

# 発育発達を追跡する縦断的データの活用

中野 貴博

## 1. 発育発達研究における縦断的データの特性

ヒトの発育発達を研究する際に用いられるデータにはさまざまな形式が想定される。その中でも、発育や発達という言葉からイメージされるデータは縦断的データではないだろうか。身体の大きさなどの量的な変化を示す発育にしても、ヒトの機能的な変化を示す発達にしても、研究の視点に「変化」が存在する以上、ある時点からある時点にかけて、どのように変化したかをデータを用いて客観的に追跡することが中心となることは、ごく自然であろう。つまり、発育発達領域の研究では、機能や特性の経時的变化を捉えることが主たる目的であるため、用いられるデータは時間軸に沿ったデータとなることが多いと考えられる。

一方で横断的データは、どのように用いられているだろうか。集団間での発育発達の違いを検討したり、要因の影響の大きさを検討したりする際には横断的研究は有効であろう。たとえば、体力や体格の世代間比較や生育環境の違いによる比較などは横断的研究で行われることが多い。また、横断的研究の成果は、縦断的研究を行う際に統制すべき影響要因を決定するのにも役に立つ。縦断的データの収集においては、多くの変数を扱いつづけることは非現実的であり、観察対象の事象を代表する限られた指標を収集し続けることになるため、代表指標を決定する際にも横断的研究の成果

が役に立つというわけである。多くの横断的研究に基づいて縦断的研究の適切な指標とデータ収集計画が構築されるため、横断的研究の成果を踏まえていない縦断的研究は、多くの影響変数が無視できない状態となり、得られた関係性や因果関係の精度は下がり、限定的な解釈をせざるを得なくなる可能性が高い。

ここで、縦断的データに強く影響してしまう要因の存在について記述していると考えられる2つの記述を紹介する。松浦<sup>1)</sup>は、「縦断データは発育発達現象を直接現すデータではあるが、得られたデータには種々の条件による変動もまた含まれている」と記している。さらに、氏家<sup>2)</sup>は、「発達研究の多くは、複数の要因が時間経過の中で相互に関連しあってそれぞれの要因を変化させているような現象（トランザクションと呼ばれる）を対象としている」と記している。

これらの記述からもわかるように、縦断的研究を行う際に、研究者は種々の条件や要因をできる限り統制することが必要である。しかしながら、それは決して簡単なことではないため、縦断的データが、そのような要素を含んでいるデータ形式であることをよく理解したうえで、結果の検討や考察をする必要がある。さもなければ、われわれは縦断的研究の結果を過大もしくは過小に評価してしまう恐れがあることを理解すべきである。

## 2. 発育発達を追跡するための縦断的データの収集

最初に縦断的データについて少し確認をする。

筆者：名古屋学院大学スポーツ健康学部

縦断的データとは、同一の個人に対して複数の時点にわたって収集されたデータのことであり、同じ対象者（グループ）に対する繰り返しの測定が原則となる。ここで、時系列データという表現もあるが、これは、厳密には等間隔の測定を原則としており、データのもつ「規則性」や「変動」、それに基づく「予測モデル」の作成などが主な関心となる。その意味で、毎分毎秒測定されるような生体情報のデータなども時系列データである。一方、発育発達領域では、このような厳密な時系列データを収集することは容易ではない。フィールドでデータを収集するような場合には、データ収集時点間における影響要因を統制することがきわめて難しくなるため、予測などの精度は高まらず、縦断的な「変化」を検討することを目的としたデータ収集が中心に行われている。もちろん、健康診断等における身長・体重データのように定期的に得られるデータを用いることができる場合には「変動」や「予測」も重要な視点になる。また、個人レベルではなく、集団レベルで同じ条件の集団を繰り返し測定するような混合縦断データもある種の縦断的データと考えられる。このようなデータは、検討の視点によっては、後述するコホート分析と限りなく類似した物になるであろう。

次に、縦断的データ収集の難点について示す。莊島ら<sup>3)</sup>は、あるセミナーにおいて、1) 多くの時間的・人的・経済的コストを要すること、2) 統計分析の方法が複雑・高度になりやすいこと、3) 作業への慣れ（練習効果）や記憶などがデータの歪み（バイアス）を与える可能性を考慮する必要があること、4) データの欠損（missing）が生じやすいこと、の4点を縦断的データ収集の難点としてあげている。筆者もこれまでに多くの縦断的データを収集・分析してきたが、1)～3)に関しては、避けることが難しいと感じている。さらに、4)に関しては、同一の対象者に関して、すべての時点でのデータが揃っていないと1つの縦断的データとして意味をなさないため、サンプルサイズが研究開始当初に比べて大幅に減少してしまうことが大きな問題となる。研究期間が長くなればなるほど、この問題は大きくなるため、開始当初は、最終的なデータ数の倍程度の対象者を用意し

て始めるぐらいの心持ちが必要になることがほとんどである。さらに、対象者への負担も横断的研究に比べ大きくなるため、協力・同意を得る段階で特殊な対象者だけが選択されてしまう恐れもある。このように、縦断的データは非常に多くの情報量を有し、発育発達研究の立場では有用である一方で、さまざまな注意と労力をはらって収集しなければいけないことをよく理解する必要がある。

では、どのようなアプローチによって、これらの問題を少しでも回避することができるであろうか。筆者が考える点で最も重要なことは、欲張らないことである。消極的な表現に聞こえるかも知れないが、過剰な変数や過剰な期間の縦断的データの収集は、非現実的であるし、難点2)の統計分析の複雑さに関しても、加速度的にその難しさを増していく。言い換えると、収集する変数の絞り込みを事前段階でしっかり行うことが大切である。すでに関係性がはっきりしている変数は対象としないことや、集団の環境として把握しておけばよいようなデータは、個々人からの繰り返しデータ収集には含まないなどの工夫をすることで、大幅な負担軽減や欠損データの減少が期待できる。さらに、縦断的データの関心の中心が「変化」であることを念頭に変数構成をすることも忘れてはいけない。

もう一点、心がけるとよいと思われる点は、各データ収集時点での分析検討をこまめに行うことである。研究期間全部のデータ収集を待って解析を行うのではなく、適宜、集計等を行うことで対象者の継続意識も高められるし、協力機関等との問題意識の共有にもつながる。このような地道な努力が縦断的データの収集を可能にすると考えている。

### 3. 縦断的データの分析視点

既述のとおり、縦断的データにおいてはデータの「変動」や「規則性」、「予測」といった視点が考えられる。ただし、これらは等間隔にデータが得られた時系列データでないと、分析精度が下がることも記した。そのような場合は、「規則性」や「予測」とまでは行かないが、「変化」つまり、体

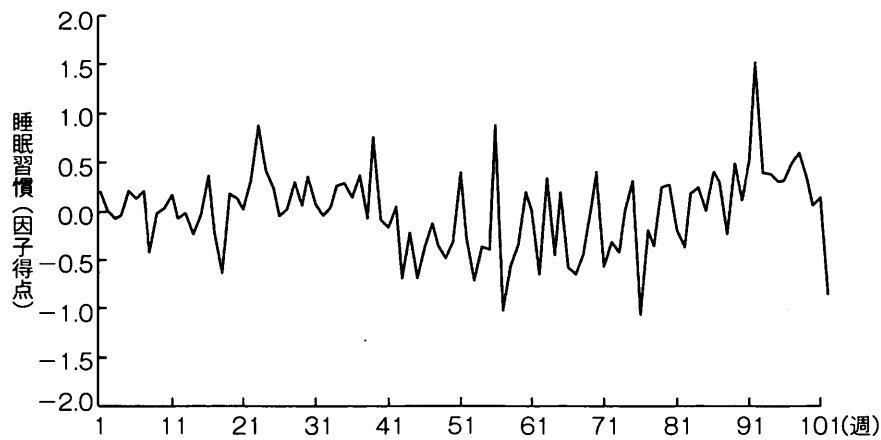


図 1 睡眠習慣因子の変動

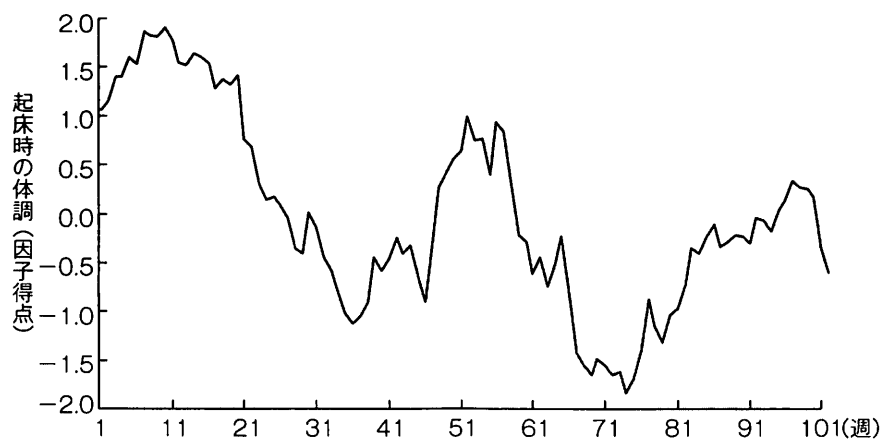


図 2 起床時の体調因子の変動

力が伸びた（発達した）、低下したなどを各時点において検討することで、個々人の発育発達を追跡していくことができる。以降では、これらの視点で分析をする際に適用可能な方法にはどのようなものがあるかを示す。

まず、時系列データとみなせる水準の縦断的データにおける「変動」や「規則性」、「予測」について考えてみる。「変動」の検討の第一歩は折れ線グラフや平滑化された近似曲線の描画であろう。平滑化の方法はさまざまあり、最小二乗法に基づく多項式回帰や移動平均を用いた方法などが代表である。曲線の式を微分すれば極値が算出できるため、対象とする事象の発達のピークを求めることが可能になり、さまざまな介入時期に対する示唆を検討できると思われる。大澤<sup>4)</sup>の論文などは、まさにこの方法で発育発達期の子ども達における体力トレーニングの至適な開始年齢を提案している。この他にも平滑化スプラインやベイズ

的方法も提案されている<sup>5)</sup>。また、身長が発育速度曲線に関しては、発育学の重要命題であり、さまざまな曲線推定の方法が検討されているが、これに関する詳細に関しては、他の文献を参照していただきたい。

次に「規則性」と「予測」に関して示す。特に「予測」に関する方法は非常に多く、近年のAI技術なども、この領域に含めることができるかもしれない。最初に「規則性」に関してだが、「変動」を図示した段階で規則性の存在はある程度観察できると思われるが、それを確認するうえでよい方法は周波数解析（スペクトル分析）の応用であろう。時系列データをフーリエ級数で表現し、ピリオドグラムの大きさによって最大周期を見つけることができる。ピリオドグラムの値が最も大きくなる点が最大周期であると判断される。

たとえば、図1と図2は、筆者が以前、アスリートの体調管理データを約100週にわたって毎

#### 4. コホートデータの形成と分析

最後に、コホートデータにより縦断的検討を行うことを提案したい。コホート分析はもともと人口統計学の分野で開発された方法であるが、現在では広くさまざまな分野で適用例がある。同一の条件を有する集団をコホートと呼び、そのコホートにおけるデータの変化を検討することで縦断的な検討を行っていく。最も典型的なものは出生年を同じにする出生コホートである。同じ年に生まれた集団から年々データの収集をすることで、縦断的な検討が可能になる。同一の対象者の繰り返し測定は難しいが、同一背景の集団（コホート）の横断的データ収集の繰り返しであれば、いくぶん実現可能性も高まるし、より多くのデータを得やすいというメリットもある。コホート分析では、時代の変化による時代効果、年齢の変化による年齢効果、世代の違いによるコホート効果の3つの効果を検討できる。しかしながら、これらの3つの効果は相互に影響を及ぼしあっており明確に分離することは難しい。最も典型的な方法としては、標準コホート表を作成して検討を進めていく方法である（図4）。標準コホート表とは、調査時点の間隔と年齢階級の幅が一致するように配置した表であり、縦に読めば加齢効果、横に読めば時代効果、斜めに読めばコホート効果を考察できるというわけである<sup>9)</sup>。コホート分析の視点を持つことで、発育発達の縦断的な検討が可能でデータ収集の幅を広げることが期待できる。

#### 5. まとめ

本稿では、発育発達研究における縦断的データの活用について示した。縦断的データはデータ収集における難点も存在するが、それを踏まえても非常に有効なデータ形式であり、適切な研究計画のもと、有効な分析手法の適用が望まれる。分析視点としては、理想的な縦断的データが得られた

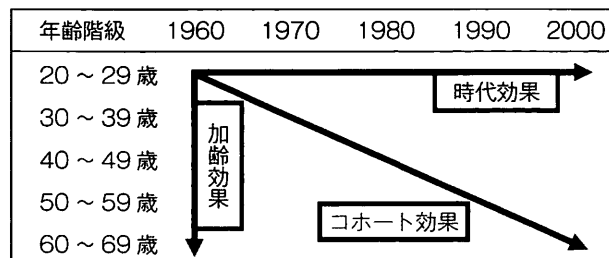


図4 標準コホート表

場合には「規則性」や「変動」の検討、それに基づく「予測モデル」の作成を目指すべきであろう。データの収集間隔や要因の統制状況に多少の難点があった場合には、「変化」に着目して、ここで示した分析手法の適用を試みていただければ幸いに思う。

#### 文 献

- 1) 松浦義行：統計的発育発達学。pp9-13, 123-153, 不味堂出版, 2002.
- 2) 氏家達夫：縦断的アプローチと横断的アプローチ。pp136-139. 田島信元, 西野泰広編著, 発達研究の技法。福村出版, 2000.
- 3) 荘島宏二郎, 宇佐美慧, 吉武尚美ほか：縦断データ分析のはじめの一步と二歩。教育心理学年報, 56: 291-298, 2017.
- 4) 大澤清二：最適な体力トレーニングの開始年齢—文部科学省新体力テストデータの解析から—。発育発達研究, 65: 25-35, 2015.
- 5) 柏木宜久, 柳本武美：平滑化法とそのソフトウェア。応用統計学, 19 (2): 79-92, 1990.
- 6) 中野貴博：個人の追跡データの因子分析。子どもと発育発達, 7: 4-8, 2009.
- 7) 高橋雄介：時間を含むデータをどう分析するか？—一人の変化・発達をとらえる統計—。子ども発達臨床研究, 7: 63-92, 2015.
- 8) 鈴木宏哉：SEMによる発育発達研究。子どもと発育発達, 7: 38-43, 2009.
- 9) 中野貴博：体格と生活習慣のコホート分析。統計情報, 54 (3): 31-35, 2005.