OECD/OCDE

2021/06/14

試験ガイドライン No.497

皮膚感作性試験のDefined Approachに関するガイドライン

ご利用に関する注意事項

本文書および本文書に掲載している翻訳は OECD による公式文書**ではありません。**本文書における翻訳と原文の間に用語ないし表現等の齟齬がある場合、**原文のみが有効とみなされます**。内容について疑義が生じた場合、もしくは研究に公的な認証を必要とする場合は、原文に準拠してください。

本翻訳を使用したことに起因する如何なる損害について、当所は一切責任を負いません。

This document and translation are NOT official work by OECD. In the event of any discrepancy between the original work and the translation, **only the text of the original work should be considered valid**. Refer to the original document when there is any doubt about the translation, or legal or official verification is needed for your study. NIHS takes no responsibility even if you get any damage by using this translation.

Original work:

OECD (2021), Test No. 497: Defined Approaches for Skin Sensitisation, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/testing-of-chemicals/previous-test-guidelines/section-4/replaced-test-guidelines/replaced-test-guideline-497-2021.pdf

©OECD(2021)

You are free to use this material subject to the terms and conditions available at http://www.oecd.org/termsandconditions/

経済協力開発機構 （OECD）の化学物質の試験に関するガイドライン Section 4 Health effects
邦訳: 国立医薬品食品衛生研究所(NIHS) 安全性予測評価部

OCDE/OECD 497

採択
2021 年 6 月 14 日

経済協力開発機構（OECD）の化学物質の試験に関するガイドライン

皮膚感作性試験の Defined Approach

© OECD, (2021)
本資料は、http://www.oecd.org/termsandconditions/に掲載されている諸条件に従って自由に使用することができる。

目次

皮膚感作性試験の Defined Approach に関する OECD のガイドライン..... 1

1. セクション 1–はじめに..... 4

1.1. 概要..... 4
1.2. 本ガイドラインに含まれる DA および使用シナリオ..... 6
1.3. 限界..... 8
1.3.1. 個々の in chemico/in vitro 情報源の限界..... 9
1.3.2. In silico 情報源の限界..... 9
1.3.3. DA の限界..... 9
1.4. 参考文献..... 11

パート I – セクション 2 –皮膚感作性ハザードの特定のための Defined Approach 13

2.1. 「2 out of 3（2o3）」Defined Approach..... 13
2.1.1. 要約..... 13
2.1.2. データ解釈手順..... 13
2.1.3. 個々の情報源の内容と限界..... 14
2.1.4. 2o3 DA 予測の信頼性..... 14
2.1.5. LLNA と比較した 2o3 DA の予測能..... 16
2.1.6. ヒトデータと比較した 2o3 DA の予測能..... 17
2.1.7. ヒトデータと比較した LLNA の予測能..... 18
2.1.8. 習熟度確認化学物質..... 19
2.1.9. DA の報告..... 19
2.2. 参考文献..... 20

パート II – セクション 3–皮膚感作性の強度分類のための Defined Approach..... 21

3.1. 「統合的試験戦略（ITS）」の Defined Approach..... 21
3.1.1. 要約..... 21
3.1.2. データ解釈手順..... 21
3.1.3. 個々の情報源の内容と限界..... 23
3.1.4. ITS DA 予測の信頼性..... 24
3.1.5. LLNA と比較した ITSv1 DA の予測能..... 27
3.1.6. LLNA と比較した ITSv2 DA の予測能..... 28
3.1.7. ヒトデータと比較した ITSv1 DA の予測能..... 29
3.1.8. ヒトデータと比較した ITSv2 DA の予測能..... 31
3.1.9. ヒトデータと比較した LLNA の予測能..... 32
3.1.10. 習熟度確認化学物質..... 33
3.1.11. DA の報告..... 34
3.2. 参考文献..... 35

補遺 1：2o3 DA に用いるために複数回実験したそれぞれの *in chemico*/*in vitro* 試験の予測モデル 36

補遺 2：*In silico* 予測を導くための DASS ITS 予測およびプロトコルに対する適用領域の規定と信頼性評価..... 39

©OECD 2021

はじめに 39
個々の情報源の適用領域 39
In chemico/in vitro 情報源 (DPRA および h-CLAT) 39
In silico 情報源 39
Derek Nexus (ITSv1) 40
QSAR Toolbox (ITSv2) 40
ITS 予測の信頼性 41
ITS に対するデータ解釈手順（DIP）の適用法 41
参考文献 44
付録 1：Derek Nexus 予測のプロトコル 45
Derek KB 2020 1.0 と Derek Nexus v.6.1.0 を用いて皮膚感作性ハザードに対する予測を導くためのプロトコル 45
付録 2：OECD QSAR Toolbox 予測のプロトコル 48
Toolbox 4.5 の DASS AW を使用して皮膚感作性ハザードの予測を導くためのプロトコル。 48
付録 3:OECD QSAR Toolbox の適用領域に関する情報 49
技術的側面 49
Toolbox の in silico 領域の算出 49
適用領域層の計算 50
1. 物理化学的特性層 50
2. 化学構造層 50
3. 作用機序層 51

©OECD 2021

1. セクション 1-はじめに

1.1. 概要

1. 皮膚感作性物質とは、国連勧告「化学品の分類および表示に関する世界調和システム」（UN GHS）(1)に定義されている通り、皮膚との反復接触後にアレルギー反応を引き起こす物質のことをいう。皮膚感作性の原因となる主要な生物学的事象に関しては、広く意見が一致している。タンパク質との共有結合に始まる皮膚感作性に関して、その化学的、生物学的機序に関する現在の知識は、有害性発現経路（Adverse Outcome Pathway: AOP）(2)の形式で要約される。これは、分子レベルの初期事象に始まり、中間の重要事象を経て、アレルギー性接触皮膚炎である有害作用に至るものである。

2. 皮膚感作性 AOP では、有機化学物質など、アミノ酸残基（システイン残基またはリジン残基）と反応する化学物質に焦点が絞られている。この場合、分子レベル初期事象（1 つ目の重要事象）は、皮膚タンパク質の求核中心と求電子性物質との共有結合である。この AOP における 2 つ目の重要事象は、角化細胞に生じ、炎症性反応や特異的細胞シグナル伝達経路による遺伝子発現の変化からなる。このようなものに、抗酸化剤/求電子物質応答配列（antioxidant/electrophile response element: ARE）依存性経路がある。3 つ目の重要事象は樹状細胞活性化であり、通常は特異的な細胞表面マーカー、ケモカインおよびサイトカインの発現により評価される。4 つ目の重要事象は T 細胞の増殖である。この場合の有害な転帰はアレルギー性接触性皮膚炎である。

3. 皮膚感作性の評価には通常、実験動物が使用されてきた。モルモットを用いる古典的方法である Magnusson and Kligman のモルモットマキシマイゼーション法（GPMT）およびビューラー法（TG 406）(3)は、皮膚感作性の誘導相および誘発相の両者を評価するものである。マウスを用いる試験である LLNA (OECD TG 429) (4) と、その非放射性修正法 3 種類、具体的には LLNA:DA (OECD TG 442A) (5)、LLNA:BrdU-ELISA および BrdU-FCM (OECD TG 442B) (6)はいずれも、誘導相のみを評価するものであり、皮膚感作性の誘導相を客観的に測定するとともに動物福祉という点でモルモット試験より優れていることから、ここに挙げた試験が受け入れられている。

4. 皮膚感作性 AOP の最初の 3 つの重要事象（KE）を対象とする *in chemico* および *in vitro* の機序に基づく試験法(OECD TG 442C、442D、442E) (7, 8, 9)を、化学物質の皮膚感作性ハザードを評価するのに使用することができる。ここに挙げた試験方法はいずれも、化学物質の皮膚感作性について結論を出したり、UN GHS（細区分 1A および 1B）に従った強度の細区分に関する情報提供には、動物データを単独で十分に置き換えることはできないと考えられる。しかし、皮膚感作性 AOP の複数の KE を対象とする上記の *in chemico* および *in vitro* の方法で得られたデータを、*in silico* や類縁化合物の read-across での予測などの情報源とともに使用することや、試験および評価に関する統合的アプローチ（IATA）または Defined Approach（DA）と併用することが提案されている。個々の情報源から得られた結果は、物質が試験法の適用領域に該当する場合にのみ、DA で使用することができる（各試験法[TG 442C 付録 1、TG 442D 付録 1A、TG 442E 補遺 1]の「初期検討事項、適用可能性および限界[Initial Considerations, Applicability and Limitations]」の項を参照）(7, 8, 9)。

©OECD 2021

5. 複数の情報源から得られた結果をDAに併用するとよい。これにより、動物試験の予測能と同じかそれ以上の予測能を、ヒトにおける反応の予測に発揮することができる。DAは、規定のデータ解釈手順（DIP）（数学的モデル、ルールベースアプローチなど）からなり、一連の定められた情報源を用いて得たデータ（*in silico* 予測、*in chemico* データ、*in vitro* データなど）に適用して予測を導く。これは専門家による判断を必要としない。皮膚感作性の個々のDAおよびその情報源は当初、ガイダンス文書256の補遺I/II (10)に記載され、予備的評価がKleinstreuer et al. (11)によって公表されている。DAは、得られた全体的な結果の信頼性を高めるため、個々の単独の方法の限界の一部を克服することを目的とした一群の方法を使用する。DAの最終目標は、動物試験で得られた情報と同等の情報、具体的には、ハザードの特定や強度分類に使用できる情報を提供することである。

6. 試験機関には、DAの記載の通りに、試験実施前に被験化学物質に関する利用可能なすべての情報を考慮することが求められる。このような情報には、被験化学物質の特定名、化学構造、物理化学的性質などがある。このような情報を考慮して、特定のDAに基づく個々のOECD試験ガイドラインの方法が被験化学物質に適用できるかどうかを判断する。

7. *In vivo*（LLNAなど）試験、*in chemico* 試験、*in vitro* 試験、*in silico* 法、DAやこのようなものの併用に基づくハザード評価や強さの細区分を実施する際には、常に同じ原則が適用される。具体的には、当該化学物質に関して利用可能なすべての情報のほか、類縁物質に関する毒性学的データがあれば考慮する必要がある。

8. 本ガイドラインは、規制当局、バリデーション機関、非政府組織および業界の科学専門家で構成される皮膚感作性のDefined Approachに関するOECD専門家グループ（EG DASS）の意見を取り入れて策定された。

9. 本ガイドラインに3つのルールベースのDAを記載し、その規制上の目的に照らして説明する。この目的とは、皮膚感作性物質と非感作性物質との識別（1.4.パートI）を指すハザードの特定や、強さの細区分（1.6.パートIIパートII）である。パートIIに記載したDAは、ハザードの特定にも適している。DAの評価およびレビューについては、*皮膚感作性Defined Approach（DA）に関するガイドライン（GL）の付属文書*に詳しく記載する(12)。

10. DA予測、個々の情報源に関するデータ、高度にキュレーションされたLLNAデータおよびヒトパッチ予測試験（HPPT）データ、物理化学的特性のある196種類の化学物質の包括的データセットをまとめ、*皮膚感作性Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**補遺 2**として添付する(12)。196種類の化学物質のうち、168種類の化学物質にLLNA分類、66種類の化学物質にHPPT分類がある。これはいずれもEG DASSの同意を得ており、DAの性能を評価するのに使用した。データの利用可能性に照らして、このデータセットは主に化粧品成分からなるが、防腐剤、染料、食品添加物など、分野を超えて使用されるその他の種類の化学物質も含む。上記化学物質のもつ物理化学的特性が示す通り、このデータセットは化学的に多様である。低分子もあれば高分子もあり（分子量 30～512 g/mol）、疎水性物質もあれば親水性物質もあり（Log P -3.9～9.4）、固体もあれば液体もあり（融点-122～253°C）、揮発性物質もあれば非揮発性物質もある（沸点-19～445°C）。参照データベースの化学空間特性評価に関する詳細を、*皮膚感作性Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第4項**に記載する(12)。

©OECD 2021

11. 別の DA を、今後の審査や承認のあとに本ガイドラインに組み込む可能性がある。リスクアセスメントに使用されることがある point of departure など、感作性の強さの定量的尺度となる DA を今後、このガイドラインの新たなパート II に含める可能性がある。

1.2. 本ガイドラインに含まれる DA および使用シナリオ

12. 現在、本ガイドラインに記載されている DA は、以下の通りである:

- *In chemico*（KE1）データおよび *in vitro*（KE2/KE3）データに基づく皮膚感作性ハザードの特定のための「2 out of 3（2o3）」によるアプローチ(13, 14)。パート I 参照。
- 専門家グループ（EG DASS）の知見を用いて開発した DIP による *in chemico*（KE1）データおよび *in vitro*（KE3）データに基づく UN GHS 強度分類並びに *in silico*（Derek Nexus）予測(14, 15)の統合的試験戦略（ITSv1）。パート II の強度分類（Potency Categorisation）を参照。
- 専門家グループ（EG DASS）の知見を組み込んで開発した DIP による *in chemico*（KE1）データおよび *in vitro*（KE3）データと *in silico*（OECD QSAR Toolbox）予測に基づく UN GHS 強度分類の修正統合的試験戦略（ITSv2）。パート II の強度分類（Potency Categorisation）を参照。

13. 本ガイドラインに記載した DA は、バリデーション済みの OECD 試験法（DPRA、KeratinoSens™、h-CLAT）に基づいている。このような試験法は、バリデーションフェーズで技術移転性、実験室内再現性および実験室間再現性が明らかにされている(7, 8, 9)。

14. ITS DA（ITSv1 および ITSv2）は、*in silico* 情報源も利用する。このようなものに、Derek Nexus v 6.1.0（ITSv1）や OECD QSAR Toolbox v 4.5（ITSv2）がある。Derek Nexus（以下 Derek と称する）は、構造アラートを用いて皮膚感作性を予測する専門知識ベースのツールである。一方、OECD QSAR Toolbox（以下 OECD QSAR TB と称する）は、類縁化合物ベースの read-across アプローチか、化学物質が感作性物質であるかどうかを予測するプロファイラーが特定したタンパク質結合の構造アラートを使用するコンピュータツールである。

15. 本ガイドラインに記載されている DA はいずれも、感作性物質（UN GHS 区分 1）を非感作性物質から区別するための国固有の要件に対応するために用いられる。ただし、感度および特異度は異なる（各 DA のそれぞれの説明のなかで詳述する）。

16. また、ITS DA（ITSv1 および ITSv2）を用いて、化学物質を 3 つの UN GHS 強度区分（区分 1A=強い感作性物質、区分 1B=その他の感作性物質、区分に該当しない[NC=区分外]）に分類することができる

17. 個々の情報源の既知の限界および適用領域を用いて、本ガイドラインに記載された DA が生成する予測のそれぞれに信頼性を割り当てるワークフローを設計した。信頼性の高い予測を行うためには、基礎データが各試験ガイドラインの基準を満たさなければならない（TG 442C 付録 1 を参照、TG 442D 付録 1A 参照、TG 442E 補遺 1 参照(7, 8, 9)）。ハザードの特定や強度に関する信頼性が高い DA 予測を、確定的なものと考える。信頼性の低い DA 予測を、ハザードの特定や強度について確定的でないと考える（詳細については **2.1.4 項**および **3.1.4 項を参照**）。それでも、この「確定的

©OECD 2021

でない」との予測を、エビデンスの重み付けアプローチや IATA とのからみで、他の情報源（例:試験系への曝露の説明、既存の *in vivo* データ、臨床データ、read-across、その他の *in vitro/in chemico/in silico* データなど）と共に考慮するとよい。

18. 本ガイドラインに記載されている DA について、感作性物質と非感作性物質とを区別する性能を、168 種類の化学物質（GHS 皮膚感作性区分 1 が 135 種類、区分外が 33 種類）を用いて評価した。これについては、DPRA、KeratinoSens™、h-CLAT、Derek、OECD QSAR TB 予測のほか、EG DASS の同意を得ている LLNA 参照データに基づく分類を利用する（このほかの詳細は、*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 2.1 項**および**補遺 3** を参照のこと）(12)。強度分類（細区分 1A、1B、または「区分外」[NC]）に基づく UN GHS 分類を予測する ITS DA の性能を評価する目的で、156 種類の化学物質（38 種類が 1A、85 種類が 1B、33 種類が NC）を使用した。この理由は、12 種類の化学物質について、LLNA データに基づいて十分な信頼性で細区分 1A または 1B の強度に割当てできなかったことによる。構造組成が不明な混合物および植物エキスを、キュレーションされた LLNA 参照データから除外した。

19. 皮膚感作性ハザードの予測に関する LLNA 参照データと比較した 3 種類の DA（信頼性の高い予測のみ）の性能は、バランス精度（感度および特異度の平均[BA]）が 80～84%、感度 82～93%、特異度 67～85%であった（**表 1.1** 参照）。なお、陰性の参照化学物質の数が少なかったことから、特異度の指標は感度の指標よりも不確定である。詳細なパフォーマンス統計量をパート I（2o3 DA）およびパート II（ITS DA）に報告する。強度分類に基づく UN GHS 分類（信頼性の高い予測のみ、細区分 1A、1B または NC）に関する ITSv1 DA および ITSv2 DA の性能を、LLNA 参照データと比較したところ、全体の精度はともに 71%、全体のバランス精度は 78%（ITSv1）または 77%（ITSv2）であり、細区分ないし NC の予測のバランス精度は 72～81%（ITSv1）または71～80%（ITSv2）であった。強い感作物質（1A）であるのに非感作性（NC）と誤って予測されたものも、その逆もなかった。詳細なパフォーマンス統計量を、パート II および*皮膚感作 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**に記載する(12)。

20. このほか、このガイドラインに記載されている DA の、感作性物質と非感作性物質とを区別する性能を、2o3 について 66 種類の化学物質（感作性物質 55 種類、非感作性物質 11 種類）または化学物質 1 種類の試験データがないことから 65 種類の化学物質を用いて評価した。ここに挙げた化学物質のヒト予測パッチテスト（HPPT）データに基づく分類は、EG DASS の同意を得ている（詳細については、*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 2.2 項**および**補遺 4** を参照のこと）(12)。強度分類（細区分 1A、1B または NC）に基づく UN GHS 分類を予測する ITS DA の性能を評価するのに、63 種類の化学物質を使用した（1A が 21 種類、1B が 31 種類、NC が 11 種類）。これは、3 種類の化学物質は、ヒト参照データに基づいて強度細区分 1A や 1B に十分な信頼性で割当てできなかったことによる。構造組成が不明な混合物および植物エキスを、キュレーションされたヒト参照データから除外した。

21. 皮膚感作性ハザードの予測に関するヒト参照データに対する DA の性能（信頼性の高い予測のみ）は、バランス精度が 69～88%、感度 89～94%、特異度 44～88%であった（**表 1.1** 参照）。なお、陰性の参照化学物質の数が少なかったことから、特異度の指標は感度の指標よりも不確定である。詳細なパフォーマンス統計量をパート I（2o3 DA）およびパート II（ITS DA）に報告する。UN GHS 皮膚感作性強度分類（信

©OECD 2021

頼性の高い予測のみ、細区分 1A、1B または NC）に関する ITSv1 DA および ITSv2 DA の性能を、ヒト参照データと比較したところ、全体のバランス精度は 72% (ITSv1)または 73% (ITSv2)、細区分ないし NC の予測のバランス精度は 68～79% (ITSv1)または 69～79% (ITSv2)であった。詳細なパフォーマンス統計量を、パート II および*皮膚感作 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**に記載する(12)。

22. LLNA とヒト参照データセットとの間で、ハザードに関して 56 種類の化学物質、皮膚感作性の強度に関して 47 種類の化学物質が一致していた。また、ヒト参照データに対する LLNA の性能を、上記化学物質を比較の根拠に用いて評価した。ヒト参照データに対する LLNA の皮膚感作性ハザード予測性能は、バランス精度が 58%、感度 94%、特異度 22%であった。なお、陰性の参照化学物質の数が少なかったことから、特異度の指標は感度の指標よりも不確定である。UN GHS 強度分類に関する LLNA の性能を、ヒト参照データと比較したところ、全体のバランス精度は 64%であり、細区分ないし NC の予測のバランス精度は 59～73%であった。ヒト参照データのなかに、強い感作物質（1A）であるのに、非感作性（区分外）と誤って DA や LLNA によって予測されたものも、その逆もなかった。詳細なパフォーマンス統計量をパート I およびパート II に報告する。

表 1.1.本ガイドラインに掲載されている DA の概要

DA/方法	情報源	能力（ハザードないし強度）	LLNA と比較したハザードに関する性能	ヒトと比較したハザードに関する性能	LLNA と比較した強度に関する性能 (正確性)	ヒトと比較した強度に関する性能 (正確性)
2o3 DA	DPRA、KeratinoSens™、h- CLAT	ハザード	84% BA、82%感度、85%特異度	88% BA、89%感度、88%特異度	-	-
ITSv1 DA	DPRA、h-CLAT、DEREK Nexus v6.1.0	ハザード、強度	81% BA、92%感度、70%特異度	69% BA、93%感度、44%特異度	70% NC、71% 1B、74% 1A	44% NC、77% 1B、65% 1A
ITSv2 DA	DPRA、h-CLAT、OECD QSAR Toolbox v4.5	ハザード、強度	80% BA、93%感度、67%特異度	69% BA、94%感度、44%特異度	67% NC、72% 1B、72% 1A	44% NC、80% 1B、67% 1A
LLNA (比較のために記載)	*in vivo*	ハザード、強度	-	58% BA、94%感度、22%特異度	-	25% NC、74% 1B、56% 1A

注：ハザード性能に関して、感度は真陽性率、特異度は真陰性率であり、バランス精度（BA）は感度と特異度の平均値である。参照データが不均衡であったことから、特異度の指標は感度の指標よりも不確定である。強度性能に関して、正確性は各 UN GHS 細区分内の正しい分類率を示す。参照データが不均衡であったことから、正確性の指標は、NC 化学物質のような小さいクラスのほうが不確定である。統計量は、確定的な DA 予測のみを示す。これは、最初のガイドライン採択時に入手しうるデータであることを指す。

1.3. 限界

23. **表 1.1** に、本ガイドラインに掲載されている DA の概要、使用した情報源、ハ

©OECD 2021

ザードないし強度の予測の有無を示すほか、LLNA およびヒト参照データに基づく性能の要約を示す。LLNA（OECD TG 429）を比較のベースとして**表 1.1** に示す。詳細を、本ガイドラインのパート I およびパート II のほか、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(12)。

24. DA にわかっている限界とその構成要素を以下にまとめる。

1.3.1. 個々の in chemico/in vitro 情報源の限界

25. 使用者は、それぞれの試験ガイドラインに規定されている通りに、個々の *in chemico/in vitro* 試験法の限界に言及する必要がある。このようなものは、新たなデータが入手可能になるたびに改訂されるため、定期的に参照する必要がある。必ず、各試験ガイドライン（TG）の最新版を使用すること。例えば、金属、無機化合物、UVCB、混合物などの化学物質は、特定の試験方法の適用領域にない場合がある。個々の試験結果が境界域にある場合（**補遺 1**）、確定的でないとの DA 予測になる可能性がある。それぞれの DA における個々の *in chemico/in vitro* 試験法の限界に関する考察を、**第 2.1.4 項（図 2.1）**および**第 3.1.4 項（図 3.1）**に記載する。

1.3.2. In silico 情報源の限界

26. DA には、情報源として *in silico* ツールを組み込んだものがある。このようなツールは、自動で read-across 予測または(Q)SAR 予測を実行することができる。(Q)SAR は、構造活性相関（SAR）モデル（構造アラート、エキスパートシステム）と、定量的構造活性相関（QSAR）モデル（統計ツール）の両方を含む。(Q)SAR モデルは、OECD 原則（OECD Principles for the Validation, for Regulatory Purposes, of (Q)SAR Models）を満たし、QSAR モデルに関わる様式（QMRF）(15)[1]に記載する必要がある。OECD QSAR バリデーションの原則の 1 つに、定義された適用領域がある。定義された適用領域は、これをはずれると信頼性の低い予測が得られる可能性があるという限界点を示すものである（モデルに含まれる記述子のトレーニングセットの範囲や、トレーニングセットに含まれる化学構造の種類）。どの *in silico* モデルも、複数の定義された適用領域があり、この適用領域はそれぞれ、バリデーションで確立された信頼性尺度と関連している。DIP に応じて、適用領域外の化学物質は、確定的でないと判断される信頼性の低い DA 予測になることがある。皮膚感作性の DA に *in silico* ツールが含まれる場合、使用者は個々の *in silico* ツールの限界と適用領域について言及する必要がある。本ガイドラインで採り上げる DA のうちの 2 種類、具体的には ITSv1 および ITSv2 はそれぞれ、*in silico* ツールである Derek および OECD QSAR TB に依存しており、その限界と適用領域の詳細を本ガイドラインの**補遺 2** に示す。

1.3.3. DA の限界

27. DA の限界は、*in chemico/in vitro/in silico* 情報源にある個々の限界に基づいている。DA 予測の信頼性を判断する目的で個々の情報源の限界を使用する場合の詳細を、**第 2.1.4 項**および**第 3.1.4 項**とそれぞれの試験ガイドライン（TG 442C 付録 1、TG 442D

[1] QMRF は、DASS に関して他の *in silico* モデル予測の報告のために、若干修正されている。修正 QPRF が、健康への影響に関する OECD 試験ガイドラインに関連するスプレッドシートおよびソフトウェアの OECD サイトに掲載されている https://www.oecd.org/env/ehs/testing/section4software.htm.。

©OECD 2021

付録 1A、TG 442E 補遺 1）(7, 8, 9)に示す。

28. 本ガイドラインの対象となる DA の評価にあたり、LLNA データに関して、DPRA（TG 442C）、KeratinoSens™（TG 442D）、h-CLAT（TG 422E）や提案されている DA は、Log P > 3.5 の被験化学物質の感度が低かった（詳細については、*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 3.1.4 項**および**補遺 5** を参照）(12)。また、LLNA 試験はヒト参照データより、このような化学物質に偽陽性が多く、裏付けとなる機構的情報が示されている（詳細については、*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 3.2 項**および**補遺 6** を参照）(12)。全体として、Log P > 3.5 である参照化学物質の数や分析結果は、確実な結論を導き出すには不十分である。しかし、TG 442E によれば、Log P > 3.5 である物質の h-CLAT 結果が陰性であると考えるべきではなく、**第 2.1.4 項**および**第 3.1.4 項**に記載されている通り、この限界を DA に適用する。

29. 2o3 DA では、信頼性の低い領域を規定するために、DA の 3 つの KE に対する個々のアッセイに境界域（BR）を規定している（詳細については、本ガイドラインの第 **2.1.4 項**および**補遺 1**、*皮膚感作性に関する Defined Approach[DA]のガイドライン[GL]の付属文書*の**第 3.3 項**および**補遺 7** を参照）(12)。このような BR や個々のアッセイの限界（Log P > 3.5 の化学物質の h-CLAT 結果陰性など[TG 442E による]）にあてはまる陽性ないし陰性の試験結果は、信頼性が低く、確定的でないとの 2o3 DA 予測になる可能性がある。

30. それでも、確定的でないとの DA 予測を、エビデンスの重み付けアプローチや IATA とのからみで、他の情報源（例:試験系への曝露の説明、既存の *in vivo* データ、臨床データ、read-across、その他の *in vitro/in chemico/in silico* データなど）と共に考慮するとよい。

©OECD 2021

1.4. 参考文献

1. United Nations (UN) (2019). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Eighth revised edition, New York and Geneva, United Nations Publications. Available at: [https://unece.org/ghs-rev8-2019]
2. OECD (2012). Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm]
3. OECD (1992). OECD Guidelines for the Testing of Chemicals No. 406. Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
4. OECD (2010). OECD Guidelines for Chemical Testing No. 429. Skin sensitisation: Local Lymph Node assay. Organisation for Economic Cooperation and Development, Paris. Available at :(oecd-ilibrary.org).
5. OECD (2010). OECD Guidelines for Chemical Testing No. 442A.Skin sensitisation: Local Lymph Node assay: DA. Organisation for Economic Cooperation and Development, Paris. Available at:(oecd-ilibrary.org)..
6. OECD (2018). OECD Guidelines for Chemical Testing No. 442B. Skin sensitisation: Local Lymph Node assay: BrdU-ELISA or –FCM. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
7. OECD (2020). OECD Guideline for the Testing of Chemicals No. 442C: *In chemico* Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key even on covalent binding to proteins).*In chemico*. Paris, France: Organisation for Economic Cooperation and Development. Available at: (oecd-ilibrary.org).
8. OECD (2018), OECD Key Event based test Guideline 442D: *In vitro* Skin Sensitisation Assays Addressing AOP Key Event on Keratinocyte Activation. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
9. OECD (2018). OECD Key event-based test Guideline 442E: *In vitro* Skin Sensitisation Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
10. OECD (2016). Series on Testing & Assessment No. 256: 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, Annex 1 and Annex 2. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
11. Kleinstreuer N, Hoffmann S, Alepee N, et al. (2018). Non-Animal Methods to Predict Skin Sensitization (II): an assessment of defined approaches. Crit Rev Toxicol Feb 23:1-16. doi: 10.1080/10408444.2018.1429386

©OECD 2021

12. OECD (2021). Series on Testing and Assessment No. 336: Supporting document to the Guideline (GL) on Defined Approaches (DAs) for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].

13. Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R. (2012). Putting the parts together: combining *in vitro* methods to test for skin sensitizing potentials. Regul Toxicol Pharmacol, 63:489-504.

14. Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods, Regul Toxicol Pharmacol, 71:337-51.

15. ECHA (2008). see "CHAPTER R.6 – QSARS AND GROUPING OF CHEMICALS" *in* Guidance on Information Requirements and Chemical Safety Assessment. European Chemicals Agency [Guidance on Information Requirements and Chemical Safety Assessment - ECHA (europa.eu)].

©OECD 2021

パート I – セクション 2–皮膚感作性ハザードの特定のための Defined Approach

31. 本ガイドラインのパート I は、ハザードの特定、具体的には、感作性物質と非感作性物質とを区別することのみを目的とする DA に適用される。ハザード特定のための DA の概要を以下に示す。そのほかの詳細情報を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

2.1. 「2 out of 3（2o3）」Defined Approach

2.1.1. 要約

32. 2 out of 3（2o3）DA は、動物試験を使用せずに、化学物質の皮膚感作性ハザードを特定（UN GHS 区分 1 対 UN GHS NC）するためのものである。現時点でデータ解釈手順（DIP）は、感作性物質の強度に関する情報をもたらすようにはできていない。

33. 2o3 DA に組み込まれている一群の試験法は、OECD が正式に記載している皮膚感作性をもたらす AOP の最初の 3 つの KE のうち 2 つ以上をカバーする。ここでいう 3 つの KE は次の通りである。KE1: タンパク質との結合（ペプチド結合性試験[DPRA; OECD TG 442C]）(2)、KE2: 角化細胞活性化（KeratinoSens™; OECD TG 442D）(3)、KE3: 樹状細胞活性化（ヒト細胞株活性化試験[h-CLAT; OECD TG 442E]）(4)。

34. DIP は、AOP の最初の 3 つの KE のうち 2 つ以上を検討する方法で、ふたつに同じ結果を認めれば最終分類が決まる。2o3 DA を、EG DASS が合意したキュレーションされた LLNA 参照データのある化学物質 168 種類について比較したところ、正確性 83%、バランス精度 84%であった（**表 2.1** 参照）。さらに、2o3 DA を、EG DASS が合意したキュレーションされたヒト参照データのある化学物質 65 種類について比較したところ、ハザード特定に関して LLNA の正確性およびバランス精度を上回った（**表 2.1～2.2** 参照）。なお、参照データが不均衡であったことから（陽性が陰性より多い）、バランス精度の指標は、ヒトデータとの比較では特に不確定である。

2.1.2. データ解釈手順

35. 2o3 DA のデータ解釈手順（DIP）は、専門家の判断を必要としない、透明性のあるルールベースアプローチである(4, 6, 7)。このアプローチでは、KE1～3（DPRA、KeratinoSens™、h-CLAT）にマッピングされた国際的に認められた動物不使用アッセイの最大 3 つを順不同で段階的に試験を実施することによって、皮膚感作性ハザードを予測する。2 つの KE について試験を実施し、この 2 つの試験結果が同じであれば、化学物質は試験結果の通りに、感作性物質または非感作性物質と予想される。最初の 2 つの試験の結果が同じでなければ、3 つ目の KE の試験を実施する。全体での結果は、2 つに一致して認めた結果となる。ここで、**第 2.1.4 項**に記載した通り、得られた予測に対する信頼性を考慮する。

36. 2o3 DA の性能は、各試験法の境界域を考慮に入れることによって影響を受けることがわかっている。これについては、以下の**第 2.1.4 項**、詳しくは、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 3.3 項**および**補**

©OECD 2021

遺 7 を参照のこと(1)。2o3 DA データ解釈手順の変更なしで、2o3 DA の予測を導くためのデシジョンツリーを**第 2.1.4 項**の**図 2.1** に示す。

2.1.3. 個々の情報源の内容と限界

37. DA の個々の情報源は、OECD の KE に基づく皮膚感作性試験ガイドラインに含まれる試験法（OECD TG 442C、442D、442E）であり（2, 3, 4）、そのなかにプロトコルが詳しく記載されている。

38. このような試験ガイドライン（TG）の試験を以下に記載する。2o3 DA はここに挙げたものからなる。

- ペプチド結合性試験(DPRA; OECD TG 442C; KE1) (2):皮膚感作性物質は一般に求電子物質であり、タンパク質の求核部分と反応する。DPRA は、システイン残基またはリジン残基を含む 2 種類のペプチドの共有結合による減少率を測定する。被験化学物質によるシステイン含有ペプチドおよびリジン含有ペプチドの平均減少率が 6.38%を超える場合（共溶出がある場合はシステインのみの減少率が 13.89%超）に陽性と判断する。ペプチド減少率の結果が境界域にある場合、OECD TG 442C および**補遺 1** に規定されているように、追加の試験を実施する必要がある。

- KeratinoSens™アッセイ(*in vitro* 皮膚感作性試験:ARE-Nrf2 ルシフェラーゼ試験法; OECD TG 442D; KE2) (3); この角化細胞は、Nrf2-Keap1 経路を介して感作性物質と反応するレポーター遺伝子を有する。溶媒対照と比較して、ルシフェラーゼ誘導が 1.5 倍を超え、生存率が 70%を超える被験化学物質を陽性と判断する。ルシフェラーゼ誘導について境界域の結果が得られた場合、**補遺 1** に記載されている追加の検査を実施する。

- ヒト細胞株活性化試験(h-CLAT; OECD TG 442E; KE3) (4):抗原提示細胞の活性化は、CD86 ないし CD54 のアップレギュレーションを特徴とする。1.5 倍を超える CD86 発現誘導ないし 2 倍を超える CD54 発現誘導を認め、生存率が溶媒対照の 50%超である場合に、h-CLAT を陽性と判断する。CD54 ないし CD86 の発現誘導について境界域の結果が得られた場合、**補遺 1** に記載されている追加の検査を実施する。

39. 個々の *in chemico* および *in vitro* 試験法にある現時点での限界（溶解度に関する限界など）が、それぞれの試験ガイドライン(TG 442C, 付録 1; TG 442D 付録 1A; TG 442E 補遺 1)およびそのなかに記載されているバリデーション試験(2, 3, 4)に記載されている。

2.1.4. 2o3 DA 予測の信頼性

40. 各情報要素を用いうるかどうかの最初の判断は、各試験ガイドライン（TG 442C, 付録 1; TG 442D 付録 1A; TG 442E 補遺 1） (2, 3, 4)に記載した通り、*in chemico* および *in vitro* 法の限界（溶解度の理由により個々の方法に確定的な結果が得られない物質など）によって決まる。また、検査結果は変動の影響を受けやすく、このような変動は、特に（分類）カットオフ値に近い場合、言い換えると境界域にある場合に、検査結果の不確定性が高くなる。DA の結果の信頼性が低い領域を規定するため、2o3 DA の 3 つの KE に関する個々のアッセイの結果について境界域（BR）が定義されて

©OECD 2021

いる（本文書の**補遺 1** のほか、*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 3.3 項**および補遺 7 を参照のこと）(1)。各分析法のバリデーション試験データに基づく具体的な境界域は以下の通りである。

- DPRA BR:平均ペプチド減少率:4.95%～8.32%、システインのみの減少率（リジンペプチドが共溶出する場合）:10.56%～18.47%;
- KeratinoSens™ BR:Imax:1.35 倍～1.67 倍;
- h-CLAT BR:RFI CD54:157%～255%; RFI CD86:122%～184%。

41. 個別情報源のそれぞれについて境界域（BR）を組入れた予測モデル（PM）を、本ガイドラインの**補遺 1** に記載する。

42. 参照データベースに報告されているように、実験が 1 回のデータについては、上述のようにペプチドの平均減少率またはシステインのみの減少率に関する境界域に基づいて、DPRA の境界例を特定する。反復実験を実施した場合は、**補遺 1、図 1.1** の PM を適用するものとする。

43. KeratinoSens™ アッセイの予測モデルは複数の実験を要する。反復実験で、KeratinoSens™に陽性、陰性または境界域の最終アウトカムのうちいずれを生じるかの評価に関しては、**補遺 1、図 1.2** の PM を適用するものとする（TG 442D に記載されている PM から採り、2o3 DA に使用して境界例の結論を得る）。この予測モデルは、TG 442D に記載されている予測モデルと同じ決定カットオフ値に基づいて、2o3 DA で使用する第 3 のアウトカム（境界域）を導く。したがって、元の予測モデルで陰性である場合には陰性か境界域にしかならず、元の予測モデルで陽性である場合には陽性か境界域にしかならない。

44. h-CLAT の予測モデルには複数回の実験が必要である。反復実験で、h-CLAT に陽性、陰性または境界域の最終アウトカムのうちいずれを生じるかの評価に関しては、**補遺 1、図 1.3** の PM を適用するものとする（TG 442E に記載されている PM から採り、2o3 DA に使用して境界例の結論を得る）。この予測モデルは、TG 442E に記載されている予測モデルと同じ決定カットオフ値に基づいて、2o3 DA で使用する第 3 のアウトカム（境界域）を導く。したがって、元の予測モデルで陰性である場合には陰性か境界域にしかならず、元の予測モデルで陽性である場合には陽性か境界域にしかならない。

45. このような BR にあてはまる陽性および陰性の試験結果のほか、*in chemico/in vitro* 試験ガイドラインの限界のために確定的でないとの結果は、信頼性が低い。例えば、Log P > 3.5 の化学物質に h-CLAT の陰性結果を得た場合（TG 442E (4)による）の信頼性は低く、以下に述べるように 2o3 DA の結果に影響を及ぼす。

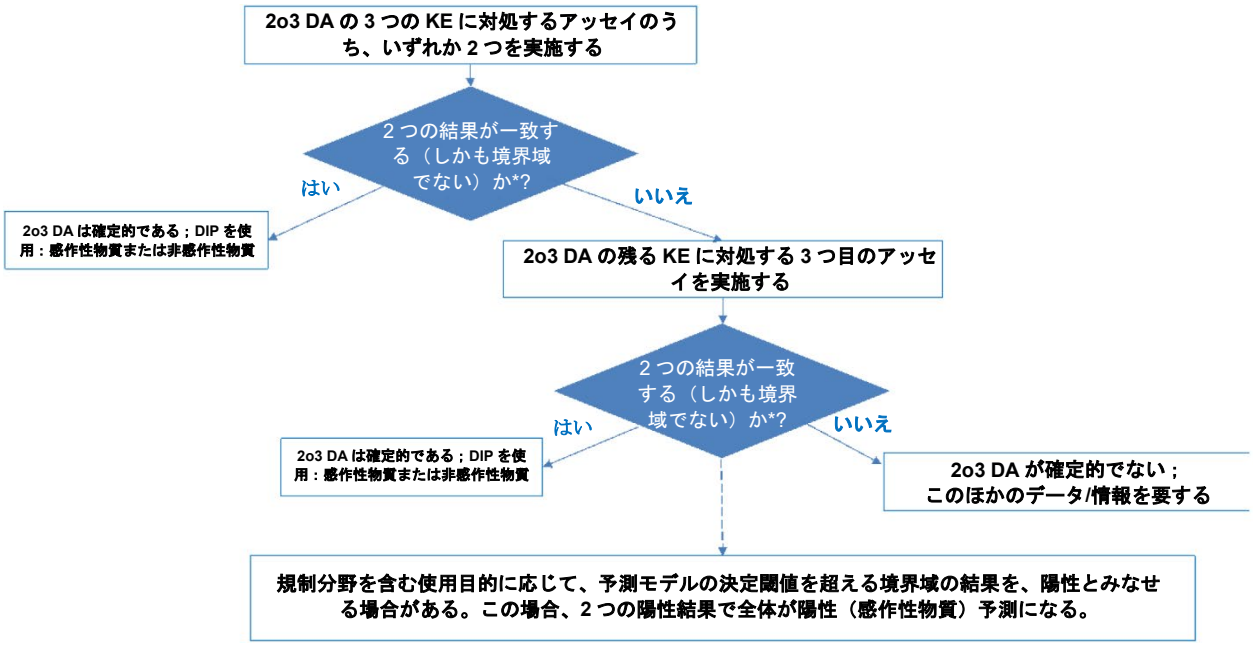
- 2o3 DA 試験法の 1 つの結果がその試験法の BR に該当する場合でも、2o3 DA を構成する他の 2 つの試験法の結果が一致し、信頼性が高い場合（結果がそれぞれの BR にあてはまらない）には、2o3 DA 予測を行うことができる。
- 同様に、Log P > 3.5 の化学物質に h-CLAT 陰性の結果を得た場合でも、2o3 DA を構成する他の 2 つの試験方法の結果が一致し、信頼性が高い場合（結果がそれぞれの BR にあてはまらない）には、2o3 DA 予測を行うことができる。
- しかし、2o3 DA 試験方法の 1 つの結果がその試験方法の BR に該当するか、Log P > 3.5 の化学物質に h-CLAT の陰性結果を得ている場合で、しかも、2o3 を構成する他の 2 つの方法が一致した信頼性の高い結果を示さない場合、2o3

©OECD 2021

DA 予測は「確定的でない」となる。それでも、このような「確定的でない」との予測を、エビデンスの重み付けアプローチやIATAとのからみで、ほかの情報源と重要性に基づくアプローチおよび/または IATA と他の情報源と共に考慮するとよい。規制分野を含む使用目的に応じて、予測モデルの決定閾値を超える境界域の結果を、なおも陽性とみなせる場合がある。この場合、2 つの陽性結果で全体が陽性（感作性物質）予測になる。

46. 境界域判断と、2o3 DA 予測の信頼性に対する境界域判断の影響を、**図 2.1** に示す。ハザード特定に関する信頼性の高い DA 予測を、確定的なものと判断する。信頼性の低い DA 予測を、ハザードの特定に関して確定的でないと考える。それでも、この「確定的でない」との予測を、エビデンスの重み付けアプローチや IATA とのからみで、他の情報源と共に考慮するとよい。

図 2.1.境界域の結果を考慮に入れた、2o3 DA のためのデシジョンツリー



注：境界域の結果は、**補遺 1** に示すワークフローに基づいて判断する。
*情報要素の使用は、各試験ガイドライン(TG 442C 付録 1; TG 442D 付録 1A; TG 442E 補遺 1)に記載されている限界によって決まる。例えば、Log P > 3.5 の化学物質に h-CLAT の陰性結果を得た場合（TG 442E (4)に記載された限界に基づく）、2o3 DA を構成する他の 2 つの試験法の結果が一致し、しかも境界域でない場合にのみ、2o3 DA 予測を行うことができる。

2.1.5. LLNA と比較した 2o3 DA の予測能

47. 2o3 DA の予測能が、LLNA を用いて得た（**表 2.1** 参照）、EG DASS が合意した通りにキュレーションされたデータに基づいて報告されている（*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の第 2.1 項および**補遺 3** を参照のこと）。上述の通りに、境界域解析を適用し、2o3 DA 予測に信頼性を割り当てた。確定的（信頼性高）予測に対し、LLNA 参照データと比較したパフォーマンス統計量を報告する。このほかに、確定的でない（信頼性低）結果を記載する。特定の

©OECD 2021

化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

表 2.1.LLNA 参照データと比較した 2o3 DA のハザード特定性能

	LLNA	
2o3 DA	非感作性物質	感作性物質
非感作性物質	22	19
感作性物質	4	89
確定的でない	7	27

LLNA データと比較した DA 性能 (N=134)	**2o3**
正確性（%）	83%
感度（%）	82%
特異度（%）	85%
バランス精度 (%)	84%

注：正確性は正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。報告されている性能は、DPRA、KeratinoSens™および h-CLAT に基づく。統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する。

48. 2 o 3 DA 予測に対する BR 解析の適用および高信頼性／低信頼性の指定を、上記の**第 2.1.4 項**および**補遺 1** に記載した通りに実施する。詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 3.3 項**および**補遺 7** に記載する(1)。

49. 参照データが不均衡であったことから、特異度の指標（LLNA 陰性化学物質 26 種類に基づく）は感度の指標（LLNA 陽性化学物質 108 種類）よりも不確定である。

2.1.6. ヒトデータと比較した 2o3 DA の予測能

50. このほか、2o3 DA の予測能が、EG DASS が合意した通りにキュレーションされた、ヒトパッチテスト(HPPT)のデータ（**表 2.2** 参照）に基づいて報告されている（*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 2.2 項**および**補遺 4** を参照のこと）(1)。上述の通りに、境界域解析を適用し、2o3 DA 予測に信頼性を割り当てた。確定的（信頼性高）予測に対し、ヒト参照データと比較したパフォーマンス統計量を報告する。また、確定的でない（信頼性低）結果を記載する。特定の化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

表 2.2 ヒト参照データと比較した 2o3 DA のハザード特定性能

	ヒト	
2o3 DA	非感作性物質	感作性物質

©OECD 2021

非感作性物質	7	5
感作性物質	1	42
確定的でない	3	7

ヒトデータと比較した DA 性能 (N=55)	**2o3**
正確性（%）	89%
感度（%）	89%
特異度（%）	88%
バランス精度 (%)	88%

注：正確性は HPPT データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。報告されている性能は、DPRA、KeratinoSens™および h-CLAT に基づく。統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

51. 2o3 DA 予測に対する BR 解析の適用および高信頼性／低信頼性の指定を、上記の**第 2.1.4 項**および**補遺 1** に記載した通りに実施する。詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 3.3 項**および**補遺 7** に記載する(1)。

52. 参照データが不均衡であったことから、特異度の指標（ヒト陰性化学物質 8 種類に基づく）は感度の指標（ヒト陽性化学物質 47 種類）よりも不確定である。

2.1.7. ヒトデータと比較した LLNA の予測能

53. 上に記載した DA のパフォーマンス統計量に関する比較の根拠を得るため、EG DASS が合意した通りにキュレーションされたヒトパッチテストのデータに基づいて、LLNA の予測能を報告する（**表 2.3** 参照）。特定の化学物質に関するデータおよび詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

表 2.3.ヒト参照データと比較した LLNA のハザード特定性能

	ヒト	
LLNA	非感作性物質	感作性物質
非感作性物質	2	3
感作性物質	7	44

ヒトデータと比較した LLNA 性能 (N=56)	**LLNA**
正確性（%）	82%
感度（%）	94%
特異度（%）	22%
バランス精度 (%)	58%

注:正確性はヒト HPPT データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

©OECD 2021

54. ヒト HPPT データと比較した確定的 2o3 DA 予測のハザード特定性能は、正確性 89%、感度 89%、特異度 88%、バランス精度 88%であり、どの尺度も、ヒト HPPT データと比較した LLNA の性能と同じかこれより優れている。

55. これまでに記載した通り、参照データが不均衡であったことから、特異度の指標は感度の指標よりも不確定である。

2.1.8. 習熟度確認化学物質

56. 2o3 DA は、単純なルールベースデータ解釈手順であり、専門家の判断を必要としない。個々の情報源（KE 1～3）のための習熟度確認化学物質が、それぞれのガイドラインに規定されている(2, 3, 4)。個々の情報源のための習熟度は、DA の習熟度を示す。

2.1.9. DA の報告

57. DA 適用の報告は、OECD GD 255 (8)に記載されたテンプレートに従い、少なくとも以下の要素を記載する必要がある。

- 被験化学物質識別情報（化学名、構造式、組成、異性体、不純物と入手可能であればその量、CAS 番号、バッチ番号、ロット番号、その他の関連する識別子）
- 対応するガイドライン（OECD TG 442C、442D、442E）に準拠して実施された個々の試験報告書。なお、各試験報告書の化学的特定名が上記と一致しなければならない
- **補遺 1** に記載されている通り、境界域の結果を決定するために 2o3 DA に用いられる個々の予測モデルの適用
- DA 適用の結果（皮膚感作性物質か、非皮膚感作性物質または確定的でない結果からなるハザードの特定）
- 2o3 DA からの逸脱または適合
- 結論

©OECD 2021

2.2. 参考文献

1. OECD (2021). Series on Testing and Assessment No. 336: Supporting document to the Guideline (GL) on Defined Approaches (DAs) for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
2. OECD (2020). OECD Guideline for the Testing of Chemicals No. 442C: *In chemico* Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key even on covalent binding to proteins) *In chemico*. Paris, France: Organisation for Economic Cooperation and Development. Available at: (oecd-ilibrary.org).
3. OECD (2018). OECD Key Event based test Guideline 442D: *In vitro* Skin Sensitisation Assays Addressing AOP Key Event on Keratinocyte Activation. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
4. OECD (2018). OECD Key event based test Guideline 442E: *In vitro* Skin Sensitisation Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
5. OECD (2016). Series on Testing & Assessment No. 256: 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, Annex 1 and Annex 2.. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
6. Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R. (2012). Putting the parts together: combining *in vitro* methods to test for skin sensitizing potential. Regul Toxicol Pharmacol, 63:489-504.
7. Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods, Regul Toxicol Pharmacol, 71:337-51.
8. OECD (2016). Series on Testing & Assessment No. 255: Guidance Document On The Reporting Of Defined Approaches To Be Used Within Integrated Approaches To Testing And Assessment. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].

©OECD 2021

パート II – セクション 3–皮膚感作性の強度分類のための Defined Approach

58. 本ガイドラインのパート II では、「化学品の分類および表示に関する世界調和システム」（GHS）に準拠し、強い感作性物質である UN GHS 細区分 1A か、その他の（中等度から弱い）皮膚感作性物質である細区分 1B かのどちらかに割り当てる Defined Approach について記載する。このような DA を、ハザードの特定、すなわち感作性物質（UN GHS 区分 1）と非感作性物質（区分外；NC）の識別にも使用できる。現時点では、本ガイドラインのこの項は ITSv1 DA および ITSv2 DA からなる。このほかの詳細情報を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

3.1. 「統合的試験戦略（ITS）」の Defined Approach

3.1.1. 要約

59. この Defined Approach は、化学物質の UN GHS に従って、皮膚感作性ハザードの可能性と強度細区分（細区分 1A および 1B）を予測するための統合的試験戦略（ITS）として構築された。

60. ITS DA は、有害性発現経路（AOP）の重要事象（KE）1 および 3 に対処する試験法を使用しており、皮膚感作性の *in silico* 予測を組み込んでいる。ペプチド結合性試験（Direct Peptide Reactivity Assay [DPRA]；OECD TG 442C）(2)を用いて、タンパク質との結合(KE1)を定量的に評価する。ヒト細胞株活性化試験(h-CLAT; OECD TG 442E) (3)を用いて、樹状細胞の活性化（KE3）を定量的に評価する。皮膚感作性の *in silico* 予測には、Derek Nexus（ITSv1）または OECD QSAR Toolbox（ITSv2）のいずれかを用いる。

61. EG DASS が合意した通りにキュレーションされた LLNA 参照データに照らして、167 種類の化学物質を用いたハザード特定および 155 種類の化学物質を用いた UN GHS 細区分について、ITSv1 DA を評価したところ、LLNA と同等の正確性を得た（**表 3.2～3.3** を参照）。ITSv1 DA の性能を、EG DASS が合意した通りにキュレーションされたヒト参照データのある 64 種類の化学物質について比較したところ（**表 3.4～3.5** 参照）、ハザードでも強度分類でも、同じヒトデータの予測における LLNA の正確性を上回った。

62. EG DASS が合意した通りにキュレーションされた LLNA 参照データに基づいて、167 種類の化学物質を用いたハザード特定および 153 種類の化学物質を用いた UN GHS 細区分について、ITSv2 DA を評価したところ、LLNA と同等の正確性を得た（**表 3.6～3.7** を参照）。ITSv2 DA の性能を、EG DASS が合意した通りにキュレーションされたヒト参照データのある 64 種類の化学物質について比較したところ（**表 3.8～3.9** 参照）、ハザードでも強度分類でも、同じヒトデータの予測における LLNA の正確性を上回った。

3.1.2. データ解釈手順

63. ITS DIP は、h-CLAT (3)および DPRA (1)のほか、Derek Nexus v 6.1.0 (2020, Lhasa

©OECD 2021

Limited, https://www.lhasalimited.org/products/derek-nexus.htm)または OECD QSAR TB v 4.5 (https://www.oecd.org/chemicalsafety/oecd-qsar-toolbox.htm)のいずれかから得た定量結果に割り当てられたスコアを使用して、化学物質を UN GHS 区分 1A（強い感作性物質）、区分 1B（その他の感作性物質）または区分外（非感作性物質）に区別する（**表 3.1**）。

64. DIP を、ITS の初版（4）から改訂し、1A 感作性物質のカットオフ値をスコア 7 からスコア 6 に変更した。これにより、強い感作性物質を検出する DA の能力を最適化し、*in silico* 予測が得られない化学物質に対する ITS の適用可能性を拡げた。DIP はまた、ECETOC 区分 [2]を用いていた最初の公表バージョンから、UN GHS 細区分を用いるように改訂された。

65. h-CLAT および DPRA の定量結果を、**表 3.1** に示すように 0 から 3 のスコアに変換する。h-CLAT では、10 µg/mL および 150 µg/mL のカットオフ値に基づく最小誘導閾値（MIT）を 0 から 3 のスコアに変換する。DPRA では、OECD TG 442C (2)に記載されている反応性クラスの閾値に基づき、システインおよびリジンの平均ペプチド減少率を 0～3 のスコアに変換する。リジンペプチドのみに共溶出が起きる場合は、システインペプチドのみの減少率を 0～3 のスコアに変換する。*In silico* 予測（Derek または OECD QSAR TB）では、陽性の結果にはスコア 1 を、陰性の結果にはスコア 0 を割り当てる（それぞれのプロトコルの詳細は**補遺 2** を参照）。ここに挙げたスコアが評価されている場合、個々のスコアを合計して求めた 0 から 7 にまたがる総スコアを用いて感作性の有無（ハザードの特定；UN GHS 区分 1 対 UN GHS NC）および強度（UN GHS 区分 1A、区分 1B および NC）を予測する。皮膚感作性物質を特定するための陽性基準（UN GHS 区分 1）を、総スコア 2 以上に設定している。最新の DIP では、総スコアで 3 つのランクに分かれ、スコア 6～7 は強い感作性物質（UN GHS 区分 1A）、スコア 2～5 は中等度/弱い感作性物質（UN GHS 区分 1B）、スコア 1 または 0 は区分外（非感作性物質）となる。

[2] ECETOC Technical Report 087 (2003), Contact Sensitisation:Classification According to Potency.https://www.ecetoc.org/publication/tr-087-contact-sensitisation- classification-according-to-potency/で入手可能。

©OECD 2021

表 3.1.ITSDefined Approach の概略 DA は、ここに示すように、OECD TG 442E および 442C に基づくアッセイと *in silico* 構造ベース予測に基づく、単純なスコアベースのシステムである。

スコア	h-CLAT MIT μg/mL	DPRA によるシステインおよびリジンの平均減少率	DPRA によるシステイン減少率*	*In silico* (ITSv1:DEREK; ITSv2:OECD TB)
3	≤10	≥42.47	≥98.24	
2	>10、≤150	≥22.62、<42.47	≥23.09、<98.24	
1	>150、≤5000	≥6.38、<22.62	≥13.89、<23.09	陽性
0	計算できない	<6.38	<13.89	陰性

強度	総スコア
UN GHS 1A	6〜7
UN GHS 1B	2〜5
区分外	0〜1

出典:Takenouchi (5)

注 : UN GHS 1A は強い感作性物質に該当し、UN GHS 1B はその他の（中等度から弱い）感作性物質に該当する。区分外は非感作性物質とされる。*リジンペプチドが共溶出する場合に、システインのみの減少率閾値を用いる。

3.1.3. 個々の情報源の内容と限界

66. 個々の *in chemico* および *in vitro* の情報源は、既存の KE ベースの OECD 試験ガイドライン（OECD TG 442C、442E）であり(2, 3)、そのなかにプロトコルが詳しく記載されている。

67. 上記試験ガイドラインの下記のアッセイを明らかにし、ITS DA に組み込む。

- ヒト細胞株活性化試験（h-CLAT; OECD TG 442E; KE3） (3):抗原提示細胞の活性化は、CD86 ないし CD54 のアップレギュレーションを特徴とする。1.5 倍を超える CD86 発現誘導ないし 2 倍を超える CD54 発現誘導を認め、生存率が溶媒対照の 50%超である場合に、h-CLAT を陽性と判断する。実験的濃度-反応曲線から、1.5 倍ないし 2 倍の CD86 ないし CD54 を誘導する濃度中央値を求め、この 2 つの値のうち低い方の値を最小誘導閾値 MIT とする。

  MIT = どちらか低い方(EC150 CD86、EC200 CD54)

  被験化学物質に、**表 3.1** に示す MIT 閾値に基づいて強度スコアを割り当てる。

©OECD 2021

- ペプチド結合性試験(DPRA; OECD TG 442C; KE1) (2):皮膚感作性物質は一般に求電子物質であり、タンパク質の求核部分と反応する。DPRA は、システイン残基またはリジン残基を含む 2 種類のペプチドの共有結合による減少率を測定する。被験化学物質によるシステイン含有ペプチドおよびリジン含有ペプチドの平均減少率が 6.38%を超える場合（共溶出がある場合はシステインのみの減少率が 13.89%超）に陽性と判断する。ペプチド減少について境界域の結果が得られた場合は、OECD TG 442C に記載されている追加の検査を実施する。**表 3.1** に示す平均ペプチド減少率閾値に基づいて、被験化学物質を強度スコアに割り当てる。

68. 個々の *in chemico* および *in vitro* 試験法にある限界が、それぞれの試験ガイドライン（TG 442C, 付録 1、TG 442E 補遺 1） (2, 3)に記載されている。

69. ITSv1 に関する *in silico* 情報源の予測は、知識ベースのエキスパートソフトウェアツールである Derek に基づく。Derek は、皮膚感作性を含む数種の毒性エンドポイントに対するアラートからなる。Derek （Derek Nexus v.6.1.0, 2020、Lhasa 社）は、構造特性に基づいて（具体的には、ハプテンが直接的または代謝/自己酸化によって皮膚タンパク質と求電子結合する可能性を有するかどうか）、アラートを発する。アラートにはいずれも、確からしさのレベルが記載されている。確からしさが certain、probable、plausible のアラートを生じた化学物質を陽性と判断する。「misclassified feature も unclassified feature もない非感作性物質」として陰性予測された化学物質を陰性と判断する (https://www.lhasalimited.org/products/skin-sensitisation-assessment- using-derek-nexus.htm#Negative%20Predictions)。ITSv1 に用いる *in silico* 適用領域を明らかにする方法と、Derek 予測を導くためのプロトコルが、本ガイドラインの**補遺 2** に記載されている。

70. ITSv2 の *in silico* 情報源の予測は、皮膚感作性ハザード予測(OECD QSAR TB v4.5)を生じる自動ワークフローである OECD QSAR TB によってもたらされる。タンパク質結合アラート（自己酸化生成物および皮膚代謝物の生成）がないか、標的化合物のプロファイリングを行い、次に、タンパク質結合アラートのプロファイリングを行う。親化合物またはその生物的代謝物にタンパク質結合アラートを認めた場合、実験的皮膚感作データのある類似化合物を特定するために同じアラートを使用する。タンパク質結合アラートが特定されない場合は、構造的プロファイラーを使用して類似化学物質を特定し、read-across を用いるか、適切な類似化合物が自動的に特定されなかった場合はプロファイラーの結果を介して直接的にデータギャップを埋める。ITSv2 に用いる *in silico* 適用領域を明らかにするための方法と、OECD QSAR TB 予測を導くためのプロトコルが本ガイドラインの**補遺 2** に記載されている。

3.1.4. ITS DA 予測の信頼性

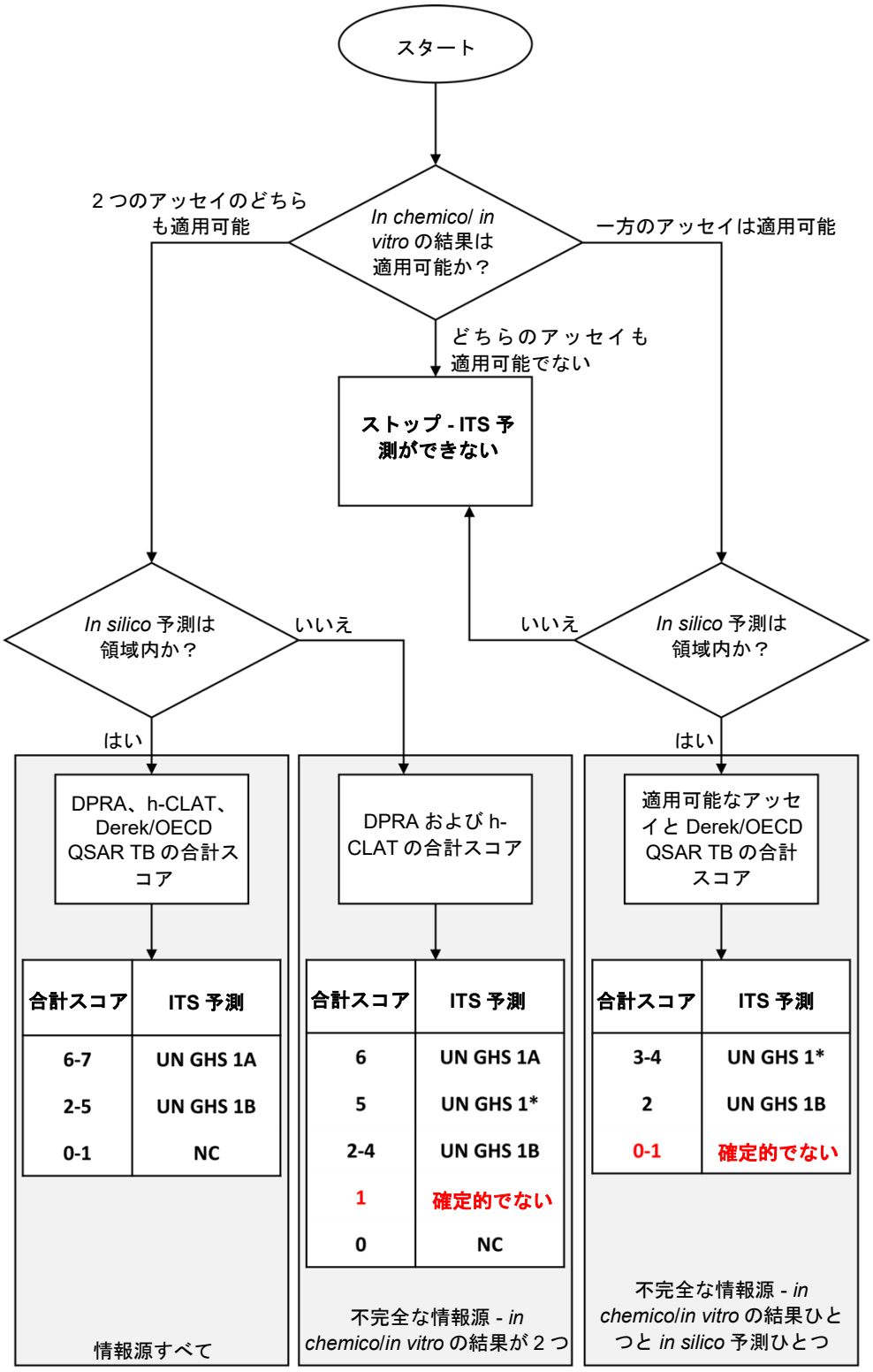
71. **図 3.1** のフローチャートに示すように、ITS DA 予測の信頼性のレベルを、個々の情報源の総 DA スコアと適用領域に基づいて割り当てる。全情報要素を使用できるかどうかの最初の判断は、TG 442C 付録 1 および TG 442E 補遺 1 (3) に記載したように、*in chemico* および *in vitro* の方法の限界や（例えば、溶解度が低いか、現時点で信頼できないと判断される Log P > 3.5 の化学物質について h-CLAT 結果が陰性であるために、個々の方法で確定的な結果が得られない物質）、*in silico* 予測の適用領域（**補遺 2**）によって決まる。**図 3.1** のフローチャートに示すように、不完全な情報源（*in chemico/in vitro* の結果が 2 つしかない、あるいは、*in chemico/in vitro* の結果がひとつと *in silico* 予測がひとつ）を用いて DA 予測を得ることができる。

©OECD 2021

72. ハザードの特定および強度に関する信頼性の高いDA予測を、確定的と考える。信頼性の低いDA予測を、ハザードの特定や強度に関して確定的でないと判断する。それでも、この「確定的でない」との予測を、エビデンスの重み付けアプローチやIATAとのからみで、他の情報源と共に考慮するとよい。適用領域および信頼性に関する判断を含む詳細を**補遺2**に示す。

©OECD 2021

図 3.1.ITS DA 予測に信頼性を割り当てるためのデシジョンツリー



***ハザードに関しては確定的、強度に関しては確定的でない**

©OECD 2021

3.1.5. LLNA と比較した ITSv1 DA の予測能

73. Derek を用いた ITSv1 の予測能が、EG DASS が合意した通りにキュレーションされた LLNA のデータに基づいて報告されている（**表 3.2～3.3** 参照）（*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 1.1 項**および**補遺 3** を参照のこと）(1)。**図 3.1** に示すワークフローを適用して、ITSv1 DA 予測に信頼性を割り当てた。ITSv1 DA 予測に対する確定的/確定的でない指定については、**補遺 2** に詳しく記載する。確定的予測に対し、LLNA 参照データと比較したパフォーマンス統計量を報告する。このほかに、確定的でない結果を記載する。特定の化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

表 3.2.LLNA 参照データと比較した ITSv1 DA のハザード特定性能

	LLNA	
ITSv1 DA	非感作性物質	感作性物質
非感作性物質	21	11
感作性物質	9	118
確定的でない	3	6

LLNA データと比較した DA 性能 (N=159)	**ITSv1**
正確性（%）	87%
感度（%）	92%
特異度（%）	70%
バランス精度 (%)	81%

注：正確性は LLNA データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。統計量は信頼性の高い予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

74. 参照データが不均衡であったことから、特異度の指標（LLNA 陰性化学物質 30 種類に基づく）は感度の指標（LLNA 陽性化学物質 129 種類）よりも不確定である。

表 3.3.LLNA 参照データと比較した、UN GHS 1A/1B 細区分に基づく ITSv1 DA 強度分類性能

	LLNA		
ITSv1 DA	NC	1B	1A
NC	21	11	0
1B	9	55	10
1A	0	12	28
確定的でない	3	7	0

全体での正しい分類 71%

©OECD 2021

LLNA 参照データと比較した ITSv1：UN GHS 1A/1B 細区分に基づく統計量

性能（N=146）	NC (N=30)	1B (N=78)	1A (N=38)
正しい分類（%）	70%	71%	74%
過小予測（%）	該当せず	14% (NC)	0% (NC); 26% (1B)
過大予測（%）	30% (1B); 0% (1A)	15% (1A)	該当せず

注：統計量は信頼性の高い予測のみを示し、確定的でない予測を灰色で示す。クラス内性能（感度、特異度、バランス精度）に関する詳細については、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**を参照のこと(1)。

75. ITSv1 DA 予測に対する信頼性高／低の指定を、上記の**図 3.1** の記載および**補遺 2** の詳述の通りに適用する。

3.1.6. LLNA と比較した ITSv2 DA の予測能

76. OECD QSAR TB を用いた ITSv2 の予測能が、EG DASS が合意した通りにキュレーションされた LLNA のデータに基づいて報告されている（**表 3.4～3.5** 参照）（*皮膚感作性 Defined Approach[DA]に関するガイドライン[GL]の付属文書*の**第 2.1 項**および**補遺 3** を参照のこと）(1)。**図 3.1** に示すワークフローを適用して、ITSv2 DA 予測に信頼性を割り当てた。ITSv2 DA 予測に対する信頼性高／低の指定が、**補遺 2** に詳しく記載されている。信頼性の高い予測に対し、LLNA 参照データと比較したパフォーマンス統計量を報告する。このほかに、確定的でない結果を記載する。特定の化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。**表 3.4.** LLNA 参照データと比較した ITSv2 DA のハザード特定性能

表 3.4.LLNA 参照データと比較した ITSv2 DA のハザード特定性能

	LLNA	
ITSv2 DA	非感作性物質	感作性物質
非感作性物質	20	9
感作性物質	10	117
確定的でない	3	9

LLNA データと比較した DA 性能 (N=156)	**ITSv2**
正確性（%）	88%
感度（%）	93%
特異度（%）	67%
バランス精度 (%)	80%

注：正確性は LLNA データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

77. 参照データが不均衡であったことから、特異度の指標（LLNA 陰性化学物質 30 種類に基づく）は感度の指標（LLNA 陽性化学物質 126 種類）よりも不確定である。

©OECD 2021

表 3.5.LLNA 参照データと比較した、UN GHS 1A/1B 細区分に基づく ITSv2 DA 強度分類性能

	LLNA		
ITSv2 DA	NC	1B	1A
NC	20	9	0
1B	10	54	10
1A	0	12	26
確定的でない	3	10	2

全体での正しい分類 71%

LLNA 参照データと比較した ITSv2：UN GHS 1A/1B 細区分に基づく統計量

性能（N=141）	**NC (N=30)**	**1B (N=75)**	**1A (N=36)**
正しい分類（%）	67%	72%	72%
過小予測（%）	該当せず	12% (NC)	0% (NC); 28% (1B)
過大予測（%）	33% (1B); 0% (1A)	16% (1A)	該当せず

注：統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。クラス内性能（感度、特異度、バランス精度）に関する詳細については、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**を参照のこと(1)。

78. ITSv2 DA 予測に対する確定的/確定的でない指定を、上記の**図 3.1** の記載および**補遺 2** の詳述の通りに適用する。

3.1.7. ヒトデータと比較した ITSv1 DA の予測能

79. Derek を用いた ITSv1 の予測能が、EG DASS が合意した通りにキュレーションされたヒト予測パッチテストのデータに基づいて報告されている（**表 3.6～3.7** 参照）。ITSv1 DA 予測に対する信頼性高／低の指定が、**補遺 2** に詳しく記載されている。信頼性の高い予測に対し、ヒト参照データと比較したパフォーマンス統計量を報告する。このほかに、確定的でない結果を記載する。特定の化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

©OECD 2021

表 3.6.ヒト参照データと比較した ITSv1 DA のハザード特定性能

	ヒト	
ITSv1 DA	非感作性物質	感作性物質
非感作性物質	4	4
感作性物質	5	51
確定的でない	2	0

ヒトデータと比較した DA 性能 (N=64)	**ITSv1**
正確性（%）	86%
感度（%）	93%
特異度（%）	44%
バランス精度 (%)	69%

注：正確性はヒト HPPT データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

80. 参照データが不均衡であったことから、特異度の指標（ヒト陰性化学物質 9 種類に基づく）は感度の指標（ヒト陽性化学物質 55 種類に基づく）よりも不確定である。

表 3.7.ヒト参照データと比較した、UN GHS 1A/1B 細区分に基づく ITSv1 DA 強度分類性能

	ヒト		
ITSv1 DA	NC	1B	1A
NC	4	4	0
1B	5	24	7
1A	0	3	13
確定的でない	2	0	1

全体での正しい分類 68%

ヒト参照データと比較した ITSv1：UN GHS 1A/1B 細区分に基づく統計量

性能（N=60）	**NC (N=9)**	**1B (N=31)**	**1A (N=20)**
正しい分類（%）	44%	77%	65%
過小予測（%）	該当せず	13% (NC)	0% (NC); 35% (1B)
過大予測（%）	56% (1B); 0% (1A)	10% (1A)	該当せず

注：統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。クラス内性能（感度、特異度、バランス精度）に関する詳細については、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**を参照のこと(1)。

81. ITSv1 DA 予測に対する確定的/確定的でない指定を、上記の**図 3.1** の記載およ

©OECD 2021

び**補遺 2** の詳述の通りに適用する。

82. 参照データが不均衡で、化学物質数が少ないことから、正確性の指標は、NC 化学物質のような小さいクラスのほうが不確定である。

3.1.8. ヒトデータと比較した ITSv2 DA の予測能

83. OECD QSAR Toolbox を用いた ITSv2 の予測能が、EG DASS が合意した通りにキュレーションされたヒト予測パッチテストのデータに基づいて報告されている（**表 3.8～3.9** 参照）。ITSv2 DA 予測に対する信頼性高／低の指定が、**補遺 2** に詳しく記載されている。確定的予測に対し、ヒト参照データと比較したパフォーマンス統計量を報告する。このほかに、確定的でない結果を記載する。特定の化学物質に関する DA 予測および詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す(1)。

表 3.8.ヒト参照データと比較した ITSv2 DA のハザード特定性能

	ヒト	
ITSv2 DA	非感作性物質	感作性物質
非感作性物質	4	3
感作性物質	5	50
確定的でない	2	2

ヒトデータと比較した DA 性能 (N=62)	**ITSv2**
正確性（%）	87%
感度（%）	94%
特異度（%）	44%
バランス精度 (%)	69%

注：正確性はヒト HPPT データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

84. 参照データが不均衡であったことから、特異度の指標（ヒト陰性化学物質 9 種類に基づく）は感度の指標（ヒト陽性化学物質 53 種類）よりも不確定である。

表 3.9.ヒト参照データと比較した、UN GHS 1A/1B 細区分に基づく ITSv2 DA 強度分類性能

	ヒト		
ITSv2 DA	NC	1B	1A
NC	4	3	0
1B	5	24	6
1A	0	3	12
確定的でない	2	1	3

全体での正しい分類 70%

©OECD 2021

ヒト参照データと比較した ITSv2：UN GHS 1A/1B 細区分に基づく統計量

性能（N=57）	NC (N=9)	1B (N=30)	1A (N=18)
正しい分類（%）	44%	80%	67%
過小予測（%）	該当せず	10% (NC)	0% (NC); 33% (1B)
過大予測（%）	56% (1B); 0% (1A)	10% (1A)	該当せず

注：統計量は確定的な予測のみを示し、確定的でない予測を灰色で示す。クラス内性能（感度、特異度、バランス精度）に関する詳細については、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**を参照のこと(1)。

85. ITSv2 DA 予測に対する確定的/確定的でない指定を、上記の**図 3.1** の記載および**補遺 2** の詳述の通りに適用する。

86. 参照データが不均衡で、化学物質数が少ないことから、正確性の指標は、NC 化学物質のような小さいクラスのほうが不確定である。

3.1.9. ヒトデータと比較した LLNA の予測能

87. DA のパフォーマンスに関する比較の根拠を得るため、LLNA の予測能が、EG DASS が合意した通りにキュレーションされたヒト予測パッチテストのデータに基づいて報告されている（**表 3.10～3.11** 参照）。特定の化学物質に関するデータおよび詳細を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*の**第 5 項**および**補遺 2** に示す (1)。

表 3.10.ヒト参照データと比較した LLNA のハザード特定性能

	ヒト	
LLNA	非感作性物質	感作性物質
非感作性物質	2	3
感作性物質	7	44

ヒトデータと比較した LLNA 性能 (N=56)	LLNA
正確性（%）	82%
感度（%）	94%
特異度（%）	22%
バランス精度 (%)	58%

注：正確性はヒト HPPT データに関する正しい分類率、感度は真陽性率、特異度は真陰性率、バランス精度は感度と特異度の平均値である。このほかの性能特性を、*皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書*に記載する(1)。

88. ヒトデータと比較した確定的 ITSv1 DA 予測のハザード特定性能は、正確性 86%、感度 93%、特異度 44%、バランス精度 69%であり、どの尺度も、LLNA の性能と同じかこれより優れている。

89. ヒトデータと比較した確定的 ITSv2 DA 予測のハザード特定性能は、正確性 87%、感度 94%、特異度 44%、バランス精度 69%であり、どの尺度も、LLNA の性能と同じかこれより優れている。

©OECD 2021

90. これまでに記載した通り、参照データが不均衡であったことから、特異度の指標は感度の指標よりも不確定である。

表 3.11.ヒト参照データと比較した、UN GHS 1A/1B 細区分に基づく LLNA の強度分類性能

このほかの性能特性を、皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書に記載する(1)。

	ヒト		
LLNA	NC	1B	1A
NC	2	3	0
1B	6	17	7
1A	0	3	9

全体での正しい分類 60%

ヒト参照データと比較した LLNA：UN GHS 1A/1B 細区分に基づく統計量

性能（N=47）	**NC (N=8)**	**1B (N=23)**	**1A (N=16)**
正しい分類（%）	25%	74%	56%
過小予測（%）	該当せず	13% (NC)	0% (NC); 44% (1B)
過大予測（%）	75% (1B); 0% (1A)	13% (1A)	該当せず

注：クラス内性能（感度、特異度、バランス精度）に関する詳細については、皮膚感作性 Defined Approach（DA）に関するガイドライン（GL）の付属文書の **第 5 項** *を参照のこと(1)。*

91. ヒトデータと比較した確定的 ITSv1 DA 予測の強度細区分に関する性能は、全体での正しい分類 68%、NC の正確性 44%、1B の正確性 77%、1A の正確性 65%で、どの尺度も、LLNA の性能と同じかこれより優れている。

92. ヒトデータと比較した確定的 ITSv2 DA 予測の強度細区分に関する性能は、全体での正しい分類 70%、NC の正確性 44%、1B の正確性 80%、1A の正確性 67%で、どの尺度も、LLNA の性能と同じかこれより優れている。

93. これまでに記載した通り、参照データが不均衡で、化学物質数が少ないことから、正確性の指標は、NC 化学物質のような小さいクラスのほうが不確定である。

3.1.10. 習熟度確認化学物質

94. ITS DA は、単純なルールベースデータ解釈手順であり、専門家の判断を必要としない。個々の *in chemico* および *in vitro* 情報源（KE1 および KE3）のための習熟度確認化学物質が、それぞれのガイドラインに規定されている(OECD TG 442C, 442E) (2, 3)。Derek および OECD QSAR Toolbox からなる *in silico* 情報源の選択肢に関するプロトコルの詳細が、このガイドラインの**補遺 2** に記載されている。Derek Nexus v6.1.0 および OECD QSAR Toolbox v4.5 に関して習熟度が明らかにされている。Derek Nexus v6.1.0 および OECD QSAR Toolbox v4.5 はそれぞれ、ITSv1 DA および ITSv2 DA に使用するようデザインされたソフトウェアバージョンである。個々の情報源のための習熟度は、DA の習熟度を示す。

©OECD 2021

3.1.11. DA の報告

95. ITS DA の報告は、OECD GD 255 (6)に記載されたテンプレートに従い、少なくとも以下の要素を記載する必要がある。

- 被験化学物質識別情報（化学名、構造式、組成、異性体、不純物と入手可能であればその量、CAS 番号、バッチ番号、ロット番号、その他の関連する識別子）
- 対応するガイドライン（OECD TG 442C、442E）に準拠して実施された個々の試験報告書。なお、各試験報告書の化学的特定名が上記と一致しなければならない
- QPRF (7)を介した報告など、*in silico* 予測**(補遺 2)**と結果に使用するプロトコルの説明。
- DA 適用の結果（ハザードの特定と UN GHS 区分に基づく強度分類または確定的でない結果）
- ITS DA からの逸脱
- 結論

©OECD 2021

3.2. 参考文献

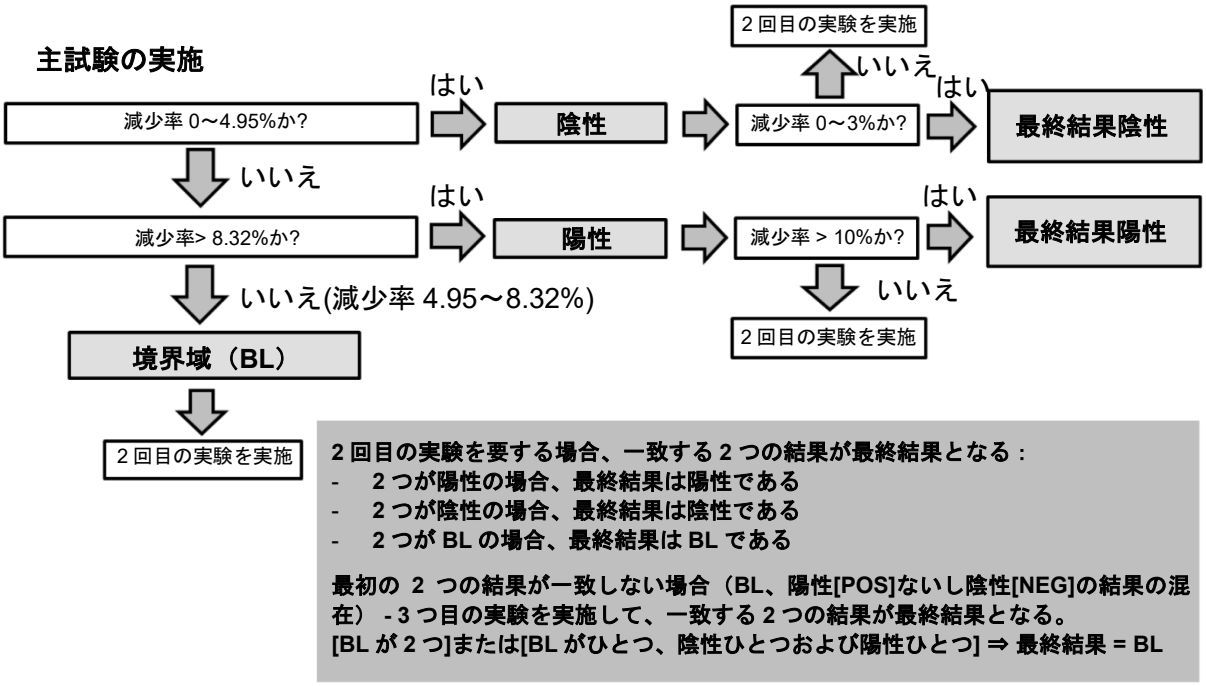
1. OECD (2021). Series on Testing and Assessment No. 336: Supporting document to the Guideline (GL) on Defined Approaches (DAs) for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
2. OECD (2020). OECD Guideline for the Testing of Chemicals No. 442C: *In chemico* Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key even on covalent binding to proteins). *In chemico*. Paris, France: Organisation for Economic Cooperation and Development. Available at:(oecd-ilibrary.org).
3. OECD (2018). OECD Key event based test Guideline 442E: *In vitro* Skin Sensitisation Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: (oecd-ilibrary.org).
4. OECD (2016). Series on Testing & Assessment No. 256: 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, Annex 1 and Annex 2. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
5. Takenouchi O, Fukui S, Okamoto K, Kurotani S, Imai N, Fujishiro M, Kyotani D, Kato Y, Kasahara T, Fujita M, Toyoda A, Sekiya D, Watanabe S, Seto H, Hirota M, Ashikaga T, Miyazawa M. (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, 35:1318-32.
6. OECD (2016). Series on Testing & Assessment No. 255: Guidance Document On The Reporting Of Defined Approaches To Be Used Within Integrated Approaches To Testing And Assessment. ENV/JM/HA(2016)28. Organisation for Economic Cooperation and Development, Paris. Available at: [https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm].
7. ECHA (2008). see "CHAPTER R.6 – QSARS AND GROUPING OF CHEMICALS" *in* Guidance on Information Requirements and Chemical Safety Assessment. European Chemicals Agency *[*Guidance on Information Requirements and Chemical Safety Assessment - ECHA (europa.eu)]

©OECD 2021

補遺 1：2o3 DA に用いるために複数回実験したそれぞれの *in chemico/in vitro* 試験の予測モデル

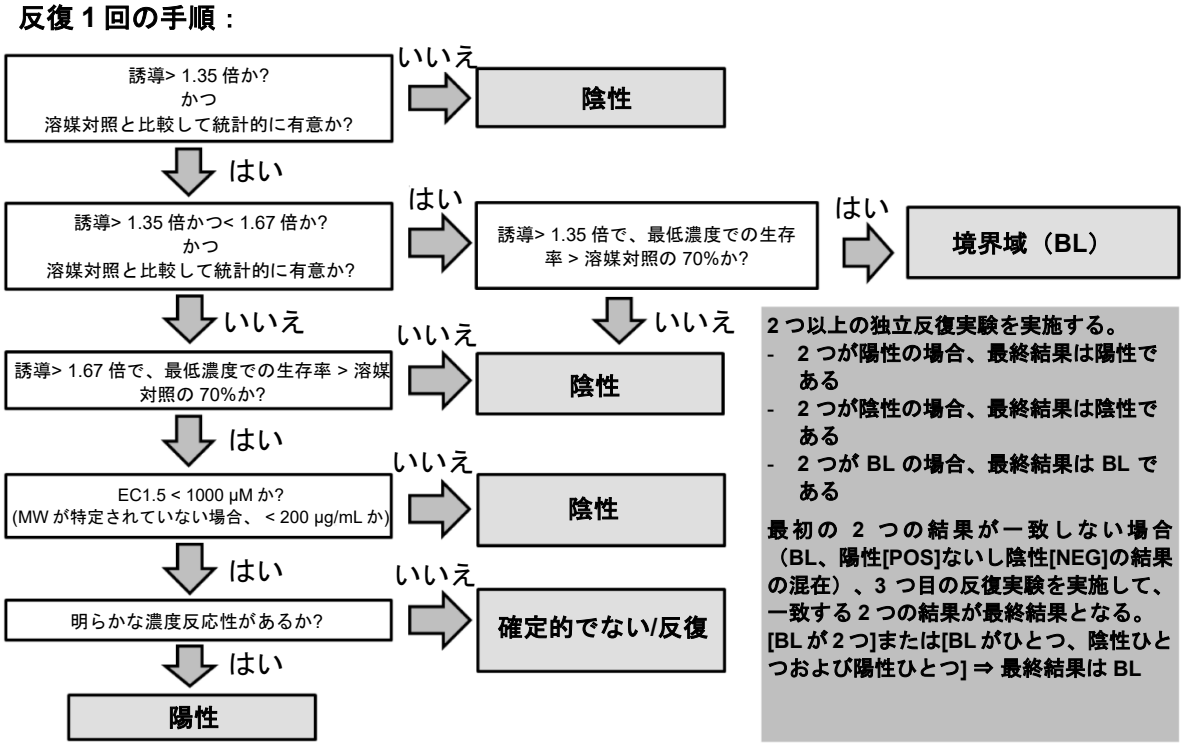
96. h-CLAT および KeratinoSens™のそれぞれの予測モデルは複数回の実験を必要とする（独立反復）。2o3 DA のなかで予測することを目的として、予測モデルを適合させて、それぞれの実験における境界例を決定するのに使用した。以下に挙げるこのような適合（図 1.2 および 1.3）を上記方法に使用することにより、それぞれの試験の最終結論を得る。

97. DPRA では、平均減少率が 3～10%である場合（システインのみの減少モデルを用いる場合は 9～17%）に反復実験を実施する必要がある。この適合には、図 1.1 のフローチャートを用いて、反復実験および 2o3 DA の境界域評価を決定する。



補遺 1、図 1.1.2o3 DA の境界域の結果に結論をもたらす、境界域と反復実験を考慮した DPRA 予測モデル（平均減少率）のフローチャート。陽性分類の当初の閾値は 6.38%であり、この閾値近辺に統計学的に求めた境界域は 4.95%～8.32%である。システインのみの予測モデルに同じフローチャートを適用するには以下の閾値を用いる：3%の代わりに 9%、>10%の代わりに>17%、4.95%の代わりに 10.56%、>8.32%の代わりに>18.47%。

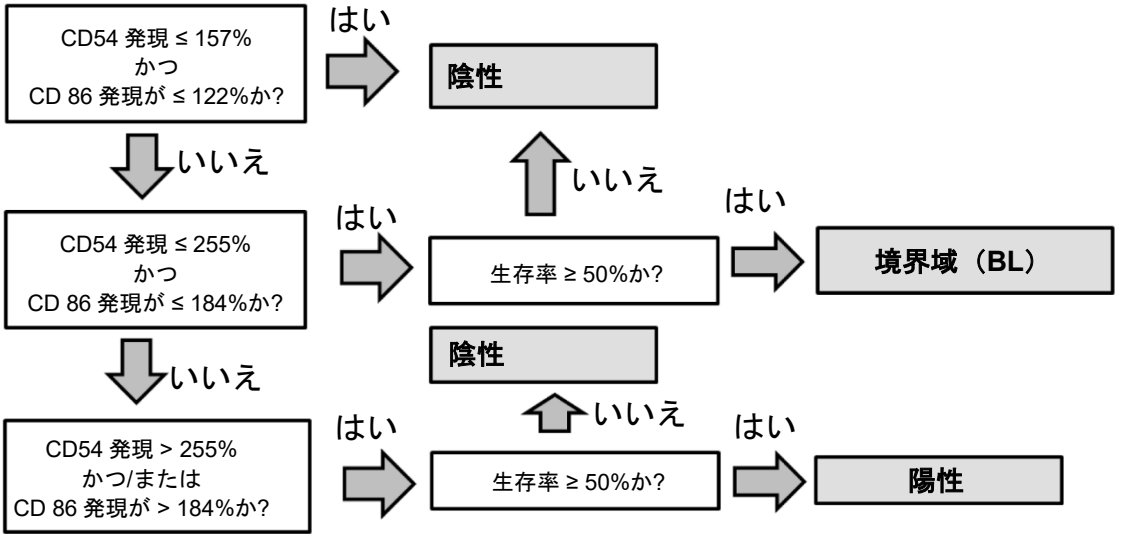
©OECD 2021



補遺 1、図 1.2.2o3 DA の境界域の結果に結論をもたらす、境界域と反復実験を考慮した KeratinoSens™予測モデルのフローチャート。陽性分類の当初の閾値は 1.5 倍の誘導であり、この閾値近辺に統計学的に求めた境界域は 1.35 倍～1.67 倍である。注：442D では、独立実験（independent run）のことを「反復（repetition）」と呼んでいるが、442C および 442E では「実験（run）」と呼んでいる。このような用語は同じものを意味する。

©OECD 2021

実験 1 回の手順：



2 つ以上の独立実験を実施する。
- **2 つが陽性の場合、最終結果は陽性である**
- **2 つが陰性の場合、最終結果は陰性である**
- **2 つが BL の場合、最終結果は BL である**

最初の 2 つの結果が一致しない場合（BL、陽性[POS]ないし陰性[NEG]の結果の混在）、3 つ目の反復実験を実施して、一致する 2 つの結果が最終結果となる。[BL が 2 つ]または[BL がひとつ、陰性ひとつおよび陽性ひとつ] ⇒ 最終結果は BL

補遺 1、図 1.3.2o3 DA の境界域の結果に結論をもたらす、境界域と反復実験を考慮した h-CLAT 予測モデルのフローチャート。陽性分類の当初の閾値は CD86 発現 150%であり、この閾値近辺に統計学的に求めた境界域は 122%～184%である。また、CD54 発現に関しては 200%であり、この閾値近辺に統計学的に求めた境界域は 157%～255%である。

©OECD 2021

補遺 2：*In silico* 予測を導くための DASS ITS 予測およびプロトコルに対する適用領域の規定と信頼性評価

はじめに

98. ***皮膚感作性 Defined Approach に関するガイドライン***の**第 3.1 項**に記載されている通り、ITSDefined Approach(DA)は、3 つの情報源に基づく。これは具体的には、2 つの *in chemico/in vitro* アッセイ(DPRA; OECD TG 442C [OECD, 2015]および h-CLAT; OECD TG 442E [OECD, 2018])とひとつの *in silico* ツール(Derek Nexus [ITSv1]または OECD QSAR Toolbox[ITSv2]のどちらかによる予測[これ以降 *in silico* と呼ぶ])である。それぞれの情報源に対し、個々のアッセイないし予測の結果に基づいてスコアを付け、これを合計して DA 予測を導く。

個々の情報源の適用領域

In chemico/in vitro 情報源 (DPRA および h-CLAT)

99. 各アッセイの技術的、化学的なタイプの限界を考慮し(OECD TG 442C および OECD TG 442E 試験ガイドラインに記載されている通り[OECD, 2015, 2018])、個々のプロトコルに準拠して試験しうる被験化学物質を、DPRA ないし h-CLAT の *in chemico/in vitro* 領域内にある（言い換えると適用可能）と判断する。技術的ないし化学空間的な限界がなく、アッセイで得た結果を考慮できないとする理由がない場合に、*in chemico/in vitro* の結果を適用可能と判断する。

In silico 情報源

100. ITS DA は、化学構造に基づく *in silico* 情報源を使用する。このような *in silico* 情報源は、化学物質の molecular representation に依存する。これは通常、化学構造の描画か、Simplified Molecular-Input Line-Entry System (SMILES)ないし IUPAC 国際化学物質識別子(InChi)の記入によって入力する。ひとつの化学物質が、いくつかの CAS 番号または EC 番号で表される場合があるため（異なる塩形態や混合物中の主要成分として存在する立体化学的な違いによる）、可能であれば正確な構造を規定することが重要である。US EPA CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard)や NIH PubChem (https://pubchem.ncbi.nlm.nih.gov/)などの情報源が、SMILES や InChi フォーマット用の化学名や化学構造をマッピングするのに有用である。分子構造に基づく *in silico* 予測に用いる物質の最低純度レベルについて、利用可能な指針を参照するとよい。[34]

[3] OECD (2017), *Guidance on Grouping of Chemicals, Second Edition*, OECD Series on Testing and Assessment, No. 194, OECD Publishing, Paris, https://doi.org/10.1787/9789264274679-en.

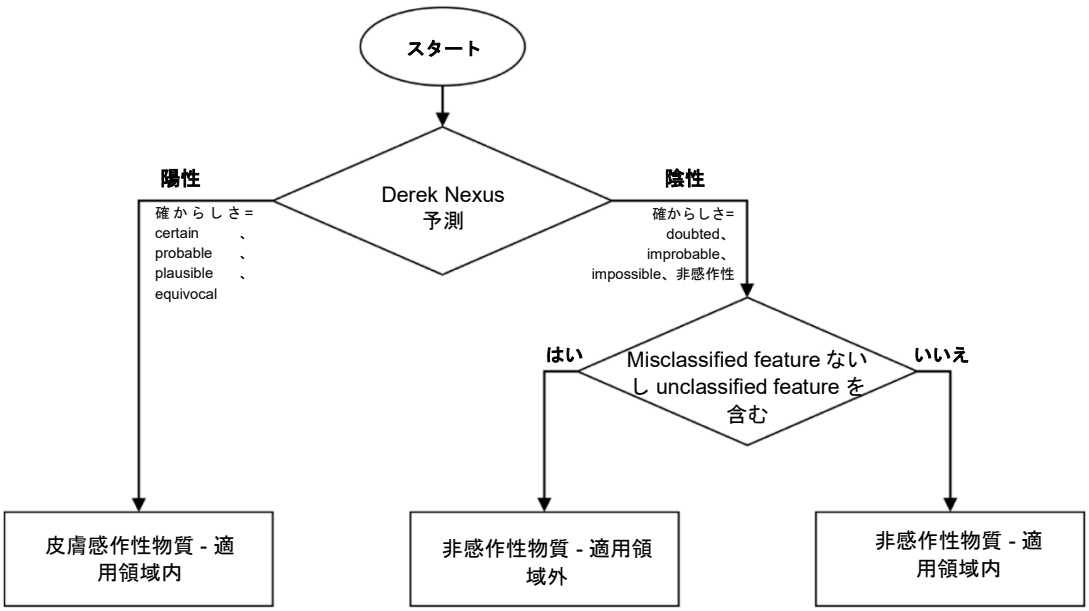
[4] ECHA (2008) CHAPTER R.6 – QSARS AND GROUPING OF CHEMICALS *in* Guidance on Information Requirements and Chemical Safety Assessment.European Chemicals Agency [Guidance on Information Requirements and Chemical Safety Assessment - ECHA (europa.eu)

©OECD 2021

Derek Nexus (ITSv1)

101. ITSv1 では、Derek Nexus v 6.1.0 の皮膚感作性の予測を使用する。Derek Nexus (Derek)予測を実行するためのプロトコルを、本文書の**付録 1** に規定している。陽性の予測（確からしさ= certain、probable、plausible または equivocal）のいずれも、適用領域内と考える。Misclassified feature も unclassified feature もない場合に、陰性予測（確からしさ= doubted、improbable、impossible または非感作性）もまた適用領域内にあると考える。Misclassified feature のある非感作性予測は、Derek がアラートしなかった既知の感作性物質にのみ認められているフラグメントが存在することを示している。Unclassified feature のある非感作性予測は、公表データにはないフラグメント（ただし、Derek がこれを独自開発データのなかに認めている可能性がある）の存在を示している(Chilton et al., 2018)。Misclassified feature や unclassified feature を含む予測には、通常、エキスパートレビューが推奨される。しかし、DA に求められる固定データ解釈手順がエキスパートレビューを許さない場合、このようなものを ITSv1 の使用に関して領域外と考えるのが妥当である（**図 A2.1**）。

図 A2.0.1.ITSv1 に使用する Derek Nexus 皮膚感作性予測のための適用領域



QSAR Toolbox (ITSv2)

102. ITSv2 には、QSAR Toolbox 自動ワークフローの「皮膚感作性の Defined Approach（Skin sensitisation for defined approaches）」(Yordanova et al., 2019)による皮膚感作性予測を使用する。QSAR Toolbox 予測を実行するためのプロトコルが、本文書の**付録 2** に記載されている。

103. DASS AW 予測を実行する際に、予測の適用領域の計算は、Toolbox が自動で行い、これは化学構造、物理化学的特性、作用機序の 3 層からなる。表 A2.1 に記載するように、それぞれの個々の予測で検討する適用領域層は、予測の種類および結果に応じて異なる。この 3 層の詳しい説明およびその選択の根拠を、本文書の**付録 3** に記載する。Toolbox の結果が適用範囲内にあれば、DA に適用可能と考えられる。

©OECD 2021

表 A2.1.QSAR Toolbox 自動ワークフロー「皮膚感作性の Defined Approach[Skin sensitisation for defined approaches]」予測の適用領域層。

Toolbox DASS AW の結果		適用領域層		
		化学構造	物理化学的特性	作用機序
陽性	Read-across	考慮しない	考慮しない	**考慮する**
	プロファイリング	考慮しない	考慮しない	**定義により満たす**
陰性	Read-across	考慮しない	考慮しない	**考慮する**
	プロファイリング	**考慮する**	**考慮する**	**定義により満たす**

ITS 予測の信頼性

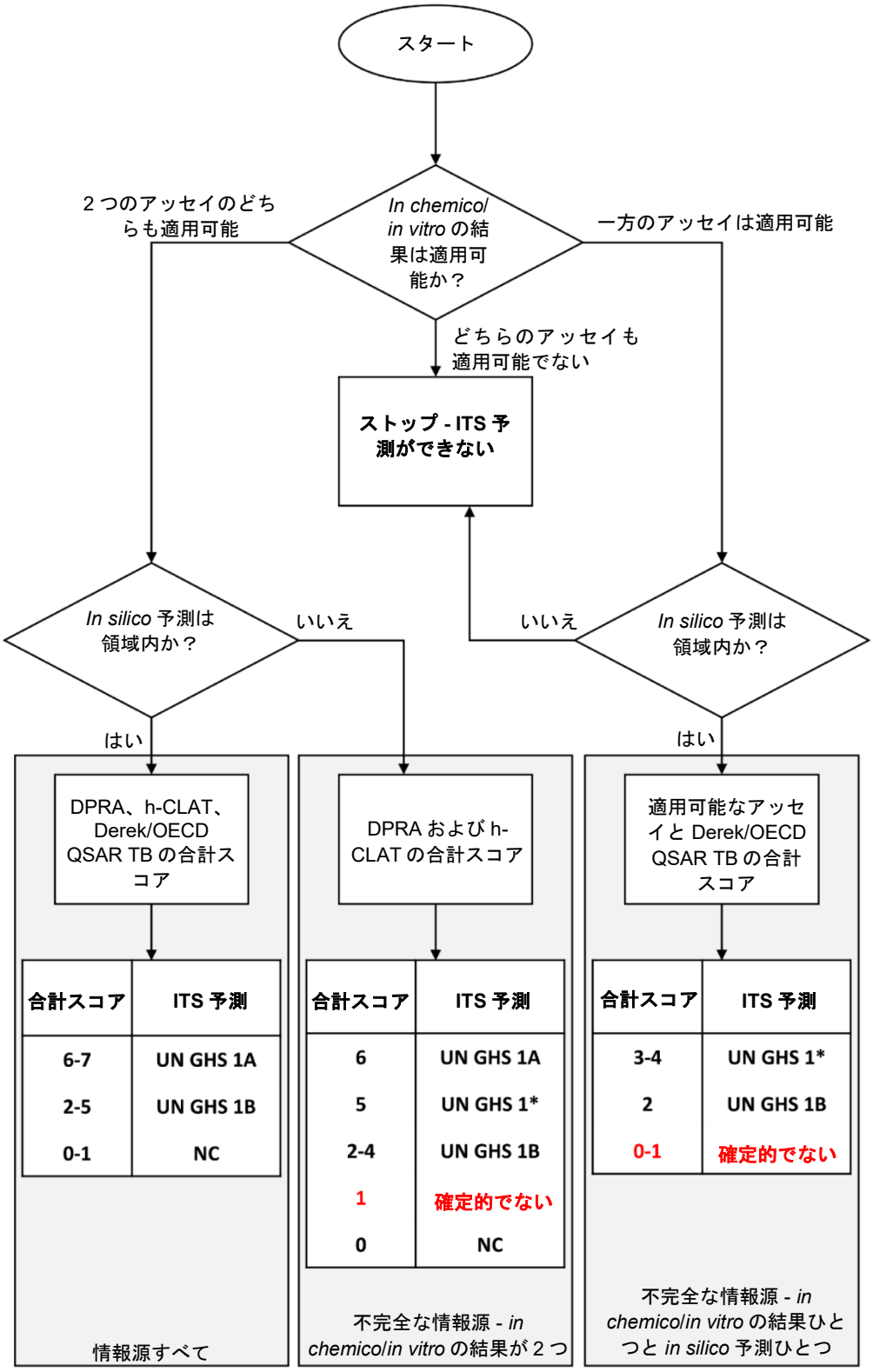
104. ITS DA に使用した個々の情報源の適用領域を評価する。これが、ハザードの特定や強度に関する ITS 予測を確定的と判断するか（信頼性高）または確定的でないと判断するか（信頼性低）を決定する。

ITS に対するデータ解釈手順（DIP）の適用法

105. ITS は元来、3 つの情報源（DPRA、h-CLAT および *in silico* ツール[Derek Nexus または OECD QSAR Toolbox]）を利用するために開発された。3 つの情報源を適用しうる場合、確定的な ITS 予測を行うことができる。場合によっては、適用しうる結果のある 2 つの情報源があれば、確定的な ITS の予測が可能である（**図 A2.2**）。

©OECD 2021

図 A2.0.2.ITS のデータ解釈手順のワークフロー。



***ハザードに関しては確定的、強度に関しては確定的でない**

©OECD 2021

106. 個々の情報源の適用可能性に応じて、ITS DA に 3 種類のシナリオが考えられる（図 A2.2 および表 A2.2 を参照）。シナリオ 1 では、3 つの情報源すべてが適用可能である。シナリオ 2 および 3 では、2 つの情報源のみが適用可能である。詳細を以下に示す。

107. シナリオ 1:情報源である *in chemico/in vitro* の結果がすべて適用可能であり、考慮に入れることができ（個々のアッセイそれぞれについて記載している通り）、*in silico* 予測が領域内にある。得られた ITS DA 予測は確定的で、信頼性が高い。

108. シナリオ 2:*in silico* 予測が領域外であるが、*in chemico/in vitro* 法は領域内にあり、予測は確定的である（*in chemico/in vitro* 法が適用可能）。

- 合計 DA スコアが 0、2、3、4 または 6 で、*in silico* 予測が *in silico* 領域外にある:*In chemico/in vitro* の 2 つの結果に基づいて、DA の結論を得ることが可能である。*In silico* 予測が異なる DA 予測を導かないことから、予測は確定的である。
- 合計 DA スコアが 5 で、*in silico* 予測が *in silico* 領域外にある：ハザード特定に関して DA の結論を得ることが可能である（ハザード特定に関する陽性 DA 予測は確定的である）。強度に関して DA の結論を得ることはできない(強度に関する DA 予測は確定的でない)。
- 合計 DA スコアが 1 で、*in silico* 予測が *in silico* 領域外である：DA の結論を得ることができない。ハザード特定および強度に関する DA 予測は確定的でない。

109. シナリオ 3:*in chemico/in vitro* 法のひとつが領域外にあるか、その方法の結果を考慮できない（適用できない）：

- ひとつの *in chemico/in vitro* 予測および *in silico* 予測に基づく合計 DA スコアが 2 である：DA の結論を得ることが可能である。もう一方の *in chemico/in vitro* 法の結果が異なる DA 予測を導かないことから、UN GHS 1B であるとの DA 予測は確定的である。
- ひとつの *in chemico/in vitro* 予測および *in silico* 予測に基づく合計 DA スコアが 3 または 4 である：ハザード特定に関する DA の結論を得ることが可能である（ハザード特定に関する陽性の DA 予測は確定的である）。強度に関する DA の結論を得ることができない（強度に関する DA 予測は確定的でない）。
- ひとつの *in chemico/in vitro* 予測および *in silico* 予測により、合計 DA スコアが 0 または 1 である：DA の結論を得ることができない。ハザード特定および強度に関する予測は確定的でない。

©OECD 2021

表 A 2.2.ITS の適用領域および信頼性。

シナリオ	合計スコア [5]	ITS 予測	信頼性	信頼性に関する判断を含む DA 予測
1	0-1	NC	高	区分外(NC)の確定的予測
	2-5	UN GHS 1B	高	UN GHS 1B の確定的予測
	6-7	UN GHS 1A	高	UN GHS 1A の確定的予測
2	0	NC	高	NC の確定的予測
	1	確定的でない	低	陽性または陰性の確定的でない予測
	2-4	UN GHS 1B	高	UN GHS 1B の確定的予測
	5	UN GHS 1	高	ハザード特定に対する確定的な陽性予測
			低	強度に対する確定的でない予測
	6	UN GHS 1A	高	UN GHS 1A の確定的予測
3	0-1	確定的でない	低	陽性または陰性の確定的でない予測
	2	UN GHS 1B	高	UN GHS 1B の確定的予測
	3-4	UN GHS 1	高	ハザード特定に対する確定的な陽性予測
			低	強度に対する確定的でない予測

参考文献

Chilton, M. L., Macmillan, D. S., Steger-Hartmann, T., Hillegass, J., Bellion, P., Vuorinen, A., Etter, S., Smith, B. P. C., White, A., Sterchele, P., De Smedt, A., Glogovac, M., Glowienke, S., O'Brien, D., & Parakhia, R. (2018). Making reliable negative predictions of human skin sensitisation using an in silico fragmentation approach. *Regulatory Toxicology and Pharmacology*, *95*, 227–235. https://doi.org/10.1016/j.yrtph.2018.03.015

OECD. (2015). *Test No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA)*. OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264229709-en

[5]適用可能/領域内にある情報源からのみ、合計スコアを算出した。

©OECD 2021

OECD. (2018). *Key Event Based Test Guideline 442E: In Vitro Skin Sensitisation Assays Addressing The Key Event On Activation Of Dendritic Cells On The Adverse Outcome Pathway For Skin Sensitisation*. OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris.

Yordanova, D., Schultz, T. W., Kuseva, C., Tankova, K., Ivanova, H., Dermen, I., Pavlov, T., Temelkov, S., Chapkanov, A., Georgiev, M., Gissi, A., Sobanski, T., & Mekenyan, O. G. (2019). Automated and standardized workflows in the OECD QSAR Toolbox. *Computational Toxicology*, *10*, 89–104. https://doi.org/10.1016/j.comtox.2019.01.006

付録 1：Derek Nexus 予測のプロトコル

110. ITSv1Defined Approach の *in silico* 情報源として、Derek Knowledge Base (KB) 2020 1.0 と Derek Nexus v.6.1.0 を用いて皮膚感作性ハザードに対する予測を導くため、以下のプロトコルを使用する。

Derek KB 2020 1.0 と Derek Nexus v.6.1.0 を用いて皮膚感作性ハザードに対する予測を導くためのプロトコル

単一化学物質

1. Nexus を立ち上げる
2. 以下の選択肢のいずれかを使用して構造を入力する：
    a. キャンバスに描画することによって、構造をマニュアルで入力する
    b. Go to File>Open Structure により、ファイル（ファイルには、.mol、.sdf、.smi、.csv、.cdx などがある）から構造ひとつを入力する
    c. Go to File>Type Chemistry により、SMILES、InChi または MOL ファイルを入力またはペーストする
    d. Go to File>New Structure により、構造を描画して構造を入力する
3. 予測を設定する
    a. Go to Prediction>Derek Prediction>Derek Prediction Setup
4. 処理制約の適用
    a. 知識ベース
        i. Nexus v6.1.0 では、必ず Derek KB 2020 1.0 を選択する
        ii. 最近リリースされたものに関しては、提供されているデフォルトの Derek KB を使用する
    b. 認識
        i. 「Perceive tautomers」および「Perceive mixtures」が選択されていることを確認する
        ii. 「Match alerts without rules」が解除されていることを確認する
    c. 種
        i. 「mammal」を選択する
    d. エンドポイント
        i. 「Deselect all」をクリックしたあと、「Skin sensitisation (ALL)」を展開して「Photoallergenicity」と「Skin sensitisation」を表示させる「Skin sensitisation」を選択する
    e. 構造特性
        i. logP、logKp のボックスが「Overwrite」されていることを確認し、平均分子量が解除されて Derek Nexus が算出した値が用いられているか、そうでなければ「Overwrite」ボックスを確認し、独自の値を入力する。
5. 予測を導く

©OECD 2021

a. 「Start Prediction」をクリックする
b. アラートが発生した場合：知識ベース、エンドポイント、種、推論レベル、アラート生成、EC3 予測（該当する場合）、対応する例（該当する場合）が、prediction navigator に表示される。
    i. 確からしさ(certain、probable、plausible、equivocal)をクリックすると、確からしさのレベルをもたらす推論ルールが表示される。
    ii. prediction navigator のアラートをクリックして、alert match、説明画像、コメント、バリデーションコメント、エンドポイント、参考文献、パターン、アラートに関する例を表示する。
c. アラートが生じない場合、陰性予測となる：知識ベース、エンドポイント、種、negative prediction reasoning（非感作性物質）および negative prediction overview （misclassified feature または unclassified feature のありまたはなし）が、prediction navigator に表示される。
    i. Negative prediction overview (「No misclassified or unclassified features」、「Contains misclassified/unclassified features」)をクリックして、陰性予測に関する情報を表示させる。類似最近傍を利用して misclassified feature を表示させることができる。
d. Derek の確からしさを用いて、各化合物を陽性または陰性に分類することができる(certain、probable、plausible または equivocal でアラートが生じた場合は陽性に分類し、doubted、improbable、impossible でアラートが生じるか、misclassified feature や unclassified feature がない非感作性の陰性予測を陰性に分類する)。
    i. Misclassified feature ないし unclassified feature のある非感作性の陰性予測は信頼性が低く、ITSv1 に使用しない。
    ii. 複数のアラートがあるか、混合物中の構造が別の確からしさを示す場合、最も控え目な分類を適用する（陽性 > 陰性）。
    iii. Derek による陽性の結果を ITSv1 の 1、陰性結果を 0 とスコア付けする。

複合化学物質

1. Nexus を立ち上げる
2. 構造を入力する
    a. Go to File>Open Structure により複数の構造を含むファイル(ファイルには、.mol、.sdf、.smi、.csv、.cdx などがある)を入力する
    b. 予測時に使用する構造特性にマップされるファイルからフィールドを選択する（名称、平均分子量、LogP、LogKp）。変更せずにいると、Derek が設定した値を使用することになる。
3. バッチ予測を設定する
    a. Go to Prediction>Derek Prediction>Derek Batch Setup
4. 処理制約の適用
    a. 知識ベース
        i. Nexus v6.1.0 では、必ず Derek KB 2020 1.0 を選択する
        ii. 最近リリースされたものに関しては、提供されているデフォルトの Derek KB を使用する
    b. 認識
        i. 「Perceive tautomers」および「Perceive mixtures」が選択されていることを確認する

©OECD 2021

ii. 「Match alerts without rules」が解除されていることを確認する

c. 種
   i. 「mammal」を選択する

d. エンドポイント
   i. 「Deselect all」をクリックしたあと、「Skin sensitisation (ALL)」を展開して「Photoallergenicity」と「Skin sensitisation」を表示させる「Skin sensitisation」を選択する

e. レポートの設定
   i. ディレクトリ - デフォルトディレクトリまたはマップとして任意の場所に残す。
   ii. タイプを選択 - バッチのためのレポートを選択する（左側アイコン）
   iii. フォーマットを選択 - 希望するファイルタイプを選択する（Excel など）
   iv. デザインを選択 - 希望するデザインを選択する（表形式レポートなど）
   v. ファイル名 - 希望するファイル名を入力

f. レポート表示オプション
   i. 「Show predictions of at least impossible」が選択されていることを確認する
   ii. 「Show Negative Predictions」を選択する
   iii. 「Filter All Nearest Neighbours by Misclassified Features」を選択する
   iv. 「Show Open Likelihood」を選択する
   v. 「Show Rapid Prototypes」を選択する

5. バッチ予測を導く

a. 「Start Batch Prediction」をクリックする
   i. バッチ予測が終了し、プロンプトが表示されたら、「Open Report Directory」を選択する

b. Derek の確からしさを用いて、各化合物を陽性または陰性に分類することができる (certain、probable、plausible または equivocal でアラートが生じた場合は陽性に分類し、doubted、improbable、impossible でアラートが生じるか、misclassified feature や unclassified feature がない非感作性の陰性予測を陰性に分類する)。
   i. Misclassified feature ないし unclassified feature のある非感作性の陰性予測は信頼性が低く、ITSv1 に使用しない。
   ii. 複数のアラートがあるか、混合物中の構造が別の確からしさを示す場合、最も控え目な分類を適用する（陽性 > 陰性）。

c. Derek による陽性の結果を ITSv1 の 1、陰性結果を 0 とスコア付けする。

©OECD 2021

付録 2：OECD QSAR Toolbox 予測のプロトコル

111. ITSv2Defined Approach の *in silico* 情報源として使用される皮膚感作性の Defined Approach（DASS AW）に対する自動ワークフローと OECD QSAR Toolbox v.4.5 を利用し、以下のプロトコルを用いて皮膚感作性ハザードの予測を導くことができる。

Toolbox 4.5 の DASS AW を使用して皮膚感作性ハザードの予測を導くためのプロトコル。

ステップ 1：「Input module」に化学物質を入力する。構造の入力には SMILES が推奨される。（CAS 番号など、ほかの識別子を入力に用いる場合は、Toolbox が内部データベースに基づいて SMILES を割り当てる。この場合、ユーザーは、Toolbox が正しい構造を特定し、予測に使用するよう徹底する。）

ステップ 2：「Data gap filling module」にジャンプし、「Automated」ボタンをクリックする。「EC3 from LLNA or Skin sensitization from GPMT assays for defined approaches」を選択し、OK をクリックする。実装したロジックとスキームが表示される。

ステップ 3：Run ボタンをクリックする - または、キーボードの F5 キーを押して、「Yes」で確定する。ワークフローが自動的に実行される。

ステップ 4：Read-across の結果、物質が「陽性」または「陰性」と予測されると、この予測は結果の前に「R」を付けてデータマトリックスに表示される(“R:Negative のように)。プロファイリングの結果、物質が「陽性」または「陰性」と予測されると、この結果は、カスタムプロファイラー名「Skin sensitization for DASS」の隣に表示される。

ステップ 5：DASS の自動ワークフローの領域に対する物質の属性が、自動的に決定され、表示される。

©OECD 2021

付録 3:OECD QSAR Toolbox の適用領域に関する情報

技術的側面

112. DA ITS v.2 が利用する Toolbox 予測は、OECD QSAR Toolbox v.4.5 の DASS 自動ワークフローを用いて算出される。このワークフローは、以下に挙げる Derek 皮膚感作性の適用領域の自動計算を含む。

Toolbox の in silico 領域の算出

113. ITSDefined Approach 自動ワークフロー（DASS AW）で使用する QSAR Toolbox 皮膚感作性予測の適用領域は、同じ自動ワークフローのトレーニングセットの物質に基づいて規定される。トレーニングセット(TS)は、LLNA ないし GPMT 皮膚感作性実験データ[6]（全物質のリストを QSAR Toolbox で閲覧可能）がある 2268 種類の物質からなる。以下に挙げる OECD QSAR Toolbox データベースの一部が TS 物質である。

- 皮膚感作性
- REACH 皮膚感作性（正規化）データベース

114. 正しく予測されたトレーニングセット物質に基づいて、Toolbox は 3 層の適用領域を自動的に計算する:これは 1)物理化学的特性層; 2)化学構造層および 3)作用機序層である。Toolbox 予測アプローチ（read-across 予測またはプロファイリング予測）および予測結果（陽性または陰性）に応じて、ここに挙げた層の 1 つ以上を考慮に入れて、特定の予測の Toolbox 領域全体を明らかにする。

115. 各種 Toolbox 予測で考慮する適用領域層を、以下の表にまとめる：

Toolbox DASS AW の結果		適用領域層		
		化学構造	物理化学的特性	作用機序
陽性	Read-across	考慮しない	考慮しない	**考慮する**
	プロファイリング	考慮しない	考慮しない	**定義により満たす**
陰性	Read-across	考慮しない	考慮しない	**考慮する**
	プロファイリング	**考慮する**	**考慮する**	**定義により満たす**

116. 別の領域層の使用に関する理由と根拠：

1. 陽性予測（read-across によるもの、プロファイリングによるもののいずれも）：アラートがひとつあること（陽性 Toolbox 予測が作用機序領域内にあると判断するための要件）は、予測が Toolbox 領域内にあると判断するのに十分である。アラートを発する物質を、領域内にあると判断する。その理由は、実験で皮膚感作性を認める toxicophore を有するからである。予測が Toolbox *in silico* 領域内にあると判断するのに、これ以上のチェックは不要である。

[6]ひとつの物質に複数のデータポイントがある場合、最も控え目なシナリオを採用する。

©OECD 2021

2. Read-across による陰性予測：化学構造領域および物理化学的特性領域は考慮しない。その理由は、Toolbox はすでに、トレーニングセットの別の物質と、ある程度似ていることを認めているからである。このような物質は、read-across に適した類縁化合物として選択される要件を満たす（この要件については、DASS AW の記載のなかで詳しく説明されている）。

3. プロファイリング予測による陰性予測：Toolbox 予測で最高の信頼性レベルを確保するため、全領域層を考慮に入れる。一層厳格な要件が必要である主な 2 つの理由を下に挙げる：1.アラートがないことは、感作性がない証拠にならない。2.陰性予測を採用すると、偽陰性予測の場合にヒト健康保護レベルが低下する可能性がある。

適用領域層の計算

1. 物理化学的特性層

物質の 4 つの物理化学的特性、すなわち logKow、分子量、蒸気圧および水溶解度を考慮に入れる[7]。選択した物理化学的特性の変動幅を、DASS AW が正しく予測しているトレーニングセットの物質に基づいて決定する。

QSAR Toolbox で求めた当該物理化学的特性値が下記表に示した変動幅に収まる場合、その物質は DASS AW の物理化学的特性領域内にあると判断する。なお、Toolbox に導入されている EPISuite モデルを用いて計算した物理化学的特性値の範囲が、既存の試験法の範囲より広い場合がある。

物理化学的特性	算出した物理化学的特性の範囲
Log Kow	-9.66 ÷ 18.6
分子量	16 Da～2290 Da
蒸気圧*	0 Pa～3.45 x 10^7 Pa
水溶解度	2.48 x 10^{-15} mg/L ～1.00 x 10^6 mg/L

*EPIWIN 蒸気圧(アントワン法)を算出に用いる。

2. 化学構造層

化学構造層は、DASS AW で正しく予測されている TS 物質の化学構造特性から求めた[8]原子中心フラグメント（ACF）に基づいて規定される。

ACF は、ACF のための下記 Toolbox デフォルト値に基づいて規定される：

- 全原子間距離 = 1

[7] 物理化学的特性の計算には QSAR Toolbox を使用する。

[8]正しく予測された TS 被験化学物質から抽出された ACF はいずれも“good space”である。“bad space”は、誤って予測された被験化学物質にある ACF から形成される。ACF に対するデフォルトの QSAR Toolbox 設定を使用する。Good space および bad space を形成する ACF と補助ファイルを入手できる。

©OECD 2021

- ヘテロ原子間距離 = 1
- C (sp^3)フラグメント抽出 = あり
- 全芳香環組入れ = なし

それぞれの物質について、以下の値を計算する：

- 正しいフラグメントの%：トレーニングセットのうち、正しく予測された構造を生じる ACF の割合
- 誤ったフラグメントの%：トレーニングセットのうち、誤って予測された構造を生じる ACF の割合
- 未知フラグメントの%：トレーニングセットにない ACF の割合

ACF の 100%が正しいフラグメントに属する場合、物質が DASS AW の化学構造領域内にあると判断する。

3. 作用機序層

（非）生物的活性化なしおよびありで予測した物質の皮膚タンパク質との相互作用能を考慮に入れる。このような相互作用を予測するのに、Toolbox エンドポイント別プロファイラーである *Protein binding for skin sensitization by OASIS* と、2 つの代謝シミュレータ – *Autoxidation simulator* および *Skin metabolism simulator* を使用する。

（非）生物的活性化なしまたはありで、物質が「*Protein binding for skin sensitization by OASIS*」アラートを生じる場合、作用機序領域で陽性予測と考える。

エキスパートレビューができない物質の場合、作用機序領域で陰性予測と考える。このようなものを、（非）生物的活性化なしまたはありにおける ITS「*trigger Protein binding for skin sensitization by OASIS（OASIS に基づく皮膚感作性に関するタンパク質結合誘発）*」での使用に関して領域外であると考えるのが妥当である。

117. なお、プロファイリングの結果得た予測は、作用機序層要件を定義により満たす。その理由は、プロファイラーによる陽性 Toolbox 予測は、アラートの存在によって正確に得られるためである。2 つ以上の情報源(*in chemico/in vitro ないし in silico)* で、被験化学物質を検討できないか、結果/予測を考慮できない場合、DA を適用できない。

©OECD 2021