＜教育報告＞

# 上腕周囲長によるGrowth Monitoringの妥当性 ータンザニアにおける縦断研究ー

田　中　　あ　ゆ　子

# Validity of mid-upper-arm circumference for growth monitoring; longitudinal study in Tanzania

Ayuko TANAKA

Body weight is one of the most effective indices that evaluate growth and nutritional status in infancy. Mid-upper-arm circumference (MUAC) is often used as an index instead of weight where there are no scales, when the age of the children are not clear in such places as a refugee camp. But it is not discussed enough whether the method is valid or not. Therefore it is necessary for us to clarify its validity immediately. This study is to examine the validity of growth assessment using MUAC. We analyzed these data not only cross-sectional, in which we ordinary analyze, but also longitudinally. Physical growth is a phenomenon, which occurs along with time series, so longitudinal analysis is effective method of study. We measured body weights and MUACs of 555～711 children every month (total number was 3850) from July in 1999 to December in 1999, at Ghonghona village, Dodoma region, Tanzania. The results are as follows. 1) Sensitivity of screening for malnutrition, which is under 80% of standard weight, was low, 2) the phase was different between weight-gain and MUAC-gain, 3) the fluctuation in MUAC-gain was about one month slower than that of weight-gain. Above all, MUAC can not be used for area diagnosis, group evaluation, and screening of malnutrition. And it wasn't suitable for evaluating the individual growth and nutritional status longitudinally. That is because it was difficult to make a rapid detection of growth delay; the stagnation of growth gain and aggravation of the nutritional status. This study suggests that 1) Before recommending MUAC instead of body weight, we should work for the wide permeation of the weight scales, 2) the manuals of growth monitoring have to mention about validity and problems concerning current methods of measurements, and 3) WHO and UNICEF should provide proper information.

**Supervisor** : Noriko KATO

## I. 緒　言

開発途上国と呼ばれる国々の5歳未満児のうち，年間1200万人に及ぶ死亡の半数以上に栄養不良が関与している．栄養不良を識別し，死や罹患の危険を予測・予防する一つの手段がGrowth Monitoringである．Growth Monitoringの指標として最も有効かつ普及しているのは体重であるが，体重計のない場所や難民キャンプ等，児の年齢が不明な地域では，体重に代わる有効な指標として上腕周囲長を代用し，UNICEFやWHOのマニュアルでも紹介されている．一方，アメリカ公衆衛生協会によるGrowth Monitoringの指標に関する信頼性の評価では，5段階評価の“3”であるが，重度栄養不良のスクリーニングには有用だとされている．また，UNICEFやWHOの採択する発育基準値が改定された際，上腕周囲長は改定の対象外となり発育基準値として記載されなかった．このように，上腕周囲長を指標とするGrowth Monitoringの妥当性の明確な結論は得られていない．

本研究の目的は，Growth Monitoringが乳幼児の健康管理の重要な役割を担っているタンザニアにおいて，上腕周囲長を指標とするGrowth Monitoringの妥当性について，従来行われてきた横断的分析に，発育という時系列的現象の把握・検討に有用な縦断的分析をおこない検討することである．なお，本研究では，発育計測値を一時点で観察された

指導教官：加藤 則子（母子保健学部）

データとして分析する方法を横断的分析，継続して観察された同一例の時間的変化に着目して分析する方法を縦断的分析とした．

## II．研究方法

### II-1．調査地概要

タンザニアはアフリカ大陸東海岸に位置し，国土面積は94.5万km$^2$，人口3279万人，１人当りGNPは約240米ドルである．調査地のあるドドマ州は，タンザニアのほぼ中央に位置し，人口は162万人である．保健統計資料によると，1歳未満乳児死亡率は90（出生1000対），５歳未満児死亡率141（出生1000対）であり，世界187カ国中30位と高く，また５歳未満児の栄養不良の割合は，低体重27％，消耗症６％，発育阻害42％とされ，乳幼児のGrowth Monitoringを積極的に実施している．

### II-2．調査期間

調査期間は，1999年７月から同年12月である．調査地は半乾燥地帯に属し，毎年10月から11月の乾期は水不足に悩まされ，衛生環境が悪化し健康状態が低下するとされている．したがって，この期間は乳幼児の栄養状態の低下が推測され，栄養不良の識別あるいは予測についての検討に必要と考え，この期間を含む６ヵ月間を調査期間とした．

### II-3．対象

対象は，同クリニック乳幼児定期健康診査を受診した５歳未満乳幼児である．調査対象として登録した児数は836人，うち男児426（51.0％）人，女児410（49.0％）人であった．バンツー語系のgogo族が全対象児の96.6％を占めた．なお，同村の推定人口は7772人（1988年度人口調査で得た値に年間人口増加率を乗じて算出されている），５歳未満乳幼児の推定人口は1554人（人口の20％として算出されている）である．同クリニック現地スタッフによる把握では，同村５歳未満人口の95％以上が受診しているとのことであった．

### II-4．調査項目

Ⅱ-4-1）形態計測

定期健康診査を受診した全ての児に対し，体重及び上腕周囲長の計測を行った．計測は全て習熟した同一の験者が行った．上腕周囲長は左上腕の肩甲骨肩峰突起と尺骨肘頭突起を結ぶ線の中間点の周径を計測した．２回計測し平均値を計測値とした．体重計は現地で使用している計測器を使用し，上腕周囲長を計る巻尺は日本から持参した．

Ⅱ-4-2）対象に関する情報

計測値の他に以下のデータを収集した．生年月日，性別，本調査開始までの発育経過及び健康状態のデータを各児のGrowth monitoring cardから収集した．また，部族，調査時の健康状態，Growth monitoring cardの記入漏れデータは，児の母親に対し聞き取り調査を行って得た．以下，分析方法，結果，考察については，Ⅲ.横断的分析，Ⅳ.縦断的分析の各項でそれぞれ述べる．

## III．横断的分析

### III-1．横断的分析方法：栄養不良のスクリーニング精度の検討

上腕周囲長を指標とする栄養不良のスクリーニング精度について，体重による評価を真の評価と見なして検討を行った．まず，体重及び上腕周囲長の個々のデータについて，各標準値に対する計測値の百分率（以下，対標準比とする）を求めた．体重の標準値は，タンザニアのGrowth monitoring cardの標準値Harvard Standardを，上腕周囲長はWHO，UNICEF等がグローバル・スタンダードとしているWolanskiの標準値を用いた．得られた対標準比について，タンザニアで利用されている栄養状態のカット・オフ・ポイントにしたがって，「栄養状態良好」，「中等度栄養不良」，「重度栄養不良」に分類した．

### III-2．横断的分析結果：上腕周囲長による栄養不良のスクリーニング精度

上腕周囲長による栄養不良のスクリーニング精度，つまりカット・オフ・ポイントを「栄養状態良好」と「栄養不良」の境界においた場合では，敏感度35.4％，特異度97.0％，偽陽性率3.0％，偽陰性率64.6％，陽性適中度88.9％，陰性適中度68.7％であった．次に，重度栄養不良のスクリーニング精度，つまりカット・オフ・ポイントを「中等度栄養不良」と「重度栄養不良」の境界値においた場合は，敏感度94.0％，特異度96.6％，偽陽性率3.4％，偽陰性率6.0％，陽性適中度32.6％，陰性適中度99.9％であった．

### III-3．横断的分析結果の考察

横断的分析結果より，上腕周囲長による栄養不良のスクリーニングは，敏感度が低く偽陰性率が高いことから，栄養不良を見逃す確率が高いと考察した．また，重度栄養不良のスクリーニングでは，敏感度及び特異度は高かったが，重度栄養不良の有病率が低いため陽性適中度が低く，重度栄養不良と判定された児に占める真の重度栄養不良児の割合は低かった．以上のことから，上腕周囲長による栄養不良のスクリーニングは，重度栄養不良のスクリーニングを含め，その妥当性は低いと考察した．

## IV．縦断的分析

### IV-1．縦断的分析方法

Ⅳ-1-1）体重及び上腕周囲長の発育増加量の比較

個々の縦断的データを用い，体重及び上腕周囲長の発育増加量について検討を行った．体重及び上腕周囲長の１ヵ月間の発育増加量を求め，一元配置分散分析を行い，各時期の月間増加量の差について検討を行った．等分散性の検討にはLevene検定を，分散分析後の多重比較検定にはTukey HSD検定を用いた．

Ⅳ-1-2）体重及び上腕周囲長の調整増加量の比較

月齢による発育増加量が異なるため，体重及び上腕周囲

長の平均発育増加量は，各計測月の月齢構成による影響を受ける．Ⅳ-1-1）の検討において，年齢構成に大きな違いがないことを確認したが，議論を正確にするため月齢構成による影響を調整し同様の検討を行った．各時期の月齢に対する月間増加量の直線回帰式を求め，各計測点において仮に30月齢とした場合の月間増加量を求め，これを調整増加量とした．

Ⅳ-1-3）体重と上腕周囲長の発育増加量の交差相関

体重及び上腕周囲長の発育増加量について両者の変動の差異を検討するため，両者の発育増加量の交差相関を求め，時系列分析を行った．Ⅳ-1-1)，2）で求めた体重及び上腕周囲長の発育増加量，調整増加量，欠損値を持たない児の発育増加量の各平均値について，０次から±３次の交差相関係数を求めた．これは，体重と上腕周囲長の計測時期を前後に１ヵ月から３ヵ月ずらして組み合わせた場合の相関係数のことである．そして，＋１次，０次，－１次の交差相関係数に着目し，体重と上腕周囲長の増加量について時系列分析を行い，発育増加量の変動の時間的なズレについて検討を行った．

Ⅳ-1-4）縦断データごとの発育増加量の交差相関

Ⅳ-1-3）における検討の信頼性を高めるため，欠損データを持たない児322人を対象に，個々の縦断データの体重と上腕周囲長の発育増加量の交差相関係数を求めた．まず，個々の平均発育増加量の違いにより生じる，データとデータのつなぎ目の段差を除くため，個々の発育増加量から個々の平均発育増加量を減じた縦断データを算出した．そして，平均発育増加量を減じた体重と上腕周囲長の発育増加量の縦断データについて交差相関係数を求めた．

### IV-2．縦断的分析結果

Ⅳ-2-1）体重及び上腕周囲長の発育増加量の変動

体重の各時期ごとの平均発育増加量は0.263±0.391kgで，分散分析の結果，各時期の増加量の間に有意な差を認めた（$F=29.96$, $p<0.01$）．多重比較の結果，９－10月の増加量が最も小さく，他の全ての時期と比べ有意に小さかった（$p<0.001$）．上腕周囲長の各時期ごとの平均発育増加量は0.273±0.627cmで，散分析の結果，各時期の増加量の間に差を認めた（$F=9.92$, $p<0.001$）．多重比較の結果，10－11月増加量が最小値を示し，７－８月，８－９月増加量に比べ有意に小さかった（$p<0.001$）が，９－10月，11－12月とは有意な差は認めなかった．また，調整増加量においても，体重及び上腕周囲長の各時期の増加量に有意な差を認め（$F=34.531$ $p<0.001$, $F=10.463$ $p<0.001$），多重比較の結果においても，体重では９－10月が最も小さく，上腕周囲長では10-11月が最も小さく，同様の傾向を示した．

Ⅳ-2-2）体重と上腕周囲長の発育増加量の交差相関

体重と上腕周囲長の平均発育増加量の交差相関係数は，１次が０次より大きくその差は1.17であった．平均調整増加量の交差相関係数は，１次が－１次及び０次より大きく１次と－１次の差は1.23であった．また，欠損データを持たない児の平均発育増加量の交差相関係数も１次が－１次より大きく，その差は0.949であった．以上のように，平均発育増加量，平均調整増加量，欠損データを持たない児の平均発育増加量の交差相関係数は，１次が－１次より大きい値を示した．また，欠損データを持たない児を対象とした，体重及び上腕周囲長の発育増加量の交差相関係数は，１次が－１次より0.102大きかった．

### IV-3．縦断的分析結果の考察

縦断的分析結果より，上腕周囲長と体重の発育増加量の変動にはズレがあり，上腕周囲長は体重に比べ約１ヶ月遅れ，また，その変動は体重に比べ緩やかであると考察した．このことから，上腕周囲長によるGrowth Monitoringでは，発育遅延（発育増加量の停滞）あるいは栄養状態の悪化を迅速にとらえることが困難であり，個々の発育及び栄養状態の経過に対する観察・評価には適さないと考察した．

## VI. 考　察

横断及び縦断的分析結果を総合的にみると，上腕周囲長によるGrowth Monitoringは，栄養不良のスクリーニング精度が低く，発育遅延あるいは栄養不良の早期把握も困難であるため，地域診断にも個々の発育及び栄養状態の経過観察にも適切でないと考察した．栄養不良の削減には，栄養不良の早期発見と早期治療が肝要である．とくに，タンザニアを含む発展途上国と呼ばれる国々の医療状況を考えると，栄養不良に陥る前の予測と対処が望まれる．事実，タンザニアにおける重度栄養不良の治癒率は非常に低く，手後れになるケースが多い．したがって，上腕周囲長によるGrowth Monitoringは，体重によるそれに比べ栄養不良の発見が遅れることから，その妥当性は低いといえる．また，アメリカ公衆衛生協会による信頼性の評価も高くはなく，UNICEFやWHOの採択する発育基準値としても記載されなかった事実を鑑みると，マニュアルの内容を改正することなく，体重の代替として上腕周囲長によるGrowth Monitoringを推奨し，普及している状況へ疑念を抱くとともに，根本的な改善の必要があると考察した．

## VII. 提　言

上腕周囲長によるGrowth Monitoringの妥当性は低く利用は避けるべきである．そして，上腕周囲長による代替ではなく，体重計の配備に努めるべきであり，WHOやUNICEF等はフィールドに対して明確な情報提供をし，徒に利用を促すようなマニュアルを改正する必要がある．

＜教育報告＞

# 中国黒竜江省における人口動態統計の現状と出生体重統計の活用に関する研究

康　　文　江

# Study on vital statistics and birth weight statistics in Heilongjiang Province of China

Wenjiang KANG

Mortality rate, infant mortality rate and age-adjusted mortality rate have been used as indicators for evaluating health status of a particular community. In China, the Population Census was carried out in years of 1953, 1964, 1982 and 1990, and the total population all over the country became to be clear. Using these values, the mortality and age-adjusted mortality were calculated as a country value. However, considering that a main management section for vital statistics is uncertain and that many lacks of birth and/or death registration exist in this country, it would be doubtful that health indicators presented as rates calculated from vital statistics reflect real health status. A precise rate will be able to be calculated after a concrete registration system would be established.

On the other hand from birth statistics that is one of vital statistics, birth weight distribution and gestational distribution are available. These distributions have been reported as new type health indicators by many authors. The four delivery hospitals in a large city of Heilongjiang province in China were selected, and four years delivery records (1995-1998) were copied and analyzed. Birth weight (BW) distribution and gestational period (GP) distribution were different between delivery − A belonging to a university and delivery − D belonging to a small community government. For BW distribution in this period, delivery − A has a high proportion of the <2500g-birthweight in both male and female; and delivery − D has a high proportion of the >3500g-birth weight. In reference to GP distribution, delivery − D has less pre-term (<37 weeks) babies and more post-term (=42 or >42 weeks) babies. These phenomena seem to indicate a difference of characteristics with the both deliveries. Using two kinds of adjusted MBW (mean birth weight) and methods in terms of IUGL (intra-uterine growth level) and ML (maturity level), there are differences among four deliveries when compared to the standard population (from data of Japan from 1969-1997). The delivery − A made a role similar to Japan in relation to both of IUGL and ML. But delivery − D is not similar to Japan.

In order to obtain a representative health indicator that can be implemented quickly, it will be necessary for a sample survey for birth weight in deliveries selected by stratified sampling similar to the Patient Survey in Japan thought to be appropriate. For this reason, it will be necessary to obtain a balanced selection of Heilongjiang province delivery hospitals and their delivery records. Moreover, information of sex, GP, BW, malformation, mother's age, mother's occupation, delivery method and ethnical background records is necessary for a detail analysis. For a precise analysis, a standard recording form is requested in terms of the above information. In examining the birth records, %LBW size, number of pre-term babies/post-term babies, BW distribution, and adjusted MBW or IUGL, ML indicators derived from a standard population can be used to research with confounding factors of ethnical background.

**Supervisor**: Toru DOI

指導教官：土井　徹（保健統計人口学部）

## I. 緒言

死亡率，乳児死亡率，調整死亡率など人口動態統計の主として死亡を用いた諸率は，その地域の保健水準を表す指標として活用されてきている．中国でも1953年，64年，82年，90年に国勢調査（人口普査）を実施して分母となる人口数を把握し，国としての死亡率，調整死亡率を算出している．しかし，これらの人口動態統計の信頼性は決して高いとは云えず，国際比較には疑問が提示されている．さらに日本の場合，人口並びに人口動態統計の数値は日本全地域の全数を男女別かつ年齢別に把握できるのに対し，中国の場合は全地域ではなく定点観測の数値から推測して全国値としたもの，男女合計しかないもの，年齢別になっていないもの等がある．従って中国の１地域の保健水準を検討する上で人口動態統計の主として死亡を用いた諸率は有効かどうか疑問である．中国の場合，数値の信頼性の低さの１つの原因は中国の統計制度にあり，もう１つは届け出漏れにあると考えられる．また定点観測からはずれている地域では数値そのものが存在しない恐れがある．一方，人口動態統計の１つである出生統計には出生児の体重が記載される場合が多く，2500g未満の割合（低出生体重児の割合，％LBW）を地域の保健水準に使用する例も少なくない．ただし，この割合も事象の有無を分子にとった「率」であるので，より小さな地域の保健水準として使用する場合には変動が大きくなる弱点を持っている．そこで，出生体重分布を保健水準に使用する例も見られるが，中国の場合，出生体重分布は公表されていない．このような場合には，分娩施設を選択して，その分娩記録の転記により出生体重統計を作成することしかできない．しかし今までの所，中国の１地域，あるいは１施設においても保健水準という観点から出生体重統計を扱った報告は無い．

そこで本研究は中国の１省である黒竜江省を対象地域として，まず人口動態統計の把握状況，数値の信頼性等の現状を検討し，次に人口動態統計の１つである出生体重統計をとりあげ，その利用可能性を探ることを目的とした．

## II. 資料および方法

**1**）中国の人口動態統計に関する統計制度を調べるとともに，中国全土ならびに黒竜江省で過去に実施された人口動態統計に関する調査にどのようなものがあるかを調べた．

**2**）これらの統計調査から死亡率，調整死亡率，SMR，乳児死亡率の算出ならびにその年次推移，地域間格差の比較が可能かどうか検討した．

**3**）人口動態統計の１つである出生統計に関して，黒竜江省の大都市に設置されている分娩施設の中から４つの分娩施設（大学病院からＡとＢ施設，市立病院からＣ施設，区立病院からＤ施設）を選択し，その分娩記録（1995-98年）に記載されている事項を調べ，新たに出生体重統計を作成した．その上で，施設別年次別男女別に，出生体重分布については500g刻みの度数分布表，低出生体重児の割合（%LBW），平均出生体重，標準偏差を算出し，一方在胎期間別（早期産・正期産・過期産の３区分）についてはその割合，平均出生体重，標準偏差を算出し，施設間の比較を行った．

**4**）出生体重を日本と比較するために，調整MBWの手法を適用し，標準集団とした日本の値との有意差の検定を行った．また，日本のMBWの年次推移と比べ，黒竜江省の４施設のMBWがどの時期と類似しているかを検討した．

## III. 結果

**1**）日本では出生・死亡・死産に関する届け出はすべて市区町村が窓口であり，その後の小票の行き先も全て同じく最終的に厚生省に集積するのに対して，中国では公安部，衛生部にそれぞれの形で届いている．

**2**）過去に公表された乳児死亡率を年次順に見ると，対象地域が異なっていたり，調査方法が異なるため一概には言えないが，中国全土で見ると横ばい状況なのに対し，黒龍江省では死亡状況が改善してきていることが窺える．1995年の日本を標準集団に取った場合のSMRを見ると，黒竜江省と中国のSMRは日本の県別のSMRのばらつきより大きく離れて，まだ高い位置にあり，日本と比べまだ約20年遅れている．黒竜江省と中国で公表されている数値自体に矛盾は見出せないが，これらの数値を基にする限り現在でも黒竜江省の死亡状況には改善の余地が大きいといえる．

**3**）４施設の出生体重の分布では，1995-98年合計でみると平均出生体重は男女ともにＤ，Ｃ，Ｂ，Ａの順に高くなっており，その年次推移は施設Ａでは男女ともに減少傾向，施設Ｄでは男女ともに増加傾向にあった．2500g未満の割合は男では施設Ａ，Ｂ，Ｃ，Ｄの順に大きく，施設Ａの6.5％に対し，施設Ｄは1.7％であった．女ではＢ，Ａ，Ｃ，Ｄの順に大きく，施設Ａの5.8％に対して，施設Ｄが2.2％であった．一方，3500g以上の割合は男女ともＤ，Ｃ，Ｂ，Ａの順に大きくなっており，施設Ｄでは男の場合は52.6％を占め，女の場合は40.3％を占めていた．また在胎期間の分布では，1995-98年合計で見ると施設Ａに比べて施設Ｄは，男女ともに早期産は少なく，過期産が多かった．早期産については男の場合，施設Ａは6.5％であり，施設Ｄでは1.7％であった．女の場合は，施設Ａが3.8％であり，施設Ｄは2.1％であった．過期産について男の場合は，施設Ａの1.9％に対して施設Ｄでは6.4％であった．女の場合は施設Ａが2.1％，施設Ｄは6.0％であった．

**4**）４施設の出生体重と在胎期間の分布の状況を日本と比較するために，２種類の調整MBW及びその大小を見るためのIUGL（子宮内発育水準）とML（成熟水準）を使用し，標準集団を1969年から97年までの日本に取ったときの調整MBW，IUGLとMLを算出した．その有意差検定を行った結果を見ると，有意差のある年が多い施設と少ない施設とが

あり，４施設間で違いが見られた．施設Ａが子宮内発育水準，成熟水準ともにこの間の日本と近い役割を果たしている（母親の健康管理指導にも重点を置く）のに対し，施設Ｄは日本と異なっている状況（栄養状況の改善による低出生体重児の回避に重点を置く）にあり，施設Ｂ，Ｃは成熟水準のみこの間の日本と近い役割を果たしていると考えられた．

**5**）出生体重ならびに在胎期間の分布が，その両方をこみにした場合，日本のどの時期と類似しているのかを見るために，４施設の４年間のＭＢＷを単位にして類似性尺度としての$\sqrt{(IUGL^2+ML^2)}$と$\sqrt{(m-M)^2}$両指標の平均（補足としての中央値）を算出した結果を見ると，４施設における出生体重ならびに在胎期間の分布は日本の1972年頃と類似しているといえたが，この４施設では正期産がまだ少なく，日本とは異質な印象が感じられた．

## IV. 考察

地域健康指標として，死亡率，乳児死亡率等従来の伝統的な死亡統計を用いた健康指標を補うものとして出生体重分布の公衆衛生学的意義が検討されている．出生体重は人種・遺伝・母親の体格などの生物学的要因のほか，母親の妊娠中の健康管理状況など母親が属する集団の社会・経済条件に大きく左右されるので，その地域（集団）の出生体重分布が保健水準を反映していることは先行研究で指摘されている．しかし，問題なのは出生体重分布のどの指標を使用するかという点である．％LBWはよく用いられている指標だが，これは率を用いる指標なので例数がある程度揃わなければ議論しにくいという欠点を持っている．一方，平均値（平均出生体重MBW）は，その数値の持つ意味がわかりにくい．すなわちMBWが高ければ保健水準が高いといえるのかという問題である．１つの地域を取り上げた場合，社会・経済条件の上昇とともに，保健水準も向上し，MBWも増加しはじめると考えられている．ＭＢＷの増加は主としてLBWの減少によるものと理解できるが，日本のＭＢＷは1976年まで増加した後，減少の一途をたどっている．この要因は出生体重分布自体の変化（過体重児の減少）によるもので，社会・経済条件が高いままで安定し，母親の出産に対する意識が自主的になったためとの報告もあり，MBWという１つの指標で保健水準を長期的に評価するのは難しい．しかし，本研究の黒竜江省のように社会・経済条件が発展しつつある社会ではMBWの大小が保健水準をある程度反映しているといえる．実際，日本のＭＢＷの統計数値の年次推移を見ると，1969-76年まで上昇しているが，この上昇の過程ではＭＢＷと％LBW，乳児死亡率間に0.8以上の強い相関が見られている．また４施設間のMBWの相違が分娩施設の属性によって合理的に説明ができ，分娩施設の全数あるいは適切な標本調査が実施できればMBWの有用性が確認できると思われる．分娩施設ごとの出生体重の分布はその施設の役割を表現できるので，母子保健施策にも役立つことが推察できる．もしも地域住民が多く集まる分娩施設を選択すれば，その地域の保健水準を評価できることが示唆される．

一方，標準集団を1969年から1997年までの日本に取ったときのIUGLとMLの有意差検定結果は，施設Ａがこの間の日本と近い役割を果たしていることに対して，施設Ｄは日本と異なっている状況にあることを示した．また，1969-1997年の日本のＭＢＷとの類似性を検討した結果は，日本の1972年あたりの状況と最も類似していることが示された．このように，出生体重統計は保健水準の国際比較にも役立てることが可能である．

## V. 結論および提言

中国及び黒竜江省において，国際比較のために人口動態統計（特に出生数と死亡数）を使った諸率を使用することの信頼性は疑わしく，その年次推移の検討にも疑問が残る．この因は実数の正確な把握がまだ困難なことにあり，社会状況の改善を待たなければ信頼性は薄い．出生体重分布と在胎期間分布は施設の所属による相違があらかじめ明確であることから，出生体重統計を作成し解析することは保健水準の国際比較が可能な分析方法であると考えられる．

早急に実施できる保健指標算出の方策として分娩施設を選択した出生体重の標本調査が考えられる．すなわち，日本の患者調査のように黒竜江省内の分娩施設を均一に選択し，その分娩記録を使用する方法である．しかしそのためには分娩記録への記載事項の標準化を計らなければならない．性・在胎期間・体重はもちろんのこと，奇形の有無・母の年齢・母の職種・分娩方法ならびに中国の特徴として民族の記載も必要である．このような分娩記録から％LBWの大小，早期産・過期産の多寡，体重分布を検討するとともに，標準集団を黒竜江省全体あるいは日本などに設定した調整MBWあるいはIUGL，MLといった指標で，民族をはじめとする交絡要因に配慮した検討が期待できる．

＜教育報告＞

# 発展途上国の結核対策における喀痰塗抹検査の質に関する研究
## ―フィリピン国セブ州を例にして―

佐　藤　准　子

# A study on quality of sputum smear microscopy for National Tuberculosis Control Program in developing countries
## - the cases of Cebu province in the Philippines-

Setsuko SATO

The study investigated the quality of sputum smear microscopy on National Tuberculosis Control in Cebu province, the Philippines. The analysis was conducted for cross-checked slides (total number = 9,917 in 1,997) by 4 validators at provincial health office and 40 medical technologists at peripheral laboratories.

As a result, the quality of cross-checked slides showed a comparatively good agreement (kappa=0.64), but some of the laboratories indicated very high false positive or false negative rates. The QC system was structurally established as a part of district health system, however it was recommended that additional duty for a national reference laboratory needed to strengthen the microscopic reading of validators.

The qualities of smear preparation were classified as "dissatisfied" by national standard for more than 90% of total slides. Laboratories with acceptable label in quality of "evenness" were only 4 out of 64(6.3%), and those in "staining" were only 5 out of 64(7.8%).

From logistic regression analysis, false positive in cross-checking was caused by "sputum quality", "cleanness of slides" and "type of examination", and false negative was caused by "thickness". As for the laboratories, high false positive rates attribute to "encouragement by validators", "Med Tech's understanding importance of overall management", and high false negative rates attribute to "satisfaction for supervision" and "asking problems in supervision".

To improve the quality of sputum smear microscopy and its system, it was important not to increase the frequency of supervisory visits but to understand sufficiently the needs of medical technologists through the observing examination as well as discussion at each laboratories, and to give an encouragement and a supportive advise for improvement.

**Supervisor**: Nobuyuki HYOI

## I. はじめに

近年の経済成長ならびに医学教育の整備に伴い，発展途上国においても臨床検査は診断に必要不可欠となってきており，臨床検査サービスの量は拡大してきたものの，その質は必ずしも保証できる状況とは言いがたい．特に地域住民の保健医療サービスを担う末端検査室における臨床検査の質の向上を図るためには，そこで検査業務を担当する検査技師に精度管理の重要性を広く認識させるとともに，地域における外部精度管理システムを確立し，さらに，そのシステムが機能するため効率的効果的な管理運営も含めて検討することが重要である．

一方，結核は，成人における単一病原体における主な死因という点で，公衆衛生上，最重要課題の感染症の一つである．結核対策は，感染の予防，発病の予防，早期発見および治療からなり，その中でも感染性の患者を正しく発見し，効果的な化学療法を患者管理の下で実施することが重要である．WHOは近年効果的な結核対策として，直接監視下短期化学療法（Direct Observed Treatment, Short-Course：以下DOTS）の推進を積極的に展開し，その中で

指導教官：兵井　伸行（保健統計人口学部）

も「喀痰塗抹検査による患者発見」を重要な要素の一つとしている．特に，末端検査室における喀痰塗抹検査は，患者を早期発見するための重要な診断であり，また，治療経過評価手段としても重要な柱であることから，塗抹検査の質の確保は結核対策の成否を左右すると思われる．しかし，これまで塗抹検査結果の質に影響する要因の分析はほとんど検討されてこなかった．

本研究では，発展途上国での地域レベルでの結核対策における塗抹検査の質の向上について検討することを目的とし，フィリピン国セブ州を対象に研究を行った．

## II. 対象の概要

フィリピン国は日本の８割の広さに7000以上の島々からなり，人口７千３百万人，国民一人あたりのＧＮＰはUS$916，近年保健医療状況が改善されてきたと言うものの，結核の年間推定発生数22.4万人と，世界７番目に結核患者数が多く，厳しい状況である．今研究で調査対象地域としたセブ州は，フィリピン国の南部に位置し，人口約300万人（フィリピン全人口の4.1％），48地方自治体からなる州である．

## III. 方法

### 1．調査方法

＜末端検査室の精度管理調査＞

1997年１月～12月にセブ州衛生部管轄下全末端保健医療施設（Rural Health Unit：以下RHU）の末端検査室にて検査された喀痰塗抹標本を対象とし，州衛生部評価者らにより精度管理サーベイとして肉眼ならびに鏡検下にてクロスチェックされた評価結果を分析した．評価項目は，標本作成技術，ならびに喀痰塗抹鏡検の質（末端検査技師の結果と評価者による盲検下での再鏡検結果の比較）とした．なお，喀痰塗抹鏡検の質については，州衛生部評価者と末端技師との判断が陽性・陰性と不一致であった場合，リファレンスラボ（National reference TB laboratory region7）の技師によるアンパイア判定を最終判断結果として用いた．

＜州衛生部評価者の鏡検技術の評価＞

州衛生部評価者と末端検査室技師との判断が陽性・陰性不一致で，リファレンスラボにて最終判断された標本を対象とし，その最終結果と評価者および末端検査室技師の結果とその要因を比較した．

＜意識調査＞

2000年８月から10月まで，現地にて末端検査室の全検査技師ならび喀痰塗末検査全評価者を対象にアンケート調査を実施した．その内容は，経験年数，各検査室の環境，喀痰塗抹検査に対する意識，評価者らによるスーパービジョンの現状ならびに期待することなどとした．

### 2．解析方法

＜精度管理調査結果の評価＞

標本作成評価については，①喀痰の質，②染色状況，③汚れの付着，④大きさ，⑤均質さ，⑥厚さの６指標を用いた．政府の達成目標は各々90％以上とされていた．鏡検結果報告のスケールについては，①陰性（300視野中抗酸菌なし），②１＋（300視野中抗酸菌１－５個），③２＋（300視野中抗酸菌２－24個），④３＋（300視野中抗酸菌25個以上），⑤４＋（毎視野中抗酸菌多数）５段階とした．

末端検査室における喀痰塗抹鏡検結果と州衛生部評価者らによる喀痰塗抹鏡検の質については，その陰性から強陽性まで５段階判定での一致の評価を$\kappa$一致係数を用い，また，陰性ならびに陽性判定二群での感度，特異度，偽陽性率，偽陰性率を算出した．

＜州衛生部評価者および末端検査室技師に対するアンケート調査＞

州衛生部評価者および末端検査室技師の意識について，アンケート調査各質問項目に対する分析を行った．また末端検査室技師については，スーパービジョンに対する満足の有無と関連されると想定される項目について有意差検定を実施した．

＜塗抹鏡検に影響する要因分析＞

偽陽性とそれらに影響を及ぼすと推定される諸要因（検査目的，標本作製に関する６指標）を説明変数，真・偽陽性を従属変数とし，多重ロジステイック回帰解析を用いて，調整オッズ比とその95％信頼区間を算出した．なお，偽陰性についても，同様に解析した．

また，末端検査室に対して，各検査室を「目標達成群（偽陽性率５％以下)」「高偽陽性群（偽陽性率５％以上）」の二群に分類し，「高偽陽性群」へ影響を及ぼすと推定される諸要因との関連について，要因のカテゴリーに応じてMann-Whitney検定又はKruskal-Wallis検定を用いた検討した．さらに，有意とされた因子を同様に検討し，さらに，有意とされた因子を説明変数，目標達成群・高偽陽性群を従属変数とし，多重ロジステイック回帰解析を用いて，調整オッズ比とその95％信頼区間を算出した．また，偽陰性についても，同様に解析した．統計解析には SPSS for windows 8.0.1Jを用いた．

## IV. 結果

### 1．精度管理調査結果の評価

対象となった全標本数は9,917枚，最終鏡検結果は，陰性：6,336枚（63.9％），１＋：398枚（4.0％），２＋：1,067枚（10.8％），３＋：2,018枚（20.3％），４＋：98枚（1.0％）であった．また，クロスチェックの結果は，真陰性6,079枚，真陽性5,097枚，偽陰性116枚，偽陽性257枚，感度96.8％，特異度95.4％，偽陰性率1.9％，偽陽性率6.8％，$\kappa$一致係数0.64であった．末端検査室ごとで見ると，偽陽性率を目標（５％以下）達成していた検査室は34ヶ所（53.1％），偽陰性率を目標（２％以下）達成していた検査室は42ヶ所（65.6％）存在した．

喀痰塗抹標本作製技術について政府は90％以上を達成目標としていたが，「喀痰の質」が目標を達成していた検査室は28ヶ所（43.4％），「染色状態」は目標達成５ヶ所（7.8％），「汚れ」は目標達成17ヶ所（26.6％），「大きさ」は目標達成

9ヶ所（14.1％），「均質さ」は目標達成4ヶ所（6.3％），「厚さ」は目標達成6ヶ所（9.4％）と標本作成技術レベルに問題があることが指摘された．

またこれら標本作成技術の質を一元配置分散分析で検討した結果，末端検査室技師（n=40）によって，あるいは末端検査室（n=64）によって各々に有意なばらつきがあることが明らかとなったが，両者のばらつきにはほとんど差がなかったことから，標本作成技術向上のためには末端検査技師を対象に，検査技術習得のための自覚ならびに訓練の必要性が明らかになった．

**2．末端検査室に従事する技師に対するアンケート調査の検討**

末端検査技師に対するアンケート調査の回収率は100％（40名中40名）で，各技師の平均経験年数は10.4±4.6年，一人の技師が担当していたRHU数は平均1.6±0.7であった．検査室管理運営面に関して，試薬の供給状況は「リファレンス検査センターから定期的に受領可能」の回答が100％であった．また，スーパービジョンに関して，「満足」と回答したのが22名で，その年間訪問回数が「0－2回」10名（45.5％），「3－5回」8名（36.4％），「9回以上」4名（18.2％）に対し，「不満足」5名は「0－2回」5（100.0％）であった．

次に，満足度とスーパービジョンの内容との関連について検討した結果，スーパービジョンの実施内容として「ロジステイックの点検」，また期待することとして「マネージメントのアドバイス」の項目が5％で，また「標本作成のアドバイス」「鏡検のアドバイス」の項目が1％で有意差が認められた．

**3．州衛生部評価者に対するアンケート調査の検討**

回収率は100％（評価者4名中4名）であった．検査技師としての経験年数は平均23±4.1年，各評価者の担当RHU数は15.8±4.1であった．評価者としての仕事については，標本再鏡検の仕事量は「多い」が3人（75％），また4人とも，NTPにおける評価者としての任務は「とても重要／重要」，自分の仕事に対して「とても興味あり／興味あり」と回答していた．スーパービジョンで理想とする訪問回数は「毎月」2名（50％），「各月」2名（50％）であった．

**4．評価者の鏡検技術に関する検討**

中央レベルのアンパイアの最終判断を，州の評価者および末端検査室技師の報告と比較した結果，全体で見ると，末端技師らよりも評価者らの方が，$\kappa$一致係数が高いこと，偽陽性率ならびに偽陰性率が低いことより，鏡検力が高いことが認められた．しかし，評価者ごとに分析すると，評価者間で鏡検力に格差が著しいことが明らかとなった．

**5．塗抹鏡検に影響する要因について**

＜喀痰塗抹標本の鏡検に対する検討＞

多重ロジステイック回帰解析を用いて偽陽性の標本に影響する要因を検討した結果，「検査目的」では『診断のため』と比較して『治療経過判定のため』が2.83倍（p=0.000），「喀痰の質」では『適切な標本』と比べて『不良』が1.88倍（p=0.000），「標本の汚れ」では『汚れなし』と比べて『汚れあり』が2.00倍（p=0.000）偽陽性を招きやすいことが明らかとなった．また，偽陰性の標本に影響する要因を検討した結果，「標本の厚さ」で『厚い標本』が4.06倍（p=0.005）偽陰性になりやすいことが認められた．

＜末端検査室に対する検討＞

多重ロジステイック回帰解析を用いて，高偽陽性率の末端検査室に影響する要因を検討した結果，スーパービジョンの際に望むことについては，「マネージメントのアドバイス」を『望む者』に対して『望まない』が9.36倍（p=0.055），「励まし」を『あまり望まないもの』に対し『望む』が6.73倍（p=0.085）偽陽性を招きやすいことが明らかとなった．また，高偽陰性率の検査室に対しては，「スーパービジョンに対する技師の満足度」では『満足』に対し『不満足』が4.23倍（p=0.054），「評価者が悩みの相談にのってくれるか」では『いつも／しばしば』に対し『ほとんどない／一度もない』が2.38倍（p=0.036）偽陰性を示しやすいことが明らかとなった．

## V．考察とまとめ

**1．** セブ州の喀痰塗抹検査の精度管理調査について，検査室ごとに見ると，$\kappa$一致係数，偽陽性率ならびに偽陰性率に問題がある検査室が存在し，検査室により塗抹検査の質に格差があることが明らかになった．セブ州全体の検査の質向上のためには，先ずこれらの検査室の鏡検力の向上に向けた対策を重点的に行う必要が認められた．

**2．** 塗抹標本作成技術を評価した結果，政府目標を達成していた検査室は，「均質さ」では4検査室（6.3％），「染色状況」では5検査室（7.8％），「厚さ」では6検査室（9.4％）のみ，最も良い指標であった「喀痰の質」でも28（43.4％）と過半数を満たさず，標本作成技術の指導を強化する必要性が示された．

**3．** 評価者4名の鏡検の質について，全体としては末端検査技師らよりも鏡検の質は高かったものの，評価者ごとに分析すると評価者間の鏡検力に格差があり，そのうち2名が評価者として問題であることが明らかとなった．リファレンスラボは，定期的に既知の標本を評価者に配布し，その質を評価するなど，評価者としての鏡検力の確保に向けて機能する必要性が認められた．

**4．** 途上国の地域臨床検査サービスにおいて，試薬・検査機器具など消耗品の供給や，その質が問題となりがちであるが，対象地域では定期的に安定供給されており，供給元も中央のリファレンスラボからで染色試薬の質も確保されていた．

**5．** 塗抹標本鏡検に影響する要因として，偽陽性に「喀痰の質」，「標本の汚れ」，「検査目的」，偽陰性に「標本の厚さ」が有意と指摘されたことから，これらの指標を中心に標本作成技術の向上に向け，適切な指導をすることが効果的

だと示唆された．

**6.** 高偽陽性率検査室の技師は，検査のマネージメントのアドバイスを望まず，賞賛を望む傾向であることが明らかとなった．偽陽性率が高い場合，スーパービジョンの際も注意や指導が中心となりがちだが，少しでも標本作成技術の上達あるいは鏡検力の向上が認められた場合には賞賛し，末端技師の志気を促すことも重要であると示唆された．

**7.** 高偽陰性率検査室の技師は，スーパービジョンに対して不満があり，悩みの相談をあまりしない技師ほど偽陰性になりやすいことが明らかとなった．現場での観察による具体的な指導，悩みの相談，励ましなど個々の末端技師のニーズを把握し，それに応じたいわゆる「サポーティブスーパービジョン」が検査の質の向上に結びつくことが認められた．

## VI. 提言

フィリピン国セブ州以外でこの精度管理事業を推進していくにあたり，その地域の保健医療システムに基づいた体制を整備するとともに，評価者の質の確保を含めた継続的な人材育成，ならびにサポーティブスーパービジョンの具体的なあり方を検討し提案することが望まれた．