

検査による異常値検出の考え方 ～基準値とは何か？ とくに フィールド健診データについて

中澤 港

<minato-nakazawa@people.kobe-u.ac.jp>

妥当性を考えてみる

- ・ ストレス評価の指標として(採血だとそれ自体が与えるストレスも問題)
 - 質問紙… POMS-J (過去1週間の緊張-不安等, 5件法)など。自覚の有無, 読解力, 性格等に影響される。中長期的。
 - 心拍ゆらぎ… 心電図のR波間隔のポアンカレプロットの長軸と短軸の比が小さい(即ち, R波間隔の自己相関が小さい)とストレス・緊張状態。短期。
 - 唾液中アマラーゼ(酵素法)… 血中ノルエピネフリンと直接神経作用により数分以内に分泌亢進。ごく短期。個人差, 日内変動大。
 - 唾液中コルチゾール (ELISA 法)… 短期間で反応。血中遊離コルチゾールと相関。日内変動(起床時高値で午前中減少)あり
 - 唾液中クロモグラニンA (CgA) (ELISA 法)… 運動負荷には反応せず, 精神的ストレスにはコルチゾールより鋭敏。ノルエピネフリン支配らしい。メカニズム不明
- ・ 栄養素摂取量 = $\Sigma(\text{食物摂取量} \times \text{その食物の栄養素含量})$
 - 24時間思い出し法… 調査時点で先立つ24時間以内に食べたものだけ。無意識に食べているものは得にくい。定量性は中程度。
 - 食物摂取頻度調査票 (FFQ)… 食習慣。定量性は低い(半定量式も)
 - 直接秤量法… 食物摂取は正確。調査の同意を得るのが難しいので代表性が問題。かつ, 微量栄養素は, 同じ食物でも土壌により含有量が異なるので, その土地の食品成分表がないと, 食物摂取量が正確でも無意味

精度 (precision) を保つために

- ・ 同一サンプルを繰り返し測定(または duplicate や triplicate で同時に測定)して, CV が小さいことを確かめる
- ・ CV (Coefficient of Variation) は, 標準偏差を平均値で割った値(通常は 100 を掛けて%表記)
- ・ 発想としては, 測定値が真値 \pm 測定誤差の結果であり, 測定誤差が平均ゼロの正規分布に従うと考え, 誤差の標準偏差が測定値そのものに比べて十分小さい(例えば5%未満)なら測定値は信用できると考える
- ・ 異なるサンプルの測定値のばらつきを示すのに CV を使うのは誤用。標準偏差そのものを見るべき。
 - サンプルから母集団におけるデータのばらつきを推定するには不偏標準偏差(不偏分散の平方根)を用いる
- ・ サンプルから母集団の平均値を繰り返し推定した場合, その平均値がどの程度ばらつくか(即ち平均値の標準偏差)を示すのが標準誤差。サンプルから得られた不偏標準偏差をサンプルサイズの平方根で割った値になる

参考文献

- ・ Stephen Jay Gould "The mismeasure of man" W.W.Norton & Co., 1981. (邦訳: スティーブン・J・グールド『人間の測りまちがいがいー差別の科学史(上・下)』河出文庫)
- ・ David L. Streiner & Geoffrey R. Norman "Health measurement scales: A practical guide to their development and use. 3rd Ed." Oxford University Press, 2003.
- ・ 佐藤俊哉『宇宙怪人しまりす 医療統計を学ぶ』, 岩波科学ライブラリー, 2005.
- ・ 村上道夫, 永井孝志, 小野恭子, 岸本充生『基準値のからくり: 安全はこうして数字になった』講談社ブルーバックス, 2014.
- ・ 村山航 妥当性概念の展開
http://koumurayama.com/koujapanese/validity.ppt
- (2012) 妥当性. 教育心理学年報, 51: 118-130.

信頼性 (reliability) と妥当性 (validity)

- ・ 信頼性は安定性, 再現性 (test-retest reliability) や測定者間一致度 (inter-observer concordance) や項目間一致度 (Cronbach の α 係数等) で示される
 - システム誤差がなければ信頼性と妥当性は一致
- ・ 本質的に直接測定が不可能な場合(質問紙によるストレス評価など), 3種の妥当性を確保する必要あり。
 - 内容的妥当性 (content validity): 専門家の判定
 - 基準関連妥当性 (criterion validity): 併存 / 予測について既存指標との相関が高いこと
 - 構成概念妥当性 (construct validity): 収束的妥当性 (convergent validity) や弁別的妥当性 (discriminant validity) とも関連する
- ・ 村山 (2012) は, 妥当性 = 構成概念妥当性であり, 他の「 \triangle \square 妥当性」は構成概念妥当性を検証するための方法・証拠のタイプと述べている
 - 内容的証拠(専門家が妥当と判断), 収束的証拠(類似概念を知るためのテスト指標と高い相関), 弁別的証拠(異なる概念を知るためのテストと低い相関), etc.
- ・ 実用上は侵襲性も考慮。ストレス評価の場合, POMS-J のような質問紙は非侵襲, 心拍や唾液や光トポグラフィや脳波は低侵襲, 血液は高侵襲。一般に, 高侵襲な指標ほど真に測りたいものに近いので妥当性が高いことが多いが, そこまで高い妥当性が必要でなければ低侵襲な方が倫理的に良い

基準値のいろいろ

- ・ 対象の状態を示す値についての基準値(正常範囲)
 - 本講義で説明する
- ・ 環境基準, 曝露基準など外的基準値～それぞれ考え方は違う
 - ダイオキシンの耐容一日摂取量
 - 産業保健における有機溶媒の2つの基準値
 - ・ 厚労省
 - ・ 産業衛生学会
 - 一般人の人工放射線曝露 1 mSv

測定値そのものの「正しさ」

- ・ まずは「正しく」測れていることは大前提。以下3つがともに達成できている必要がある。
 - 妥当性 (validity): 測りたいモノを測れていること。違うモノを測っていないか? 数値以外の保証が必要 (例: 肥満を BMI で正しく評価できる? ハンマー投げの室伏広治氏は 187cm, 99kg なので BMI が 28.3 だが肥満でない / 非接触式で額や耳の穴で測定する赤外線体温計はいつも正しいか?)
 - 正確さ (accuracy): システム的なズレが少ないこと。ゼロ点があてはまらない測定は不正確。標準物質 (reference material) の測定結果が certified range に入っているか
 - 精度 (precision): 偶然誤差が少ないこと。繰り返して測ったときのズレが少ないこと。CV が小さいことで保証。
- ・ この3つのうち1つでも不十分だと, 測定値は「正しくない」かもしれないので無意味

正確さ (accuracy) を保つために

- ・ ゼロ点調整の重要性
 - 電子天秤で秤量するとき, 試薬皿を載せた状態で tare しておかないと, 試薬皿の重さだけ少なく量ってしまう
 - 原点を通る検量線を描きたいとき, ブランクで吸光度がゼロになるように調整する(このときのブランクは水ではなく, 対象物質の濃度がゼロで試薬は入っている, 試薬ブランク)
- ・ 標準物質 (NIST Standard Reference Materials 等; <https://www.nist.gov/srm>)
 - 通常, 自分が測りたい物質と同様の物質を使う一測定結果が certified range に入っているかどうかを確認する (certified range 自体の正しさは複数の reference labo での測定で相互保証)
- ・ 正しい検量線を作るには
 - 標準添加法: 共存物質が影響するとき, 試料溶液を分けて標準希釈系列を添加し混合したものの吸光度を測定し, 添加濃度を横軸にとつて検量線を描く。ゼロ点調整は水ブランク。吸光度ゼロに相当する添加濃度(マイナスになっている)が試料の濃度(試料は最初からそれだけその物質を含んでいると考える)。
 - 内標準法: これも共存物質の影響を除くため, 測定物質と似ていて測定対象でない物質を内標準物質として標準試料と未知試料に添加し, 内標準物質と標準試料の吸光度比を検量線の縦軸にとる

基準値って何?

- ・ 検査の目的の1つは, 異常の検出
- ・ 何らかの測定値から異常を検出するには, 「この範囲なら異常でない」値(正常値または基準値)を知る必要がある
 - 範囲外の原因 = 測定ミス, でなければ異常
- ・ 例) 東京大学医学部附属病院検査部「血液検査の参考基準値表」(注: 最近は基準値というのが普通)
<http://lab-ky.umin.jp/patient/ketueki.pdf>
には, 項目により2種類の基準値
 - 基準範囲は, 健康な成人の検査値から, その 95% が含まれる範囲(例: CK は男性 55-210, 女性 44-166 U/L)
 - (疫学研究などから得られた)臨床的に診断・治療・予後に関わる判断を下す閾値である臨床判断値(例: PSA 4ng/mL を超えると前立腺がん要精検, 等)

「正常値」=「基準範囲」決定の4パタン

- パタン1: 健康な大勢の人々について予め測定しておき、そのうち多くの人が入る範囲を「正常」とする。外れたら「異常」
 - 「多くの人が入る」は、平均 $\pm 2SD$ とか $3SD$ とか、2.5パーセンタイルから97.5パーセンタイルまで、など統計学的考え方による
 - より小さな標本内の人々が、母集団の代わりとしての大標本の平均 $\pm 3SD$ 外ということは滅多に起こらない(正規分布なら上下各0.135%のみ)ので、たぶん異常であろうと推論できる
- パタン2: 健康な大勢の人々の測定値と病気と診断されている人の測定値の分布を比べて、2群を分ける閾値を探索し、閾値より健康な人々の側にくる値を正常とする
- パタン3: 先進国(とくに米国)の集団のパタン1の正常範囲が人類の正常範囲であると見なす
- パタン4: 機能に障害が出る濃度以下を欠乏、機能に障害が出る濃度以上を過剰とみなす(4' 投与に反応したら欠乏とみなす)

パタン1: 成長曲線(母子手帳等に掲載)

図1 乳幼児(男子)身体発育曲線(体重)

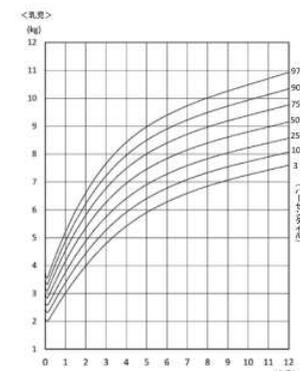
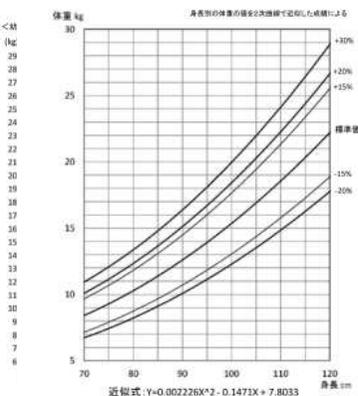
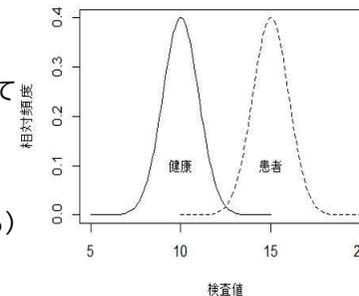


図8 幼児(男子)の身長体重曲線



パタン2は分布がはっきり違えば有効

- 一般の健康な人々(以下健康者群)と患者群の測定値が右下図のような分布なら、12.5より大きければたぶん患者で、健康な可能性はさきめて低いといえる
- しかしもっと分布が重なっていたら、どこを正常上限にしたらいかが決定困難
- 有名な解法の1つがROC曲線(受診者または受信者動作特性曲線とも呼ばれる)



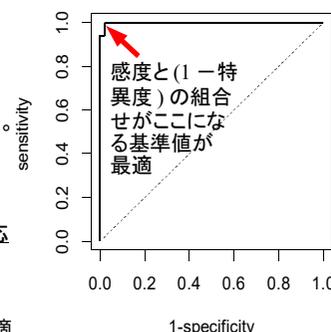
ROC (Receiver Operating Characteristics)

- <http://oku.edu.mie-u.ac.jp/~okumura/stat/ROC.html>
- 三重大学奥村先生の非常にわかりやすく詳しい解説

ROC 曲線の描き方

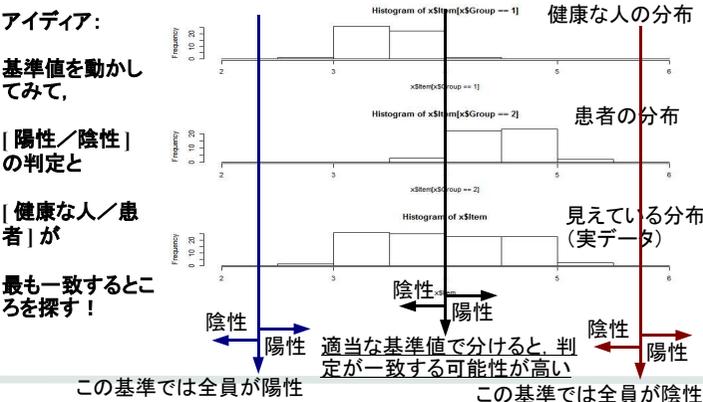
- 健康と分かっている人と患者と分かっている人について、同一の検査の結果による陽性/陰性の組合せから、クロス集計表ができる

	患者	健康な人
陽性	a	b
陰性	c	d
- a/(a+c)を感度(sensitivity), d/(b+d)を特異度(specificity)という。どちらも1に近いほど高性能
- 実際にスクリーニングなどで検査した後で陽性の人に確定診断をしたとき、陽性と出たうち真に患者である確率(上の表で言えばa/(a+b))は陽性反応的中率。稀な病気では低くなる
- 基準値を左(最小値)から右(最大値)に動かすと、感度は1 \rightarrow 0、特異度は0 \rightarrow 1へ変化する。途中のどこかに最適基準値がある



パタン3: ウォーターローによる低栄養の分類で、先進国の基準値を使うこと

- Weight for Height (身長に対する体重)が80パーセンタイル未満だと「痩せ(wasted)」と評価される。短期的な低栄養を示すとされる
- Height for Age (年齢に対する身長)が90パーセンタイル未満だと「短軀(stunted)」と評価される。長期的な低栄養を示すとされる
- 両方該当すると wasted and stunted であり、栄養状態にかなり問題がある可能性があるとして評価される \rightarrow ハイリスク
- 一般には、栄養指導や食糧援助のニーズ評価などに役立つ
- 欠点: 遺伝的な違いを無視している。とくに米国 NCHS のデータを途上国に使うと(パタン3)、遺伝的に成長が遅いだけでも、wasted や stunted の割合が高い(成人男性の身長が160cmに達しないようなところでは、80%が stunted という報告もある……意味ある?)
- 途上国の子供の栄養状態の評価を目的とするなら、米国 NCHS データより WHO の成長曲線(WHO Anthro)を使うべき。
<http://www.who.int/childgrowth/software/en/>



パタン4

- 例) 血液のヘモグロビン濃度は、WHOの基準では授乳も妊娠もしていない成人女性120g/L、成人男性130g/L以上が正常値、それ以下は貧血とされている。
- 110g/Lでもほとんど機能障害は起こらないが、鉄欠乏性貧血であれば鉄剤投与によってヘモグロビン濃度は上昇するし、葉酸欠乏なら葉酸投与、タンパク欠乏であればタンパク投与によりヘモグロビン濃度は上昇する
- 投与して増加するということは、投与しない状態では不足していたのだという論理。
- 欠点: 低鉄血症適応仮説(Susan KentやOppenheimerが提唱)では、マラリア流行地の人は血清鉄を低く抑えることにより感染が重症化しないように適応している可能性。
 \rightarrow これって本当に「不足」? 異常?

「正常値」(基準範囲)を参照して異常値検出=ハイリスクアプローチ向け

- 栄養失調というハイリスクの子どもを検出し、適切に栄養補給することで、子どもの死亡率は低下する。子ども全部に栄養補給するより、効率よく死亡率を低下させられる
- 先進国における高血圧は、確かに高い方が心血管疾患や脳血管疾患のハイリスクだが、1つの分布。
 - むしろ、ポピュレーションアプローチによって、集団全体の血圧分布を低い方にずらす方が有効。この場合は、正常範囲に入っているかどうかよりも、分布の位置の方が重要
- 日本のメタボリックシンドロームの基準(男性の腹囲85cm以上など)では、予備軍を入れると中年男性の半分以上が異常になってしまう。ハイリスクアプローチには不適
 - 基準値を変えるべきという指摘も多くの専門家が何度もしているが、今のところ採用されていない。

遺伝や環境が違えば「正常値」も違う

- 異なる生活環境で遺伝的バックグラウンドも異なる人々について、同じ表現型を「正常」といえるのか? 欧米中心主義に過ぎないのではないのか?
 - (例)日本人やピマインディアンにおける $\beta 3$ アドレナリンレセプター多型であるTrp64Argは「異常」か? 長期間の適応の結果だとしたら、「異常」とはいえないのでは?
 - (例)ピグミーやネグリの低身長は「異常」?
- 異なる環境で暮らす人々について、「正常」は一つしか無いのか?
 - (例)マラリア蔓延地帯でのHbSやサラセミアは「異常」か? 貧血は「異常」か?

体格と栄養状態の概況

- 秤量法で推定された摂取エネルギーは成人男性1人一日当たり2500～3000 kcal, タンパク質も十分, 脂質摂取は先進国の半分。鉄摂取が日本の3～10倍あるが, 村によっては貧血あり。
- マラリア患者が多いと言われる村では, Hbが6g/dLの人がいた。冷凍保存した血清を使って, 抗マラリア抗体価と鉄栄養状態を調べた際, 血清鉄濃度は, 当初 ICP-AES の測定値を利用したが査読で reject された
→血清中には溶血して赤血球から出てきた鉄があり, それを除去処理して公定法で測らないと妥当性がない
- 成長はやや遅く, 小柄だが筋肉質
- 詳しくは, 鈴木継美『パプアニューギニアの食生活: 塩なし文化の変容』中公新書を参照

血清生化学検査値

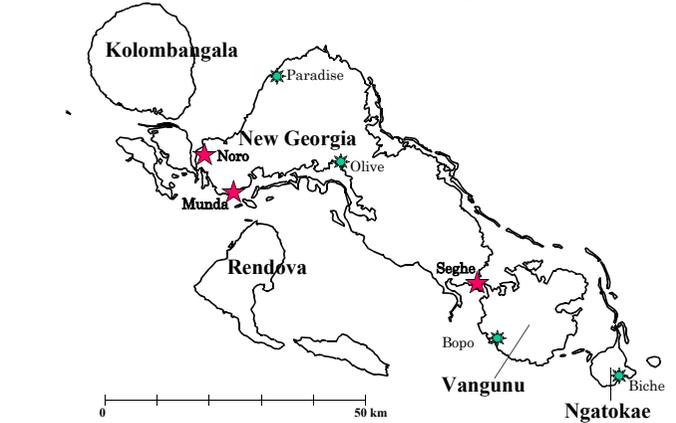
日本の基準値に比して高いが必ずしも異常でない

Table 1. The 10th and 90th percentiles of serum biochemical data

Indices*	N	Male		Female		Total			
		10th	90th	10th	90th	10th	90th		
γGTP (I U/L)	231	9.0	23.5	288	7.4	17.1	519	7.7	20.0
ALP (I U/L)	231	41.6	79.2	288	36.0	77.1	519	37.1	78.5
AST (I U/L)	231	11.9	35.4	288	10.5	31.0	519	10.7	32.7
ALT (I U/L)	231	6.5	23.2	288	6.3	19.5	519	6.3	20.9
LDH (I U/L)	231	158.0	260.0	287**	158.0	252.0	518	158.0	256.0
CK (I U/L)	231	99.0	528.0	288	72.0	315.0	519	79.0	454.0
BUN (mg/L)	231	69.0	219.0	288	64.0	214.0	519	67.0	217.0

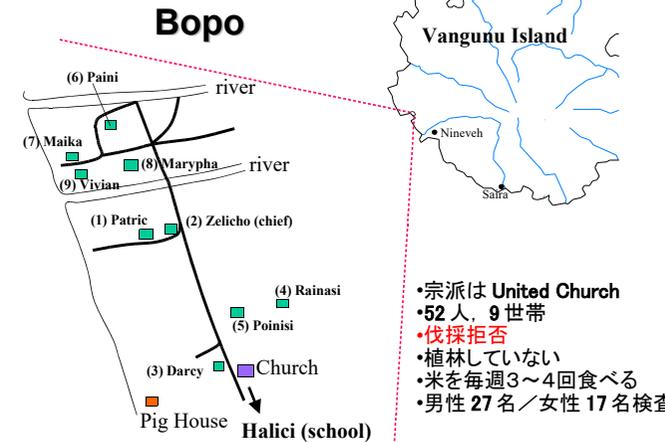
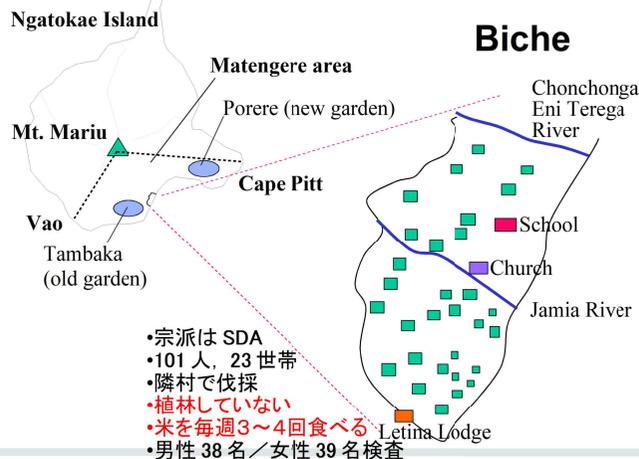
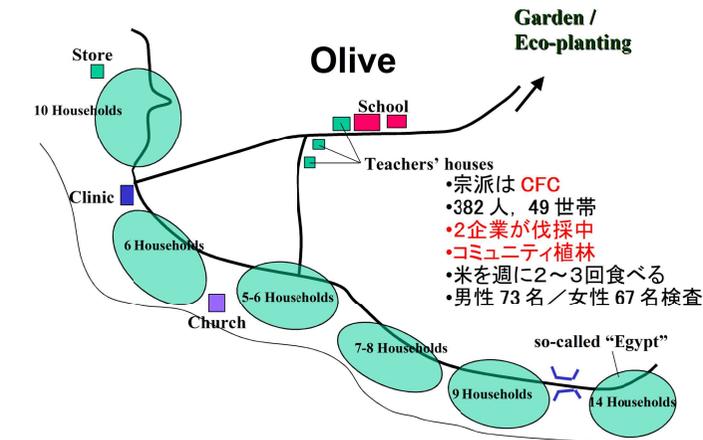
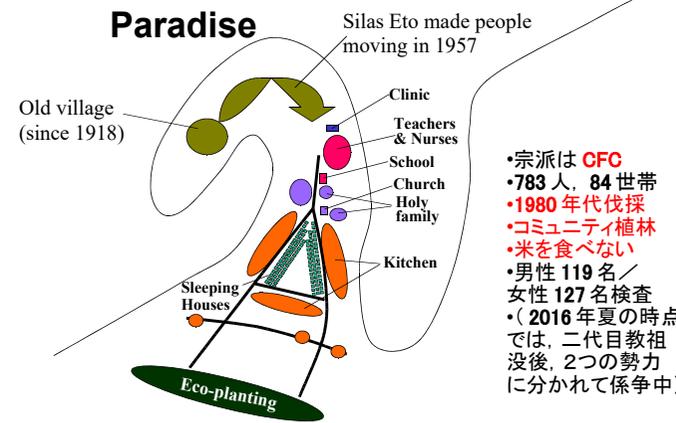
* Abbreviations used in the all tables: γGTP, γ-glutamyl transpeptidase; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine amino-transferase; LDH lactate dehydrogenase; CK, creatine kinase; BUN blood urea nitrogen.
** One of the samples had insufficient content.

ソロモン諸島ウェスタン州4村落の尿検査



対象4村落の概況

項目	パラダイス村	オリヴェ村	ビチエ村	ボボ村
言語族	クサゲ	ロヴィアナ	マロヴォ	ヴァングス
村の成立	1958年	1972年	100年以上前	1924-26年
キリスト教宗派	CFC	GFC	SDA	UC
部族集団数	4	3	1	1
現在人口	783	382	101	52
世帯数	84	49	23	9
森林伐採	1980年代	2企業が実施中	1990年代+隣村	なし
植林	CW / 個人	CW / 個人	なし	なし
米食	なし	2-3回/週	3-4回/週	3-4回/週
尿検査対象者	成人のみ	成人のみ	全員	全員
男性対象者数	119/140	73/79	38/44	27/29
女性対象者数	127/139	67/68	39/44	17/21

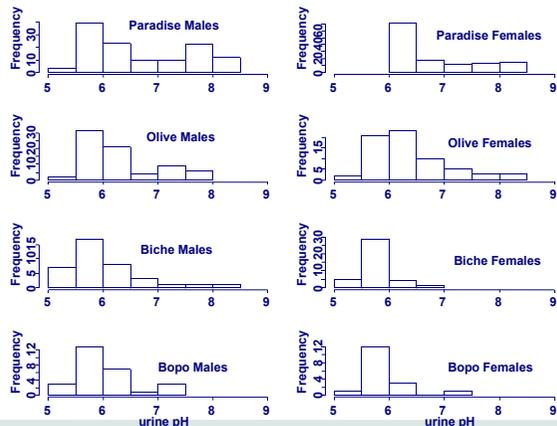


性・村落別陽性(±は含まない)出現頻度と尿 pH (平均 ± SD)

検査項目	パラダイス村		オリヴェ村		ビチエ村		ボボ村	
	男性	女性	男性	女性	男性	女性	男性	女性
グルコース	1	3	1	1	0	0	0	0
ビリルビン	0	0	1	0	0	0	0	0
ケトン体	1	0	0	0	1	0	0	0
潜血	2	7	3	8	0	1	2	1
タンパク	1	3	0	0	1	0	1	1
ウロビリノーゲン	0	0	1	0	0	1	0	0
亜硝酸塩	1	2	1	2	0	0	0	1
白血球	3	7	5	5	0	4	0	0
pHレベル	6.9±1.0	6.9±0.9	6.5±0.7	6.6±0.7	6.2±0.8	6.0±0.4	6.2±0.7	6.1±0.5
検査総数	119	127	73	67	38	39	27	17

¹ 女性は月経血による偽陽性の可能性がある。

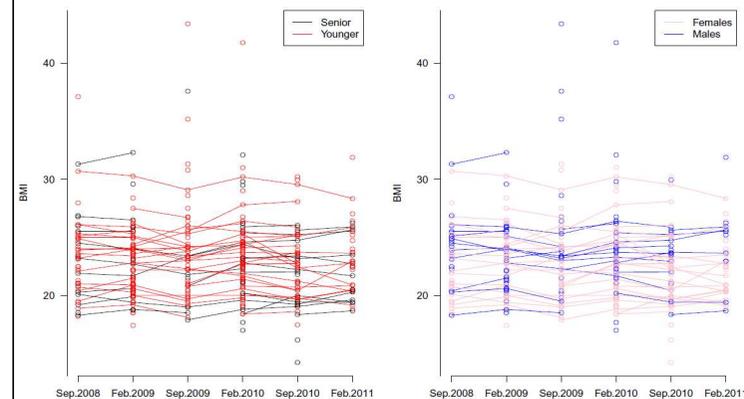
性・村落別尿中 pH レベルの分布



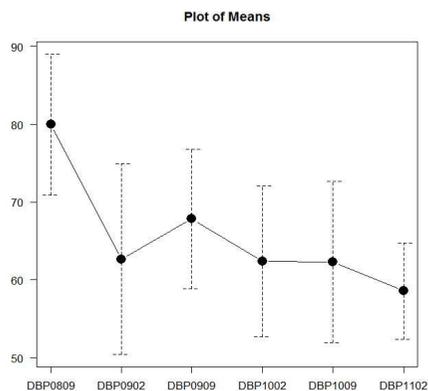
健診項目と方法

- マラリアの積極的患者調査及び血液中 Hb, 随時血糖値
 - 指先穿刺。マラリアは塗沫標本の鏡検, Hb はアムコ社の Hb201+, 血糖値はニプロフリースタイルフラッシュで測定
- 生体計測
- 血圧
 - オムロン社自動血圧計による
- スポット尿の分析
 - N-Multistix® SG-L (Siemens Healthcare Diagnostics K.K., Tokyo) 迅速検査, CARDY イオンメータ C-121 で塩分, C-131 (Horiba, Ltd., Kyoto) でカリウムイオン
- 聞き取り: 24 時間思い出し食事調査, 蚊帳利用, 発熱頻度, 人口学的情報
- ソロモン諸島と群馬大学の倫理審査委員会により承認

個人を線でつなぐと変動がわかる



平均 DBP (収縮期血圧)の推移



DBP の反復測定分散分析

Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

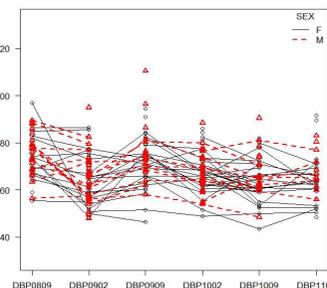
	SS	num Df	Error SS	den Df	F	Pr(>F)
(Intercept)	155105	1	672.22	5	1153.6713	4.16e-07
Time	1746	5	2094.69	25	4.1674	0.006847

Mauchly Tests for Sphericity

Test statistic	p-value
Time	0.0053433

Greenhouse-Geisser and Huynh-Feldt Corrections for Departure from Sphericity

GG eps	Pr(>F[GG])	HF eps	Pr(>F[HF])	
Time	0.56326	0.02791	1.3645	0.006847

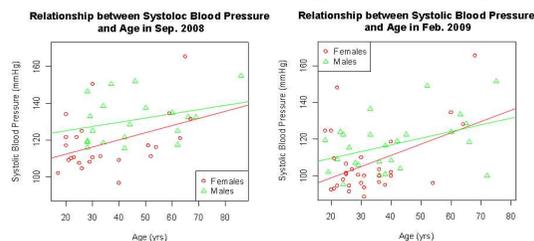


2008 年以降の塩分摂取, 肥満, 血圧

- 背景としての INTERSALT 研究
 - Intersalt Cooperative Research Group, BMJ 297: 319-328, 1988.
 - 世界 52 集団の尿中 Na/K 比はブラジル Yanomamo の 0.01 ± 0.04 から中国 Tianjin の 7.58 ± 2.41 まで大きくばらついている。これは主として塩分摂取の違いによる
 - 4 集団 (Yanomamo, Xingu, PNG, Kenyan) は Na 排出量が極めて低く, 血圧も低く, 加齢しても血圧がほとんど上昇しない (傾き $0.052-0.206$; cf. 先進国で高齢者で血圧上昇するのと対照的)。
 - 尿中 Na/K 比は BMI や収縮期血圧と弱い正の相関を示す
- 背景としての Papua New Guinea, ギデラの研究
 - Inaoka T, 1990. In: Ohtsuka R and Suzuki T [Eds.], "Population Ecology of Human Survival", Univ. Tokyo Press, pp.159-174.
 - PNG ウェスタン州低地に居住するギデラ語を話す人々の尿中 Na/K 比は, 1981 年に比べ, 1986 年には 2 倍から 3 倍に上昇した。このとき彼らの塩分摂取は増加していた

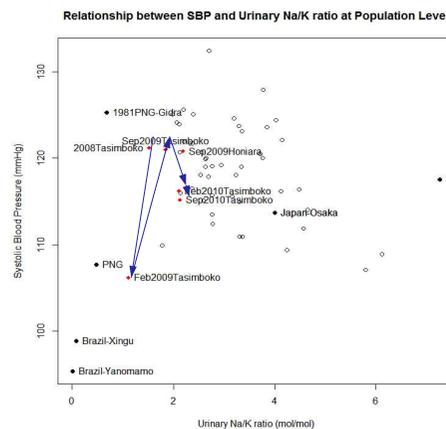
年齢と収縮期血圧 (SBP) の関係

- 収縮期血圧 (SBP) は低くなかった
 - 2008 年 9 月は, 131 ± 13 mmHg (男性) と 119 ± 16 mmHg (女性)
 - 2009 年 2 月は, 117 ± 15 mmHg (男性) と 107 ± 18 mmHg (女性)
- 2008 年も 2009 年も, 男女とも, 年齢との相関は弱いけれども統計学的に有意だった (有意水準を 5% として)
- SBP は加齢とともに緩やかに上昇した
 - 2008 年 9 月の傾きは, 0.36 ± 0.16 (男性) 0.61 ± 0.23 (女性)
 - 2009 年 2 月の傾きは, 0.23 ± 0.17 (男性) 0.39 ± 0.19 (女性)



尿中 Na/K 比と SBP の関係 ~ 集団間比較

- 尿中 Na/K 比 (塩分摂取の指標) と SBP (収縮期血圧) の関係
- INTERSALT データに, タシンプコのデータを重ね描き
- アマゾン集団と先進国の中間
- 全体の傾向から比べると塩分摂取の割に血圧が高め
- 経時的には一定の傾向なし



まとめ

- ・ まずは測定値そのものが正しくなくては無意味。妥当性, 正確さ, 精度がすべて達成されるように。
- ・ 「正常値」「基準範囲」には幅がある。考え方も主なものだけでも 4 通り。その幅は環境条件と遺伝条件両方の影響を受けて変動する (測定方法や機器による違いもある) ので, 個人の検査値の異常を検出するには, その集団固有の「正常範囲」を設定すべき。
- ・ 健康指標を使って集団の健康状態を評価するのは公衆衛生学や疫学の常套手段だが, 正常値が集団固有のものである以上, その解釈には注意が必要。複数の指標値を組み合わせた多角的評価が必要。