

第2章 臨床評価のための アトラス



臨床評価のためのアトラス

なぜ臨床評価のために 画像スケールを開発する のか？

美容業界においては近年、施した処置の効果について評価が行われるようになりましたが、この評価は皮膚科分野でもまた大きな関心を集めています。実際、美容整形外科、皮膚科または美容治療の分野は長い間、効果の測定にほとんど関心を示していませんでした。というも、この分野の治療では期待される結果が確実に目に見える、あるいは見えるはずであると考えられていたためです。

効果が明らかに目で確認できる場合は、判断が揺らぐ余地はありません。しかし、有効性と効果に対する客観的証拠は常に求められており、過去30年間にわたって、機器による測定から心理感覚的アプローチを通じた臨床試験まで、さまざまな方向性の評価ツールや方法論が生み出されてきました[47,48]。このような要求はまた、厳正にかつ広く行われている評価方法で客観的に実証される、効能効果の訴求(クレーム)の発展からも生まれてきました。これらの方法では、機器ツール(顔の狭い領域での高精度な測定)と臨床ツール(顔の広い領域での視覚情報の定量化)の両方を同時に使用し、さまざまな情報を組み合わせて総体的な評価を行っています。

ロレアルリサーチ&イノベーションにおける過去の4つの研究は、欧州系[30]、アジア系[31]、アフリカ系[32]、インド系[33]の男性と女性の顔の皮膚の老化に関する一連のスキンエイジングアトラスとして結実しました。さらに、このアトラスシリーズの第5巻では、欧州系とアジア系の女性の顔と身体における日光曝露習慣の影響や皮膚老化の臨床徴候について論じ、定量化しました。日本人を研究対象にした本書は、『スキンエイジングアトラス第2巻アジア系編』、『スキンエイジングアトラス第5巻フォトエイジングアトラス フェイス&ボディ編』を基に、新しい診断スケールと知識を加えて構成されています[31,34]。

本書は日本人女性に焦点を合わせ、網羅的で正確な皮膚老化の臨床的特徴を構築するため、顔と首および手の老化の詳細な徴候を列挙しています。結果として、この特定のプロトコルに基づいた診断スケールは、高い精度と再現性で、日光曝露の影響を定量的に診断することを可能にしました[46]。本書は検証された直線的な画像スケールを提供することで、内因性か外因性にかかわらず、あらゆる種類の皮膚老化の徴候の評価を、信頼性が高く、標準化され、再現性のあるプロセスの中で実施することを可能にします。本書はこれまでのシリーズと同様のレイアウトを採用しています。各徴候の記述は顔の上部から始まり、首、そして最後に手に至ります。それぞれの徴候では、その特徴、評価すべき正

臨床評価のための アトラス

確な位置、評価すべきパラメータについて簡単に説明しています。

本書で使用されている徴候は、主に3つのグループに分けることができます。

- シワと皮膚の表面の状態: 目尻のシワのような表情線やほうれい線などの皮膚組織の下垂によるシワと肌の質感の変化
- タルミと下垂: 輪郭のぼやけなどの、皮膚組織のゆるみ
- 色素沈着異常: 色素沈着や微小血管の不均一性または色素異常症による皮膚の色の均質性の喪失

ある一つの民族集団の身体の特徴、または環境に関する習慣の影響を正確に記述している本シリーズのような参考文書は、さまざまな面において大きな価値を持っています。まず、化粧品の効能・効果を実証するプロセスにおいて、評価対象の選択基準やそれに伴う外見を限定することが可能になります。また大規模集団の幅広い加齢の臨床徴候が示されているため、類型学的(タイポロジー)アトラスとしての正確さが高くなっています。これにより、評価対象をより正

確に設定でき、かつ測定の際のばらつきが大幅に減少して効果・効能の識別力が向上します。これらの要素は、規制当局が研究の信頼性を判断する際の極めて重要な事項です。評価対象についての標準化された基準は、多施設研究または疫学的研究で最も重要な要素であり、異なる国での結果を比較可能にします。また、長期的なタイポロジー研究では、これらの診断スケールの使用によって、調査対象集団を正確に記述することが可能になります。また、本書のもう1つの価値は、再現性が高く、信頼できる標準化された臨床評価ツールを提供していることであり、専門家や臨床医がそれぞれの結果を比較できるようにします。

多くの文献で、評価を向上させるための画像チャートの使用が取り上げられています[26,29,49-53]。臨床医の目視による評価は実際、皮膚の観察には非常に正確かつ敏感ですが、治療効果を長期にわたって定量的に評価するには、評価ツールとしての基準スケールが必要です。このような基準がない場合は、評価者の臨床評価の記憶、位置付け、比較、標準化が非常に難しく、研究の質に疑いを生じさせることになるでしょう[54,55]。

臨床評価のための アトラス

臨床スケールの設定

『スキンエイジングアトラス』の制作過程については、既に発刊されている巻[30-34]に述べられています。それぞれの臨床徴候について最も程度が小さいものをグレードを0とした、最大10段階の直線的な定量的スケールを代表的な写真で表しており、すべての画像は、臨床症状の研究のために撮影した写真を用いています。

このスケールは、画像解析による皮膚の老化徴候に関する2つの研究プロジェクトの結果を集約したものです。上海(中国)と東京(日本)のロレアル リサーチ&イノベーションセンターで行われたこの研究には、18歳から80歳までの健康な女性ボランティア250名にそれぞれの地域でご協力いただきました。

被験者の撮影は、一定の温度(21±1°C)と湿度(50±5%)を保った部屋で、特別にデザインさ

れた撮影装置の標準条件で行いました。顔の両側2mの距離にフラッシュライトをそれぞれ配置し、Nikon 105mmレンズを装着したNikon D300デジタルカメラを用い、ISO感度200、シャッタースピード1/60秒、最小口径f/32で撮影しました。それぞれの被験者について、鼻を中心とした正面全顔1パターンと、鼻梁を基準として45度と90度の角度で左右両側から4パターン、計5パターンの撮影を行いました。撮影中、すべての被験者は白色の綿キャップを着用し、臨床徴候の評価でそれぞれの徴候を全顔で判断する際に、毛髪のボリュームや色などが影響を及ぼさないようにしました。各画像は被験者の氏名と番号で管理し、撮影終了後は専門現像所で光沢紙にプリントしました。

手の撮影では、皮膚の質感と均質性を担保するために、特別な撮影台を作成しました。スタンドRS3XA(Kaiser)の上に24/120mm対物レンズを装着したNikon D300デジタルカメラと環状フ



図2 - 撮影条件を統一した手の甲の画像

臨床評価のための アトラス

ラッシュSB-29s(Nikon)を設置したものです。被験者は椅子に座り、手の平と前腕の手首との角度が0度になるよう撮影台の上に両手を置き、その左右の手の甲をプロの写真家が撮影しました。例を図2に示します。

臨床スケールを設定するプロセスは、過去に出版したアトラスと同一です。重症度と直線的に対応するスケールを作成するため、撮影したすべての写真は5名の専門家が観察し、徴候とその部位を特定し、項目ごとに最も代表的な写真を選択しました。この段階で、専門家メンバーが評価基準の定義や基準の評価法を決定しました。また、ここで取り上げられた多数の画像から、さまざまな徴候について8段階から12段階のグレードを作成しました。次に、各々の徴候について定義された基準に沿って、各グレードが明確な進行を示すよう、最適な画像を選択しました。そして、後述する「一対比較法」に基づいた方法で検証し、最終的なグレード数を確定し、画像を選択しました。画像の最終確定前に、評価基準に影響のない、小さな傷などを画像編集ソフト(Photoshop®)を使って除去し、該当部位を中心にリフレーミングし、臨床評価に必要な情報のみに絞り込みました。このリフレーミングは生体に現れた徴候には影響を与えず、評価の対象のみを分離することを目的としています。

臨床スケールの検証

写真によるスケールの直線性の検証は、臨床医が治療効果を正確に評価するために必須です。適切な検証方法として一対比較法を採用し、特に専門知識を持たない16名からなる判定メンバーにより、照明、位置、較正方法などに関して完全に標準化された条件で検証を行いました[56,57]。

ある臨床徴候について、合計N個のグレード写真から得られた2枚の写真のあらゆる組み合わせを、高解像度24インチディスプレイ(1920×1200ピクセル、測色計で較正)に提示します。この画像を判定者一人ひとりが、ディスプレイから1m離れた位置で観察し、それぞれの組み合わせごとに進行度が高い方の写真を選択します。この選択には二肢強制選択法が採用されており、判定者はどちら画像の進行度が高いかを判定し、高い方はスコア1、低い方はスコア0とします。これにより、 $N(N-1)/2$ 個のランダムなペア画像が目視によって検討されます。それぞれの臨床徴候については、すべての判定者に同じ情報が与えられるように、オペレーターがその定義を読み上げ、またディスプレイには評価すべき箇所を正確に示すための視覚的な補助を提示しました。データの取得には特別に開発され

臨床評価のための アトラス

たソフトウェアを用い、結果はその後の統計分析に適した表形式に変換しました。

この手順では、より進行した徴候を示す画像がより高いスコアを得ることになり、各画像が得たスコアの合計点に基づいて進行度の順番が確立されます。グレードがN段階ある場合、0からN-1の範囲にレベル分けされることになります。この段階で、各観察者の間で結果が一致していれば、観察者がグレードの定義を完全に理解しており、その適用において大きな差異がないことを意味します。

明らかな偏差が生じていなければ、分散分析とRyan-Einot-Gabriel-Welschテスト(多重比較)を行い、それぞれの臨床スケールにおいて異なるグレードの間隔の有意性を確認します。

エイジングアトラス作成のための専門家グループは、各臨床スケールの二つの主要な情報項目についてさらに確認を行います。一つ目は、グレード別に分けた画像の順序が間違っていないか、間違った場所に配置されていないかの確認です。二つ目は、「正しい」順序で配置されているが区別が不明確、すなわち重複したグレードの確認です。この方法論の利点は、専門家グループによる調整(グレードの再編成または削除)後、すぐに新たに統計分析を行い、その結果を検証あるいは棄却することが可能な点です。

すべてのプロセスが成功裡に完了すれば、この画像スケールが顔の領域と徴候に基づいて臨床医が評価・検討するために適正であると判断します。

臨床評価のための アトラス

臨床スケールの応用と 使い方について

『スキンエイジングアトラス』は、主に二つの場面で利用できます。一つ目は、臨床医が患者の皮膚を評価する場合です。評価は正確な手順に従って適切な照明の下、診察室で行わなければなりません。光の当たり方により、評価すべき徴候の見え方が大きく左右されるため、患者がいる位置も重要になります。患者に対する医師の位置もまた大切で、研究手順書中に記述しなければならないポイントです。診断中の患者の顔の表情、視線の方向およびヘアスタイルなども、適切な診断を行うには重要な要素です。顔や首の症状はそれぞれ適切なスケールを用いて、続けて評価します。そして1回または数回の撮影を行います。これは後に同じ被験者を再度評価する際、位置合わせするためにも役立ちます。

二つ目は、類型学的(タイポロジー)研究においてです[58,59]。この場合、特別にトレーニングを受けた15名の専門家が、色校正されたスクリーンにランダムな順序で映し出される画像を評価します。これらの画像は『スキンエイジングアトラス』作成時に採用された条件と正確に同じ条件下で撮影され、アトラス内のグレード写真と範囲が等しくなるような位置でリフレーミングされています。これにはいくつかの利点があります。まず、評価者は評価すべき点以外の徴候に影響さ

れません。次に、15名の専門家が評価することにより、平均評価が得られ、主観的要因が除去され、評価の質が向上します。さらに、それに続く統計処理によって、専門家の評価間に一貫性があるかどうかを検証されます。各専門家の評価の信頼度は、多数の重複画像を入れ込むことで確認することができます。

状況によっては、直接臨床評価を行うほうが、この画像評価に基づくアプローチよりも効果的かもしれませんが、このような画像評価は、盲検法アプローチや有効性試験における測定時期のランダム化を可能にするため、皮膚科、美容治療または外科治療の効果を確立する研究で効果的です。

見かけの年齢について

私たちの外見はすべての年齢において必ずしも同じように見えるわけではありません。ある年齢において顔に現れる老化の徴候は、その位置、度合い、現れ方が人によってさまざまです。しかし、これらの徴候によって、全体としての見かけの年齢が特徴付けられます。

同じ民族グループに属する60名(未婚・既婚、職業を問わず、年齢18歳-60歳)に、同じ民族に属するさまざまな年齢の被験者の全顔写真を提示

臨床評価のための アトラス

し、「この女性は何歳だと思いますか?」と質問しました。同じ画像について得られた60名の回答を平均し、見かけの年齢を算出しました。これらをもとに、多変量解析により、被験者の各臨床徴候がどのような重みで見かけの年齢に影響を与えているかを示す指標を算出しました。このような指標は、人が他者の見かけの年齢を判断する

ために選択した、さまざまな臨床徴候に関する重み付けされたスケールの値の総和であり、年齢の判断基準(顔のどの部分に注目するか、徴候の進行度、年齢層における認識の変動など)をよりよく理解し、化粧品の効果、皮膚科治療、外科治療の効果について評価するために役立ちます。

