

【論文】

「正常値 normal value」から「基準値 reference value」への転換の
狙いは何であったのか——検査値の判断基準をめぐる歴史的分析古俣めぐみ¹

1. はじめに

2013年、日本人間ドック学会と健康保険組合連合会が共同して、人間ドック受診者のデータをもとに新たな「基準範囲 (reference interval²)」(後述)、つまり検査値の判断基準の一種を作成することとなった。これについて2014年4月4日に厚生労働省でプレスリリースが行われたのだが、当日の夜以降、このことはテレビ番組や新聞、週刊誌などで大きく取りあげられた。新聞には“人間ドック学会、健康基準を緩和”、“健康基準広げます”、“男性の中性脂肪高くても健康”などと記載され、全国的に大きな反響を呼んだ。その後、日本医師会や日本医学会の専門の医師から、健康基準が緩和されたと安易に報じるべきではないという批判もなされた³。

この一連の事態の背景には、「基準範囲」が「健康基準」を示しているとマスコミが(あるいはもっと広く非専門家が)認識している一方で、専門家による定義のうえでは、「基準範囲」は「健康基準」を示しているものでも、「健康」と「病気」あるいは「正常」と「異常」を分けるためのものでもないというギャップがある。三井田(2020)によれば、基準範囲とは、健常と考えられる集団から測定で得られた検査値の中央95%が分布する範囲であると定義される。現在では、米国臨床検査標準協会(Clinical and Laboratory Standard Institute;

1 東京大学大学院 総合文化研究科 科学史科学哲学研究室 博士後期課程/早稲田大学 現代政治経済研究所 研究補助者 Email: megk0515@gmail.com

2 直訳すれば「参照範囲」あるいは「参考範囲」になるが、日本では「基準範囲」の訳語が用いられてきたため、本稿もそれに倣う。「参照範囲」という訳語を用いるべきだという意見も出されている(三井田2020)。

3 以上の記述は渡辺(2015)を参照した。

CLSI) の定義にしたがい、一定の条件を満たす「基準個体⁴」をある集団から抽出し、その試料から得られた測定値を用いて基準範囲を設定することが推奨されている⁵。これは「検査値を解釈する“健常”の目安(物差し)として利用される」(佐藤・市原 2020、827 ページ)のものであり、「検査値がこの基準範囲内にあるから健康であるとか、あるいは範囲外であるから病気であるとか、病気ではないというものではない」(渡辺 2015、552 ページ)⁶。

ここで歴史を遡れば、基準範囲はもともと「基準値 (reference value⁷)」の概念が提唱されるなかで定義されたもので⁸、この「基準値」とは、それ以前に慣習的に用いられていた「正常値 (normal value)」(あるいは「正常範囲 normal range」)に対する批判から生まれてきたものであった。正常値がいつどのようにして用いられるようになったのかは定かではないが、1968 年に正常値を批判

4 日本臨床検査医学会ガイドライン作成委員会編 (2018) によれば、「健常者のうち当該検査値に影響を及ぼす生理的変動や病態変動の要因を持たない個体」(p. 12)を指す。CLSI の指針においては、問診票を用いた健康状態からの逸脱のチェックによって基準個体としての条件を満たすかを判断することになっているが、必ずしも明確な健康の基準が存在しているわけではない。

5 なお、基準個体を抽出する過程で潜在的に異常値を示す可能性のある個体を除去できるため、基準個体の母集団は厳密に健常とみなされた集団である必要はない。

6 なお検査値の判断基準としては、基準範囲の他にもう一つ、臨床判断値という指標が広く用いられている。三井田 (2020) によれば臨床判断値とは、特定の病態の診断・予防や治療・予後について判定するための基準と定義され、診断閾値 (diagnostic threshold)、治療閾値 (therapeutic treatment threshold)、予防医学閾値 (prophylactic threshold) の大きく三つに分けられる。前川 (2020) によれば、まず診断閾値(カットオフ値とも呼ばれる)とは、特定の疾患や病態を診断するための閾値であり、症例対照研究によって疾患群と非疾患群の検査値の分布を調べ、最も高い判別能を有するとされた値に設定される。治療閾値とは、治療介入の必要性を示す限界値であり、臨床医学における長期的な経験則や症例集積研究で設定される。予防医学閾値(健診基準値とも呼ばれる)とは、生活習慣病などの予防医学的に扱われる疾患のリスクを判別するための閾値であり、コホート研究によって得られた検査値と発症率の関係から、当該疾患の専門家の合意(学会などのコンセンサス)によって設定される。臨床判断値のうち、特に予防医学閾値は、基準範囲と混同されやすい。

7 前述の「基準個体」から得られた(個々の)検査値を指す。直訳すれば「参照値」あるいは「参考値」になるが、日本では「基準値」の訳語が用いられてきたため、本稿もそれに倣う。なお reference value の訳語については「基準個体値」を用いるべきであるという提案がなされている(「基準範囲」に関わる用語についての検討ワーキンググループ 2019)。

8 「基準値」概念が提唱された当初は、検査値の判断基準の一つとして「基準値」を用いるべきだと主張されていたが、その後「基準範囲」を用いるべきだとされるように変化したと考えられる。その詳しい経緯に関しては今後の検討課題としたい。

するかたちで基準値の概念が提唱され、世界中で受容されるようになった（この点は第2節で詳しく論じる）。

臨床検査の専門家は、しばしばこの正常値から基準値への転換を次のような流れとして描写する。すなわち、“正常値は、検査値がその値に近ければ正常、その値から外れたら異常、という誤解を招くものであったために、基準値の概念が提案された”というものである。以下に例を三つ挙げる。

“正常値”という用語は今でも使用されている。しかし、正常値が正常を意味付ける用語と認識され、それから外れると正常ではなく異常であると誤解されることから、“基準範囲”という用語が生まれてきた背景がある。（前川 2020、1262–1263 ページ）

基準範囲は以前「正常範囲」と呼ばれていたが、「正常」と「異常」を必ずしも区別する値ではないので、この名称に改められた。すなわち、(1) 定義から「基準个体」の5%は異常値をとる、(2) 「基準个体」の選別で潜在病態を完全には除外しきれない、(3) 個人の「基準範囲」は狭く、基準範囲内でも個人別にみると異常な場合がある、などの理由から「正常範囲」という用語を用いると誤解を生じやすいためである。（市原 2015、718–719 ページ）

基準範囲が使われる以前は「正常範囲」という言葉が採用されていた。しかし、この用語は「検査値がその範囲に入っていれば“正常”、範囲外であれば“異常”」という意味に誤解されやすい。そこで、国際臨床化学連合(IFCC)が正常範囲という言葉廃止して「基準値」「基準範囲」に変えることを提案し、1995年発行の米国臨床検査標準化協議会(CLSI:旧NCCLS)による「基準範囲設定に関する指針」において採用された。しかし、現在でも基準範囲を正常範囲と誤解し、検査値を判断する境界値(threshold)として捉える傾向がある。（佐藤・市原 2020、827 ページ）

しかし、正常値あるいは正常範囲が基準値へ、そしてその後基準範囲へと変化

していくときに、本当にこの部分が論点になっていたのだろうか。言いかえれば、正常値概念が批判されて基準値概念が登場してくる際に、“正常値は正常／異常の基準であるという誤解を招くため、そのような誤解を招かないような指標を作ろう”と意図されていたのだろうか。もしそうでないとしたら、基準値は実際にはどのようなものとして考案されたのだろうか。これが本稿の問いである。

この問いに答えるために、以下のように分析を進める。まず第2節では、基準値が提唱され国際的に受容されるまでの経緯を概観する。第3節では、基準値は考案された当初どのようなものとして構想されていたのかを分析する。第4節では、基準値はどのような点で正常値と異なるものであると規定されていたのかを、正常値に対する批判から検討する。第5節では最初の問いに対する本稿の回答を提示し、今後の課題を述べる。

なお以下では、正常値 (normal value) や基準値 (reference value) といった用語について、訳語ではなく元の英語をそのまま記すこととする。日本語にしまうと議論が追いつらい箇所があるためである。

2. reference value の登場の経緯

本節では reference value の提唱から受容までの流れを概観する。どの時点を受容とみなすかについては、国際臨床化学連合 (International Federation of Clinical Chemistry; IFCC) による勧告の策定をその目安と捉えることとする。市原(2020)によればこの勧告は、用語の概念を整理して妥当な基準範囲を設定するための一定の方向性を示したという意味で大きな意義をもつものであり、国際的な組織によって基準範囲の理論構築が行われた初の成果であった。また渡辺(2015)によれば、この勧告が reference value という用語を導入したことによって WHO がそれを認知し、reference value が世界に広がっていくきっかけとなった。これらの記述を踏まえて本稿では、IFCC による勧告の策定を reference value という用語や概念が広く普及する契機とみなし、そこに至るまでの経緯を検討の対象とする⁹。

9 以下、Alström(1980)、Siest and Henny(2013)、Ozarda(2020)の記述を参照した。なお、reference

1968年までの間、検査値の判断基準としては慣習的に normal value¹⁰ が用いられていた。しかし normal value のもつ曖昧さなどへの批判意識から、フィンランドの医師・臨床化学者 Ralph Gräsbeck と化学者 Nils-Erik Saris の二人は、1968年の秋、それに代わるものとして reference value という新たな概念を考案した。その直後の冬に彼らは短い論文を執筆し、そこで normal value に代わって reference value の語を導入することを提案した (Gräsbeck and Saris 1969)。また彼らは、1969年のスカンジナビア臨床化学・臨床生理学学会の第12回大会でも reference value の導入を提案した。その結果、学会は Gräsbeck を委員長として reference value 委員会を組織した。

1970年には、IFCCが「reference valueの理論に関する専門家パネル」(Expert Panel on the Theory of Reference Value; EPTRV)を組織した。最初の座長は T. R. Whitehead が務めたが、一定以上の任期を終えると座長が交代する必要があったため、1976年には Gräsbeck が新たな座長となった。しかし、EPTRV はその後しばらく出版物を公刊することができずに批判を浴びた。その理由は、reference value や health といった語の意味をどう定義するかや、検査の試料の採取などの実践的な部分の詳細について合意に至らなかったためであった。

1972年には、第8回臨床化学国際会議において「reference valueの理論」と題されたシンポジウムが開催され、R. Dybkaer と Gräsbeck が世話人を務めた (Dybkaer and Gräsbeck 1973)。同時期には世界各地で reference value に関する小規模な研究グループや委員会が作られていった¹¹。reference value の概念は、特に

value の提唱から現在までの国際的な流れと国内における流れを包括的に論じているものとして市原 (2020) がある。

10 あるいは normal range という語が用いられることもあった。この value や range、あるいは interval といった語がどのように用いられていたのか、またどのような変遷を辿っていったのかについては今後の検討課題としたい。

11 Alström(1980)によれば、アメリカ、フランス、カナダ、デンマークのグループがこの時期活動していた。アメリカのグループは、数値の個人内変動についてや数値の統計的処理、臨床化学の分析における目標設定における原則を明らかにした。フランスでは reference value 委員会が Siest によって始められ、reference value の概念や用語、統計的原則など様々なトピックに関する文書が公表された。カナダのグループにおいては、PetitClerc と Munan が中心となって、Sherbrooke 大学で reference value が収集された。デンマークとアメリカのチームでは、Winkel と Statland が中心となり、個人内 reference value を設定したり教科書の作成に貢献したりした。また Siest and Henny(2013)によれば、1970年から2000年の間に reference value の研究に関わってい

Sunderman(1975)の公表以降、世界中の多くの臨床化学者に広まっていった。

1975年には、スカンジナビアの reference value 委員会が、成人の reference value の収集に関して、最初の勧告案を作成した (Alström et al. 1975)。この勧告案は、Gräsbeck が編集を行っていた IFCC ニュースレターの 1977 年 2 月号で紹介された。またこのニュースレターには、reference value の簡潔な紹介記事が掲載された (Gräsbeck 1977)。

1978年には、メキシコシティ臨床化学会議の後に Benson によってシンポジウムが開催された (Benson and Rubin 1978)。また同年には、IFCC の EPTRV から reference value の理論に関する予備的勧告が出された (Gräsbeck et al. 1978)。この年に EPTRV の座長が Gräsbeck から Solberg に交代したが、Gräsbeck は非公式なアドバイザーとしてその後も EPTRV に関わりつづけた。

1979年には臨床検査室基準国際委員会 (National Committee for Clinical Laboratory Standard; NCCLS) が、小児からの皮膚穿刺血の採取に関する基準案を作成した (Meites et al. 1979)。その翌年には、臨床検査医学における reference value に関する国際シンポジウムが開催された (Gräsbeck and Alström 1981)。さらに 1983年には、Gerard Siest と Joseph Henny が統計学者 Eugene K. Harris とともに会合を開き、統計的知識に基づいた reference value の発展を試みた。

1978年に予備的勧告を発表していた EPTRV は、1987年から 1991年にかけて、最終版となる承認勧告六つを公表した (Solberg 1987a, PetitClerc and Solberg 1987, Solberg and PetitClerc 1988, Solberg and Stamm 1991, Solberg 1987b, Dybkaer and Solberg 1987)。また EPTRV のメンバーは、個々人の論文でも reference value の普及に努めた¹²。

以上見てきたように、1968年に reference value の概念が提唱されて以降、各々の国から国際的組織まで様々なレベルで reference value に関する研究グループや委員会が作られ、シンポジウムなどで検討が重ねられていったことがわかる。その一つの到達点とみなされうるのが、IFCC の EPTRV が公表した六つの承認

たグループは次の四つに分けられる。(1)SFBC (Société Française de Biologie Clinique) およびフランス語圏のグループ、(2)スカンジナビアのグループ、(3)アメリカ・イギリスのグループ、(4)スペインのグループ。各グループはそれぞれで論考を発表しているが、それらの内容は必ずしも同じではない。

12 その中で最も著名なものともみなされているのは Gräsbeck and Solberg(1981)である。

勧告であるといえる。

第3節と第4節では、この承認勧告に加えて、EPTRVの座長を務めた Gräsbeck と Solberg が、承認勧告の中で扱われた話題の一部を提示・評価した論文である Solberg and Gräsbeck(1989) を参照し、reference value および normal value がそれぞれのどのようなものとして捉えられていたのかを分析する。

3. reference value はどのようなものとして構想されたか

前述のように「reference value の理論に関する承認勧告」は全部で六つある。パート1では reference value や関連用語を定義するとともに、その後の勧告で扱われる、reference value の生成・取り扱い・使用といった話題の導入を行っている。パート2から4は reference value の生成に関するものであり、パート2は適切な reference individual の集団の選択について、パート3は試料採取のための個体の用意と標準化された条件における分析について、パート4は適切な品質管理下での決められた検査手法による試料の分析について、それぞれ述べている。パート5では reference value の統計的分析と reference limit のパラメトリック推定・ノンパラメトリック推定¹³を扱っている。パート6は observed value との比較における reference value の用い方についてである¹⁴。以下では reference value がどのようなものとして構想されたのかを明らかにするために、他のパートの基礎となっているパート1「reference value の概念」(Solberg 1987a) の特に第2節から第4節の内容を、途中で適宜 Solberg and Gräsbeck(1989) を参照しつつ詳しく見ていく。

パート1の第2節「定義」では、よく使われる用語体系がしばしば曖昧であるという指摘がなされ、各用語の明確な定義が試みられている。八つの用語の定義は次のようになされている。

13 パラメトリック推定とは、母集団が何らかの分布（多くの場合正規分布）に従っているという前提が成立するときに用いる手法のことである。ノンパラメトリック推定とは、母集団が何らかの分布に従っているという前提が成立しないとき、あるいは母集団の分布に対して仮定を設けないときに用いる手法である。

14 以上の記述は Solberg and Gräsbeck(1989) を参照した。

- reference individual とは「決められた基準を用いて、比較のために選択された個体 (individual)」であり、「その個体の健康状態 (state of health) を定義することが通常重要である」(Solberg 1987a, p. 113)。これはつまり、reference value を獲得するために、ある基準に即して選ばれた個人のことを指すと考えられる。そして個人を選ぶ際には、その人の健康状態がどのようなものであるかをきちんと定めたり把握したりしておく必要があると述べられている (詳しくは後述)。
- reference population とは「全ての可能な reference individual を含む」もので、集合論の用語で言えば「全ての可能な reference individual の集合」である。また「通常未知数の成員をもち、したがって仮説的な存在」である。なお「一人の成員のみを含む」こと、つまり「ある個体が彼自身や他の個体の reference の役目を果たす」こともありうる (Solberg 1987a, p. 113)。これは、前述のようにある基準に即して選ばれうる人々全てを含む集団、つまり母集団を表していると考えられる。
- reference sample group とは「reference population を代表するために選ばれた適切な数の reference individual」、言い換えれば「reference population の部分集合」である。また「sample group の reference individual は可能であれば reference population からランダムに取りだされるべきである」(Solberg 1987a, p. 113)。つまりこれは母集団から選びだされた標本のことであろう。またその選びだし方については、ランダムであること (無作為抽出) が望ましいとされている。
- reference value とは「reference sample group に属している個体に関する、特定の種類の量の観察あるいは測定によって得られた値」である。なお後述の reference limit と「混同されるべきではない」(Solberg 1987a, p. 113)。すなわち reference value とは、ある基準に即して選ばれた人々に、特定の観察や測定を行った結果として得られる個々の数値であるといえる。「個々の」数値であるという点が後述の reference limit とは異なっている。
- reference distribution とは「reference value の統計的な分布」である。そして「reference population から得られた reference distribution に関する仮説は、reference sample group の reference distribution と適切な統計的手法を用

いて検定される (tested)」。また「reference population の仮説的な reference distribution の母数は、reference sample group の reference distribution と適切な統計的手法を用いて推定される (estimated)」(Solberg 1987a, p. 113)。これは、ある基準に即して選ばれた人々全体の集団に関する情報 (統計的な分布や特性値) を、標本の分布から求めることができると述べていると考えられる。

- reference limit は「reference distribution から導かれ、記述的な目的のために用いられる」。なお、「定められた割合の reference value が、定められた確率をもつ限界 (limit) よりも小さくなる、あるいは限界 (limit) と等しくなるように reference limit を定義することは一般的な方法である」。また「reference limit は reference value を記述するもので、解釈的な目的で用いられる様々な種類の decision limit (discrimination value) とは区別されるべきである」(Solberg 1987a, p. 113)。少しわかりにくいのがこれはつまり、得られた個々の reference value のうち定められた割合 (たとえば 95%) のものが、ある値以上あるいは以下 (または、ある値より大きいあるいは小さい) に含まれるような値のことであり、そしてこのような値は予め定められた確率 (たとえば 90%) の信頼区間をもつということであろう。また、ここでいう記述と解釈が何を意味しているのかに関しては後述する。
- reference interval とは「二つの reference limit の間の (between)、またそれを含む (and including)、区間である」(Solberg 1987a, p. 114)。
- observed value とは「特定の種類の量の値であり、観察あるいは測定によって獲得され、医学的判断を行うために生成される」ものであり、「reference value、reference distribution、reference limit、reference interval と比較される」(Solberg 1987a, p. 114)。つまりこれは患者から得られた検査結果の数値であり、予め獲得されていた reference value などの指標と比較することで医学的判断を行うものであると考えられる。ここでいう比較の意味については後述する。

これらの用語の関係は次のような図にまとめられている (図 1)。

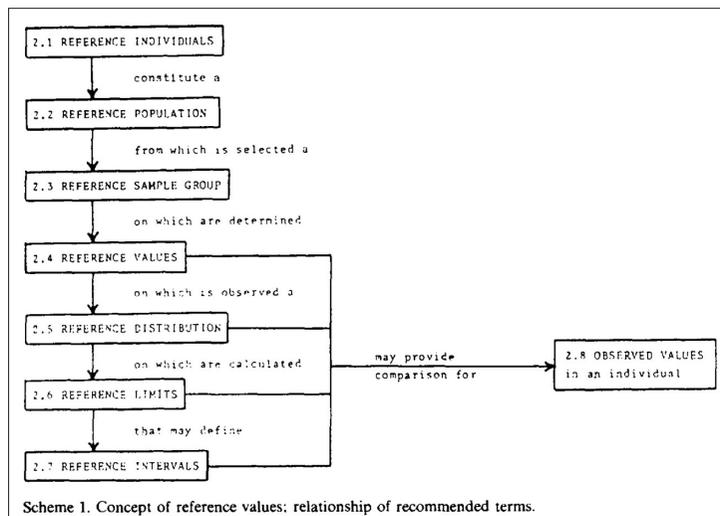


図1 各用語の関係図。Solberg 1987a, p. 114 より転載。reference individual が集まって reference population を構成する。reference population から選択された対象によって reference sample group が作られる。reference sample group から得られた（個々の）値が reference value である。reference value から reference distribution が得られ、そこから reference limit が計算され、それによって reference interval が定義される。reference value、reference distribution、reference limit、reference interval は observed value と比較される。

ここで reference limit について「reference value を記述するもので、解釈的な目的で用いられる様々な種類の decision limit (discrimination value) とは区別されるべきである」と定義されていたが、ここでいう記述と解釈とはどのような意味なのだろうか。承認勧告の中で扱われた話題の一部を提示・評価した論文である Solberg and Gräsbeck(1989) では、次のように説明されている。

記述的というのがここでのキーワードである。observed value が reference limit との関係で占めている位置にのみ基づいて臨床的判断を行うことはできない。しかし observed value が reference value の中で典型的かどうかは述べることができる。したがって reference limit は、decision limit あるいは discrimination limit と明確に区別されるべきである〔後略〕。(Solberg and Gräsbeck 1989, p. 11)

つまりここでいう「記述的」であることとは、「典型的かどうか」といったように、臨床的判断を伴わない説明であるということであろう。具体的には、observed value に対して、「病的か」などの臨床的な価値判断を行うのではなく、「典型的であるか」といった事実に基づいた説明を行うことが、ここで「記述的」と述べられていることの意味であろう。そしてその反対に「解釈的」とは、「病的か」などの臨床的な価値判断を行うという意味であり、decision limit/discrimination value はそのために用いられる指標であるとされている¹⁵。そして、「記述的」な目的のためには reference limit を、「解釈的」な目的のためには decision limit/discrimination value を用いるという役割分担の必要性がここでは指摘されているといえる。

また、observed value は「医学的判断¹⁶を行うために生成される」ものであり「reference value、reference distribution、reference limit、reference interval と比較され」と述べられていた。この部分に関しても、Solberg and Gräsbeck(1989)でより詳しい説明が与えられている。

[前略]全ての比較は二つの演算対象をもつ。解釈されるべきデータと比較のために用いられるデータ、つまりそれぞれ observation と reference である。医学的判断は、observation を一つのあるいはもっと多くの reference と比較することによって達成される。これは検査結果 (laboratory result) が解釈される場合も同様である。検査結果つまり observed value は、一つのあるいはもっと多くの reference value の集合と比較される。(Solberg and Gräsbeck 1989, pp. 4-5)

つまり医学的診断とは、それ以前に何かしらのかたちで得られていたデータと、目の前の患者から得られたデータとを比較することでなされるのである。そし

15 なお、本稿注6で説明している臨床判断値と、ここで述べられている decision limit/discrimination value とはどのように関係しているのかという点は重要であり、今後の検討課題としたい。

16 ここでいう「医学的判断」とは、前述の「記述的」なもの、つまり「典型的かどうか」といった価値判断を伴わないものであると推測される。

て「それ以前に何かしらのかたちで得られていたデータ」として適切に利用できるデータセットとして意図されているのが reference value に他ならない。

そしてそのような医学的判断の根底には、自然科学における実験をその模範とする考え方がある。

基本的なアイディアは、調査中の患者に対する適切な対照群であるような対象から値を用意すべきであり、それらの対照群は詳細に記述されるべきであるということである。これを表現する説教的な (didactic) 方法は、次のように述べることである。すなわち、病気は自然の実験であるとみなすことができうるし、また実験科学におけるように、実験の批判的な評価と反復が可能になるように陽性対照と陰性対照を十分詳細に記述すべきである。(Solberg and Gräsbeck 1989, p. 4)

ここからわかるように、患者に対して医学的判断を行う際には、自然科学における実験のように、対照群と患者との比較を行う必要がある。そしてそのような対照群は、予め定められた条件に基づいて選ばれたもの、そしてその患者と比較するのに適切なものでなければならないのである。

承認勧告パート 1 の検討に戻ろう。第 3 節「reference values の生成」では、まず、「生成しようとしている reference value の目的と用途を最初に明確に定義しなければならない」と述べられている。これは「reference value の生成のための手続きはこの目的と関連づけられているべきであり、測定される量の種類に依存する」(Solberg 1987a, p. 114) ためである。つまり reference value とは、特定の目的に応じて生成・使用されるものであって、絶対的なただ一つの基準とといったものではないといえる。

第 3 節で特に着目すべきは、健康の定義に関する部分である。勧告によれば、reference population (あるいは reference individual) の健康の定義については「どんな定義も十分なものではない」(Solberg 1987a, p. 115)。これは WHO による定義、すなわち「身体的、精神的、社会的に完全な幸福状態であって、単なる病気や虚弱の欠如ではない¹⁷⁾」に関しても同様である。また国や時代や年齢によって

17 World Health Organization, "Constitution," <https://www.who.int/about/governance/constitution> (最

もその定義は変わってくる。したがって健康とは「相対的な状態であって絶対的な状態ではない」(Solberg 1987a, p. 115)とされる。reference population (reference individual) は、予め質問表や検査を用いて健康を徹底的に評価される必要があるが¹⁸、「この文脈では、病気とは一つの「健康の状態」(‘state of health’)とみなされうる」。そのため、reference value などの用語における「reference」の語は「健康状態を修飾する言葉、たとえば、健康に関連した (health-associated)、選択されていない (unselected)、糖尿病の (diabetic) などの言葉」(Solberg 1987a, p. 115)とともに用いられるべきである。健康の定義については本稿第4節でも詳しく扱うが、ここで確認しておく必要があるのは、健康とは絶対的なものではなく相対的なものであり、病気は健康に関する状態の一つの形とみなされている点である。

承認勧告パート1の第4節「reference values の使用と提示」では、reference value の分析にあたって統計的手法を用いる際の前提が示されている。まず、「統計的手法はその基本前提を無視して用いられるべきではない」。具体的には、生物学的データは通常正規分布しているという考えは不適切であって、他の分布を前提にしている手法、あるいは分布に関係なく用いることのできる手法を用いるべきであると述べられている。また「observed value と reference value の比較の方法は、達成されるべき目的に応じて選ばれうる」(Solberg 1987a, p. 116)という点も明記されている。

第4節で重要なのは、「observed value と reference value の統計的有意差は、必ずしも医学的な有意差を示すわけではない」という言及であろう。それに続く「統計的有意性は記述的でしかない。解釈の重要性は生化学的・生理学的・臨床的判断に基づいている」という部分は、第2節に登場した記述と解釈をめぐる部分とも関連している。また「臨床検査科学者 (laboratory scientist) の中心的な役割は、関連する reference value をわかりやすいかたちで提示して、臨床医が observed value を解釈する手助けをすることである」(Solberg 1987a, p. 116)。つまり、reference value は臨床医による判断の有用な道具として用いられること

終閲覧 2022年2月19日)。

18 これは、reference value が患者から得られたデータに対する適切な対照群として機能するためには、reference population (reference individual) がどのような性質をもっているかを正確に把握しておく必要があるということであろう。

が想定されているといえる。

ここまで見てきた reference value の特徴は、以下の三つにまとめることができるだろう。一つ目は、reference value (あるいは reference limit) は記述的なもの、すなわちそれに基づいて臨床的な価値判断を行うのではなく、事実に基づいた説明を行うためのものだとみなされている点である。対して解釈的な、つまり臨床的な価値判断を行うために用いられるべき指標は decision limit/discrimination value であり、記述と解釈の役割分担がなされている。また、記述は統計の役割であり、解釈は臨床の役割であるという部分と関連して、reference value は統計的なもの、decision limit/discrimination value は臨床的なものと表現することもできるかもしれない。

二つ目は、reference value (あるいは reference distribution、reference limit、reference interval) は、observed value をそれと比較することによって医学的判断を行うための基準として意図されている点である。その根底には、実験科学を模範とした適切な比較を行うべきであるという考え方があり、reference value はいわば比較実験における対照群として詳細に記述できなくてはならない。

三つ目は、reference value は絶対的なただ一つの基準というわけではなく、特定の目的に応じて作成・使用されるものであり、だからこそ目的を示すような適切な修飾語とともに用いられるべきであるという点である。これは、reference population (あるいは reference individual) は健康に関してどのような状態なのかが定められていることが必要とされるという点とも関係する。ここで健康は絶対的なものではなく相対的なものとして捉えられており、病気も健康に関する状態の一部であるとみなされている。

4. normal value はどのように批判されたのか

次に本節では、reference value の前身である normal value に対してどのような批判がなされてきたのかを検討することで、normal value と reference value の違いを明確にする。

まず、承認勧告のパート 1 の導入部分では以下のように記述されている。

従来、臨床検査の結果は、伝統的に、しかし不適切に定義された「正常」値（'normal' value）との比較によって解釈されていた。しかしながら、生理学的・病理学的過程における生物学的変化への認識が高まったことによって、より正確で包括的な解釈が要求されている。（Solberg 1987a, p. 112）

これによれば normal value は「不適切」に定義された、不正確で非包括的な指標である。これは具体的にはどのような意味なのであろうか。承認勧告のパート5を見てみよう。

reference interval の二つの reference limit に基づいて、ある observed value は次のように階級づけできる。reference limit の下限より低いときには「非常に低い（unusually low）」、reference limit の間かあるいは reference limit のどちらかと等しいときには「普通（usual）」、reference limit の上限より高いときには「非常に高い（unusually high）」。（「異常な（abnormal）」、「病理的な（pathological）」、「正常な（normal）」といった用語は推奨できない、なぜなら「normal」の意味が数学的なものと生理学的なものとの混乱するかもしれないし、「abnormal value」は単に数学的な、非病理学的な存在でありうるためである）。（Solberg 1987b, pp. S37-S38）

つまり、normal などの語は「数学的なもの」や「生理学的なもの」など意味が複数あって、どの意味かが曖昧になってしまうために使用を控えるべきであると考えられているのであろう¹⁹。

ここまで勧告の内容を確認してきたが、勧告の中では normal value についてはあまり触れられていないことがわかった。一方、Solberg and Gräsbeck(1989)

19 ただし、パート6には次のような記述があり、勧告のなかでも用語の使用法が徹底されていないことがわかる。

検量線における変化を経時的に検出し、分析手法の特異度を管理することは必須である。このことは、それに対して reference value が生成される種類の量の濃度の、正常な（normal）レベルと病理的な（pathological）レベルの両方において、多量の正確な対照試料を必要とする。（Dybkaer and Solberg 1987, p. S10）

には冒頭にまとまった記述がある。以下ではこの部分の記述を詳しく見ていくこととする。

まず導入部では次のように述べられている。

臨床検査科学は疑いなく、臨床の領域の中で最も客観的かつ精密である。しかしながら〔中略〕臨床化学は〔中略〕科学的批判に耐えることができない側面をもっている。一つの例が「normal value」の概念であり、これはおよそ 20 年前、不明確すぎて近代的な要求を満たすことができないと考えられたものである。(Solberg and Gräsbeck 1989, p. 2)

これによれば、normal value が曖昧な言葉であると批判される背後には、臨床の領域にも「科学的」であること、具体的には客観性や精密性を備えていることが求められるようになってきている状況があることがわかる。

その続きの部分においては、normal value の歴史が端的にまとめられている。それによれば、近代的な normality 概念のルーツはプラトンのイデア概念に遡るもので、そこでは、normal とはイデアと一致することを意味していた。「平均人」概念を考案したベルギーの統計学者 Quételet²⁰ は、生物を測定したデータが

20 Quételet が『人間とその能力の発展について』で述べていることは、reference value の概念の先駆けと呼べるものだと思う。

平均人の検討は医学において非常に重要である。そのため、正常な状態 normal state/l'état normal にあるとみなされる他の架空の人物〔中略〕の状態と比較することなしに、ある個人の状態を判断することはほとんど不可能である。医者は病人に呼ばれて、彼を診察した後に、彼の脈拍が速すぎる、彼の呼吸が過度に頻繁である、などのことに気づく。そのような判断をするためには、観察された性質が平均人の、あるいは正常な状態にある人のそれと異なっているだけではなく、それらが安全性の限界を超えてさえいるということに我々が気づかなければならないのは明白である。どの医者も、そのような見積もりをするにあたっては、その科学 [= 医学] に関する既存の資料か、彼自身の経験を参考にする〔後略〕。(Quételet 1842, p. 267)

〔前略〕どの人も、自分自身を注意深く診察し、彼を構成する全ての要素と、健康な状態でそれらがその中で変動するような限界を決定するために、必要となる知識と慎重さをもつことを私は考えたい。彼は平均人のそれから多かれ少なかれ違っている表を作るだろう。それは、彼自身の場合に多かれ少なかれ異例なものや、緊急に注意を必要とするものは何でも、

「逆さにした花瓶 (vase invertée)」のように分布することに感銘を受けたが、このベル型曲線はガウス関数で適切に記述できると考えられた。臨床試験の結果はヒストグラムや分布曲線の形で容易に表すことができたため、20世紀初期に誕生した統計的な normality の概念は、臨床検査科学において特によく知られるようになった。このあまりにも単純な見方は、多くの血清成分の濃度の分布がガウス分布ではないということが発見されたことで否定された。また、血液中の脂質濃度と循環器疾患のリスクとの関係に関する研究によって、一般集団の中でよく見られる値が必ずしも望ましいものではないということも示された。これによって normal や normal value に対する批判が起こった。そして reference value の概念が提示されると、世界中の臨床化学者や血液学者に広く受け入れられた。このように normal value から reference value までの流れが提示されているのである。

続いて、normal value という言葉の曖昧さについての分析がなされている。それによれば normal value の曖昧さは、normal という語が様々な意味で用いられることに起因している。臨床検査医学では少なくとも次の三つの意味で用いられている。

1. 統計学的な意味：ある値が正規分布、つまりガウス分布のように分布しているように思われる場合、その値は normal であるとみなされる。このような normal の語の使用によって、生物学的データの分布が対称的でベル型であるかのような間違った考え方が生み出された。そのため、normal limit という語は用いないように推奨する人々が存在しており、正規分布 (normal distribution) はガウス分布 (Gaussian distribution) と言いかえるべきである。

彼が理解するのを助けるだろう。正常な状態からの逸脱の度合いや、何が最もとりわけ侵されている臓器なのかを推定するために、医者が病気の際に参考にすべきなのはこの表であろう。しかしより多くの場合、病人は彼自身や彼に特有の要素を満足に観察することができないので、医者は共通の基準に立ちもどり、患者を平均人と比較する義務がある〔後略〕。(Quételet 1842, p. 268)

この部分は Martin et al.(1980) の冒頭 (pp. 233–234) で英語訳が引用されている。日本語訳は、引用されている英語訳とフランス語の原文の両方を参照しながら筆者が行った。

2. 疫学的な意味:ある値が一般集団の中で典型的に見られる場合、その値は normal であるとみなされる。逆に、ある値が一般集団の中で非典型的である場合、その値は abnormal であるとみなされる。observed value が normal limit と比較されるときに normal と abnormal という語が使われる。この文脈では、normal は common、frequent、habitual、usual、typical と言い換えられるべきである。
3. 臨床的な意味:ある値が、病気のリスクのなさ、あるいはリスクの低さと関連づけられることを示す場合、その値は normal であるとみなされる。このようなときは healthy、nonpathological、harmless と言い換えられるべきである。

これら三つの normal の意味は、しばしば互いに衝突する。典型的な値であっても、病気のリスクのなさを示すわけではなかったり、ガウス分布でなかったりすることはありうるためである。

以上見てきたような normal value から reference value までの流れと normal の語の多義性を合わせて考えると、normal value がどのように曖昧であるとされてきたのかがより詳しく理解できる。つまり、ある時期までは（統計的な normality 概念が誕生してからしばらくの間までは）、normal という言葉が複数の意味をもっている、それぞれがある程度重なりあっているように考えられていたと思われる。しかし、血清成分の濃度分布や血液中の脂質濃度と循環器疾患のリスクとの関係などに関する研究が進展することによって、normal の意味のうち、あるものには当てはまるが別のものには当てはまらない、という状況が生まれるようになり、normal という言葉の曖昧さが顕在化していったと考えられる。

Solberg and Gräsbeck(1989)にはもう一つ、normal value に関する注目すべき記述がある。それによれば、伝統的な normal value は通常、検査室の職員や医学生などの、健康で若くて自分で移動できるような (ambulant) 人から集められているにもかかわらず、寝たきりの高齢者から得られた値を判断するために用いられており、ほとんど関係がない集団同士が比べられている。しかも試料が採取されるとき条件 (断食しているか、座っているか、試料採取の際に止血帯は用いられているかなど) は通常全くわからない。さらに、normal range が与えられても、それが観察された中の両極端の値なのか、統計的に決定された区間

なのかは曖昧である。これはつまり、normal value や normal range は標準化された定義や手続きのもとで存在していたわけではなかったために、observed value との比較に用いるという目的にはそぐわないと考えられるようになったということであろう。

ここまで、normal value がどのような点で批判されてきたのかを分析してきた。一言で言ってしまえば「normal value は曖昧である」という点に集約されるのだが、その曖昧さは二つの側面から説明できる。一つは、normal という言葉の意味が複数存在しており、どの意味で用いられているかがはっきりしないということである。20世紀初頭くらいまでは複数の意味がそれぞれ重なりあっている場合もあったものの、生理学や疫学の研究の進展によって、normal の複数の意味のうち、ある意味には当てはまるが別の意味には当てはまらないという状況が生まれるようになり、言葉の意味の曖昧さが顕在化していったと思われる。もう一つは、第3節で述べた「比較」という点とも関係するが、normal value や normal range という数値の定義や、数値の採取および理論化の手続きが明確に定められていないということである。臨床の領域でも「科学的」であること、すなわち客観性や精密性を備えていることが求められるようになり、observed value との比較も「科学的」に行われる必要があると考えられたなかで、様々な定義や手続きのもとで集められたデータである normal value は、比較対象として不適切であるとみなされたといえる。

5. おわりに

本稿は、normal value 概念に代わって reference value 概念が登場してくる際に、“normal value は正常／異常の基準であるという誤解を招くため、そのような誤解を招かないような指標を作ろう”という意図が果たしてあったのかどうか、もしそうでないとしたら、reference value はどのようなものとして考案されたのか、という問いに答えることを目標としていた。

本稿で取りあげた IFCC の EPTRV による承認勧告、および EPTRV の元座長の二人による論考に基づけば、normal value の問題点は、normal の語の多義性、そして normal value に関する統一された定義や手続きの欠如にあった。そして

normal value に関する問題点を踏まえた reference value の狙いは、reference value と decision limit/discrimination value との使い分けによって指標の意味を明確に規定すること²¹、医学的判断の基礎となる比較を厳密に行うために、特定の目的に応じた適切な対照群を生成することであった。reference value の登場の段階においては、“正常／異常の基準であるという誤解を招かないようにする”という内容は、中心的な論点にはなっていなかったといえる。

この結果をもとに考えると、すぐに次の疑問が思いうかぶ。臨床検査の専門家が描写していたような、“正常値は、検査値がその値に近ければ正常、その値から外れたら異常、という誤解を招くものであったために、基準値の概念が提案された”という認識は、どこから生じてきたのだろうか。この点は今後明らかにすべき課題の一つである。またこの点は、冒頭の基準範囲をめぐる混乱と無関係ではないように思われる。本稿第1節で、基準範囲をめぐる混乱の背景には、マスコミ（あるいはもっと広く非専門家）と専門家間の認識のギャップがあると指摘したが、本稿の分析を踏まえれば、次のような仮説を立てることができる。すなわち、専門家は（必ずしも事実ではないかもしれない）“normal value から reference value への変化によって、そういった数値に基づいて健康／病気あるいは正常／異常が判断されるという誤解は払拭された”と認識しているが、非専門家は“数値的な指標は健康／病気あるいは正常／異常が判断の基準となる”と感じているのではないか²²。もしそうであるとしたら、そのような非専門家の理解はどのように形成されてきたのだろうか。この点も今

21 前述の normal の三つの意味、すなわち (1) 統計学的なもの、(2) 疫学的なもの、(3) 臨床的なものとの関係で考えると、reference value は (2) に、decision limit/discrimination value は (3) に対応すると考えられる。reference value の考案は、normal の複数の意味にそれぞれ対応する指標を作ろうとする試みであるともいえるかもしれない。

22 本稿注5で言及した臨床判断値は、特定の病態の診断や治療などについて判定するための基準であり、健康／病気あるいは正常／異常の判断基準であると表現することもできるだろう。その意味で、ここで述べたような非専門家の理解が全面的に間違っているとはいえない。しかし基準範囲をめぐる混乱の原因を、非専門家による基準範囲と臨床判断値の単なる混同のみに求めることもまた妥当でない可能性があるのではないか。臨床判断値という概念の登場・発展の経緯、およびその（特に非専門家に対する）普及の様相を分析したうえで、そもそもなぜ基準範囲と臨床判断値が混同されたのか、そして混同された結果がなぜ“健康／病気あるいは正常／異常の判断基準”であるという理解に向かったのかといった点を明らかにしていくことが重要であると思われる。

後の検討課題とする。

参考文献

- Alström, T. (1980). Evolution and Nomenclature of the Reference Value Concept. In: Gräsbeck, R., and Alström, T. (eds.) *Reference Values in Laboratory Medicine: The Current State of the Art*. Wiley, Chichester. 3–13.
- Alström, T., Gräsbeck, R., Hjelm, M., and Skandsen, S. (1975). Recommendations Concerning the Collection of Reference Values in Clinical Chemistry and Activity Report by the Committee on Reference Values of the Scandinavian Society for Clinical Chemistry and Clinical Physiology. *Scandinavian Journal of Clinical and Laboratory Investigation*, 35, Suppl. 144, 1–44.
- Benson, E. S., and Rubin, M. (eds.) (1978). *Logic and economics of clinical laboratory use: proceedings of the Conference on the Logic and Economics of Clinical Laboratory Test Selection and Use, Cancun, Mexico, March 4-9, 1978*. Elsevier North-Holland.
- Dybkaer, R., and Gräsbeck, R. (1973). Theory of Reference Value. *Scandinavian Journal of Clinical and Laboratory Investigation*. 32(1), 1–7.
- Dybkaer, R., and Solberg, H. E. (1987). International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values (EPTRV), and International Committee for Standardization in Haematology (ICSH) Standing Committee on Reference Values. Approved Recommendation (1987) on the Theory of Reference Values. Part 6. Presentation of Observed Values Related to Reference Values. *Clinica Chimica Acta*. 170, S33–S42.
- Gräsbeck, R. (ed.) (1977). Reference values. *IFCC Newsletter*, Fasc. 16, 4–11.
- Gräsbeck, R., and Alström, T. (eds.) (1981). *Reference Values in Laboratory Medicine: The Current State of the Art*. Wiley, Chichester.
- Gräsbeck, R., and Saris, N.-E. (1969). Establishment and Use of Normal Values. *Scandinavian Journal of Clinical and Laboratory Investigation*, 26, Suppl. 110,

62–63.

- Gräsbeck, R., Siest, G., Wilding, P., Williams, G. Z., and Whitehead, T. P. (1978). International Federation of Clinical Chemistry, Committee on Standards, Expert Panel on the Theory of Reference Values. Provisional Recommendation on the Theory of Reference Values (1978). *Clinica Chimica Acta*, 87, 459F–465F.
- Gräsbeck, R., and Solberg, H. E. (1981). Can Reference Values be Properly Used? *Clinica Chimica Acta*, 27, 1795–1796.
- Martin, H. F., Hologgitas, J. V., Driscoll, J., Fanger, H., and Gudzinowicz, B. J. (1980). Reference Values Based on Populations Accessible to Hospitals. In: Gräsbeck, R., and Alström, T. (eds.) *Reference Values in Laboratory Medicine: The Current State of the Art*. Wiley, Chichester. 233–262.
- Meites, S., Levitt, M. J., Blumenfeld, T. A., Hammond, K. B., Hicks, J. M., Hill, G. J., Sherwin, J. E., and Smith, E. K. (1979). Skin Puncture and Blood-Collecting Techniques for Infants. *Clinica Chimica Acta*, 25, 183–189.
- Ozarda, Y. (2020). Establishing and Using Reference Intervals. *Turkish Journal of Biochemistry*, 45(1), 1–10.
- PetitClerc, C., and Solberg, H. E. (1987). International Federation of Clinical Chemistry, Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values (EPTRV). Approved Recommendation (1987) on the Theory of Reference Values. Part 2. Selection of Individuals for the Production of Reference Values. *Clinica Chimica Acta*, 170, S3–S12.
- Quételet, A. (1835). *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*. Tome 2. Paris, Bachelier, Imprimeur-Libraire.
- Siest, G., and Henny, J. (2013). The Theory of Reference Values: An Unfinished Symphony. *Clinical Chemistry and Laboratory Medicine*, 51(1), 47–64.
- Solberg, H. E. (1987a). International Federation of Clinical Chemistry, Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values and International Committee for Standardization in Haematology Standing Committee on Reference Values. Approved Recommendation (1986) on the Theory of Reference Values. Part 1. The Concept of Reference Values. *Clinica*

Chimica Acta, 165, 111–118.

Solberg, H. E. (1987b). International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values (EPTRV), and International Committee for Standardization in Haematology (ICSH) Standing Committee on Reference Values. Approved Recommendation (1987) on the Theory of Reference Values. Part 5. Statistical Treatment of Collected Reference Values. Determination of reference limits. *Clinica Chimica Acta*, 170, S13–S32.

Solberg, H. E., and Gräsbeck, R. (1989). Reference Values. *Advances in Clinical Chemistry*, 27, 1–79.

Solberg, H. E., and PetitClerc, C. (1988). International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values (EPTRV). Approved recommendation (1988) on the Theory of Reference Values. Part 3. Preparation of Individuals and Collection of Specimens for the Production of Reference Values. *Clinica Chimica Acta*, 177, S1–S12.

Solberg, H. E., and Stamm, D. (1991). International Federation of Clinical Chemistry (IFCC), IFCC Recommendation, Theory of Reference Values. Part 4. Control of Analytical Variation in the Production, Transfer and Application of Reference Values. *Clinica Chimica Acta*, 202, S5–S12.

Sunderman, F. W., Jr. (1975). Current Concepts of “Normal Values,” “Reference Values,” and “Discrimination Values” in Clinical Chemistry. *Clinical Chemistry*, 21(13), 1873–1877.

市原清志 (2020) 「基準範囲の調和化・共用化 基準範囲設定の調和化——世界の現状」『臨床検査』第 64 号、1276–1285 ページ。

「基準範囲」に関わる用語についての検討ワーキンググループ (2019) 「「基準範囲」に関わる用語についての検討ワーキンググループの報告書(2019.2.8)」 (<https://www.jslm.org/about/jslm/20190208.pdf>, 最終閲覧 2022 年 2 月 19 日)。

佐藤正一・市原清志 (2020) 「なぜ正常範囲ではなく“基準範囲”なのか “異常” は本当に異常なのか？」『medicina』第 57 巻第 6 号、827–831 ページ。

前川真人 (2020) 「基準範囲と臨床判断値、そして基準値」『臨床検査』第 64

号、1262–1267 ページ。

三井田孝（2020）「臨床判断値を考える 脂質検査の臨床判断値」『臨床検査』第 64 号、1286–1292 ページ。

日本臨床検査医学会ガイドライン作成委員会編（2018）『臨床検査のガイドライン JSLM2018 検査値アプローチ／症候／疾患』宇宙堂八木書店（<http://www.jslm.info/GL2018/pdf/GL2018.pdf>、最終閲覧 2022 年 2 月 19 日）。

渡辺清明（2015）「臨床検査の基準範囲——正しい理解の必要性について」『臨床検査』第 59 号、548–554 ページ。