN) and light gradient boosting machine (LightGBM) to evaluate the potential of skin sensitization. Physical and structural properties of chemicals and the skin sensitizer/non-sensitizers based on the classification of GHS were used as input variables. The results showed that the dual-input LightGBM model (74%) and dual-DNN model (72%) were moderately accurate when compared with the traditional approaches (Matsumura, 2020).

In addition, other algorithms and methods have been adopted to improve the classification performance. For instance, Abdallah et al. used binary crow search algorithm (BCSA) that was initially proposed by Askarzadeh in 2016 (Askarzadeh, 2016) to select the most relevant descriptors in the model development. The results gained the classification accuracy for the compound into sensitizer/non-sensitizer.

FUTURE PERSPECTIVES

The dermatology researched has shifted into a new paradigm after the introduction of artificial intelligence. Most of the applications involve in image analysis, but also include the analyses of the physiochemical properties of substances (Gomolin et al., 2020). Various applications in toxicity and environmental hazard endpoints, for instance, indicate that the great diversity of QSAR models (Chinen and Malloy, 2020). Data quality plays a critical role in model development and it is almost impossible to build a sound in silico model based on contaminated or impure data, especially for the quantitative predictive models (Cherkasov et al., 2014). The accuracy of a virtual model depends on the data quality, the lack of instruction of how to report and publish the toxicogenomic studies can
hinder the usage of data in building the computational models (FitzGerald, 2020). There are some strategies to build an in silico model. The first strategy is to collect a large number of experimental data, extended the data coverage, and the use of big data approach. Another more effective way to build a model relied on deeply understanding the biological mechanism to predict the biochemical processes and the bioactivity of the novel compounds. The model based on mechanism does not need a large amount of training set, but it needs highly accurate experimental data (Kostal and Voutchkova-Kostal, 2020). Accordingly, it is of necessity to implement data curation prior to model development by removing those assay data obtained from impurity or mixture to maintain data integrity.

In 2020, Golden et al. carried out an investigation, which compared the accuracy of eight in silico models (PredSkin, Toxtree, QSAR Toolbox, Danish QSAR database, CAESAR, REACHAcross™, TIMES-S and Derek Nexus) against human data sets. Most of the models showed the accuracies of 70–80% on human data sets, suggesting that in silico models can be a convenient and inexpensive tool to define the skin sensitization in human (Golden et al., 2021). There is no doubt that an in silico model to predict skin sensitization based on human data will be more realistic and much needed. However, the scarcity in consistent human data in the public domain has created an unsurmountable hurdle for creating a sound predictive model due to their small amount of available data and limited structural diversity (Cherkasov et al., 2014). The relatively ample amount of LLNA data makes it a better alternative since the LLNA predictions can well correlate with the human tests in most of the cases and it has been commonly recognized as the gold standard for the human skin sensitization. However, the LLNA model is inevitably susceptible to some chemotypes, suggesting that it is of necessity to develop different predictive models for different chemotypes to accommodate the variations in skin sensitization mechanism. Some problems remain unresolved when the test compounds have consisted of more than one chemotype or when the test compounds lie outside of the applicability domain of the derived model. With the effort to improve the accuracy of skin sensitization tests, Leontaridou et al. identified the borderline of DPRA, LuSens, h-CLAT and LLNA. The substances with the test results fell into the BR and another available test method was required to depict the positive/negative outcome (Leontaridou et al., 2017).

The applications of animal tests on cosmetics products have been prohibited in Europe since 2013 (European Commission, 2013a) and lately, some countries also have accepted the OECD non-animal method to test the skin sensitization (Strickland et al., 2019). Nevertheless, animal tests, especially GMPT and LLNA, are still available and required by numerous countries such as Canada, China, Brazil, Japan, and the United States (Daniel et al., 2018). Additionally, the applications of animal tests for pesticides, plant protection products, pharmaceuticals, household products, art materials, industrial chemicals, medical devices, and workplace chemicals are needed and still acceptable in many industries, even in Europe (Daniel et al., 2018). These data provide valuable resources for building some in silico models to assess the covert of skin sensitization. In addition, the skin sensitization QSAR models can be applied to not only cosmetic ingredients but also the compound, which can impact the ecosystems such as dye pollution, the effect of personal care products on aquatic species, or plasticizers (Arunalanandam et al., 2021; Funar-Timofei and Ilia, 2020; Khan et al., 2020) as well as the pharmacokinetics profiles of low molecular weight oligohydroxyalkanoates (Roman et al., 2020), suggesting that the skin sensitization models has a wide range of applications. To date, most of the published skin sensitization models are qualitative predictions, viz. binary classification of sensitizers or non-sensitizers. Nevertheless, it has been observed that a quaternary predictive model would execute better than its ternary counterpart, which, in turn, performed better than a binary one in the case of drug-induced liver injury (DILI) prediction (Weng and Leong, 2020). Accordingly, it is plausible to expect a multiple-class qualitative model to predict skin sensitization can function better than a two-class one. Predictive models based on a single type of assay data can only take into account one single pathway, suggesting that no single non-animal test can comprehensively render the whole complex skin sensitization process (Adler et al., 2011; Hartung et al., 2011). Additionally, some cosmetic or commercial products have currently contained sensitizers (Robinson et al., 2000) despite the fact that they do not trigger adverse reactions in the induction and elicitation phases when the applied doses are low. In addition, the application volume of those cosmetic products that are in direct and persistent contact with skin, e.g. cream or foundation, are different from those that can be washed or rinsed off, e.g. shampoo or body lotion. The differences in exposure dose between these groups can be up to 30 folds (Fewings and Menné, 1999; Frosch et al., 1995; Robinson et al., 2000). Therefore, qualitative in silico or non-animal models have hindered applications for those weak or moderate sensitizers in pharmaceuticals or cosmeceuticals markets, more importantly, a quantitative prediction model can be truly useful. To solve this problem, the package SpheraCosmolife, which is implemented in VEGAHub, can process various ingredients in a product, has been derived recently. It can predict the mutagenicity, genotoxicity, and skin sensitization based on the concentration and the product type, namely lotion, shampoo, shower gel, etc, recorded in the internal database (Regulation (EC) No. 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products). This software is available at https://www.vegahub.eu/download/sphera-cosmolife-download/ and is implemented in VEGA. However, some challenges still remain since they cannot predict the skin sensitization caused by metals (Biswas et al., 2020).

Based on the principles published by the International Cooperation on Cosmetic Regulation (ICCR), Gilmour et al. displayed next generation risk assessment (NGRA) framework for skin sensitizers in 2020, which can be illustrated in Figure 1 of their publication (Gilmour et al., 2020). According to four elements of risk assessment, which included consumer exposure, hazard identification, hazard characterization and establishment of a dose response, that they presented a workflow assembled three tires and integrated all relevant
information using a weight of evidence approach to predict a chemical to be a skin sensitizer or non-sensitizer (Gilmour et al., 2020).

The number of in silico models to predict skin sensitization has increased for the past few years. Those models have been built based on in vivo, in vitro, in chemico, and/or in silico data. Johnson et al. defined the rules and principles to develop the effective in silico skin sensitization models to facilitate the implementation and acceptance of in silico approaches based on the skin sensitization mechanism and the strengths/limitations of each experimental methods. The standardization of this hazard assessment framework has further strengthened the use and application of in silico tools in agencies and industries (Johnson et al., 2020).

With the effort to collect the available web portal to predict 6 of acute toxicity tests, namely acute oral toxicity, acute dermal toxicity, acute inhalation toxicity, skin irritation and corrosion, eye irritation and corrosion, and skin sensitization), Borba et al. developed a package called Systemic and Topical chemical Toxicity (STopTox), which is available at https://stoptox.mml.unc.edu/ (Borba et al., 2020b).

**CONCLUSION**

In vitro skin sensitization tests alone cannot replace human and animal tests because they only focus on one single pathway in AOP. In silico approach, conversely, has more advantages than in vitro tests since it can take into account more than one AOP key event by combining various in chemico, in vitro, and in vivo data simultaneously. To date, most of the published in silico models only classify chemicals into sensitizers or non-sensitizer. This binary classification system has severely limited the applications of weak or moderate sensitizers in commercial products. Multiple-class in silico models can be greatly useful in practical applications as exemplified by the DILI study (vide supra) and quantitative ones will be even better. With the effort to develop the quantitative models, some in silico models have been generated recently, and yet they inevitability suffer from major limitations, suggesting that these quantitative models still need to be improved to be more accurate and have a wider range of applications in evaluating the skin sensitization potential. The development of robust and accurate in silico models for skin sensitization prediction is still a long and winding path ahead of the molecular modeling community.

**AUTHOR CONTRIBUTIONS**

GT Investigation, Writing-Original Draft. C-FW: Conceptualization, Investigation, Writing-Reviewing and Editing. ML: Conceptualization, Funding acquisition, Investigation, Supervision, Writing-Reviewing and Editing.

**FUNDING**

This work was supported by the Ministry of Science and Technology, Taiwan.

**REFERENCES**

Adler, S., Baskett, D., Creton, S., Pelkonen, O., van Benthem, J., Zhang, V., et al. (2011). Alternative (Non-animal) Methods for Cosmetics Testing: Current Status and Future Prospects-2010. *Arch. Toxicol.* 85 (5), 367–485. doi:10.1007/s00204-011-0693-2

Ahlfors, S. R., Sterner, O., and Hansson, C. (2003). Reactivity of Contact Allergenic Haptens to Amino Acid Residues in a Model Carrier Peptide, and Characterization of Formed Peptide–Hapten Adducts. *Skin Pharmacol. Physiol.* 16 (1), 59–68. doi:10.1159/000068288

Ahkilani, A., McCrudden, M. T., and Donnelly, R. (2015). Transdermal Drug Delivery: Innovative Pharmaceutical Developments Based on Disruption of the Barrier Properties of the Stratum Corneum. *Pharmaceutics* 7 (4), 438–470. doi:10.3390/pharmaceutics7040438

Api, A. M., Baskett, D., and Lalko, J. (2015). Correlation between Experimental Human and Murine Skin Sensitization Induction Thresholds. *Catan. Ocud. Toxicol.* 34 (4), 298–302. doi:10.3109/15569527.2014.979425

Aptula, A. O., Patlewicz, G., and Roberts, D. W. (2005). Skin Sensitization: Reaction Mechanism Applicability Domains for Structure–Activity Relationships. *Chem. Res. Toxicol.* 18 (9), 1420–1426. doi:10.1021/tx050075m

Arulanandam, C. D., Prathiviraj, R., and Kaveriyappan, G. R. (2021). Raspberry Pi: Assessments of Emerging Organic Chemicals by the Predictive In Silico Methods. *biosRv* 2001 (2015), 426465. doi:10.1101/2021.01.15.426465

Askerzadeh, A. (2016). A Novel Metaheuristic Method for Solving Constrained Engineering Optimization Problems: Crow Search Algorithm. *Comput. Structures* 169, 1–12. doi:10.1016/j.compstruc.2016.03.001

Asturiol, D., Casati, S., and Worth, A. (2016). Consensus of Classification Trees for Skin Sensitisation Hazard Prediction. *Toxicol. Vitro* 36, 197–209. doi:10.1016/j.tiv.2016.07.014

Ball, N., Cagen, S., Carrillo, J.-C., Certa, H., Eigler, D., Emter, R., et al. (2011). Evaluating the Sensitization Potential of Surfactants: Integrating Data from the Local Lymph Node Assay, guinea Pig Maximization Test, and In Vitro Methods in a Weight-Of-Evidence Approach. *Regul. Toxicol. Pharmacol.* 60 (3), 389–400. doi:10.1016/j.yrtph.2011.05.007

Barratt, M. D., Baskett, D. A., Chamberlain, M., Payne, M. P., Admans, G. D., and Langowski, J. J. (1994). Development of an Expert System Rulebase for Identifying Contact Allergens. *Toxicol. Vitro* 8 (4), 837–839. doi:10.1016/0887-2333(94)90081-7

Baskett, D. A., Alépée, N., and Ashikaga, T., Barroso, J., Gilmour, N., Goebel, C., et al. (2014). Categorization of Chemicals According to Their Relative Human Skin Sensitizing Potency. *Dermatitis: Contact atopic*, *Occup. Drug* 25 (1), 11–21. doi:10.1097/der.0000000000000083

Bauch, C., Kolle, S. N., Fabian, E., Pachel, C., Ramirez, T., Wiench, B., et al. (2011). Intralaboratory Validation of Four In Vitro Assays for the Prediction of the Skin Sensitizing Potential of Chemicals. *Toxicol. Vitro* 25 (6), 1162–1168. doi:10.1016/j.tiv.2011.05.030

Bauch, C., Kolle, S. N., Ramirez, T., Eltze, T., Fabian, E., Mehling, A., et al. (2012). Putting the Parts Together: Combining In Vitro Methods to Test for Skin Sensitizing Potentials. *Regul. Toxicol. Pharmacol.* 63 (3), 489–504. doi:10.1016/j.yrtph.2012.05.013

Benfenati, E., Manganaro, A., and Gini, G. (2013). “VEGA-QSAR: AI Inside a Platform for Predictive Toxicology,” in Proceedings of the workshop “Popularize Artificial Intelligence 2013”. Turin, 1107.
Assessment of Defined Approaches. Crit. Rev. Toxicol. 48 (5), 359–374. doi:10.1080/10408444.2018.1429386
Kligman, A. M., and Epstein, W. (1975). Updating the Maximization Test for Identifying Contact Allergens. Contact Dermatitis 1 (4), 231–239. doi:10.1111/j.1600-0536.1975.tb03589.x
Kligman, A. M. (1966). The Identification of Contact Allergens by Human Assay. J. Invest. Dermatol. 47 (5), 393–409. doi:10.1038/jid.1966.160
Klimek-Szczykutowicz, M., Szopa, A., and Ekiert, H. (2020). Citrus Limon (Lemon) Phenomenon-A Review of the Chemistry, Pharmacological Properties, Applications in the Modern Pharmaceutical, Food, and Cosmetics Industries, and Biotechnological Studies. Plants 9, 119. doi:10.3390/plants9011019
Klopman, G. (1992). MULTICASE I. A Hierarchical Computer Automated Structure Evaluation Program. Quant. Struct. Act. Relat. 11, 176–184. doi:10.1021/qsar.19920110208
Kohtarou, Y. (2010). Generating two-class classification model for predicting chemical toxicity. U.S. Patent No. US 7725413 B2.
Kostal, J., and Voutchkova-Kostal, A. (2016). CADRE-SS, an In Silico Tool for Predicting Skin Sensitization Potential Based on Modeling of Molecular Interactions. Chem. Res. Toxicol. 29 (1), 58–64. doi:10.1021/acs.chemrestox.5b00392
Kostal, J., and Voutchkova-Kostal, A. (2020). Going All in: A Strategic Investment in In Silico Toxicology. Chem. Res. Toxicol. 33 (4), 880–888. doi:10.1021/acs.chemrestox.9b00497
Kraft, J. N., and Lynde, C. W. (2005). Moisturizers: what They Are and a Practical Approach to Product Selection. Skin Ther. Lett 10 (5), 1–8.
Kumar, S. (2005). Exploratory Analysis of Global Cosmetic Industry: Major Players, Technology and Market Trends. Technovation 25 (11), 1263–1272. doi:10.1016/j.technovation.2004.07.003
Ladzinski, B., Mistry, N., and Kundu, R. V. (2011). Widespread Use of Toxic Skin Lightening Compounds: Medical and Psychosocial Aspects. Dermatol. Clin. 29 (1), 111–123. doi:10.1016/j.det.2010.08.010
Langton, K., Patlewicz, G. Y., Long, A., Marchant, C. A., and Basketter, D. A. (2006). Structure/activity Relationships for Skin Sensitization: Recent Improvements to Derck for Windows. Contact Dermatitis 55 (6), 342–347. doi:10.1111/j.1600-0536.2006.00969.x
Lee, A., and Thomson, I. (1999). “Drug-induced Skin Reactions,” in Adverse Drug Reactions (London: Pharmaceutical Press), 262, 125–156.
Lee, S., Greenstein, T., Shi, L., Maguire, T., Schloss, R., and Yarmush, M. (2018). Improvements to Derek for Windows. Eur. J. Pharmaceutics Biopharmaceutics 60 (2), 167–178. doi:10.1016/j.ejpb.2005.03.004
Nilsson, A.-M., Bergström, M. A., Luthman, K., Nilsson, J. L. G., and Karlberg, A.-T. (2005). A Conjugated Diene Identified as a Prohapten: Contact Allergic Activity and Chemical Reactivity of Proposed Epoxyde Metabolites. Chem. Res. Toxicol. 18 (2), 308–316. doi:10.1021/tx049758c
Natsch, A., Yashikaga, T., Sakaguchi, H., Sono, S., Mugita, N., Hirota, M., et al. (2011). Predictive Performance for Human Skin Sensitizing Potential of the Human Cell Line Activation Test (H-CLAT). Contact Dermatitis 65 (3), 343–353. doi:10.1111/j.1600-0536.2011.01952.x
Natsch, A., Miyazawa, M., Kazutoshi, S., Sakaguchi, H., and Nishiyama, N. (2013). Data Integration of Non-animal Tests for the Development of a Test Battery to Predict the Skin Sensitizing Potential and Potency of Chemicals. Toxicol. Vitro 27 (2), 609–618. doi:10.1016/j.tiv.2012.11.006
Nuryia, S., Yagita, H., Okumura, K., and Azuma, M. (1996). The Differential Role of CD86 and CD80 Co-stimulatory Molecules in the Induction and the Effector Phases of Contact Hypersensitivity. Int. Immunol. 8 (6), 917–926. doi:10.1093/intimm/8.6.917
Ochoa, M. T., Loncaric, A., Krutzik, S. R., Becker, T. C., and Modlin, R. L. (2008). “Dermal Dendritic Cells” Comprise Two Distinct Populations: CD1+ Dendritic Cells and CD209+ Macrophages. J. Invest. Dermatol. 128 (9), 2225–2231. doi:10.1038/jid.2008.56
OECD (2016b). In vitro skin sensitization: U937 Cell Line Activation Test (U-SENS™). Paris, France: OECD.
OECD (2018a). In vitro skin sensitisation assays addressing the AOP Key Event on Keratinocyte Activation. Paris, France: OECD.
OECD (2017). Draft key event based test guidelines 442D: In Vitro Skin Sensitisation Assays Addressing the AOP Key Event on: Keratinocyte Activation. Paris, France: OECD.
OECD (2020). Draft OECD Guideline for the Testing of Chemicals: Key-Event-Based Test Guideline for In Chemico Skin Sensitization Assays Addressing the Adverse Outcome Pathway Key Event on Covalent Binding to Proteins. Paris, France: OECD.
OECD (2016a). 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. Paris, France: OECD.
OECD (2018b). Key Event Based Test Guideline 442Dvito Skin Sensitization Assays Addressing the AOP Key Event on Keratinocyte Activation. Paris, France: OECD.
OECD (1992). Test No. 406OECD guideline Test. Chemicals- Skin Sensitisation. Paris, France: OECD.

OECD (2010a). Test No. 429 Skin Sensitization: Local Lymph Node Assay. Paris, France: OECD.

OECD (2019). Test No. 439. In vitro Skin Irritation: Reconstructed Human Epidermis Test. Method. Paris, France: OECD.

OECD (2010b). Test No. 442A Skin Sensitization: Local Lymph Node Assay DA. Paris, France: OECD.

OECD (2015a). Test No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). Paris, France: OECD.

OECD (2015b). Test No. 442D: In Vitro Skin Sensitisation: ARE-Nrf2 luciferase test method. Paris, France: OECD.

OECD (2018c). Test No. 442E: In Vitro Skin Sensitization Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Paris, France: OECD.

OECD (2012). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, Paris, France: OECD.

Ohtake, T., Maeda, Y., Hayashi, T., Yamana, H., Nakai, M., and Takeyoshi, M. (2018). Complications of Chronic Use of Skin Lightening Cosmetics. StatPearls Publishing Copyright © 2021, StatPearls Publishing LLC.

Panico, A., Serio, F., Bagordo, F., Grassi, T., Idolo, A., De Giorgi, M., et al. (2019). Skin Safety and Health Prevention: an Overview of Chemicals in Cosmetic Products. J. Prev. Med. Hyg. 60 (1), E50–E57. doi:10.3747/jpmh2019.60.1.1080

Pirouz, C., Ovigne, J.-M., Rossouw, F., Martinetti-Zeissier, S., Gomes, C., and Cotovio, J., et al. (2015). The Myeloid U937 Skin Sensitization Test (U-SENS) Addresses the Activation of Dendritic Cell Event in the Adverse Outcome Pathway for Skin Sensitisation. Toxicol. Vitro 29 (5), 901–916. doi:10.1016/j.tiv.2015.03.009

Posadas, S. J., and Pichler, W. J. (2007). Delayed Drug Hypersensitivity Reactions. J. Am. Acad. Dermatol. 53 (6), 1016–1820. doi:10.1016/j.jaad.2007.06.021

Roberts, D. W. (2018). 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. 98, 155–160. doi:10.1016/j.yrtph.2018.07.014

Roberts, D. W., and Lepottewin, J. P. (1998). Hapten-protein Interactions. Berlin, Heidelberg: Springer. 119–111. doi:10.1007/978-3-642-80331-4_6

Roberts, D. W., and Patlewicz, G. (2018). Non-animal Assessment of Skin Sensitization Hazard: Is an Integrated Testing Strategy Needed, and if so What Should Be Integrated? J. Appl. Toxicol. 38 (1), 41–50. doi:10.1002/jat.3479

Robinson, M. K., Gerberick, G. F., Ryan, C. A., McNamee, P., White, I. R., and Baskett, D. A. (2000). The Importance of Exposure Estimation in the Assessment of Skin Sensitization Risk. Contact Dermatitis 42 (5), 251–259. doi:10.1111/j.1529-8019.2000.tb03521.x

Rollin, B. E. (2003). Toxicology and New Social Ethics for Animals. Toxicol. Pathol. 31 (Suppl. 1), 128–131. doi:10.1080/01926230390175011

Roman, D. L., Isvoran, A., Filip, M., Ostafe, V., and Zinn, M. (2020). In Silico Assessment of Pharmacological Profile of Low Molecular Weight Oligo-Hydroxyalkanoates. Front. Bioeng. Biotechnol. 8 (1352). doi:10.3389/fbioe.2020.584010

Saikakhov, R., Chakravarti, S., and Klopman, G. (2013). Effectiveness of CASE Ultra Expert System in Evaluating Adverse Effects of Drugs. Mol. Inf. 32 (1), 87–97. doi:10.1002/minf.201200081

Sakaguchi, H., Ashikaga, T., Miyazawa, M., Kosaka, N., Ito, Y., Yoneyama, K., et al. (2009). The Relationship between CD68/CD54 Expression and THP-1 Cell Viability in an In Vitro Skin Sensitization Test - Human Cell Line Activation Test (H-CLAT). Cell Biol. Toxicol. 25 (2), 109–126. doi:10.1007/s10565-008-9059-9

Sakaguchi, H., Ryan, C., Ovigne, J.-M., Schroeder, K. R., and Ashikaga, T. (2010). Predicting Skin Sensitization Potential and Inter-laboratory Reproducibility of a Human Cell Line Activation Test (H-CLAT) in the European Cosmetics Association (COLIPA) Ring Trials. Toxicol. Vitro 24 (6), 1810–1820. doi:10.1016/j.tiv.2010.05.012

Sakaguchi, H., Ashikaga, T., Ashikaga, T., Yoshiida, Y., Ito, Y., Yoneyama, K., et al. (2006). Development of an In Vitro Skin Sensitization Test Using Human Cells; Human Cell Line Activation Test (H-CLAT) II. An Inter-laboratory Study of the H-CLAT. Toxicol. Vitro 20 (5), 774–784. doi:10.1016/j.tiv.2005.10.014

Sato, K., Umemura, T., Tamura, T., Kusaka, Y., Aoyama, K., Ueda, A., et al. (2012). Skin Sensitization Study by a New Qualitative Structure-Toxicity Relationships (QSTR) Approach: K-step Yard Sampling (KY) Methods. J. Oral Tissue Engin. 9 (3), 167–173. doi:10.11223/jote.9.167

Sato, K., Umemura, T., Tamura, T., Kusaka, Y., Aoyama, K., Ueda, A., et al. (2009). Skin Sensitization Study by Quantitative. Structure-Activity Relationships (QSAR). Alternatives Anim. Test. Experimentation 14 (3), 940–946. doi:10.11232/axtes.14.940

Schweiger, E. W., and Weingberg, J. M. (2004). Novel Antibacterial Agents for Skin and Skin Structure Infections. J. Am. Acad. Dermatol. 50 (3), 331–340. doi:10.1007/s10211-003-6655-8

Schöbel, J. A. H., Koleva, Y. K., Enoch, S. J., Bajot, F., Hewitt, M., Madden, J. C., et al. (2011). Measurement and Estimation of Electroplxic Reactivity for Predictive Toxicology. Chem. Rev. 111 (4), 2562–2596. doi:10.1021/cr100098n

Serio, F., Pazzolante, G., Cozzolino, G., D’Alba, M., Bagordo, F., De Giorgi, M., et al. (2017). A New Formulation Based on Oxazalone Sunflower Seed Oil: In vitro Antibacterial and Safety Evaluation. Ozone: Sci. Eng. 39 (3), 139–147. doi:10.1080/019191512.2016.1272405

Silva, F. A. L. S., Brites, G., Ferreira, I., Silva, A., Miguel Neves, B., Costa Pereira, J. L. (2019). Skins Own Natural Moisturizing Systems. Multivariate Modeling. Springer Nature, 411–432. doi:10.1007/978-3-319-41873-5_20

Steinberg, P. (2013). “High-throughput Screening Methods,” in Toxicity Testing Editor P. Steinberg. 1st ed. (Hoboken, NJ: John Wiley & Sons).
Stevenson, M., Czekala, L., Simms, L., Tschieres, N., Larne, O., and Walele, T. (2019). The Use of Genomic Allergen Rapid Detection (GARD) Assays to Predict the Respiratory and Skin Sensitizing Potential of E-Liquids. *Regul. Toxicol. Pharmacol.* 103, 158–165. doi:10.1016/j.yrtph.2019.01.001

Strickland, J., Daniel, A. B., Allen, D., Aguila, C., Ahir, S., Bancos, S., et al. (2019). Skin Sensitization Testing Needs and Data Uses by US Regulatory and Research Agencies. *Arch. Toxicol.* 93 (2), 273–291. doi:10.1007/s00204-018-2346-4

Strickland, J., Zang, Q., Kleinsteu, N., Paris, M., Lehmann, D. M., Choksi, N., et al. (2016). Integrated Decision Strategies for Skin Sensitization Hazard. *J. Appl. Toxicol.* 36 (9), 1150–1162. doi:10.1002/jat.3281

Strickland, J., Zang, Q., Paris, M., Lehmann, D. M., Allen, D., Choksi, N., et al. (2017). Multivariate Models for Prediction of Human Skin Sensitization Hazard. *J. Appl. Toxicol.* 37 (3), 347–360. doi:10.1002/jat.3366

Suzuki, M., Hirota, M., Hagino, S., Itagaki, H., and Aiba, S. (2009). Evaluation of Toxicokinetics and Vertebrate Species Relevance in the Prediction of In Vivo Skin Sensitization. *Toxicol. Vitro* 23 (4), 687–696. doi:10.1016/j.tiv.2009.09.002

Takenuchi, O., Fukui, S., Okamoto, K., Kurotani, S., Imai, N., Fujishiro, M., et al. (2015). Test Battery with the Human Cell Line Activation Test, Direct Peptide Reactivity Assay and DEREK Based on a 139 Chemical Data Set for Predicting Skin Sensitizing Potential and Potency of Chemicals. *J. Appl. Toxicol.* 35 (11), 1318–1332. doi:10.1002/jat.3127

Takenuchi, O., Miyazawa, M., Saito, K., Ashikaga, T., and Sakaguchi, H. (2013). Predictive Performance of the Human Cell Line Activation Test (H-CLAT) for Lipophilic Chemicals with High Octanol-Water Partition Coefficients. *J. Toxicol. Sci.* 38 (4), 599–609. doi:10.2131/jts.38.599

Tarca, A. I., Carey, V. J., Chen, X.-w., Romero, R., and Drăghici, S. (2007). Machine Learning and its Applications to Biology. *PLoS Comput. Biol.* 3 (6), e116. doi:10.1371/journal.pcbi.0030116

Teubner, W., Mehlung, A., Schuster, P. X., Guth, K., Worth, A., Burton, J., et al. (2013). Computer Models versus Reality: How Well Do In Silico Models Currently Predict the Sensitization Potential of a Substance. *Regul. Toxicol. Pharmacol.* 67 (3), 468–485. doi:10.1016/j.yrtph.2013.09.007

Thysen, J. P., Linneberg, A., Menné, T., and Johansen, J. D. (2007). The Epidemiology of Contact Allergy in the General Population – Prevalence and Main Findings. *Contact Dermatitis* 57 (5), 287–299. doi:10.1111/j.1600-0536.2007.01220.x

Toropova, A. A., Toropova, A. P., Benfenati, E., Gini, G., Leszczynska, D., and Leszczynski, J. (2013). CORAL: QSAR Model of Water Solubility Based on Local and Global SMILES Attributes. *Chemosphere* 90 (2), 877–880. doi:10.1016/j.chemosphere.2012.07.035

Toropova, A. P., and Toropov, A. A. (2017). Hybrid Optimal Descriptors as a Tool to Predict Skin Sensitization in Accordance to OECD Principles. *Toxicol. Lett.* 275, 57–66. doi:10.1016/j.toxlet.2017.03.023

Tourneix, F., Alépée, N., Detroyer, A., Eilstein, J., Ez-Zoubir, M., Teissier, S. M., et al. (2020). Skin Sensitisation Testing in Practice: Applying a Stacking Meta Model to Cosmetic Ingredients. *Toxicol. Vitro* 66, 104831. doi:10.1016/j.tiv.2020.104831

Tsuji-Inoue, K., Hirota, M., Ashikaga, T., Atobe, T., Kouzuki, H., and Aiba, S. (2014). Skin Sensitization Risk Assessment Model Using Artificial Neural Network Analysis of Data from Multiple In Vitro Assays. *Toxicol. Vitro* 28 (4), 626–639. doi:10.1016/j.tiv.2014.01.003

Tung, C.-W., Lin, Y.-H., and Wang, S.-S. (2019). Transfer Learning for Predicting Human Skin Sensitizers. *Arch. Toxicol.* 93 (4), 931–940. doi:10.1007/s00204-019-02420-x

Tung, C.-W., Wang, C.-C., and Wang, S.-S. (2018). Mechanism-informed Read-Across Assessment of Skin Sensitizers Based on SkinSensDB. *Regul. Toxicol. Pharmacol.* 94, 276–282. doi:10.1016/j.yrtph.2018.02.014

Tuschl, H., Kovac, R., and Weber, E. (2000). The Expression of Surface Markers on Dendritic Cells as Indicators for the Sensitizing Potential of Chemicals. *Toxicol. Vitro* 14 (6), 541–549. doi:10.108887-2333(00)00051-5

United Nation. (2013). Globally Harmonized System of Classification and Labelling of Chemicals (GHS).

Urbisch, D., Honarvar, N., Kolle, S. N., Mehlung, A., Ramirez, T., Teubner, W., et al. (2016). Peptide Reactivity Associated with Skin Sensitization: The QSAR Toolbox and TIMES Compared to the DPRA. *Toxicol. Vitro* 34, 194–203. doi:10.1016/j.yrtph.2016.04.005

Urbisch, D., Mehlung, A., Guth, K., Ramírez, T., Honarvar, N., Kolle, S., et al. (2015). Assessing Skin Sensitization Hazard in Mice and Men Using Non-animal Test Methods. *Regul. Toxicol. Pharmacol.* 71 (2), 337–351. doi:10.1016/j.yrtph.2014.12.008

van Eijl, S., Zhu, Z., Cupitt, J., Gierula, M., Götz, C., Fritsche, E., et al. (2012). Elucidation of Xenobiotic Metabolism Pathways in Human Skin and Skin Tissue Models by Protoxic Profiling. *PLoS One* 7 (7), e41721. doi:10.1371/journal.pone.0041721

Wang, C.-C., Lin, Y.-C., Wang, S.-S., Shih, C., Lin, Y.-H., and Tung, C.-W. (2017). SkinSensDB: A Curated Database for Skin Sensitization Assays. *J. Cheminform.* 9, 5. doi:10.1186/s13321-017-0194-2

Wanibuchi, S., Yamamoto, Y., Sato, A., Kasahara, T., and Fujita, M. (2019). The Amino Acid Derivative Reactivity Assay with Fluorescence Detection and its Application to Multi-Constituent Substances. *J. Toxicol. Sci.* 44 (12), 821–832. doi:10.2131/jts.44.821

Wareing, B., Kolle, S. N., Birk, B., Alépée, N., Haupt, T., Kathawala, R., et al. (2020). The Kinetic Direct Peptide Reactivity Assay (kDPCA): Intra- and Inter-laboratory Reproducibility in a Seven-Laboratory Ring Trial. *ALTEX* 37 (4), 639–651. doi:10.14573/altext.2004291

Weng, C.-F., and Leong, M. K. (2020). Chapter Two - In Silico Prediction of Drug-Induced Liver Injury: Quo Vadis? Elsevier.

Wilm, A., Kühnl, J., and Kirchmair, J. (2018). Computational Approaches for Skin Sensitization Prediction. *Crit. Rev. Toxicol.* 48 (9), 738–760. doi:10.1080/10408444.2018.1528207

Wilm, A., Norinder, U., Agea, M. I., de Bruyn Kops, C., Stork, C., Kühnl, J., et al. (2020). Skin Doctor CP: Conformal Prediction of the Skin Sensitization Potential of Small Organic Molecules. *Chem. Res. Toxicol.* 33, 330–344. doi:10.1021/acs.chemrestox.0c00253

Zang, Q., Paris, M., Lehmann, D. M., Bell, S., Kleinsteu, N., Allen, D., et al. (2017). Prediction of Skin Sensitization Potency Using Machine Learning Approaches. *J. Appl. Toxicol.* 37 (7), 792–805. doi:10.1002/jat.3424

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ta, Weng and Leong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.