表4　抗菌製品によるアレルギー事例

原因化学物質	アレルギー症状	用途	報告年
〈四級アンモニウム塩系抗菌剤〉			
塩化ベンザルコニウム	ACD	手指殺菌剤	1990
塩化ベンゼトニウム	ACD	手指殺菌剤	1991
〈アミノ酸系抗菌剤〉			
アルキルジアミノグリシン塩酸塩（テゴ-51）	ACD	手指殺菌剤	1989
〈ビグアナイド系抗菌剤〉			
グルコン酸クロルヘキシジン（ヒビテン）	ACD	手指殺菌剤	1986, 1991
	アナフィラキシー接触じんましん	手指殺菌剤	1989
	アナフィラキシー	抗菌カテーテル	1997
〈フェノール系抗菌剤〉			
2,4,4′-トリクロロ-2′-ヒドロキシジフェニルエーテル（イルガサン DP-300, トリクロサン）	ACD	手指殺菌剤	1980
〈イソチアゾリノン系抗菌剤〉			
5-クロロ-2-メチル-4-イソチアゾリン-3-オン（CMI）（ケーソン CG）			
2-メチル-4-イソチアゾリン-3-オン（MI）	ACD	殺菌防腐剤（香粧品）	1987, 1989, 1990, 1991, 1992
2-n-オクチル-4-イソチアゾリン-3-オン（OIT, ケーソン 893）	ACD	殺菌防腐剤（塗料, 接着剤）	1992, 1996（スペイン, ドイツ）
1,2-ベンズイソチアゾリン-3-オン（BIT）	ACD	殺菌防腐剤（切削油, 塗料）	1990
〈四級アンモニウム塩系抗菌剤〉			
四級アンモニウム塩	ACD	繊維用抗菌剤(液剤)(洗濯時使用)	1996
〈アルデヒド系抗菌剤〉			
α-ブロモシンナムアルデヒド（BCA）	ACD	湿気取り（防カビマット）	1987, 1989
	ACD	靴のにおいとり（防カビシート）	1998
〈有機ヒ素系抗菌剤〉			
10,10′-オキシービス（フェノキシ）アルシン（OBPA）	ACD	椅子（PVC レザー製表地）	1997
〈ピリジン系抗菌剤〉			
2,3,5,6-テトラクロロ-4-（メチルスルホニル）ピリジン	ACD	椅子（PV レザー製表地）	1998, 2005
	ACD	デスクマット（PVC）	2000, 2002, 2005
〈アニリド系抗菌剤〉			
3,4,4′-トリクロロカルバニリド（トリクロカルバン）	ACD	白衣（襟）	1999

ACD: アレルギー性接触皮膚炎

そのため，既知，新規を問わず染料アレルゲンについて化学構造，皮膚感作性，過去のACD事例等の情報をデータベース化していき，染料アレルゲンによるACDの原因究明を効率的に行い，かつACDの再発防止を図っていく必要がある（表3）。

原因化学物質：抗菌剤

抗菌剤をタイプ別にみると，無機系抗菌剤は汗に溶けづらいことから，皮膚障害の原因となる可能性は低い。しかし，無機系抗菌剤のうち，亜鉛，銅についてはヒト・パッチテストでの陽性例が報告されており，金属アレルギー患者は要注意である。なお，銀，酸化チタンによるACDの事例報告はこれまでのところ見当たらない。一方，有機系抗菌剤については，汗等によって加工製品から皮膚へ移行する可能性が高いため，皮膚障害について注目していく必要がある。

一方，イソチアゾリノン系化合物を配合した外国製化粧品，塗料，接着剤等により，ACD等の皮膚障害が発生したことも報告されていた。1997年に，従来塗料，接着剤等に使用されてきた有機ピリジン系抗菌剤の2,3,5,6-テトラクロロ（メチルスルホニル）ピリジン（TCMSP）で加工されたビニルレザー製椅子によって，直接接触した下腿部にACDが発生した。

TCMSPのMSDSには，「皮膚感作性あり」と記載されていたが，TCMSPの皮膚感作性の強さについては，具体的に記載されていなかった。TCMSPの皮膚感作性について，GPMT法，LLNA法による検討を実施した結果，非常に強い皮膚感作性物質であることが確認できた。

1998年12月に公開された抗菌製品ガイドラインに

表5　ゴム製品によるアレルギー事例

原因化学物質	アレルギー症状	用途	報告年
〈ジチオカーバメート系加硫促進剤〉			
ジメチルジチオカルバミン酸亜鉛	ACD	医療用ゴム手袋	1989, 1991
ジエチルジチオカルバミン酸亜鉛	ACD	医療用ゴム手袋	1989
ジブチルジチオカルバミン酸亜鉛	ACD	医療用ゴム手袋	1989
エチルフェニルジチオカルバミン酸亜鉛	ACD	作業用ゴム手袋	1987
〈アミン〉			
ジメチルアミン	ACD	医療用ゴム手袋	1991
ジエチルアミン	ACD	医療用ゴム手袋	1986, 1987
ピペリジン	ACD	医療用ゴム手袋	1986, 1987
〈メルカプトベンゾチアゾール系加硫促進剤〉			
2-メルカプトベンゾチアゾール	ACD	ゴムはきもの	1982, 1983, 1990
	ACD	膝装具（ゴムベルト）	2000
2,2′-ジベンゾチアジルジスルフィド	ACD	ゴムはきもの	1983, 1990
〈チオウレア系加硫促進剤〉			
ジエチルチオウレア	ACD	膝装具（パッド）	1999
〈P-フェニレンジアミン系老化防止剤〉			
N-イソプロピル-N′-フェニル-p-フェニレンジアミン	ACD	作業用ゴム手袋	1980
	ACD	工業用ゴム製品	1990
	ACD	農作業用ゴム長靴	1996
	ACD	イヤホン（ゴムリング）	2001
N-1,3-ジメチルブチル-N′-フェニル-p-フェニレンジアミン	ACD	農作業用ゴム長靴	1996
6-エトキシ-2,2,4-トリメチル-1,2-ジヒドロキノリン	ACD	農作業用ゴム長靴	1996
〈クロロプレンゴム系接着剤，固着剤樹脂〉			
p-tert-ブチルフェノルホルムアルデヒド樹脂	ACD	靴用接着剤	1985
	ACD	テーピングテープ	1987
	ACD	スニーカー	1987
	ACD	膝装具	1990, 1992
	ACD	マーカーペン	1990
	ACD	ウェットスーツ	2000

ACD: アレルギー性接触皮膚炎

沿って，業界団体である繊維評価技術協議会（SEK），抗菌製品技術協議会（SIAA）を中心に，①製品に抗菌剤の種類（無機系，有機系，天然系）を表示する，②安全性評価のために皮膚感作性試験を新たに実施する，③消費者の声を取り入れるために消費者代表を加えた委員会を新たに設置する等が具体的に実施されてきている。

さらに，「改正・消費生活用製品安全法」（2007年5月14日施行）に基づき，重大製品事故（治療に要する期間が30日以上の負傷・疾病，死亡事故，後遺障害事故）の発生事例について，経済産業省による公表・注意喚起とともに，製造メーカーによる対象製品の社告等での公表，製造・出荷の停止，製品の回収等が規定されている。

2006年，TCMSPにより抗菌加工されたデスクマットによるACD事例が数多く発生していたことが確認されたことから，家庭用品における重大製品事故の第一号として認定され，社告等の注意喚起，製品回収・交換が実施された（表4）。

原因化学物質：ゴム添加剤

1980年以降，患者でのパッチテスト，GPMT法等によるアレルゲン検索，原因製品の化学分析，文献情報，メーカー情報等を総合して，天然ゴム・合成ゴムに配合されるゴム添加剤によるACDの原因解明を進めることにより，原因製品と原因化学物質との関連性を明らかにできた（表5）。

手術用・家庭用ゴム手袋ではジチオカーバメート（DTC）系加硫促進剤やアミン化合物，ゴムはきものではメルカプトベンゾチアゾール系加硫促進剤，工業用ゴム製品や農作業用ゴム長靴ではアミン系老化防止剤が主要な原因となっていた。

すなわち，アミン系老化防止剤のN-イソプロピル-N′-フェニル-p-フェニレンジアミン（IPPD），MBT系加硫促進剤の2-メルカプトベンゾチアゾール（MBT），2,2′-ジベンゾチアジルジスルフィド（MBTS），接着剤成分のp-tert-ブチルフェノールホルムアルデヒド樹脂（PTBP-FR）等のような既知ゴムアレルゲンだけでなく，これまで注目されてこなかったDTC系化合物のジメチルジチオカーバメート亜鉛（ZDMC），ジエチルジチオカーバメート亜鉛（ZDEC），ジブチルジチオカーバメート亜鉛

(ZDBC), エチルフェニルジチオカーバメート亜鉛 (ZEPC), アミンのジメチルアミン (DMA), ジエチルアミン (DEA), ピペリジン (PIP), アミン系老化防止剤のN-1,3-ジメチルブチル-N'-フェニル-p-フェニレンジアミン (DMBPPD), 6-エトキシ-2,2,4-トリメチル-1,2-ジヒドロキノリン (ETMDQ) のような新規ゴムアレルゲンについても注目する必要があることを明らかにできた。

また, チオウレア系加硫促進剤については, 日本では, 輸入・膝装具によるACD事例において, ジエチルチオウレア (DETU) が原因化学物質であったことを明らかにできた1例のみであった。しかし, 欧米ではかなりの数のACD事例が報告されており, しかも, DETUだけでなく, ジブチルチオウレア (DBTU), ジラウリルチオウレア (DLTU), ジフェニルチオウレア (DPTU) 等も, GPMTにより強いアレルギー性物質であることが確認されたことから, 日本においても, チオウレア系加硫促進剤の今後の用途展開に注目していく必要がある。

これらのゴムアレルゲン情報は, 日本皮膚アレルギー・接触皮膚炎学会刊行の学会誌, 「アレルゲン解説書」等に掲載されるとともに, これらのゴムアレルゲンを除去したアレルギー患者用の代替製品の開発にも活用されてきた。

ACD患者用代替品としては, 既知のゴムアレルゲンを配合していない製品が有効である。すなわち, 手袋ではポリクロロプレン (ネオプレン) ゴム, ウレタンゴム, シリコーンゴム等のゴム製品, 熱可塑性樹脂, ポリ塩化ビニル, ポリエチレン等のプラスチック製品が, 靴ではウレタン製品 (テニスシューズ等) が使用できる。また, 消費者がすぐにできるACD予防策としては, ゴム手袋の下に綿手袋を着けたり, 靴をはくとき必ず靴下を着けるようにする等, 製品が直接皮膚に触れないようにすることが, 簡単で, しかも効果的な方法として推奨できる。

2.4 安全対策

今後, 消費者, 特に有害性情報を必要とするアレルギー患者等のために, ①健康被害の原因解明を進め, 原因製品と原因化学物質の関連性を明らかにすること, ②MSDSの充実をはかり, 労働衛生上の健康被害の発生防止のために, ゴム添加剤メーカーから中間・最終製品メーカーへ, 用途, 曝露ルート・曝露レベルを考慮したリスク評価も含めた有害性情報等の製品情報を伝達できること, ③消費者にも具体的でわかりやすい製品表示を通じて, 製品情報の伝達機能を質量ともに高めていくとともに, 製品表示, 業界・メーカーのホームページ等を通じて, 幅広く製品情報を公開して, 消費者の理解度を高めていくことが重要である。

謝 辞

本稿のもとになった, 家庭用品に関する調査研究は, 当該機関の研究倫理委員会等の承認を得るとともに, 実験動物等の管理・使用に関するガイドライン等に従って, 調査研究は実施されました。その際, 多くの関係諸氏に御指導・御協力をいただきました。当所の療品部を始め関係部の皆さん, 都道府県市・衛生研究所等公的試験研究機関の関係諸氏に, 心より深謝いたします。また, 家庭用品による健康被害の原因解明に際して御協力いただいた, 学会, 病院皮膚科医, 家庭用品関連業界・メーカー等の関係諸氏に感謝申し上げます。

文 献

(1) 鹿庭正昭. 化学物質による皮膚障害 (7) 総論7. 接触アレルゲン解明の手順. 医薬ジャーナル 2000;36:5-9.
(2) 鹿庭正昭. 化学物質による皮膚障害 (18) 各論11. 接触アレルゲン解明の実際 (1) ～ゴム製品によるアレルギー性接触皮膚炎～. 医薬ジャーナル 2001;37:5-13.
(3) 鹿庭正昭. 化学物質による皮膚障害 (19) 各論12. 接触アレルゲン解明の実際 (2) ～プラスチック製品 (めがね部品) によるアレルギー性接触皮膚炎～. 医薬ジャーナル 2001; 37:5-16.
(4) 鹿庭正昭. 化学物質による皮膚障害 (57) 各論50. P-tert-butylphenol formaldehyde resin によるアレルギー性接触皮膚炎. 医薬ジャーナル 2004;40:5-13.
(5) 鹿庭正昭. 抗菌加工製品の現状と消費者への健康影響. 抗菌のすべて—ヘルスケアとメディカル・食品衛生・繊維・プラスチック・金属への展開—. 大阪：繊維社, 1997.
(6) 鹿庭正昭. 求められる製品の化学物質情報：家庭用品の化学物質情報の現状と課題. 化学物質と環境 2001;45:1-4.
(7) 鹿庭正昭. ゴム製品による健康被害の発生実態および健康被害情報の伝達の現状—アレルギー性接触皮膚炎, ラテックスアレルギーを中心に—. 日本ゴム協会誌 2004;77:213-218.
(8) 鹿庭正昭. 家庭用品に使用される化学物質による健康被害と安全対策. 国立医薬品食品衛生研究所報告 2006;124:1-20.

ORIGINAL ARTICLE

Skin Sensitization Study by Quantitative Structure-Activity Relationships (QSAR)

Kazuhiro Sato[1], Tomohiro Umemura[1], Tarou Tamura[1], Yukinori Kusaka[1], Kohji Aoyama[2], Atsushi Ueda[3], Kohichi Harada[3], Keiko Minamoto[3], Takemi Otsuki[4], Kunihiko Yamashita[5], Tatsuya Takeshita[6], Eiji Shibata[7], Kunio Dobashi[8], Satomi Kameo[9], Muneyuki Miyagawa[10], Masaaki Kaniwa[11], Yoko Endo[12] and Kohtaro Yuta[13,14]

[1]Department of Environmental Health, School of Medicine, University of Fukui, Fukui, Japan,
[2]Department of Environmental Medicine, Kagoshima University Graduate School of Medical and Dental Sciences, Kagoshima, Japan,
[3]Department of Environmental Health, Faculty of Medical and Pharmaceutical Sciences, Kumamoto University, Kumamoto, Japan,
[4]Department of Hygiene, Kawasaki Medical School, Kurashiki, Japan,
[5]Daicel Chemical Industries, Ltd., Himeji, Japan,
[6]Department of Public Health, School of Medicine, Wakayama Medical University, Wakayama, Japan,
[7]Department of Hygiene, Aichi Medical University, Nagakute, Japan,
[8]School of Health Sciences, Faculty of Medicine, Gunma University, Maebashi, Japan,
[9]Department of Public Health, Gunma University, Maebashi, Japan,
[10]National Institute of Occupational Safety and Health, Kawasaki, Japan,
[11]Division of Medical Devices, National Institute of Health Sciences, Tokyo, Japan,
[12]Research Center for Occupational Poisoning, Tokyo Rosai Hospital, Tokyo, Japan,
[13]Fujitsu Limited, Tokyo, Japan (14present: Research Center for Environmental Risk, National Institute for Environmental Studies, Tsukuba, Japan)

Abstract

In silico assessment of skin sensitization is increasingly needed owing to the problems concerning animal welfare, as well as excessive time consumed and cost involved in the development and testing of new chemicals. Skin sensitization positive/negative prediction models with discriminant function were generated and parameter analysis was discussed on the basis of QSAR technology.

Samples used in this research were selected from the list of "Maximale Arbeitsplatz-Konzentration" (MAK) and "Biologischer Arbeitsstoff-Toleranz-Wert" (BAT) values 2008, Deutschen Forschungsgemeinschaft (DFG) for positive samples (skin sensitizers) and from the classification results of the Japanese Globally Harmonized System of Classification and Labeling of Chemicals (GHS) Inter-ministerial Committee of the National Institute for Technology and Evaluation for negative skin sensitizers (controls). A total of 291 compounds (122 positive sensitizers and 169 negative sensitizers) were used in this study.

Parameters were generated from 2-D and 3-D structures of compounds. All of the approximately 800 parameters generated were reduced to 47 parameter sets and 32 parameter sets by feature selection. Various linear and non-linear discriminant analysis methods were applied using 2 parameter sets. All data analyses were performed using ADMEWORKS/ModelBuilder software.

Perfect classification ratios (100%) were achieved using Support Vector Machine and AdaBoost for 32 parameters. The highest prediction ratio of 81.44% by Leave-Ten-Out Cross-Validation was achieved with Neutral Network for 47 parameter sets. Log P was not found to be important.

This is the first QSAR model for skin sensitization from Japan. Future studies of this QSAR model are needed to improve its efficacy.

Key words: *skin sensitization, QSAR, animal study, occupational exposure limit*

Introduction

Occupational skin disorders are the most common non-traumatic occupational condition. They include contact dermatitis, contact urticaria, eczema, skin cancer and other conditions (Fedorowicz *et al.*, 2005). Among them, contact dermatitis is by far the most common form of occupational skin illness. In the United Kingdom, approximately 22% of all occupational diseases are skin diseases and 80% of them are contact dermatitis (Cherry *et al.*, 2000). Contact dermatitis was found to be the most prevalent occupational allergic disease treated by medical doctors in Japan (Sato *et al.*, 2004).

In July 2003, the United Nations published the Globally Harmonized System of Classification and Labelling of Chemicals (GHS). GHS became available in 2008 all over the world (United Nations, 3rd rev. 2009). The classification criteria for skin sensitizers in GHS include evidence from animal studies, for example, OECD Guideline 406 (the guinea pig maximization test and the Buhler guinea pig test) and Guideline 429 (local lymph node assay). According to DFG, 2009, in Germany and the European Chemical Bureau (ECB), European Union (EU), the criteria for skin sensitizers also include evidence from validated animal studies. In Japan, the criteria from the Japan Society for Occupational Health (JSOH) for skin sensitizers do not include evidence from animal studies (Japan Society for Occupational Health, 2008), an issue that may warrant revision.

Under the new European Union (EU) Registration, Evaluation and Authorization of Chemicals (REACH) rules, all chemicals in the EU that are produced or imported in quantities of more than 1 ton per annum will need to be assessed as potential human and environmental hazards, for example, in terms of their carcinogenicity; human sensitivity to such chemicals will also need to be determined. REACH calls for increased use of hazard assessment alternatives such as *in vitro* methods and QSARs. Since no *in vitro* replacement is currently available for skin sensitization, nor is expected to be ready in the near future, the use of QSAR approaches presents an attractive alternative (Patlewicz *et al.*, 2008). Furthermore, the legislative trend towards the abolition of the animal testing of cosmetic products in the seventh Amending Directive 2003/15/EC to Cosmetics Directive 76/768/EEC includes a demand for alternative evaluation procedures (Carrera *et al.*, 2009). Although reliable procedures for skin sensitization (OECD guideline) exist, their application is limited by time and cost constraints for the development of new chemicals (Golla *et al.*, 2009). Owing to this legislation and animal welfare concerns, computational techniques, such as QSARs, have recently been advanced for assessing various human toxicities (e.g., carcinogenicity, skin sensitivity).

We performed discriminant analyses for skin sensitizers and control chemicals and identified classification rates and prediction rates using ADMEWORKS/ModelBuilder (Fujitsu Kyusyu Systems Limited, Japan) (Hayashi, 2005). This is the first QSAR model for skin sensitization from Japan. Log P is an octanol/water partition coefficient. When it is large, the substance is hydrophobic and can permeate through a membrane. A substance with small Log P is hydrophilic and has difficulty permeating through a membrane. We postulated that Log P is important and therefore assigned it a high weight of the discriminant function. However, the weight of Log P was small at -0.05034 among 47 sets (Table 1). There was Log P, not remaining in 32 parameters (Table 2). Therefore, on the basis of the sample set used, these indicated that Log P is not important in skin sensitization.

Materials and Methods

Data sources for chemicals assessed

Positive data are for skin sensitizers from a list of 190 compounds as allergens, that is, Sh and Sah (Deutschen Forschungsgemeinschaft, DFG, 2008). The criteria are based on human epidemiological studies, case reports or validated animal studies (guinea pig maximization test, Buhler guinea pig test or mouse local lymph node assay). On the other hand, negative data are for 218 control compounds belonging to the group defined as 'not applicable' for skin sensitization in the results of the classification by the Japanese GHS Inter-ministerial Committee in NITE. Information concerns the status of the implementation of GHS in Japan, which means that these are reported as non-skin sensitizers. However, inorganic chemicals, organic metal chemicals and polymers are special compounds, and are not analyzed with general organic compounds in computational chemistry. Therefore, we deleted these compounds (117) and finally assessed 122 positive sensitizing compounds and 169 negative sensitizing compounds.

Parameters and discriminant function

A total of 291 compounds (122 positive sensitizing

chemicals and 169 negative sensitizing chemicals) were used. Parameters were generated from 2-D and 3-D structures of the compounds. All of the approximately 800 generated parameters were reduced to 47 parameter sets (Table 1) and further reduced to 32 parameter sets (Table 2) by feature selection (e.g. removing low incidence parameter, multicolinearity or noise parameter). Various linear and non-linear discriminant analyses including Neural Network (NN), Support Vector Machine (SVM), AdaBoost and the Iterative Least Squares linear discriminant (TILSQ) methods were applied using these 2 parameter sets. Negative coefficients in discriminant function indicate negative informa-

Table 1 47 parameter sets used in discriminant analysis

Parameter name	Weight	Average	SD
Superpendentivity index Halogen only	0.429203	32.604599	334.51001
All-path calc for substructure (-C-)	-0.32713	20.749599	31.982901
6th order cluster MC Simple	0.271849	0.031052	0.149152
Secondary sp3 carbon count	0.252074	0.405498	2.11412
Environment molecular connectivity of substructure (-O-C)	0.244973	0.235354	0.822589
Molecular distance edge between all sec quat C	-0.240368	0.478886	1.30622
Environment molecular connectivity of substructure (-C-)	0.218623	1.02915	0.970193
All-path calc for substructure (-C-C)	0.206777	6.86598	18.7416
Balabans topological index J	-0.19464	2.76624	0.6826
Environment molecular connectivity of substructure (-C(O)-)	0.178496	0.674613	0.839489
Molecular distance edge between all tert tert C	0.171595	0.189843	0.909471
Number of double bonds	-0.169725	1.43986	1.47356
7th order chain MC Simple	-0.161363	0.177073	0.30573
Environment molecular connectivity of substructure (-O-)	0.147864	0.877878	1.20036
Fractional mass of rotatable atoms	-0.143224	0.325496	0.295169
Distance weighted flexibility	-0.127034	10.1204	16.4505
Count of substructure (DMPATH) (-ester-)	-0.123993	0.309278	0.643709
Number of Chlorine atoms	0.114879	0.426117	1.03917
4th order cluster MC Valence	-0.102821	0.051807	0.144251
Shadow area 5 (normalized SHDW2)	0.102563	0.509005	0.091085
Superpendentivity index Carbons only	0.097147	69.290604	427.22699
Count of substructure (DMPATH) (-N)	-0.083129	0.171821	0.503125
All-path calc for substructure (=C=)	0.081448	16.8459	49.176701
Count of donors	0.080329	0.261168	0.765982
Environment molecular connectivity of substructure (-C)	0.079397	0.738309	0.633582
Shadow area 6 (normalized SHDW3)	-0.063777	0.509172	0.086962
Shadow area 3 (YZ plane)	0.05985	52.3074	29.987499
Non-bonded strain energy of molecule	0.056804	9.79443	12.5073
All-path calc for substructure (-O)	-0.055832	3.31443	10.7888
Mass weighted Width/Thickness	0.054246	318.558014	460.709015
Fractional mass of rigid atoms	0.053567	0.502532	0.257671
Angle strain energy of molecule	0.051041	17.335699	56.981201
FQlogP	-0.05034	2.04563	2.01295
Average E-State value over all hetero-atoms	-0.049819	6.93422	2.51043
Molecular distance edge between all primary quat C	0.048143	0.238553	0.629589
First/second moment of inertia with H	0.041074	1.2189	0.207198
Count of substructure (DMPATH) (-O-C)	-0.036965	0.178694	0.577473
4th order cluster MC Simple	0.036219	0.078392	0.166832
Number of Sulfur atoms	0.029465	0.278351	0.66554
All-path calc for substructure (-C)	0.027272	15.5115	21.1327
Intermolecular distance between Emin and Emax	-0.026139	2.69278	1.9522
Combined symmetry	-0.025709	0.831032	0.211908
Environment molecular connectivity of substructure (-N=)	0.025219	0.309299	0.812074
Count of substructure (DMPATH) (-N-)	-0.01344	0.199313	0.499793
Shadow area 4 (normalized SHDW1)	-0.012423	0.486792	0.076058
Third moment of inertia with H	0.01242	525.281006	598.361023
All-path calc for substructure (-N=)	0.009968	9.52062	28.847401
CONSTANT	0.163538		

tion of the activity. Positive coefficients indicate positive information (Fig. 1). Classification rates and prediction rates (Cross-Validation, CV) of discriminant function were also calculated. All data analyses were performed using ADMEWORKS / ModelBuilder software (Fujitsu Kyusyu Systems Limited, Japan) (Hayashi, 2005).

Results

Weight of parameter

The weight of 'secondary sp3 carbon count', that is, the number of methylene groups (-CH2-), was 0.290789 and that of 'environment molecular connectivity of substructure' (-O-C) was 0.288673, which were both highly positive values as shown in Table 2. These are considered to induce skin sensitization. On the other hand, the weights of 'all-path calc for substructure' (-C-) of -0.265617 and 'count of substructure (DMPATH)' (-ester-) of -0.233505 were highly negative values and considered to induce no sensitization (Table 2).

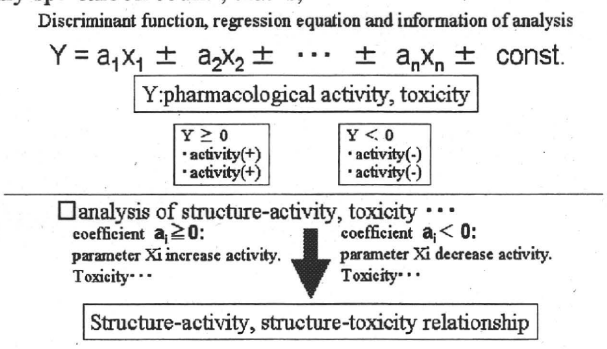

Figure 1 Format of regression equation with feedback of sign and coefficient information

Table 2 32 parameter sets used in discriminant analysis

Parameter name	Weight	Average	StdDev
Molecular distance edge between all sec quat C	−0.398936	0.478886	1.30622
Secondary sp3 carbon count	0.290789	0.405498	2.11412
Environment molecular connectivity of substructure (-O-C)	0.288673	0.235354	0.822589
All-path calc for substructure (-C-)	−0.265617	20.749599	31.982901
4th order cluster MC Simple	0.251247	0.078392	0.166832
Count of substructure (DMPATH) (-ester-)	−0.233505	0.309278	0.643709
All-path calc for substructure (=C=)	0.226793	16.8459	49.176701
4th order cluster MC Valence	−0.214339	0.051807	0.144251
Environment molecular connectivity of substructure (-C-)	0.208057	1.02915	0.970193
Count of substructure (DMPATH) (-N)	−0.2064	0.171821	0.503125
Count of donors	0.173459	0.261168	0.765982
Environment molecular connectivity of substructure (-O-)	0.173022	0.877878	1.20036
7th order chain MC Simple	−0.159527	0.177073	0.30573
All-path calc for substructure (-C-C)	0.154986	6.86598	18.7416
Average E-State value over all hetero-atoms	−0.153089	6.93422	2.51043
Molecular distance edge between all tert tert C	0.148523	0.189843	0.909471
Environment molecular connectivity of substructure (-C(O)-)	0.147898	0.674613	0.839489
Intermolecular distance between Emin and Emax	−0.136821	2.69278	1.9522
Fractional mass of rotatable atoms	−0.1306	0.325496	0.295169
Number of Chlorine atoms	0.105433	0.426117	1.03917
Third moment of inertia with H	0.10009	525.281006	598.361023
Number of double bonds	−0.096271	1.43986	1.47356
Non-bonded strain energy of molecule	0.095734	9.79443	12.5073
Shadow area 3 (YZ plane)	−0.094677	52.3074	29.987499
Distance weighted flexibility	−0.081716	10.1204	16.4505
Mass weighted Width/Thickness	0.079719	318.558014	460.709015
Shadow area 6 (normalized SHDW3)	−0.058598	0.509172	0.086962
Angle strain energy of molecule	−0.052234	17.335699	56.981201
Count of substructure (DMPATH) (-N-)	−0.028803	0.199313	0.499793
Environment molecular connectivity of substructure (-C)	0.020012	0.738309	0.633582
Count of substructure (DMPATH) (-O-C)	−0.014755	0.178694	0.577473
First/second moment of inertia with H	0.011451	1.2189	0.207198
CONSTANT	0.139817		

Discriminant analysis

We performed discriminant analysis and Leave-Ten-Out Cross-Validation (CV) of 291 (122 positive, 169 negative) chemicals assessed by Neural Network (NN), the Iterative Least Squares linear discriminant (TILSQ), Support Vector machine (SVM) and AdaBoost with 32 or 47 parameter sets (Table 3-10).

Classification rates were 90.38% to 100%. Complete classification (100%) was achieved in SVM and AdaBoost for 32 parameter sets (Table 9 and 10). The prediction rate was assessed by Leave-Ten-Out Cross-Validation (CV). CV results were 73.88% to 81.44%. The highest CV of 81.44% was achieved in NN with 47 parameter sets (Table 3). CV in SVM and AdaBoost with 32 parameter sets were 76.63% and 79.04%, respectively (Table 9 and 10), whose classification rates were 100% in both cases.

Table 3 Neural Network (NN) with 47 parameters

Class	Members	Wrong	Correct (%)
0	169	9	94.67
1	122	7	94.26
Total	291	16	94.50

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 81.44%

Table 4 The Iterative Least Squares linear discriminant (TILSQ) with 47 parameters

Class	Members	Wrong	Correct (%)
0	169	6	96.45
1	122	12	90.16
Total	291	18	93.81

Class 0: negative (control),
Class 1: positive skin sensitizer
CV: 73.88%

Table 5 Support Vector Machine (SVM) with 47 parameters

Class	Members	Wrong	Correct (%)
0	169	0	100.00
1	122	1	99.18
Total	291	1	99.66

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 78.01%

Table 6 AdaBoost with 47 parameters

Class	Members	Wrong	Correct (%)
0	169	0	100.00
1	122	2	98.36
Total	291	2	99.31

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 76.98%

Table 7 Neural Network (NN) with 32 parameters

Class	Members	Wrong	Correct (%)
0	169	9	94.67
1	122	13	89.34
Total	291	22	92.44

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation(CV): 79.04%

Table 8 The Iterative Least Squares (TILSQ) with 32 parameters

Class	Members	Wrong	Correct (%)
0	169	9	94.67
1	122	19	84.43
Total	291	28	90.38

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 75.26%

Table 9 Support Vector Machine (SVM) with 32 parameters

Class	Members	Wrong	Correct (%)
0	169	0	100.00
1	122	0	100.00
Total	291	0	100.00

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 76.63%

Table 10 AdaBoost with 32 parameters

Class	Members	Wrong	Correct (%)
0	169	0	100.00
1	122	0	100.00
Total	291	0	100.00

Class 0: negative (control),
Class 1: positive skin sensitizer
Cross-Validation (CV): 79.04%

Discussion

Since the implementation of Animal Welfare Guideline 86/609/EC in 1986, it is the declared policy of EU institutions to support the development and use of alternative methods of testing chemicals, that is, of "any method that can be used to replace, reduce or refine the use of animal experiments in biomedical research, testing or education" (Lillenblum *et al.*, 2008). However, no *in vitro* replacement is currently available for testing skin sensitization in compliance with the REACH system (Grindon *et al.*, 2007, Patlewicz *et al.*, 2008).

Therefore, several QSAR-related systems have been developed for skin sensitization. These are Toxicity Prediction Komputer-Assisted Technology, (Accelrys Inc., San Diego, CA, USA; TOPKAT) and Multi Computer-Automated Structure Evolution (MultiCASE Inc., Cleveland, Ohio, USA; M-CASE), which are both statistically based, Deductive Estimation of Risk from Existing Knowledge (Derek) for Windows (DfW. LHASA Ltd., Leeds, UK), which is knowledge-based, and Times Metabolism Simulator for Skin Sensitization (LMC, University of Bourgas, Bulgaria; TIMES-SS), which is a hybrid (Patlewicz *et al.*, 2007a, Patlewicz *et al.*, 2007b). In this study, all data analyses were performed with ADMEWORKS/ModelBuilder software (Fujitsu Kyusyu Systems Limited, Japan). ADMEWORKS/ModelBuilder is statistically based software. Of these QSAR-related systems, Derek is the most widely used expert system for predicting skin sensitizing potential (Grindon *et al.*, 2007). Fedrowicz *et al.* (2005) reported that the correct classification of QSAR predictions for guinea pig data achieves values of 73.3% and 82.9% for TOPKAT and Derek, respectively, and that the correct classification using LLNA data equals 73% for Derek. Our results of prediction rates (CV) were 73.88% to 81.44%, almost the same as those in the previous report. Although QSAR systems are still being developed and have yet to become sufficiently powerful, the use of *in silico* methods has been proposed to make predictions of skin sensitization in the first stage of a decision-tree testing strategy for skin sensitization (Grindon *et al.*, 2007). This QSAR system is thought to be at practical use level.

We postulated that skin permeability is an important factor for skin sensitivity of chemicals and that Log P could become an important factor in modeling skin sensitization. Barratt (1994) reported that the molecular volume and the octanol/water partition coefficient (Log P) were important determinants of skin permeability. In this study, solubility was highly correlated with Log P (multicolinearity). We analyzed without solubility. The weight of log P was found to be small and the weight of molecular weight was not extracted in 32 parameters (Table 1 and 2). The limited amount of experimental data available on skin permeability presumably has prevented the development of robust QSAR models for permeability (Golla *et al.*, 2009). Alternatively, in this QSAR study, the octanol/water partition coefficient could be replaced by other more important parameters, or cooperative relationships of several parameters. Many chemicals are dissolved in solvents and exposed to human skin. The skin permeability of these solvents might have larger effects on skin sensitization than Log P. Clearly, more discussions and additional tests must be carried out. Our classification rates and prediction rates were 90.83% to 100% and 73.88% to 81.44%, respectively. To improve these rates, further studies of large amounts of experimental data, revisions of feature selection according to certain hypotheses of parameters, the addition of new parameters and selecting suitable discriminant function for skin sensitization are needed.

Acknowledgement

This study was supported by a Grant-in-Aid from the Ministry of Health, Labour and Welfare, Japan (H20-Labour-009).

References

Barratt, M.D. (2000) Prediction of toxicity from chemical structure. Cell Biol Toxicol, 16, 1-13.

Carrera G.V.S.M., Gupta S. and Aires-de-Sousa J. (2009) Machine leaning of chemical reactivity from databases of organic reactions. J Comput Aided Mol Des, 23. 419-429.

Cherry, N., Meyer, J.D., Adisesh, A., Brooks, R., Owen-Smith, Y. and Beck, M.H. (2000) Surveilance of occupational skin disease: EPIDERM and OPRA. Br J Deramtol, 142, 1128-1134.

Deutsche Forschungsgemeinshaft (DFG). List of MAK and BAT Values 2008. p129-136.

Weinheim Wiley-VCH, Weinheim Fedorowicz, A., Singh, H., Soderholm, S. and Demchuk, E. (2005) Structure-activity models for contact sensitization. Chem Res Toxicol, 18, 954-969.

ECB (European Chemical Beurou ; http: //ecb. jrc.ec. europa.eu/

Golla, S., Madihally, S., Robinson R.L., and Gasem K.A.M. (2009) Quantitatve structure-property relationship modeling of skin sensitization: a quantita-

tive prediction. Toxicol in Vitro, 23, 454-465.
Grindon C., Combes R., Cronin M., Roberts D.W. and Garrod J.F. (2007) An integrated decision-tree testing strategy for skin sensitization with respect to the requirements of the EU REACH Legistration. Altern Lab Anim, 35, 683-697.
Hayashi M., Kamata E., Hirose A., Takahashi M., Morita T. And Ema M. (2005) In silico assessment of chemical mutagenesis in comparison with results of Salmonella microsome assay on 909 chemicals. Mutat Res, 588, 129-135.
Japan Society for Occupational Health (2008) Recommendation of Occupational Exposure Limits (2008-2009). J Occup Health, 50, 426-443.
Lillenblum W., Dekant W., Foth H., Gebel T., Hengstler J.G., Kahl R., Kramer P.J., Schweinfurth H. and Wellin K.M. Alternative methods to safety studies in experimental animals: role in the risk assessment of chemicals under European Chemical Legislation (REACH). Arch Toxicol, 82, 211-236.
Patlewicz, G., Aptula A.O., Uriarte, E., Roberts, D.W., Kern, P.S., Gerberick, G.F., Kimber, I., Dearman R.J., Ryan, C.A. and Basketter, D.A. (2007a) An evaluation of selected global (Q)SARs/expert systems for the prediction of skin sensitisation potential. SAR QSAR, 18, 515-541.
Patlewicz, G., Dimitrov, S.D., Low, L.K., Kern, P.S., Dimitrova, G.D., Comber M.I.H., Aptua A.O., Philips, R.D., Niemela, J., Madsem, C., Wedebye, E.B., Roberts, D.W., Bailey, P.T. and Mekenyan, O.G. (2007b) TIMES-SS—A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity. Regul Toxicol Pharmacol, 48, 225-239.
Patlewicz, G., Aptula, A.O., Roberts D.W. and Uriarte, E. (2008) A minireview of available skin sensitization (Q)SARs/Expert systems. QSAR Comb Sci 27, 60-76.
Sato, K., Kusaka, Y., Suganuma, N., Nagasawa, S. and Deguchi, Y. (2004) Occupational allergy in medical doctors. J Occup Health, 46, 165-170.
The Japanese GHS Inter-ministerial Committee, Results of the classification, in NITE Information about the status of the implementation of GHS in Japan,http://www.safe.nite.go.jp/english/ghs_index.html#results accessed on Nov. 4th, 2009
United Nations. (2009) Globally harmonized system of classification and labelling of chemicals (GHS). 3rd rev., New York and Gebeva.

(Received: December 9. 2009/
Accepted: December 25, 2009)

Corresponding author:
Dr Kazuhiro Sato
Department of Environmental Health,
School of Medicine, University of Fukui,
Matsuoka, Eiheiji-cho, Yoshida-gun,
Fukui910-1193, Japan
Tel:+81-776-61-8338
Fax:+81-776-61-8107
E-mail:satokazu@u-fukui.ac.jp

ORIGINAL ARTICLE

A Respiratory Sensitization Study by a New Quantitative Structure-Activity Relationships (QSAR)

Kazuhiro Sato[1], Tomohiro Umemura[1], Taro Tamura[1], Yukinori Kusaka[1], Toshiko Ido[2], Kohji Aoyama[3], Atsushi Ueda[5], Kohichi Harada[4], Keiko Minamoto[5], Takemi Otsuki[5], Kunihiko Yamashita[6], Tatsuya Takeshita[7], Eiji Shibata[8], Kunio Dobashi[9], Satomi Kameo[10], Muneyuki Miyagawa[11], Masaaki Kaniwa[12], Takahiko Yoshida[13], Tetsuhito Fukushima[14] and Kohtaro Yuta[15*]

[1]Department of Environmental Health, School of Medicine, University of Fukui, Fukui, Japan, [2]Department of Dermatology, School of Medicine, University of Fukui, Fukui, Japan, [3]Department of Environmental Medicine, Kagoshima University Graduate School of Medical and Dental Sciences, Kagoshima, Japan, [4]Department of Environmental Health, Faculty of Medical and Pharmaceutical Sciences, Kumamoto University, Kumamoto, Japan, [5]Department of Hygiene, Kawasaki Medical School, Okayama, Japan, [6]Daicel Chemical Industries, Ltd., Hyogo, Japan, [7]Department of Public Health, School of Medicine, Wakayama Medical University, Wakayama, Japan, [8]Department of Hygiene, Aichi Medical University, Aichi, Japan, [9]School of Health Sciences, Faculty of Medicine, Gunma University, Gunma, Japan, [10]Department of Public Health, Gunma University, Gunma, Japan, [11]National Institute of Occupational Safety and Health, Kanagawa, Japan, [12]Division of Medical Devices, National Institute of Health Sciences, Tokyo, Japan, [13]Department of Health Science, School of Medicine, Asahikawa Medical University, Asahikawa, Japan, [14]Department of Hygiene and Preventive Medicine, School of Medicine, Fukushima Medical University, Fukushima, Japan, [15]Fujitsu Limited, Tokyo, Japan (*present: In Silico Data Ltd., Chiba, Japan)

Abstract

New respiratory sensitization positive/negative prediction models with discriminant functions were generated and parameter analyses were discussed on the basis of QSAR technology. Samples used in this research were selected from the list of European Chemical Bureau (ECB): R42, R42/43 for positive samples (respiratory sensitizers) and from the classification results of the Japanese Inter-ministerial Committee for negative respiratory sensitizers (controls). A total of 214 compounds (61 positive sensitizers and 153 negative sensitizers) were used in this study. Parameters were generated from 2-D and 3-D structures of compound. All of the approximately 800 parameters generated were reduced to 12 parameter set by feature selection. Various linear and non-linear discriminant analysis methods were applied using the parameter set. All data analyses were performed using ADMEWORKS/ ModelBuilder software. Perfect classification ratios (100%) were achieved using Iterative Least Squares (ILS) and AdaBoost. The highest prediction ratio of 97.2% by leave-one-out cross-validation was achieved with Support Vector Machine (SVM). This model is applicable to initial prediction of respiratory sensitization.

Key words: respiratory sensitization, quantitative structure-activity relationships (QSAR), animal study

Introduction

In occupational health, occupational skin disorders are the most common non-traumatic occupational condition. They include contact dermatitis, contact urticaria, eczema, skin cancer and other conditions. Among them, contact dermatitis is by far the most common form of occupational skin illness (Fedorowicz *et al.*, 2005). Respiratory sensitization is characterized by rhinitis, asthma, hypersensitivity pneumonia and eosinophilic pneumonia (Sato *et al.*, 2007). The prevalence of respiratory sensitization is low compared with that of contact dermatitis. However, severe and occasionally fatal reactions have been reported in respiratory sensitization (Fabbri *et al.*, 1988).

In July 2003, the United Nations published the Globally Harmonized System of Classification and Labeling of Chemicals (GHS). GHS became available in 2008 all over the world (United Nations, 2009). The classification criteria for sensitizers in GHS include evidence from animal studies. According to DFG, 2009, in Germany and the European Chemical Bureau (ECB), European Union (EU), the criteria for skin sensitizers also include evidences from validated animal studies. Moreover, under the new European Union (EU) Registration, Evaluation and Authorization of Chemicals (REACH) rules, all chemicals in the EU that are produced or imported in quantities of more than 1 ton per annum will need to be assessed as potential human and environmental hazards, for example, in terms of their carcinogenicity and human sensitivity to such chemicals will also need to be determined. REACH calls for increased use of hazard assessment alternatives such as *in vitro* methods and QSARs. Since no *in vitro* replacement is currently available for sensitization, the use of QSAR approaches presents an attractive alternative (Patlewicz *et al.*, 2008). Furthermore, the legislative trend towards the abolition of the animal testing of cosmetic products in the seventh Amending Directive 2003/15/EC to Cosmetics Directive 76/768/EEC includes a demand for alternative evaluation procedures (Carrera *et al.*, 2009). Although animal studies for sensitization exist, their application is limited by time and cost constraints for the development of new chemicals (Golla *et al.*, 2009). Owing to this legislation and animal welfare concerns, computational techniques, such as QSARs, have recently been advanced for assessing various human toxicities (e.g., carcinogenicity, sensitization). Therefore, QSAR-related systems have been extensively developed for skin sensitization. Some of them are already commercialized. Those are Toxicity Prediction Komputer- Assisted Technology, (Accelys Inc., San Diego, CA, USA; TOPKAT) and Computer-Automated Structure Evolution (MultiCASE Inc., Cleaveland, Ohio, USA; M-CASE), which are both statistically based, Deductive Estimation of Risk from Existing Knowledge (Derek) for Windowss (DfW, LHASA Ltd., Leeds, UK), which is knowledge-based, and Times Metabolism Stimulator for skin Sensitization (LMC, University of Bourgas, Bulgaria; TIMES-SS), which is hybrid (Patlewiez *et al.*, 2007a, Patlewiez *et al.*, 2007b). We also made Japanese first QSAR model for skin sensitization using ADMEWORKS/ModelBuilder software (Fujitsu Kyushu Systems Limited, Japan), which is statistically based (Hayashi *et al.*, 2005, Sato *et al.*, 2009).

To date there is no thoroughly validated method to induce and detect respiratory sensitization in animal model (Deutschen Forschungsgemeinschaft, DFG, 2008, United Nations, 2009), although characterization of sensitization potential by cytokines produced by T-cells in *in vivo* tests shows considerable promise (Dearman *et al.*, 1996). Compared with QSAR model for skin sensitization, QSAR models for respiratory sensitization are scarce (Agius *et al.* 1991, Karol *et al.* 1996, Graham *et al.* 1997, Seed *et al.* 2008, Warne *et al.* 2009, Seed *et al.* 2010). We made QSAR model for respiratory sensitization using ADMEWORKS/ModelBuilder (Fujitsu Kyushu Systems Limited, Japan). and the highest prediction rate was 97.2% by leave-one-out cross-validation with Support Vector Machine (SVM). This is the first Japanese QSAR model for respiratory sensitization.

Materials and Methods

1 Data sources for chemicals assessed

Positive data for respiratory sensitizers are from a list of 116 compounds as allergens, that is, R42 and R42/R43 (European Commission). R42 are the substances which may cause sensitization by inhalation. "*if there is evidence that the substance or preparation can induce specific respiratory hypersensitivity, where there are positive results from appropriate animal tests, or. if the substance is an isocyanate, unless there is evidence that the specific isocyanate does not cause respiratory hypersensitivity.* " R42/43 are allergens which may cause sensitization by inhalation and skin contact. On the other hand, negative data are for 222 control compounds(4 substances belonging to the group defined as 'not classified' for respiratory sensitizer, which means that these are reported as non-respiratory sensitizers and 218 substances belonging to the group defined as 'not classified' for skin sensitization, which means that these are reported as non-skin sensitizers and also defined as 'not possible' for respiratory sensitization ,which means that these are impossible to classify as respiratory sensitizer) in the results of the classification by the Japanese GHS Inter-ministerial Committee in National Institute of technology and evaluation (NITE) Information concerns the status of the implementation of GHS in Japan (NITE Information). We also regarded dermal non-sensitizers as respiratory non-sensitizers (Graham C *et al.*, 1997). However, inorganic chemicals, organic metal chemicals and polymers are special compounds, and are not analyzed with general organic compounds in computational chemistry. Therefore, we deleted these compounds (124) and finally assessed 61 positive sensitizing compounds and 153 negative sensitizing compounds (total 214 substances).

2.2 Parameters and discriminant function.

A total of 214 compounds (61 positive sensitizing chemicals and 153 negative sensitizing chemicals) were used. Parameters were generated from 2-D and 3-D structures of the compounds. All of the approximately 800 generated parameters were reduced to 12 parameter sets (Table 1) by feature selection (e.g. removing low incidence parameter, multicolinearity or noise parameter). Various linear and non-linear discriminant analyses including Linear Learning Machine (LLM), Neural Network (NN), Support Vector Machine (SVM), AdaBoost and Iterative Least Squares (ILS) linear discriminant methods were applied using 12 parameter set. Negative coefficients in discriminant function indicate negative information of the activity. Positive coefficients indicate positive information. Classification rates and prediction rates (leave-one-out cross-validation, CV) of discriminant function were also calculated. All data analyses were performed using ADMEWORKS/ ModelBuilder software (Fujitsu Kyushu Systems Limited, Japan). This study was approved by the Conflict of Interest (COI) Committee of School of Medicine, University of Fukui, Fukui, Japan.

Table 1 12 Parameters for respiratory sensitization

Parameter name	Weight	Average	Std Dev
Count of substructure (DMPATH) (-ester-)	0.38621	0.432558	0.769671
Count of substructure (DMPATH) (-N-)	0.317884	0.274419	0.636907
Fast count the substructure (=S=)	0.182767	0.153488	0.916961
Fast count the substructure (-N-)	0.084473	0.306977	0.70293
Environment molecular connectivity of substructure (-C(O)-)	0.069542	0.816819	0.864617
All-path calc for substructure (-C)	-0.373663	18.116777	22.04109
Molar refractivity environment of the substructure (-C(O)N-)	-0.251503	5.153574	17.063501
Molar refractivity environment of the substructure (-C(O)N=)	-0.197693	4.361633	16.89031
Number of double bonds	0.150684	1.776744	1.636733
Average Energy Resulting from ALL Group Energies	-0.431519	1.365581	10.02545
Bond strain energy of molecule	-0.112731	112.8868	106.937775
Mass weighted Width/Thickness	-0.066721	620.8718	479.725891
CONSTANT	-0.48619		

Results

1 Weight of parameter

The weight of 'Count of substructure (DMPATH)', that is, the number of ester bonds (R-COO-R'), was 0.386210, and that of 'Count of substructure (DMPATH)', that is, the number of nitrogen (-N-) was 0.317884, which were both highly positive values as shown in Table 1. These are considered to induce respiratory sensitization. On the other hand, the weight of 'All-path calc for substructure' (-C-) of -0.373663 and 'Average Energy Resulting from All Group Energies' of -0.431519 were highly negative values (Table 1). 'Average Energy Resulting from All Group Energies' calculates average binding energies as a sum of contributions of 10 common functional groups (Andrews binding) (Andrews PR *et al.*, 1984).

2 Discriminant analysis

We performed discriminant analyses and leave-one-out cross-validation (CV) of 214 (61 positive, 153 negative) chemicals assessed by Linear Learning Machine (LLM), Neural Network (NN), Iterative Least Squares (ILS) linear discriminant, Support Vector machine (SVM) and AdaBoost with 12 parameter set (Table 2).

Classification rates were 99.53% to 100%. Complete classification (100%) was achieved in ILS and AdaBoost. The prediction rate was assessed by leave-one-out cross-validation (CV). CV results were 95.33% to 97.2%. The highest CV of 97.2% was achieved in SVM. CV in ILS and AdaBoost were 95.33% and 96.73%, respectively, whose classification rates were 100%.

Table 2 Discriptive analyses, classification rate (%Correct) and prediction rate (Cross-Validation (CV)) with 12 parameters

Linear Learning Machine (LLM)

Class	Members	Wrong	%Correct
0	153	0	100
1	61	1	98.36
Total	214	0	99.53
CV: 95.79%			

Neural Network (NN)

Class	Members	Wrong	%Correct
0	153	0	100
1	61	1	98.36
Total	214	1	99.53
CV: 95.33%			

Iterative Least Squares (ILS)

Class	Members	Wrong	%Correct
0	153	0	100
1	61	0	100
Total	214	0	100
CV: 95.33%			

Support Vector Machine (SVM)

Class	Members	Wrong	%Correct
0	153	0	100
1	61	1	98.36
Total	214	1	99.53
CV: 97.20%			

AdaBoost

Class	Members	Wrong	%Correct
0	153	0	100
1	61	0	100
Total	214	0	100
CV: 96.73%			

Class 0: negative control, Class 1: positive respiratory sensitizer

Discussion

The paucity of a significant experimental database of respiratory sensitization potential has resulted in fewer attempts to form quantitative structure activity relationships (QSAR) for respiratory sensitization than those for skin sensitization (Rodford *et al.*, 2003). Of particular difficulty is the identification of reliable negative controls (Kimber *et al.*, 1996). We were unable to locate any clinical reports of non-sensitizing respiratory chemicals (control) with the exception of only 4 chemicals defined as 'not classified' for respiratory sensitization, which means that these are reported as non-respiratory sensitizers (NITE information). We assumed that dermal non-sensitizers are also respiratory non-sensitizers defined as 'not classified' for skin sensitization, which means that these are reported as non-skin sensitizers and also defined as 'not possible' for respiratory sensitization ,which means that these are impossible to

classify as respiratory sensitizer in the results of the classification by the Japanese GHS Inter-ministerial Committee in National Institute of technology and evaluation (NITE) Information about the status of the implementation of GHS in Japan (Graham *et al.*, 1997).

First computational methods for the analysis of the relationships between structure and activity of respiratory allergens have focused only upon electrophilic potential (Agius *et al.*, 1991). Graham C (1997) reported a SAR model for chemical respiratory allergens using CASE/MultiCase systems. The CASE/ MultiCase predicted activity in two sub models each: Units of potency and probability of activity. The sub models were validated using the leave-one-out method. The MultiCase models correctly predicted 80% of the sensitizers (sensitivity) and 98% of the nonsensitizers (specificity). The CASE models had lower sensitivities (72% and 75%) but similar specificities (98% and 95%) to MutiCase. The external data-withholding exercise used to validate the model yielded a sensitivity of 0.95 and a specificity of 0.95 (Graham *et al.*,1997).

Sensitizers were found to differ from non-sensitizers in certain physical-chemical properties, specifically in molecular weight, water solubility, Log P (octanol/water partition coefficient) or LUMO (lowest unoccupied molecular orbital) (Graham *et al.*, 1997). However, there were not these parameters in Table 1. It must be noted that respiratory sensitization does differ from skin sensitization in number of respects. Antigen formation can occur in the aqueous respiratory tract fluids, and thus water solubility is an important property to be considered. For skin sensitizers, low water solubility is necessary to enable the chemical to penetrate the skin (Rodford *et al.*, 2003). Our previous QSAR model on skin sensitization (Sato *et al.*, 2009) showed that the weight of 'Secondary sp3 carbon count', that is, the number of methylene groups (-CH2-), was 0.290789, that of 'Environment molecular connectivity of substructure' (-C(O)-) was 0.288673. These were high positive values in 32 parameter set. That of the weights of 'All-path calc for substructure' (-C-) was -0.265617 and that of 'Count of substructure (DMPATH)' (-ester-) was -0.233505. These were high negative values in 32 parameter set. That of Log P was –0.05034 in 47 parameter set. Both 'Environment molecular connectivity of substructure (-C(O)-)' and 'All-path calc for substructure (-C)' are kinds of molecular connectivity (MC) indexes (Todescini *et al.*, 2009, Randic's) which are indexes of branching and measures of the complexity of the assigned substructure ((-C(O)-: carbonyl group, (-C): methyl group). The basic algorithm of these parameters is modified based on Randic's topological parameters (Randic M, 1975). The prediction rate was assessed by leave-ten-out Cross-Validation (CV). CV results were 73.88% to 81.44%. The highest CV of 81.44% was achieved in NN with 47 parameter sets (Sato *et al.*,2009). We will perform external validation of the QSAR model for skin sensitization using 413 skin sensitizers (Coz *et al.*, 2006) according to Organization for Economic Co-operation and Development (OECD) principles of QSAR models in the near future (OECD, 2004).

In this study, 'Secondary sp3 carbon count' and 'Environment molecular connectivity of substructure' were not in Table 1. The weight of 'All-path calc for substructure' was -0.373663. That of 'Count of substructure(DMPATH)' (-ester-) was 0.386210. The absolute value of 'Average Energy Resulting from All Group Energies' was the largest in 12 parameters (-0.431519) (Table 1). 'Average Energy Resulting from All Group Energies' calculates average bind energies as a sum of contributions of 10 common functional groups (Andrews binding) (Andrews PR *et al.*, 1984). The primary use of 'Average Energy Resulting from All Group Energies' is to deduce whether a particular drug represents a good or bad match to its receptor. Negative value indicates bnding with its receptor or certain protein. Therefore, binding with its receptor or certain protein might be most important factor in respiratory sensitization (haptenation).

All approaches of these five discriminant methods (Table 2) achieved very high classification and prediction rates. Though there are not major differences in these five approaches, each model is strongly influenced by chance correlation and overfitting. The number of parameters in this study did not meet

chance correlation. The Linear Learning Machine (LLM) and Iterative Least Squares (ILS) linear discriminant models are suitable for predicting new samples because of their low probability of overfitting.

More extensive researches, e.g. further studies of large amounts of experimental data, revisions of feature selection according to certain hypotheses of parameters, the addition of new parameters and selecting suitable discriminant function for respiratory sensitization, are needed to investigate respiratory QSAR system.

Although QSAR systems are still being developed and have yet to become sufficiently powerful, the use of *in silico* methods has been proposed to make predictions of respiratory sensitization in the first stage of a decision-tree testing strategy for respiratory sensitization (Seed *et al.*, 2008). In this study, classification rates were 99.53% to 100%. Prediction rates (CV) were 95.33% to 97.2%. This QSAR system is thought to be applicable to initial prediction of the respiratory sensitizing ability of untested chemicals.

Acknowledgement

This study was supported by a Grant-in-Aid from the Ministry of Health, Labour and Welfare, Japan (H20-Labour-009).

References

Agius RM, Nee J, McGovern B, Robertson A. (1991) Structure activity hypotheses in occupational asthma caused by low molecular weight substances. Ann Occup Hyg, 35, 129-137.

Andrews PR, Craik DJ, Martin JL. (1984) Functional group contributions to drug-receptor interactions. J Med Chem 27, 1648-1657.

Carrera GV, Gupta S, Aires-de-Sousa J. (2009) Machine learning of chemical reactivity from databases of organic reactions. J Comput Aided Mol Dis 23, 419-429.

Coz CJL, Lepottevin JP. (2006) Dictonary of contact allergens: chemical structures, sources and references. In Frosch PJ, Menne T, Lepoittevin JP editors. Contact Dermatitis 4th ed. Springer-Verlag; 2006. pp943-1105, Berlin.

Dearman RJ, Basketter DA, Kimber I. (1996) Characterization of chemical allergens as a function of divergent cytokine secretion profiles induced in mice. Toxicol Appl Pharmacol 138, 308-316.

Deutsche Forschungsgemeinshaft (DFG). (2008) List of MAK and BAT Values IV Sensitizing substances. Wiley-VCH, 158-173, Weinheim.

European Commision. Annex 1 Directive 67/548/EEC on Dangerous Substances. http://ecb.jrc.ec.europa.eu/classification-labelling/clp/ visited on November 2, 2010.

Fabbri LM, Danieli D, Crescioli S, Bevilacqua P, Meli S, Saetta M, Mapp CE. (1988) Fatal asthma in a subject sensitized to a toluene diisocyanate. Am Rev Respir Dis 137, 1494-1498.

Fedrowicz A, Singh H, Soderholm S, Demchuk E. (2005) Structure-activity models for contact sensitization. Chem Res Toxicol 18, 954-969.

Golla S, Madihally S, Robinson RL Jr, Gasen KA. (2009) Quantitative structure-property relationship of skin sensitization: a quantitative prediction. Toxicol in vitro 23, 454-465.

Graham C, Rosenkranz HS, Karol MH. (1997) Structure-activity model of chemicals that cause human respiratory sensitization. Reg Toxicol Pharmacol 26, 296-306.

Hayashi M, Kamata E, Hirose A, Takahashi M, Morita T, Ema M. (2005) In silico assessment of chemical mutagenesis in comparison with results of Salmonella microsome assay on 909 chemicals. Mutat Res 588, 129-135.

Karol MH, Graham C, Gealy R, Macina OT, Sussman N, Rosenkranz HS. (1996) Structure-activity relationships and computer-assisted analysis of respiratory sensitization potential. Toxicol Lett 86, 187-191.

Kimber I, Bernstein IL, Karol MH, Robinson MK, Sarlo K, Selgrade MK. (1996) Workshop overview. Identification of respiratory allergens. Fundam Appl Toxicol 33, 1-10.

Graham C, Rosenkranz HS, Karol MH. (1997) Stucture-activity model of chemicals that cause human respiratory sensitization. Regul Toxicol Pharmacol 26, 296-306.

NITE Information about the status of the implementation of GHS in Japan, http://www.safe.nite.go.jp/english/ghs_index.html# results visited on November 2, 2010.

OECD (2004) OECD Series on Testing and Assessment. No. 49-The report from the expert group on

(quantitative) structure-activity relationships ((Q)SARs) on the principles for the validation of (Q)SARs. p206. http://appli1.oecd.org/olis/2004doc.nsf/linkto/env-jm-mono(2004)24 visited on November 2, 2010.

Patlewicz G, Aptula AO, Uriarte E, Roberts DW, Kern PS, Gerberick GF, Kimber I, Dearman RJ, Ryan CA, Basketter DA. (2007a) An evaluation of selected global (Q)SARs/expert systems for the prediction of skin sensitisation potential. SAR QSAR Environ Res 18: 515-541.

Patlewicz G, Dimitrov SD, Low LK, Kern PS, Dimitrova GD, Comber MIH, Aptua AO, Philips RD, Niemela J, Madsem C, Wedebye EB, Roberts DW, Bailey PT, Mekenyan OG. (2007b) TIMES-SS—A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity. Regul Toxicol Pharmacol 48, 225-239.

Patlewicz G, AptulaAO, Roberts DW, Uriarte E. (2008) A minireview of available skin sensitization (Q)SARs/Experts systems. QSAR Comb Sci 27, 60-76.

Randic M. (1975) Characterization of molecular branching. J Am Chem Soc 97, 6609-6615.

Randic's molecular connectivity index. http://en.wikippedia.org/wiki/Randi%C4%87_index visited on Dec 24, 2010.

Rodford R, Patlewicz G, Walker JD, Payne MP. (2003) Quantitative structure-activity relationships for predicting skin and respiratory sensitization. Environ Toxicol Chem 22, 1855-1861.

Sato K, Kusaka Y, Spaecial committee of Japanese Society of Occupational and Environmental Allergy. (2007) A proposal for guideline for prevention of allergic occupational asthma-in conformity with the globally harmonized system of classification and labeling o chemicals (GHS). Int J Immunol Pharmacol 20(S2), 51-53

Sato K, Umemura T, Tamura T, Kusaka Y, Aoyama K, Ueda A, Harada K, Minamoto K, Otsuki T, Yamashita K, Takeshita T, Shibata E, Dobashi K, Kame S, Miyagawa M, Kaniwa M, Endo Y, Yuta K. (2009) Skin sensitization study by Quantitative Structure-Activity Relationships (QSAR). AATEX 14, 940-946.

Seed MJ, Cullian P, Agius RM. (2008) Methods for the prediction of low-molecular-weight occupational respiratory sensitizers. Curr Opin Alergy Clin Immunol 8, 103-109.

Seed M, Agius R. (2010) Further validation of computer-based prediction of chemical asthma hazard. Occup Med 60, 115-120.

Todeschini R, Consonni V (2009) Molecular descriptors for chemoinformatics. Methods and Principles in Medical Chemistry. Vol.41. Mannhold R, Kubinyi H, Folkers G eds. Wiley-VCH, Weinheim.

United Nations. (2009) Globally harmonized system of classification and labelling of chemicals (GHS). 3rd rev., New York and Gebeva. http://www.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html visited on November 2, 2010.

Warne MA, Nicholson JK, Lindon JC, Guiney PD, Gartland KPR. (2009) A QSAR investigation of dermal and respiratory chemical sensitizers based on computational chemistry properties. SAR QSAR Environ Res 20, 429-451.

(Received: November 8/
Accepted: December 24)

Corresponding author:
Kazuhiro Sato, Ph.D.,
Department of Environmental Health,
School of Medicine, University of Fukui,
Shimoaizuki 23-3, Matsuoka, Eiheiji-cho,
Yoshida-gun, Fukui 910-1193, Japan
Tel.:+81-776-61-8338
Fax:+81-776-61-8107
E-mail: satokazu@u-fukui.ac.jp

Int J Immunopathol Pharmacol 2011; 24 (1S):17-21 in press

Validation study of the new criteria for sensitizer using German sensitizers of Deutschen Forschungsgemeinschaft (DFG)

Kazuhiro Sato*, Research group of classification and listing of sensitizer.
Department of Environmental Health, School of Medicine, University of Fukui, Eiheiji-cho, Fukui 910-1193, Japan
Tel: +81-776-61-8338, Fax: +81-776-61-8107, e-mali: satokazu@u-fukui.ac.jp
*Corresponding author

1 Table

Key words: [1]criteria for sensitizer, [2]animal studies, [3]globally harmonized system of classification and labeling of chemicals (GHS)

ABSTRACT

The globally harmonized system of classification and labeling of chemicals (GHS) was recommended by United Nations (UN) and became available in 2008 all over the world. The classification criteria for skin and airway sensitizers in GHS include evidences from animal studies, for example, OECD Guideline 406 (guinea pig maximization test, GPMT and Buhler guinea pig test) and Guideline 429 (local lymph node assay, LLNA). According to Deutschen Forschungsgemeinschaft (DFG) in Germany and European Chemical Bureau (ECB), the criteria for sensitizers also include evidences from validated animal studies. At present recognized and validated animal models for the testing of respiratory hypersensitivity are not available. In Japan, the criteria from the Japan Society for Occupational Health (JSOH) for sensitizers do not include evidences from animal studies. We revised the criteria for sensitizers of JSOH and adopted evidences of animal studies. We organized the research group for sensitizer in 2005 and reviewed the criteria of Germany, EU, GHS and so on (19 experts). The meetings were held twelve times and made the revised criteria for sensitizer which adopted animal studies. We tried to validate the criteria using 28 German sensitizers of DFG, which were not listed in JSOH.

We could correctly classify 24 sensitizers by our revised criteria, however, four sensitizers could not be classified at first. Therefore, we visited the secretariat of the committee of DFG in Freising, Germany to investigate the evidenced papers of these four sensitizers in October, 2008. We could find out the evidenced papers of two, however, two sensitizers could not be classified at last. We could correctly classify 26 out of 28 sensitizers. We concluded that our revised criteria were appropriate and that this validation study was successful.

There are over 30 million chemicals in the world. Many countries address classification and labeling for at least some chemicals in their countries. In the area of trade, the need to comply with multiple regulations regarding hazard classification and labeling is costly and time-consuming (confusing). These differences impact both protection and trade. In the area of protection, users in countries that don't have specific requirements may see different label warnings or data sheet information for the same chemical.

In July 2003, UN recommended and published globally harmonized system of classification and labeling of chemicals (GHS). GHS became available in 2008 all over the world (1). The classification criteria for skin sensitizers in GHS adopt evidences of validated animal studies, e.g. OECD guideline 406 (guinea pig maximization test, GPMT and Buhler test) and Guideline 429 (Local Lymph Node Assay, LLNA). The criteria for skin sensitizers of DFG in Germany (2) and European Chemical Bureau (ECB) (3) in European Union (EU) also adopt evidences of validated animal studies. There are some qualitative and quantitative relationships between the results of validated animal studies and human data (4). However, in Japan, Criteria for sensitizer of Japan Society for Occupational Health (JSOH) for sensitizers do not adopt evidences of animal studies (5) and are revised (6).

The definition of respiratory sensitization is that a respiratory sensitizer is a substance that will lead to hypersensitivity of the airways following inhalation of the substance and that at present, recognized and validated animal models for the testing of respiratory hypersensitivity are not available. Under certain circumstances, data from animal studies may provide valuable information in a weight of evidence assessment. The definition of skin sensitizer is that a skin sensitization is a substance that will lead to an allergic response following skin contact and that at present, evidence from animal studies, e.g. guinea pig maximization test (GPMT), Buhler test in OECD406, mouse local lymph node assay (LLNA) in OECD429 is usually more reliable than evidence from human exposure in reproducibility and handling (1,2). However, in cases where evidence is available from both sources, and there is conflict between the results, the quality and reliability of the evidence from both sources must be assessed in order to resolve the question of classification on a case-by-case basis (1).

JSOH Criteria of sensitizer has adopted only human evidences and has not adopted animal evidences till now (5). Therefore, we organized the research group of classification and listing of sensitizers, reviewed the criteria for sensitizer in German DFG (2), German dermatologists group (7), European Chemical Bureau (ECB) (3) and American Conference of Governmental Industrial Hygienists (ACGIH) (8) and revised the criteria for sensitizer in JSOH (5,6).

The criteria for sensitizer in German DFG (2) adopts validated animal studies and describe that at present, the sensitization potential of a substance can be best be estimated in animal studies in skin sensitization. Sensitizers in German DFG (2) are 234. German dermatologists group (7) also adopts appropriate animal studies in the criteria for skin sensitization and lists 244 skin sensitizers. The criteria of ECB (3) for sensitizer adopt appropriate animal studies and ECB lists R42 (respiratory sensitizer): 29, R42/43 (respiratory and skin sensitizer): 86, R43: 910 (skin sensitizer). In ACGIH, the designation "SEN" in the "Notation" column refers to the potential for an agent to produce sensitization, as confirmed by human or animal data and 26 sensitizers are listed (8). The list of sensitizers in JSOH (5) includes 18 airway and 32 skin sensitizers. We should expand the list of sensitizers in JSOH (5).

MATERIALS AND METHODS

Our research group held meetings twelve times to revise the criteria for sensitizer and reclassify the sensitizers of JSOH (5) from August 2005 to February 2010. Our new criteria for sensitizer are as follows (6);

Definition of sensitizers

Respiratory sensitizers are substances which can induce respiratory sensitization* by inhalation. Skin sensitizers are substances which can induce specific allergic skin sensitization by skin contact.

*rhinitis, asthma, hypersensitivity pneumonitis, eosinophilic pneumonia and so on.

Classification of sensitizers

Occupational sensitizers are recommended for respiratory tract and skin. The sensitizers are

classified into Group 1 substances which induce allergic reactions in humans, Group 2 substances which probably induce allergic reactions in humans and Group 3 substances which possibly induce allergic reactions in humans.

Occupational exposure limits

Recommendation of occupational exposure limits for the occupational sensitizers does not necessarily consider either prevention of induction or elicitation. Respiratory sensitization might be severe to human health.

Respiratory sensitizers

Group 1

There is a clear association between respiratory symptoms and occupational exposure. Case reports of positive inhalation challenge tests, of positive serological studies or positive prick tests are reported in at least two different research organizations.

and

There is at least one epidemiological study which indicate a clear association between respiratory symptoms and occupational exposure.

Group 2

There is a clear association between respiratory symptoms and occupational exposure. Case reports of positive inhalation challenge tests, positive serological studies or positive prick tests are reported in at least two different research organizations.

However, there isn't an epidemiological study.

Group 3

(1) Positive animal tests which fulfill all conditions as below are reported in at least two different research organizations.

(i) induction and elicitation are performed by inhalation, nasal or bronchial administration.

(ii) detections of elicitation are performed by either bronchial alveolar lavage fluid (BALF), cell fractionation or histopathological studies. One of respiratory function studies, detections of antibodies or analyses of cytokines are performed.

(iii) both only induction group and only elicitation group are set up as negative control.

(iv) clear positive control. is integrated.

or

(2) Positive animal test which fulfills all conditions as stated above is reported in only one research organization. Positive appropriate animal test which does not fulfill all conditions is also reported in other research organization.

*Other conditions than a set of the all stated above ((i)~(iv)) might also provide an evidence for a sensitizer to animals.

Skin sensitizers

Group 1

Case reports which show a clear association between skin symptoms and patch tests are reported in at least two different research organizations.

and

An epidemiological study which clearly indicate associations among occupational exposures, symptoms of contact dermatitis and patch tests is also reported in at least one research organization. Patch tests should be appropriately performed with controls.

Group 2

Case reports which show a clear association between skin symptoms and patch tests are reported in at least two different research organizations.

However, there isn't an epidemiological study.

Group 3

Positive appropriate animal tests; guinea pig maximization test (GPMT) and the Buehler guinea pig test (OECD Guideline 406) or Local Lymph Node Assay (LLNA) (OECD 429) are reported in at least one