

乳幼児身体発育曲線作成のための平滑化ソフト活用法に関する研究

研究分担者 加藤則子 (十文字学園女子大学 教育人文学部)

研究協力者 磯島 豪 (帝京大学医学部 小児科学講座)

研究要旨

2020年に実施予定の厚生労働省の乳幼児身体発育調査をもとに、よりよい身体発育曲線を作成するため、平滑化法に関する検討を行った。2020年調査における乳幼児身体発育曲線の作成に向けては、LMS法を採用し、具体的には国立保健医療科学院横山徹爾部長作成のSASプログラム、LMSchartmaker®、GAMLSS in R 等の方法が考えられている。この度、多用されているLMSchartmaker®のサポートが終了していることが確認されたため、R言語で動くGAMLSSパッケージの活用について検討する目的で、前回2010年に実施された厚生労働省の平成22年乳幼児身体発育調査のデータを用いて平滑化を試行した。プログラムの作動には問題なかった。男子身長、胸囲(男女)、頭囲(男女)については、2010年公表値とほぼ変わらない平滑化曲線が得られた。男子体重では、条件を変えた平滑化曲線をつなげ合わせることで、ほぼ変わらない平滑化曲線を得た。女子身長、女子体重では若干の改善の余地が残った。実際のデータセットに取り組むに当たっては、個々に出てくる問題への対応が必要となると考える。

A. 研究目的

厚生労働省の平成22(2010)年乳幼児身体発育調査において公表された乳幼児身体発育曲線は、LMS法を用いて平滑化されている。LMS法は身体計測データ等の分布の特徴をとらえて年齢軸に対する平滑化曲線を作成するためのよく使われる方法である。L(ゆがみ)M(中央値)S(ばらつき)に関する変数を、Box-Cox変換により正規化したデータを用いて算出し、それぞれを平滑化することにより平滑化分布曲線を得る。

2010年値の算出に当たっては、国立保健医療科学院横山部長作成によるSASプログラムを用いて平滑化の計算を行った。当時LMSchartmaker®がweb上で公開されていた

が、実際の平滑化作業において節点の位置などが明確に把握できる方法としてより優れているとして選択されたものである。

平成30年度に国際的に多用されているLMSchartmaker®によって試作してみたが、自由度について細かく調整しなければならない課題が残った¹⁾。

R言語によって動くGAMLSSパッケージは、国際標準になりつつあり、LMS法の考案者のTJ Cole氏も現在はこれを活用している。

GAMLSSはgeneralized additive model for location, scale and shapeの頭文字をとったもので、セミパラメトリックな回帰モデルに属し、ノンパラメトリックな平

滑化関数も含有する²⁾。Location において位置 (平均値) scale はばらつき、shape はゆがみと尖りに対応し、分布におけるその4つを推計する。目的変数である推計値は、横軸に対する関数としてあらわされる。4つの要素にそれぞれに関して logit, log, fixed, identity 等分布を設定することが可能である。これらの組み合わせにより、100 を超えるタイプの推計法が設定され、R 言語によるパッケージの一部として無料で公開されている。4つの要素の分布のタイプ設定の他に追加設定を行うことが出来、その中の一つに三次スプラインも含まれている。平滑化法に関する設定である。

100 個以上提供されている回帰のタイプの一つは、LMSChartmaker®の機能³⁾ と同じものであり、BCCG distribution と呼ばれている。BCCG は Box-Cox Cole and Green を意味する。R 内での表現は BCCG() となっている。LMSChartmaker Pro® では回帰の妥当性を検証する Q-Q plot を返してくるが、これは GAMLSS の BCCG distribution の回帰でも計算される機能である。

今年度においては、この GAMLSS パッケージを用いて平成 22 年乳幼児身体発育調査データに対する平滑化を試みた。

B. 方法

GAMLSS パッケージによって提供されているプログラムを動かすにあたっては、R 言語に関する書籍、web で提供されている情報、GAMLSS に関する書籍⁴⁾を参考にスクリプト作成を試みた。

平滑化法の演算に関して、男子体重を例として示す。

1) R 言語は無償で提供されており、CRAN の web ページからダウンロード可能であるが、国内のミラーサイト (例:統計数理研究所) からのほうが、短時間に可能である。これに合わせて、アプリケーション RStudio もインストールすると、ユーザーフレンドリな使い勝手となる。GAMLSS パッケージも無料で提供されているので、インストールしておく。

2) 演算の実際としては、まず男子の体重に関してエクセルファイルからデータを読み込む。

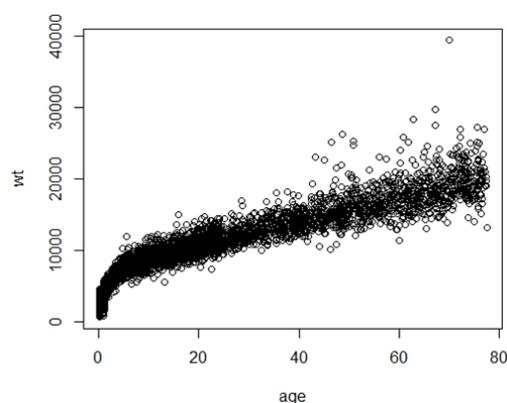
```
> library(readxl)
> mw<- read_excel
+("G:/GSyokoyama/191102/mw.xlsx")
```

age	wt
10.61190965	9390
4.73100616	7305
14.22587269	9500
22.63655031	10800
11.40041068	8840
15.37577002	8690
15.37577002	8680

mw というオブジェクトにデータが入る。

3) 入ったデータの散布図を見る。

```
> plot(wt~age, data=mw)
```



計算してほしい百分位を指定する。指定しないとデフォルトで計算される。

```
> cent <- c(3, 10, 25, 50, 75, 90, 97)
```

4) GAMLSS パッケージを呼び出す。こうすることで、`lms`, `gamlss`, `wp` などの関数を使える。

```
> library(gamlss)
```

percentile estimation をするとき
は、まず `lms` 関数を使う。

“BCCGo”, “BCPEo”, “BCTo” の 3 つのモデルを
比べる。

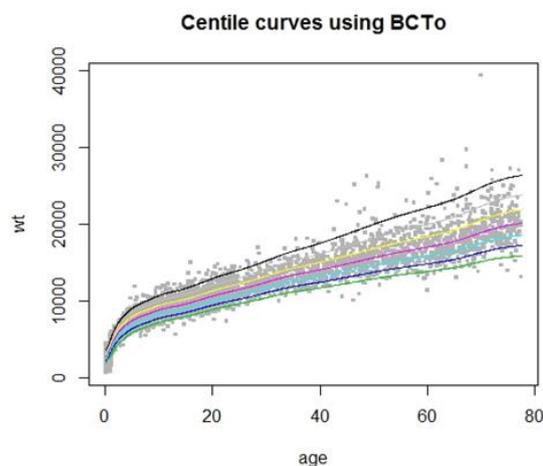
```
m0 <- lms(wt, age,
+ families=c("BCCGo", "BCPEo", "BCTo"),
+ data=mw, k=4, calibration=F,
+ trans.x=T, cent=cent)
```

BCCGo: Box-Cox Cole and Green (Box-Cox normal) Cole が開発した LMSchartmaker® と
同じ計算法である。

BCPEo: Box-Cox Power Exponential 上の
方法の Power に関して、対数変換をして計算するもの。

BCTo: Box-Cox t distribution (four parameter distribution) 中央値、分散、歪度に加えて、尖度も考慮して分布を平滑化する。

5) 平滑化が終わると、データの散布と平滑化曲線のイメージ図が返される。



最もあてはめが良かったモデルを出力する。

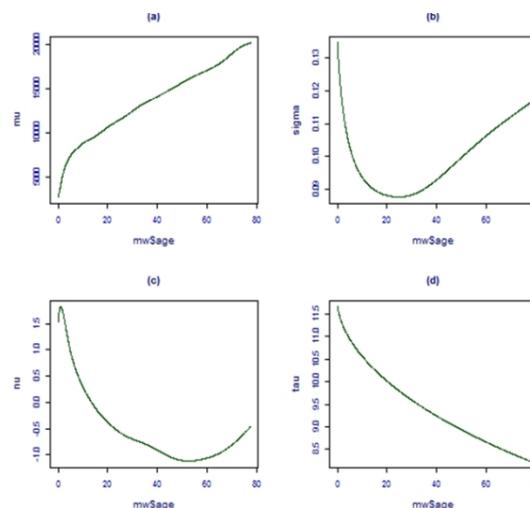
```
> m0$family
[1] "BCTo" "Box-Cox-t-orig."
尖度も考慮した平滑化が最もあてはめが  
良いことが多い。
```

パワーの値を出力する。

```
> m0$power
[1] 0.5894091
```

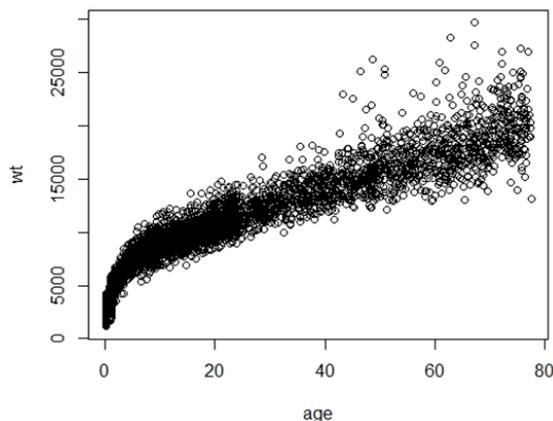
推計された μ σ λ τ の 4 つのパラメータ
を出力する。

```
> fittedPlot(m0, x=mw$age)
```



推計の内容に関する要約

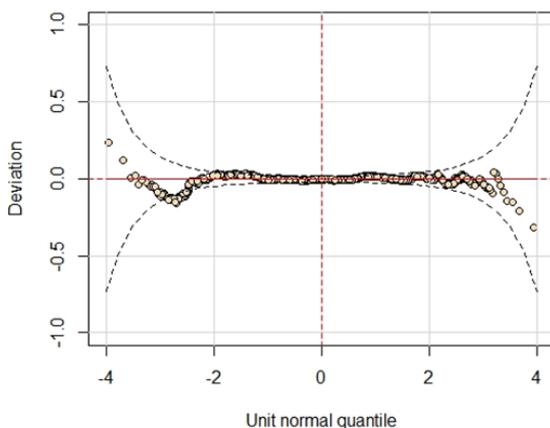
```
> plot(m0)
*****
Summary of the Quantile Residuals
mean    = 0.0001578735
variance = 1.000312
coef. of skewness = -0.001505493
coef. of kurtosis = 3.002107
Filliben correlation coefficient
= 0.9998853
*****
```



散布図を取ってみると、上下の外れ値が削除されているのが分かる。

6) モデルのあてはめについて評価するためにワームプロット(Q-Qプロット)を求め

```
> wp(m0, ylim.all=1)
```



ほぼ上下の点線の間に入っており、あてはめに大きな問題はない。

7) Outlier を除くために、 ± 3 SD 以内の値のサブセットを作成する。

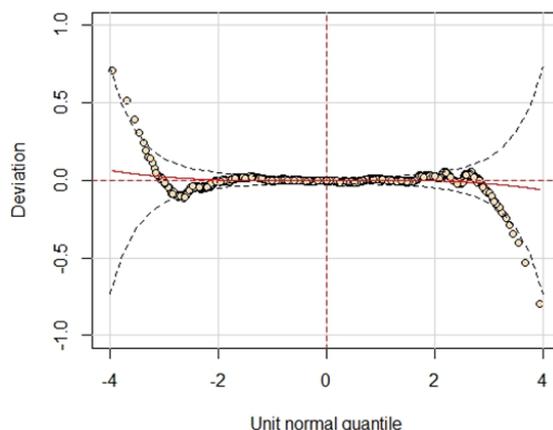
```
> mwsb <- subset(mw,
+(resid(m0) > -3) & (resid(m0) < 3))
2010年値を作成するときもこの操作を行った。
```

```
> plot(wt~age, data=mwsb)
```

このサブセットに対して BCTo による平滑化を行う。

```
> m3 <- gamlss(wt~pb(age^m0$power),
+sigma.fo=~pb(age^m0$power),
+nu.fo=~pb(age^m0$power),
+tau.fo=~pb(age^m0$power),
+family=BCTo, data=mwsb)
```

```
> wp(m3, ylim.all=1)
```



ワームプロット(Q-Qプロット)を取ると、より上下の線の間におさまり、あてはめが良くなっていることが分かる。

8) 平滑化結果を、2010年の公表値と比較するために、年月齢ごとの平滑値を読み

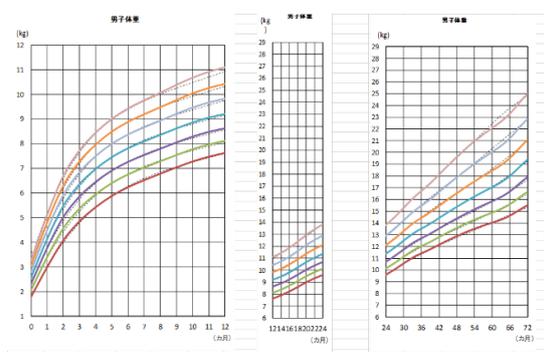
取る。ここでは、1 カ月ごとのパーセンタイル値を読む。

```
> nage<-seq(0, 72, 1)
> centiles.pred(m3, xname="age",
+ xvalues=nage, plot=T, ylab="wt",
+ xlab="age",
+ legend=F, cent=cent)
```

	age	C3	C10	C25
1	0	1776.150	2055.004	2304.835
2	1	2987.715	3401.877	3766.921
3	2	4075.933	4574.206	5020.431
4	3	4847.200	5374.412	5856.238
5	4	5418.879	5952.726	6449.751
	

(50, 75, 90, 97 パーセンタイル省略)

9) このようにして読み取った値を、2010 年公表値のエクセルのグラフに重ね書きする。



の重ね書き結果を評価し、必要があれば改善策を講ずる。男女それぞれ体重、身長、胸囲、頭囲について行う。

C. 研究結果

体重に関しては、生理的体重減少を反映して、日齢別の集計において、生後 3 日付

近までは減少、その後増加という傾向が認められたので、病院調査における生後 3, 4, 5 日と 1 カ月健診、および一般調査データを用いて平滑化することとした (2010 年値の公表時の計算に同じ)。

男子体重において、生後 3, 4, 5 日、1 カ月健診、一般調査データからの平滑化結果を図 1-1 に示す。図の上部が平滑化結果と 2010 年公表値との重ね合わせのグラフ、中段左が、元データの散布図、中段右が $\pm 3SD$ 以内のデータのための散布図、下段左が $\pm 3SD$ 以内のデータで行った平滑化における、 μ σ λ τ の平滑化曲線、下段右が同じく $\pm 3SD$ 以内のデータによる平滑化の適合度の良さを見るワームプロット (Q-Q プロット) である (以降の図に関しても同じ)。

平滑化によりおおむね 2010 年公表値と重なる平滑化が得られたが、54 カ月以降曲線が波うつことが分かった。平成 31 年 3 月 TJ Cole 氏との面談の際、就学前までのデータによる平滑化がうまくいくのは、5 歳程度までとのコメントがあり、やむを得ないことと思われたが、乳幼児身体発育調査データの特性として、生後数日に大量のデータが集積していることが平滑化を難しくしているとのコメントもかつてあった (2000 年調査時丹後氏)。

このような傾向を緩和させるため、生後 3, 4, 5 日分を使用せず、1 カ月健診と一般調査データのみで平滑化を行った。この結果を図 1-2 に示す。54 カ月以降の曲線の波うちが軽減されたが、1 カ月よりも若い月齢での曲線が得られない。図 1-1 と図 1-2 を見比べると、生後 2 カ月以降 48 か月くらいまでの間の月齢で適宜二つの平滑化曲線を連結することが考えられる。実際の運用と

しては、グラフの切れ目を考慮して、12 カ月もしくは 24 カ月における連結が妥当であろう。

女子体重に関して、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 2-1 上段のように、はなはだ揺れの大きい平滑化曲線が得られた。±3SD 以内のデータに絞っても下段のように、同様の平滑化結果が得られた。

生後 3, 4, 5 日のデータを除くと平滑化が安定する経験則から、1 カ月健診と一般調査データで平滑化を行ってみた結果を図 2-2 に示す。42 カ月以降の波うちは認められるものの、2010 年公表値とほぼ同様の平滑化が得られた。

上記の方法で得られなかった女子体重の 1 カ月未満の平滑化値を得るために、24 カ月以降に外れ値が多かったことを鑑みて、生後 3, 4, 5 日、1 カ月健診、24 カ月までの一般調査による平滑化を行った (図 2-3)。1 カ月未満の平滑化値は得られたものの、1 カ月以降で 2010 年公表値とのずれが出て、課題が残った。あらかじめデータクリーニングを徹底してから演算する必要が示唆された。

男子身長に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 3 のような結果が得られた。42 カ月以降の中央値にわずかなずれが見られたものの、2010 年公表値とほぼ良好な一致を見せた。

女子身長に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 4-1 のように、2010 年公表値との間でほぼ良好な一致を見たが、48 カ月以降の曲線の波うちに課題が残った。

平滑化曲線を安定させる目的で、1 カ月健診と一般調査データから女子身長に関し平滑化を行った (図 4-2)。48 カ月以降の波うちは改善されなかった。

男子胸囲に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 5 のような結果が得られた。3 パーセントイル曲線と、97 パーセントイル曲線の間隔がやや広め (72 カ月においてはやや下にゆがみ) であるが、大きな問題はなく、2010 年公表値とほぼ良好な一致を見せた。

女子胸囲に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 6 のような結果が得られた。97 パーセントイル曲線が 24 カ月以降若干ずれたが、大きな問題はなく、2010 年公表値とほぼ良好な一致を見せた。

男子頭囲に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 7 のような結果が得られた。1 カ月から 7 カ月まで曲線の湾曲がやや異なったが、節点の取り方の影響と考えられ、大きな問題はなく、2010 年公表値とほぼ良好な一致を見せた。

女子頭囲に関し、生後 3, 4, 5 日、1 カ月健診、一般調査データから平滑化を行ったところ、図 8 のような結果が得られた。1 カ月から 7 カ月まで曲線の湾曲がやや異なったが、節点の取り方の影響と考えられ、大きな問題はなく、2010 年公表値とほぼ良好な一致を見せた。

D. 考察

2020 年調査による成長曲線作成のための

平滑化法は、LMS 法をベースとするのが妥当であると考えられるが、実際の演算として、3つの方法の選択肢がある。ひとつは、平成 22 年乳幼児身体発育調査における乳幼児身体発育曲線作成で用いられた国立保健医療科学院横山徹爾部長作成の SAS プログラムによる算出、二つ目は LMSchartmaker®による算出、そして三つ目は R 言語で提供されている GAMLSS パッケージ活用である。

横山徹爾部長作成の SAS プログラムは、2010 年値作成の実績があるので、これを用いれば確実に平滑化曲線を提供することができる。

LMSchartmaker®はポピュラーな方法というのが強みであるが、サポートが終了しており、わが国の乳幼児身体発育調査データを用いて計算してみると、自由度についてさらなる試行錯誤が必要であることが分かっている。

R 言語で提供されている GAMLSS パッケージは、LMS 法の開発者である TJ Cole 氏も現在はこれを推奨している。R 言語のリソースも GAMLSS 等のパッケージも、無料でダウンロード可能であり、マニュアル、web での記事、書籍等を参考にすることで、活用が可能であることが分かった。

実際に 2010 年乳幼児身体発育調査データを用いて運用してみると、男子身長、男女胸囲、男女頭囲においては試行結果をそのまま平滑化値として差し支えないと考えられた。男子体重に関しては、二通りのデータセットの選び方で演算した結果を、継ぎ合わせて使用することで、平滑化曲線として活用していくことが可能と考えられた。女子体重および身長では、2010 年公表値と

平滑化曲線が似ているとはいえものの、成長曲線として利活用できる水準には、至っておらず、さらなる検討が必要である。女子体重では、使用データに極端な外れ値が多いので、クリーニングの徹底も、解決策の一つとして考えられる。

E. 結論

2020 年調査における発育曲線の作成に向けて、国立保健医療科学院横山徹爾部長作成の SAS プログラム、LMSchartmaker®、GAMLSS in R 等の方法について、それぞれ吟味していくなかで、今後国際標準となっていく GAMLSS in R による平滑化の試行を、すでに公表値が計算されている平成 22 年乳幼児身体発育調査のデータに関して行った。男子身長、男女胸囲、男女頭囲のように大きな問題がなく平滑化値が得られる項目もあったが、その他の項目に関しては、演算に当たってさらなる工夫が望まれた。

【参考文献】

1. 加藤則子,磯島豪. 乳幼児身体発育曲線作成のための平滑化ソフト活用法に関する研究. 平成 30 年度厚生労働行政推進調査事業費補助金 (成育疾患克服等次世代育成基盤研究事業)「乳幼児の身体発育及び健康度に関する調査実施手法及び評価に関する研究 (研究代表: 横山徹爾)」。分担研究報告書. 平成 31 年 3 月.
2. Stasinopoulos DM, Rigby RA. Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*. 2007; 23(7):1-46
3. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS

method and penalized likelihood.
Stat Med 1992; 11: 1305-19.

4. By Mikis D. Stasinopoulos, Robert A. Rigby, Gillian Z. Heller, Vlasios Voudouris, Fernanda De Bastiani Flexible, Regression and Smoothing Using GAMLSS in R, Taylor & Francis

F 健康危険情報

なし

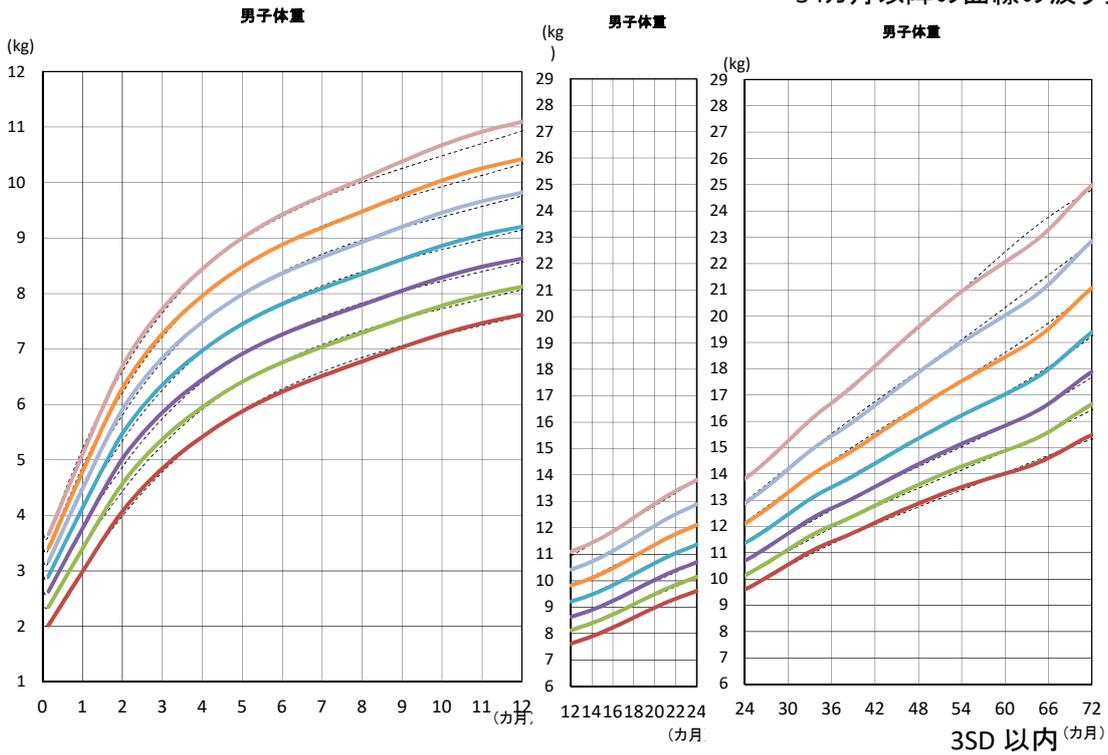
G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

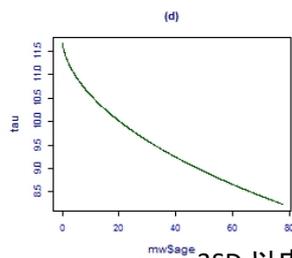
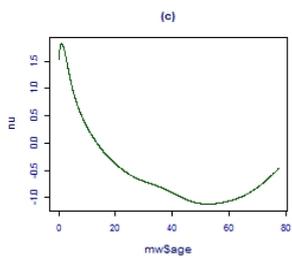
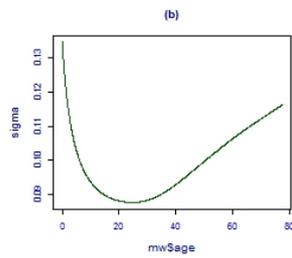
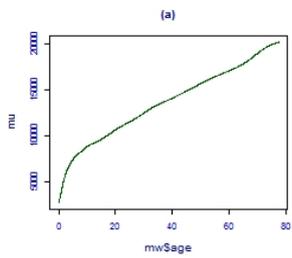
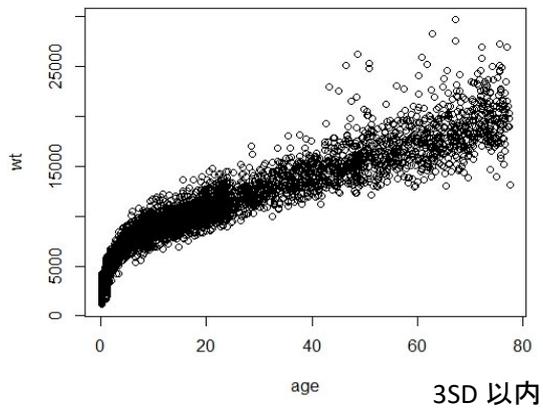
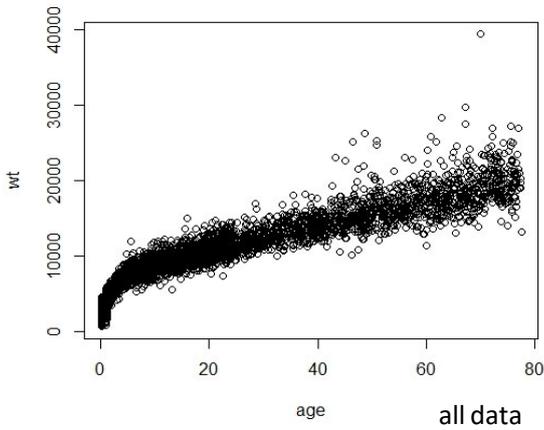
H. 知的財産権の出願・登録状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

図1-1 男子体重（生後3, 4, 5日 1カ月健診 一般調査 を使用）
54カ月以降の曲線の波うち



(点線 2010年公表値 実線 今回検討)



3SD 以内

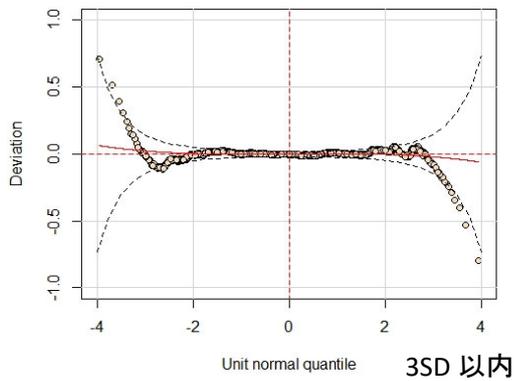


図1-2 男子体重 (1カ月健診 一般調査 を使用) 60か月～72か月の波うちがやや解消

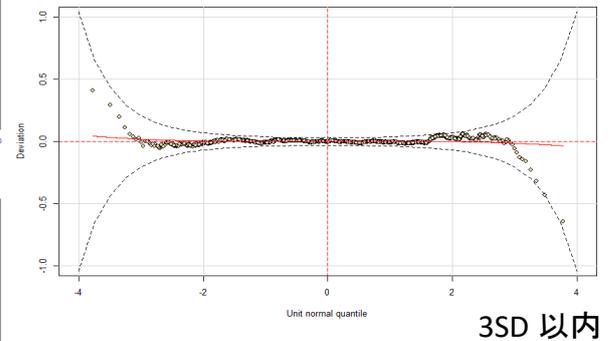
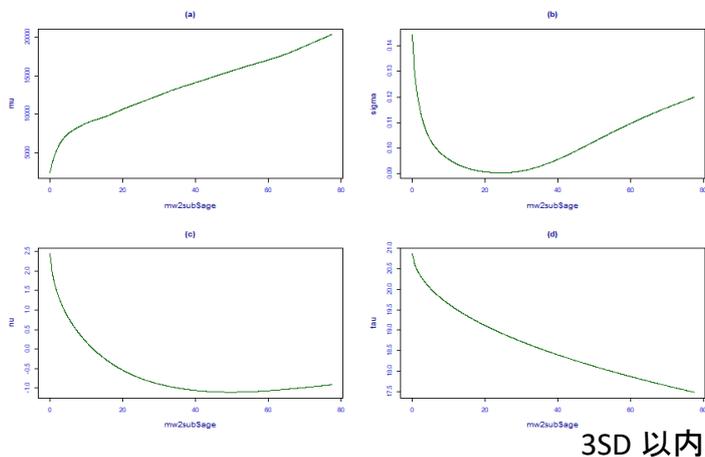
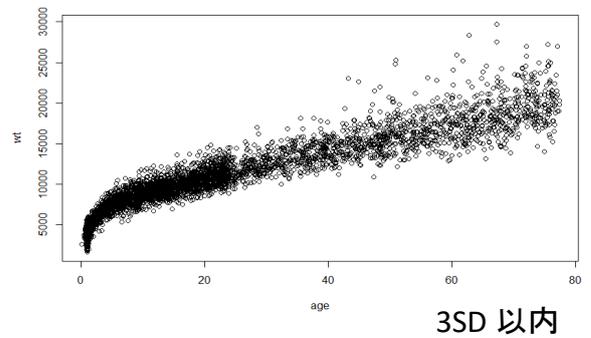
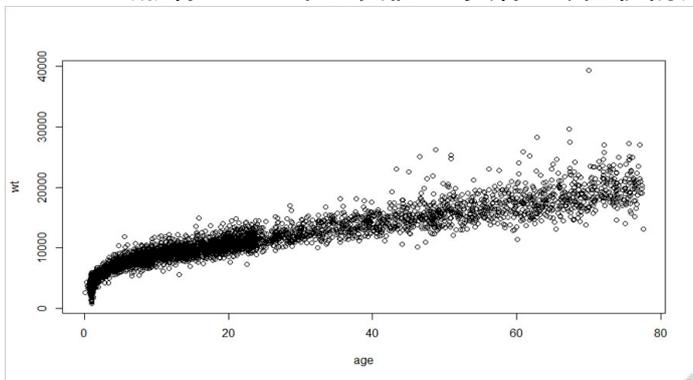
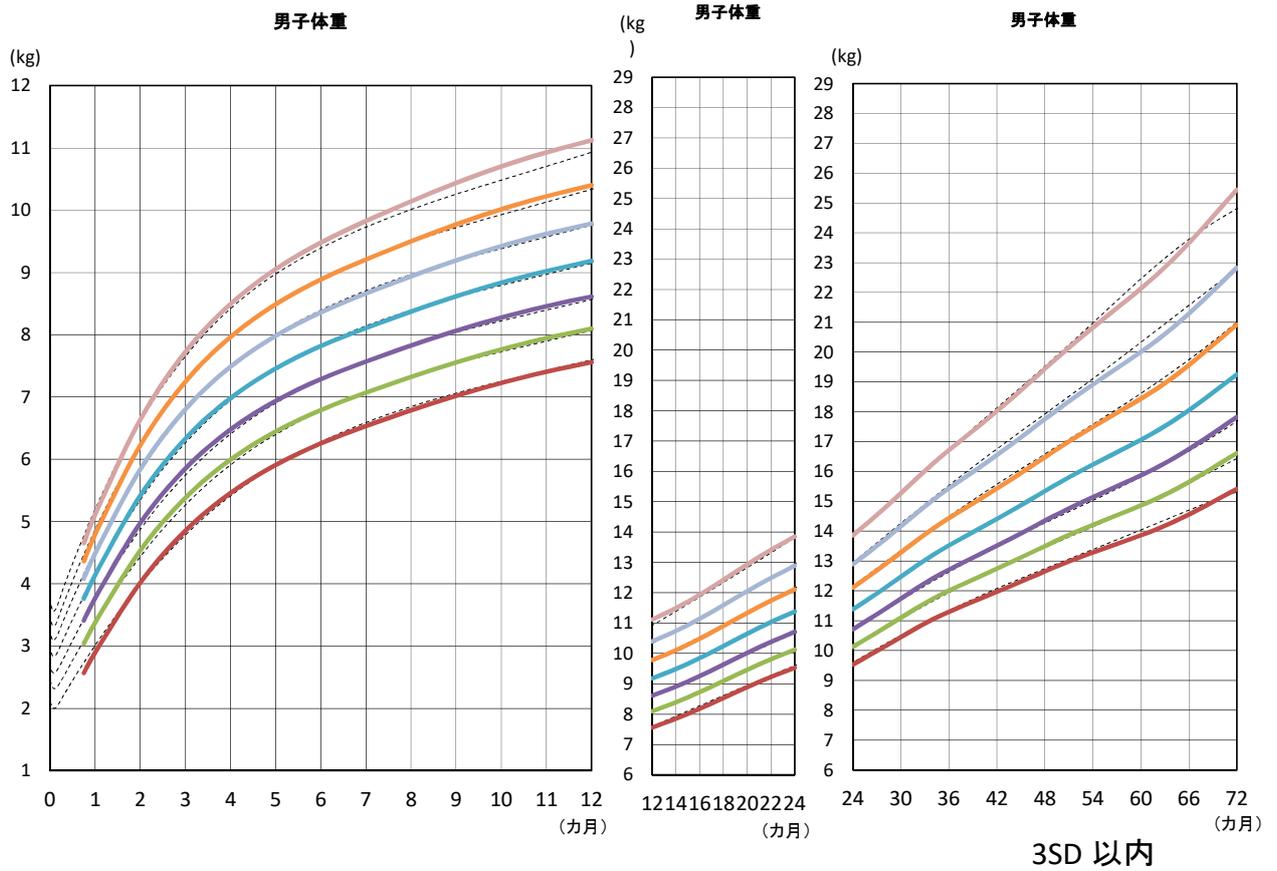
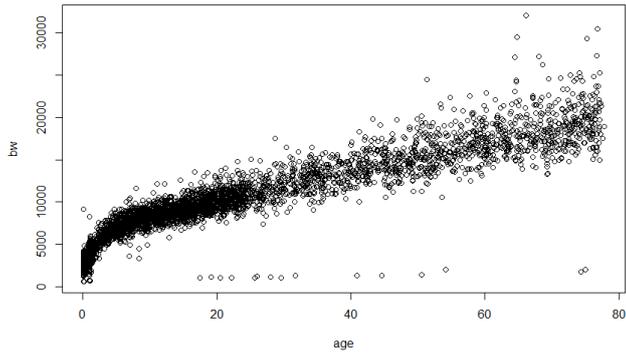
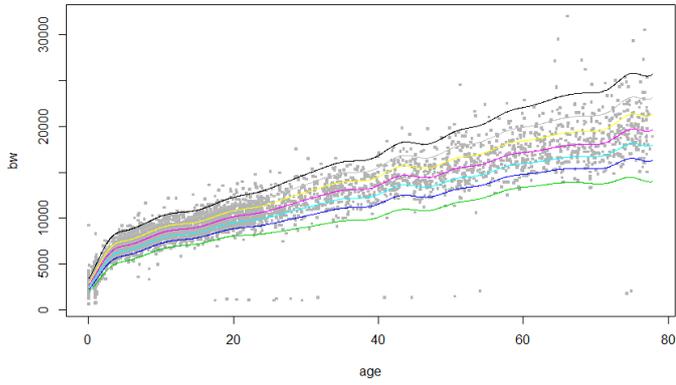


図2 - 1 女子体重（生後3, 4, 5日 1カ月健診 一般調査 を使用）



Centile curves using BCTo



±3SD 以内

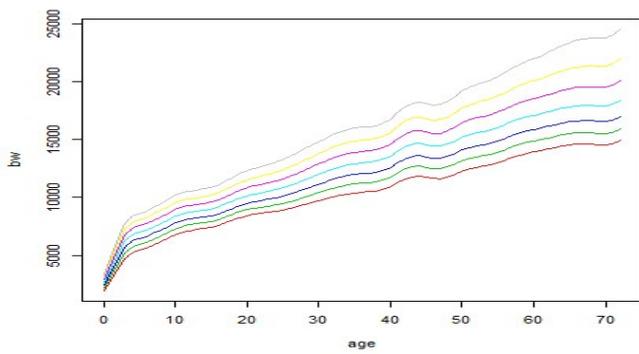
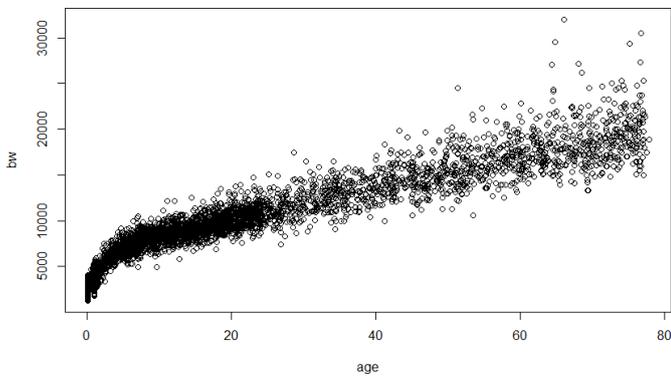


図2-2 女子体重（1カ月健診 一般調査 を使用）

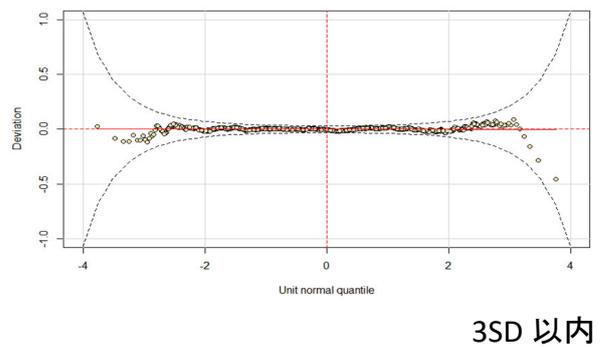
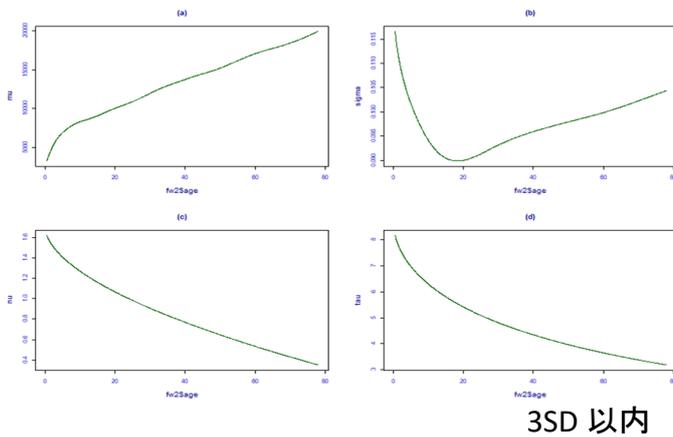
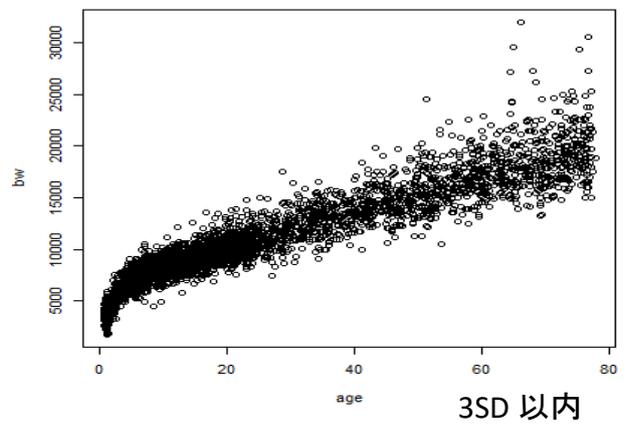
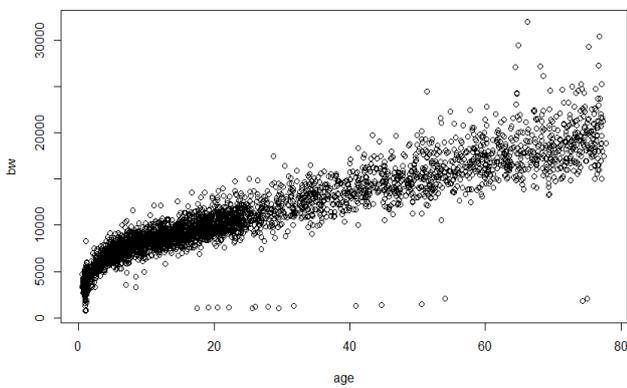
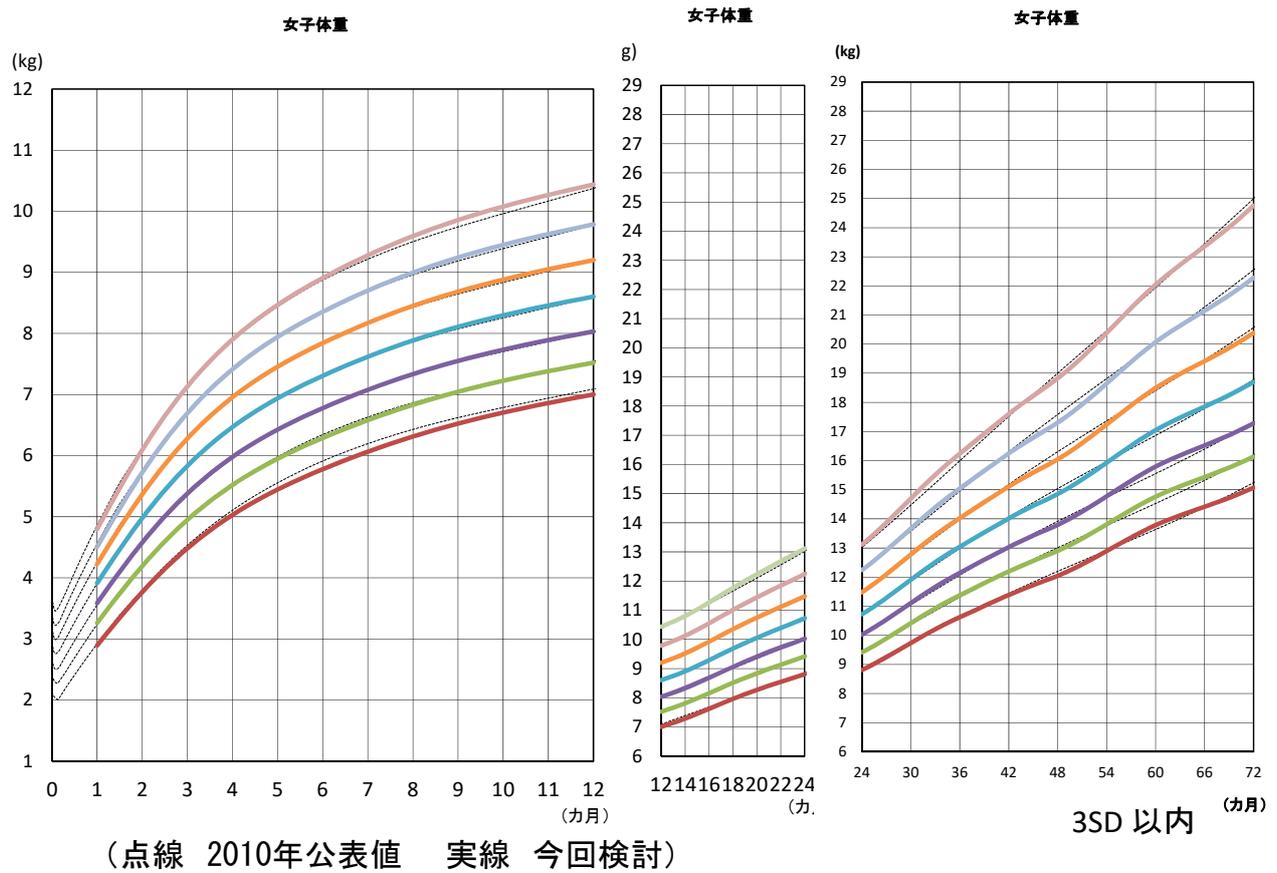


図2-3 女子体重（生後3, 4, 5日 1カ月健診 24カ月までの一般調査 を使用）

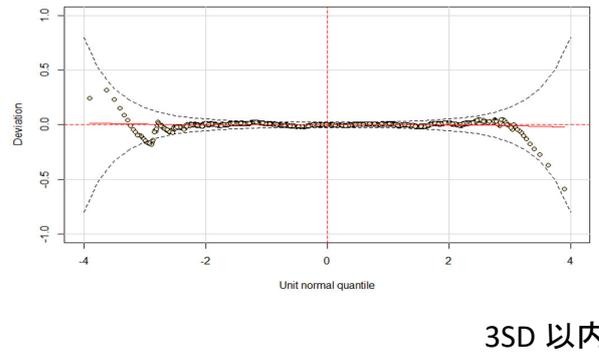
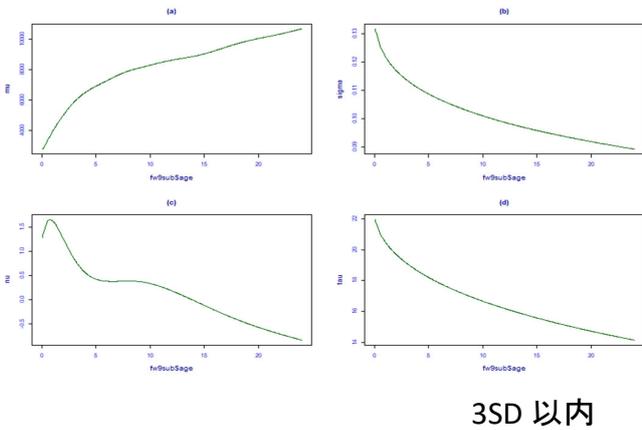
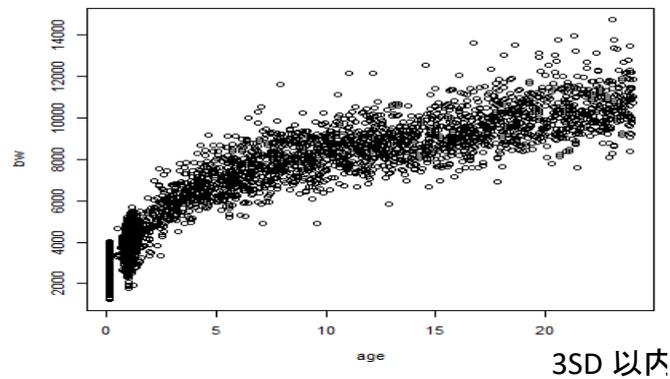
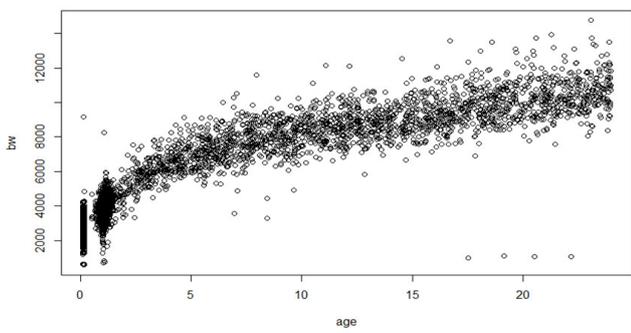
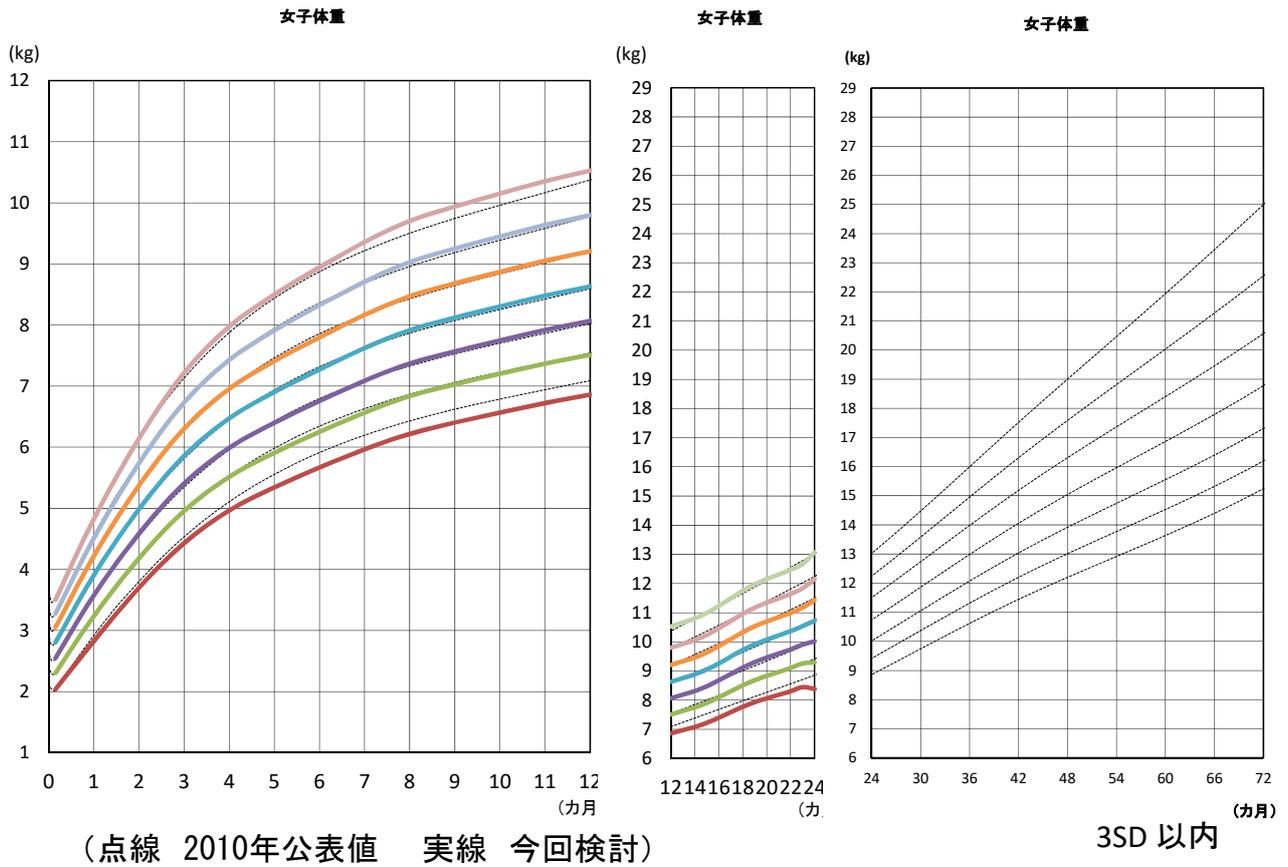
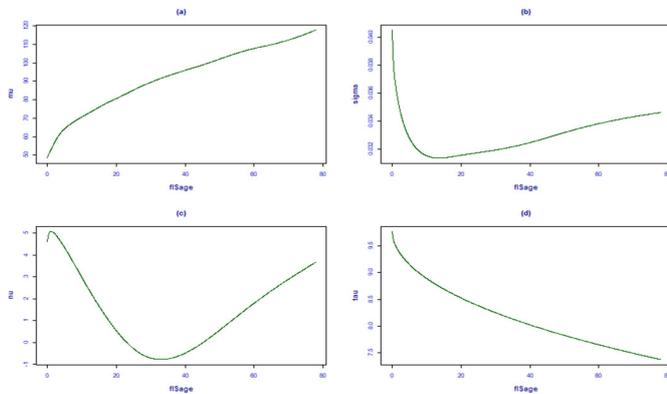
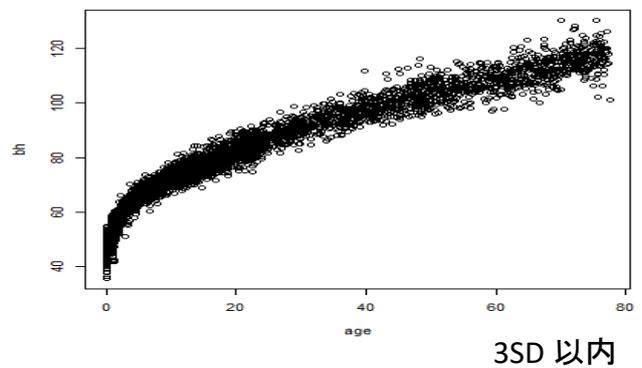
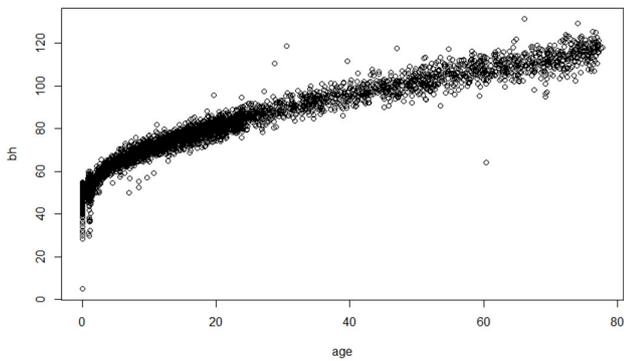
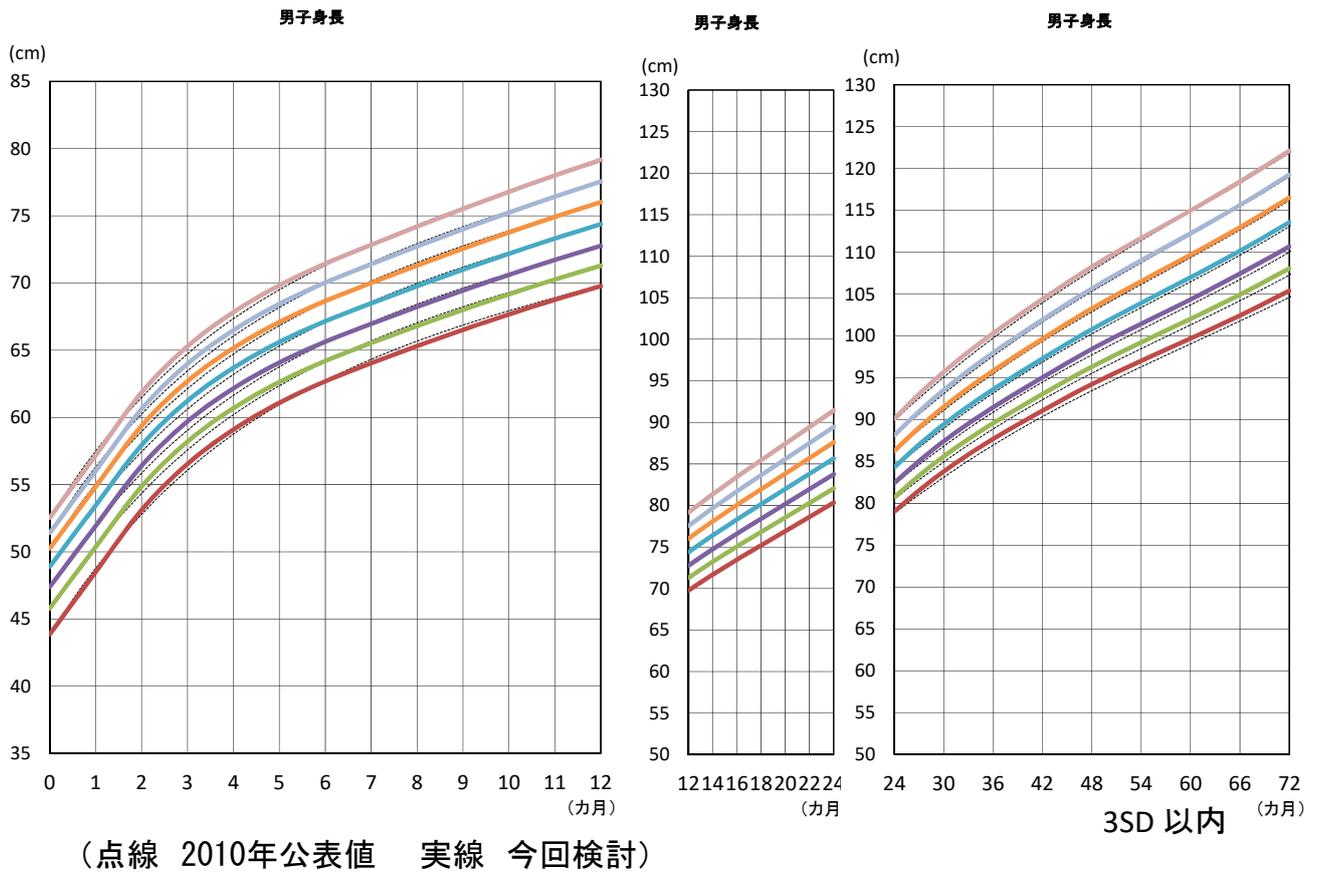
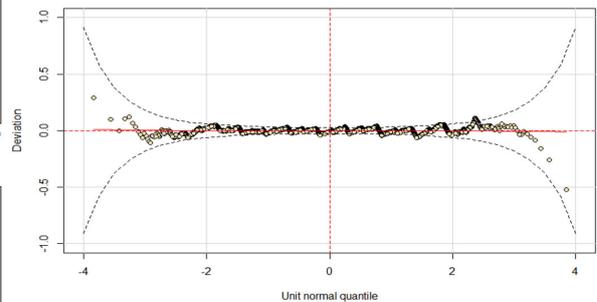


図3 男子身長（生後3, 4, 5日 1カ月健診 一般調査 を使用）

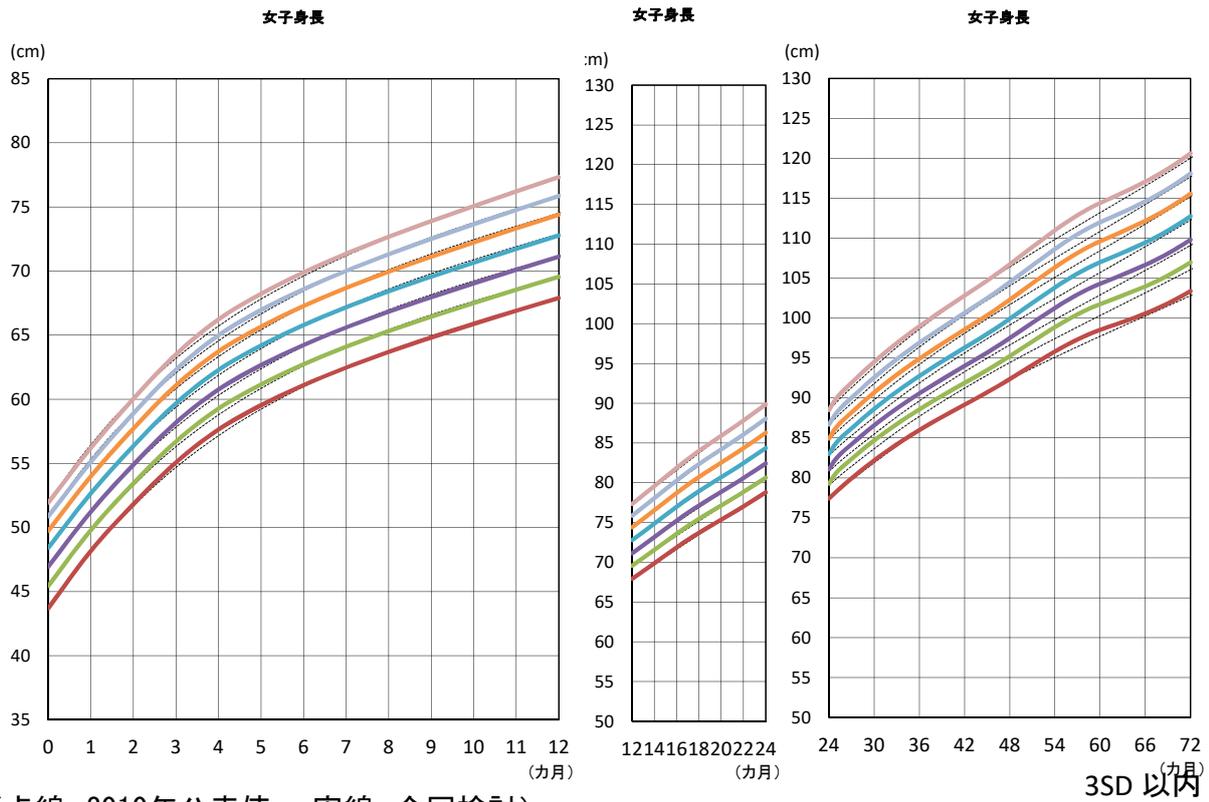


3SD 以内

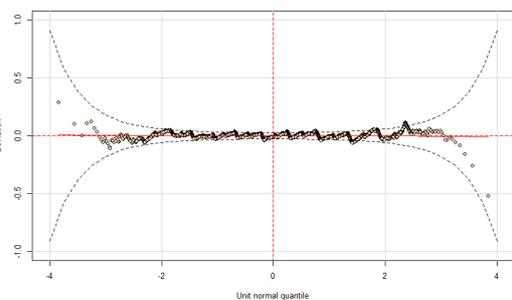
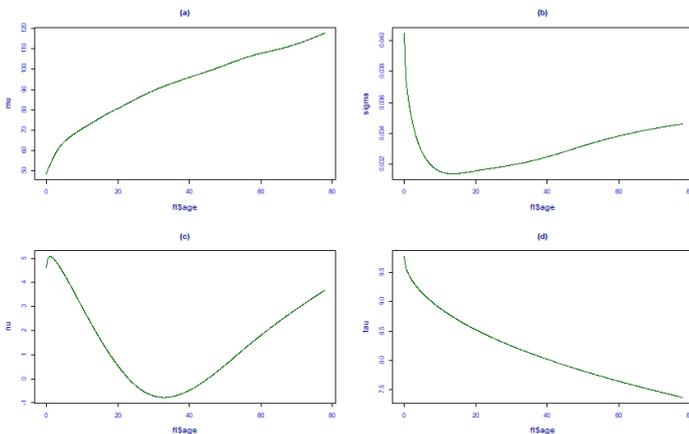
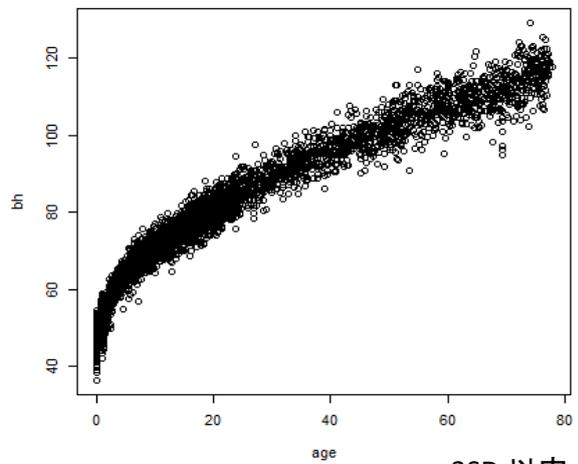
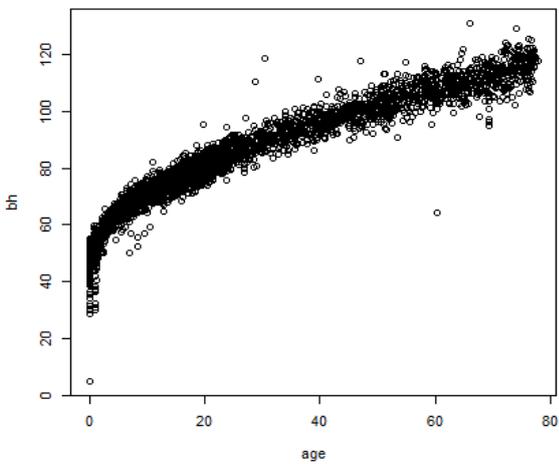


3SD 以内

図4-1 女子身長（生後3, 4, 5日 1カ月健診 一般調査 を使用）

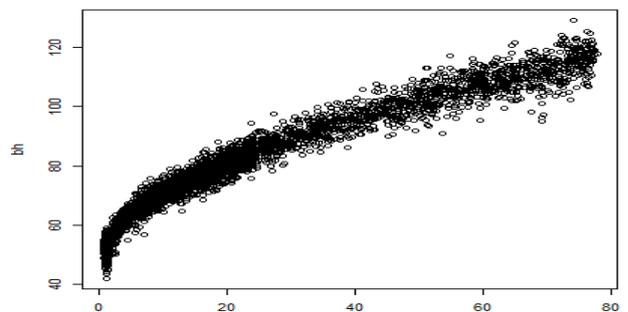
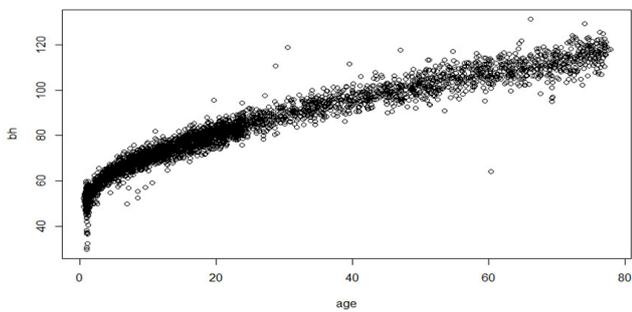
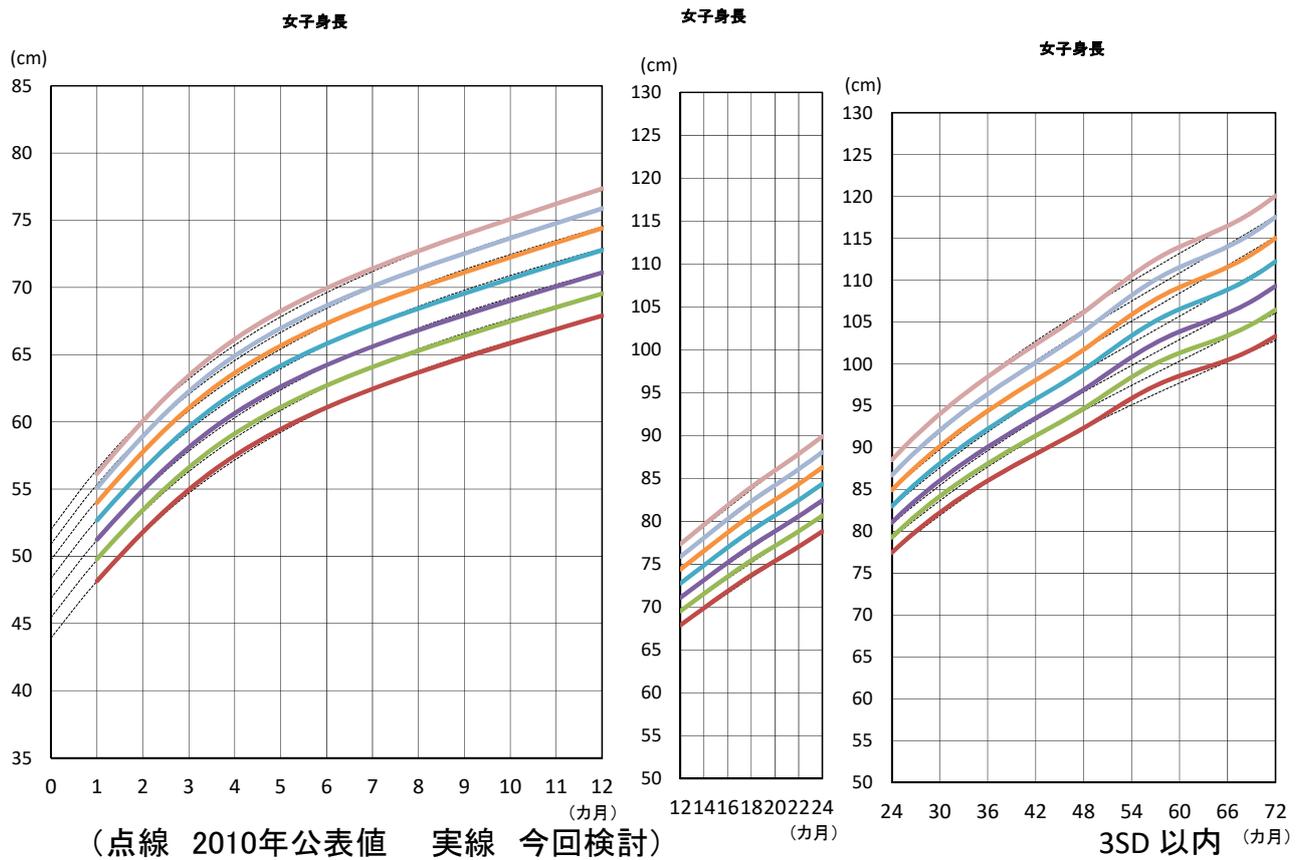


(点線 2010年公表値 実線 今回検討)

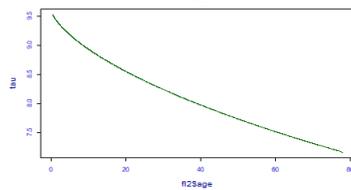
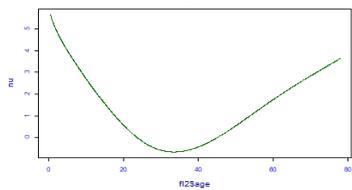
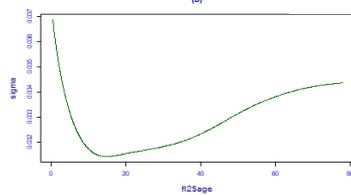
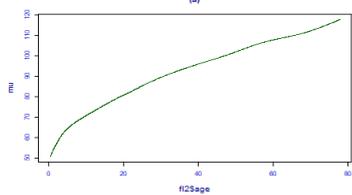


3SD 以内

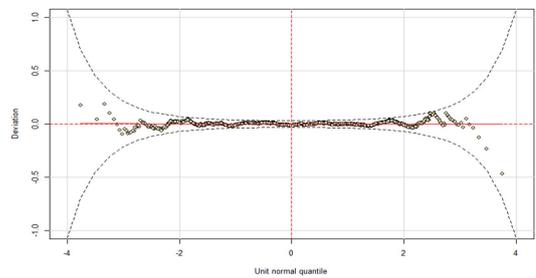
図4-2 女子身長（1カ月健診 一般調査 を使用）



3SD 以内

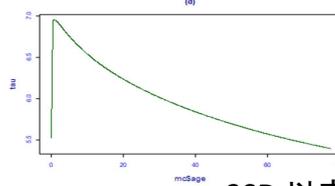
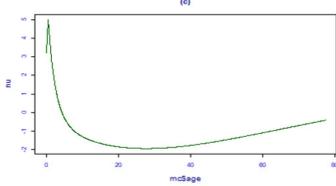
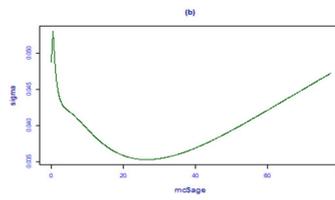
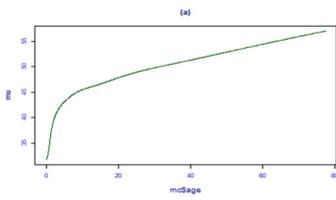
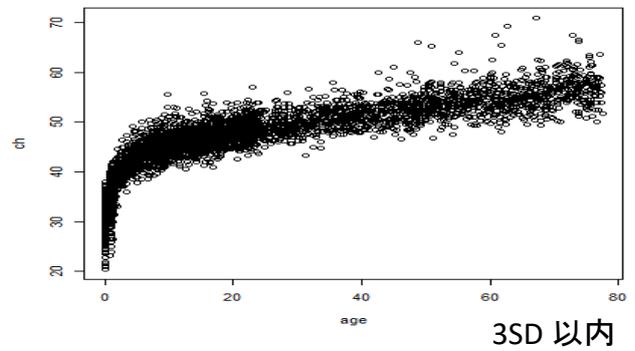
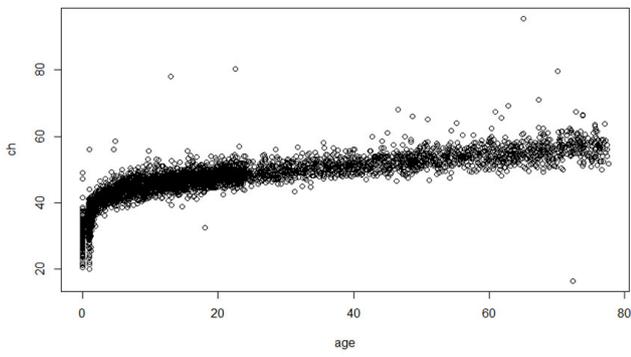
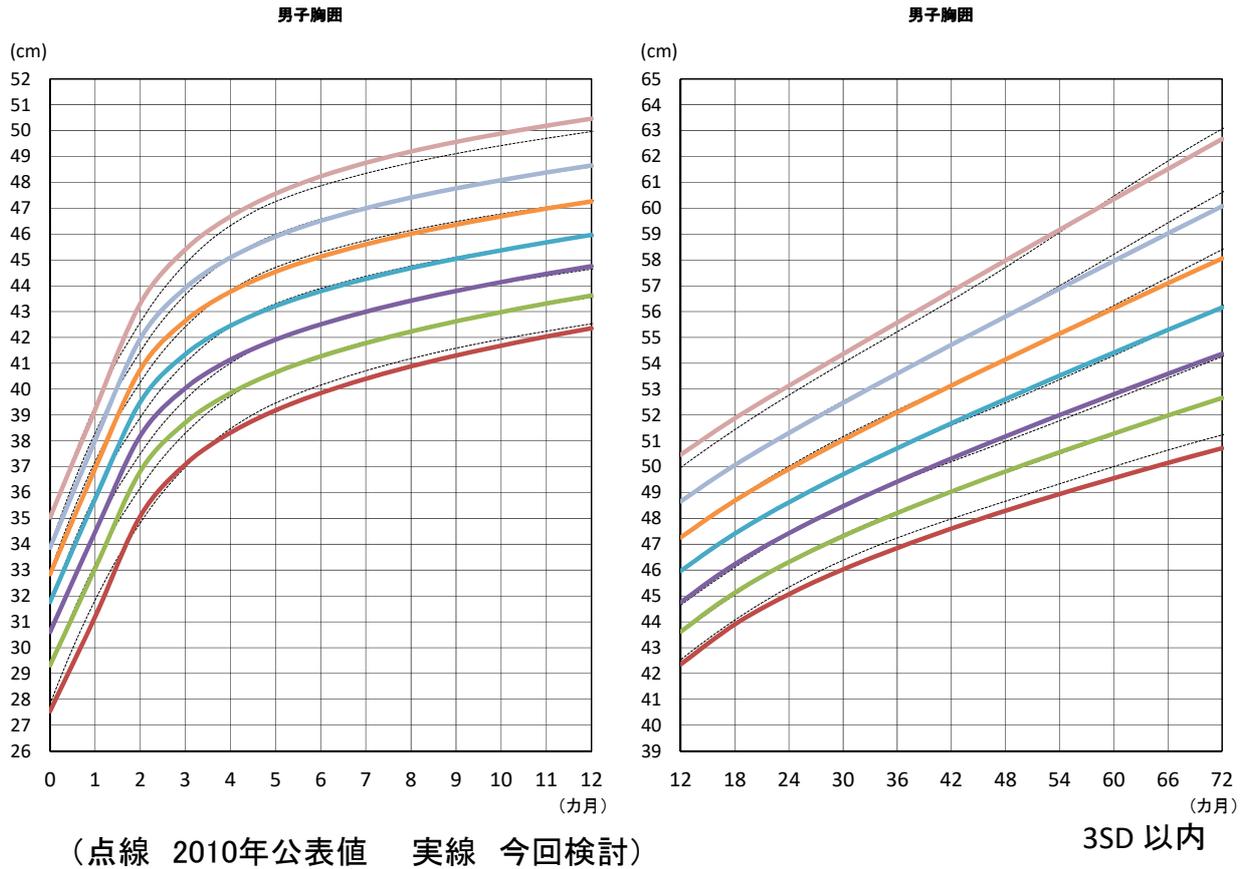


3SD 以内



3SD 以内

図5 男子胸囲 (生後3, 4, 5日 1カ月健診 一般調査 を使用)



3SD 以内

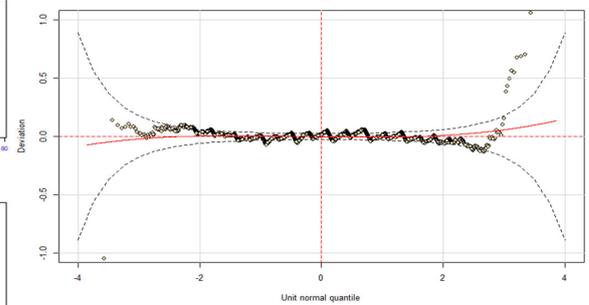


図6 女子胸囲（生後3, 4, 5日 1カ月健診 一般調査 を使用）

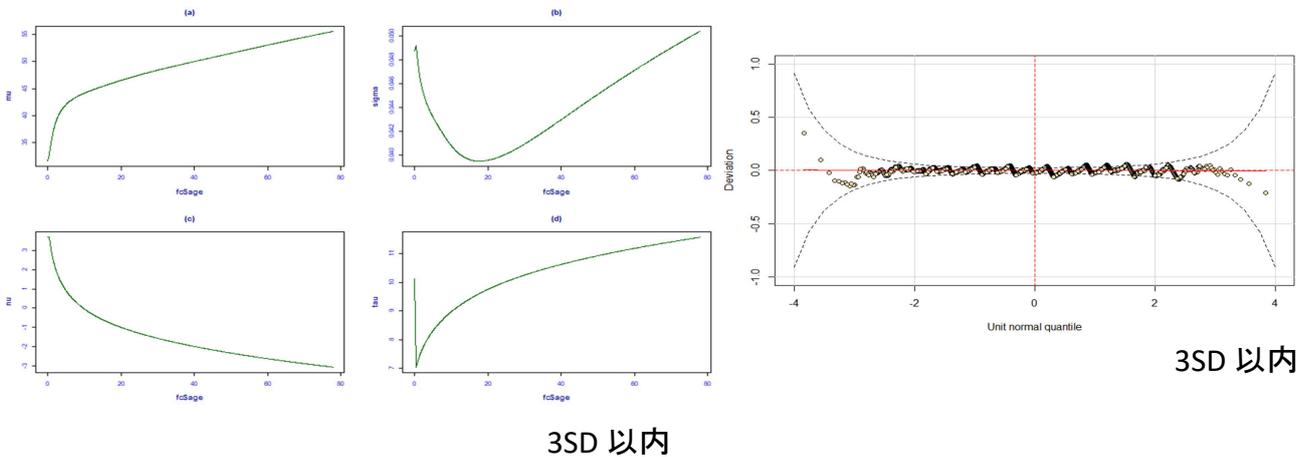
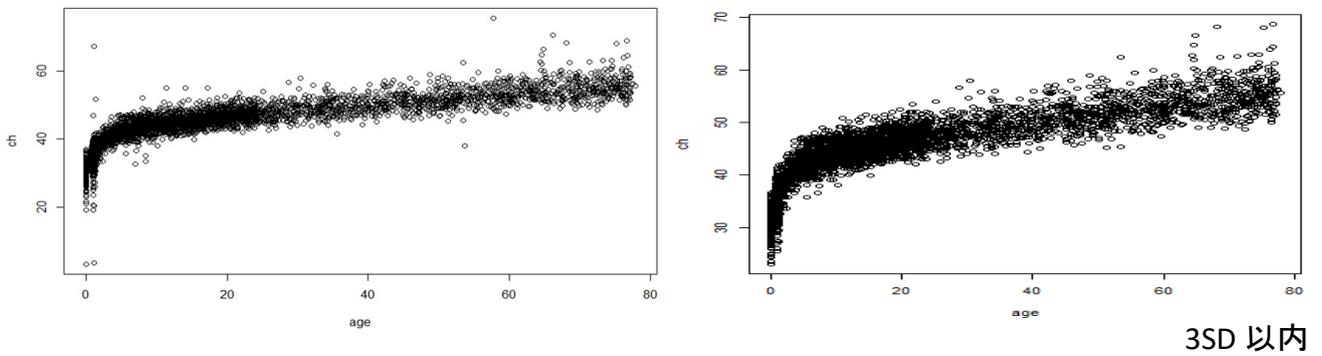
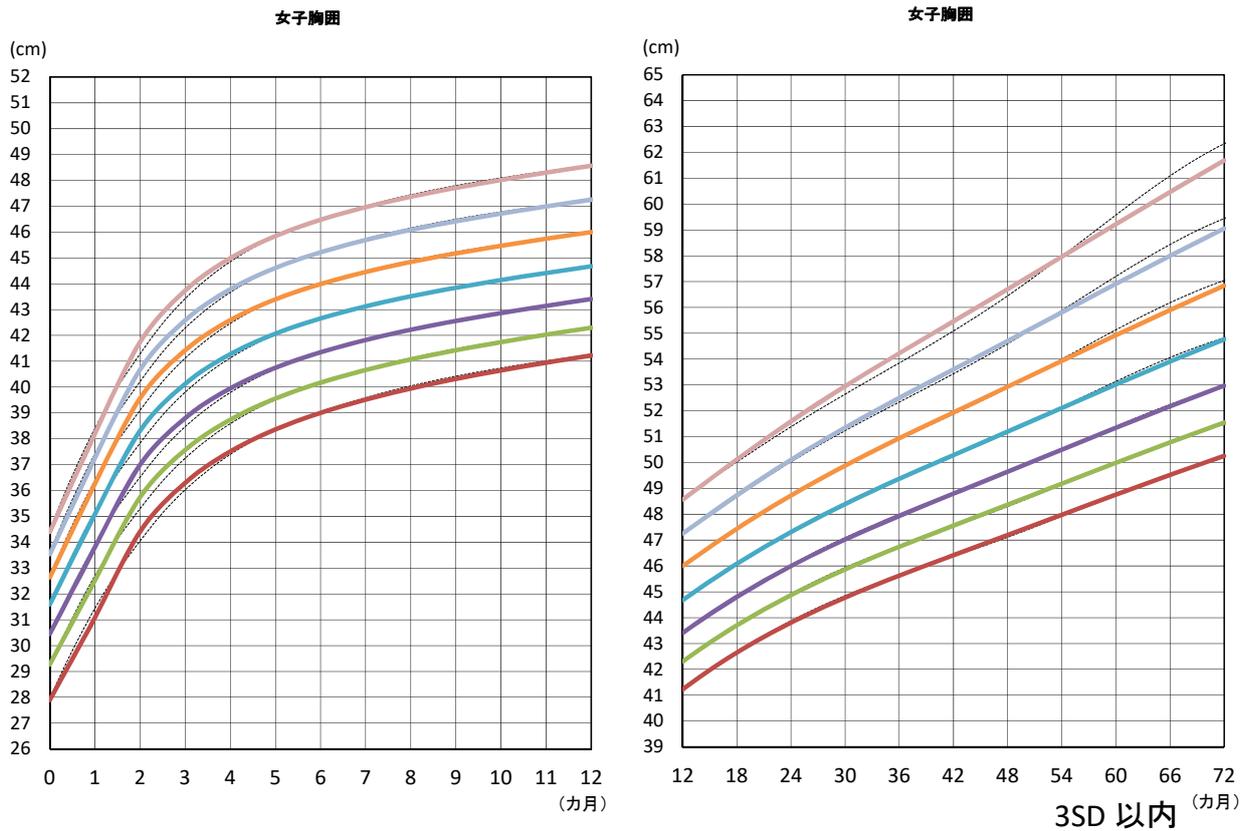
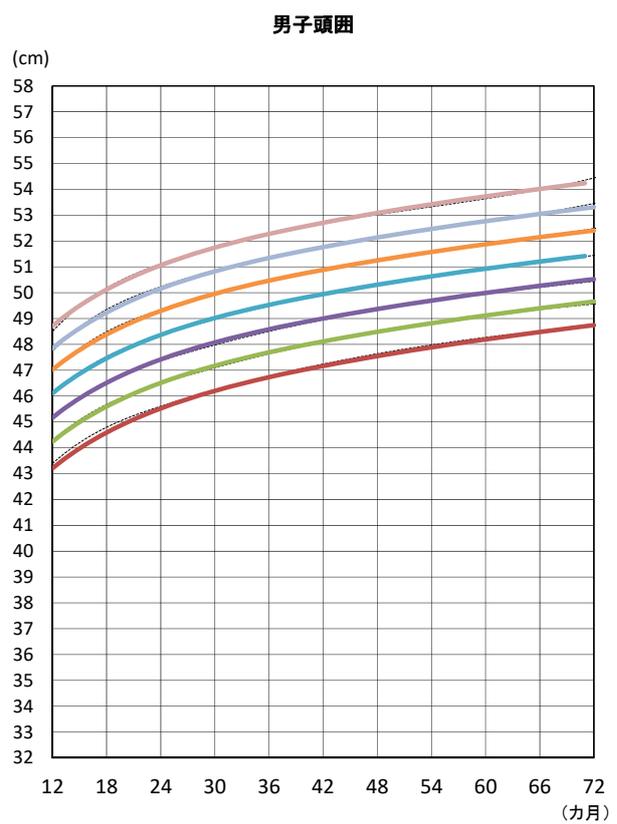
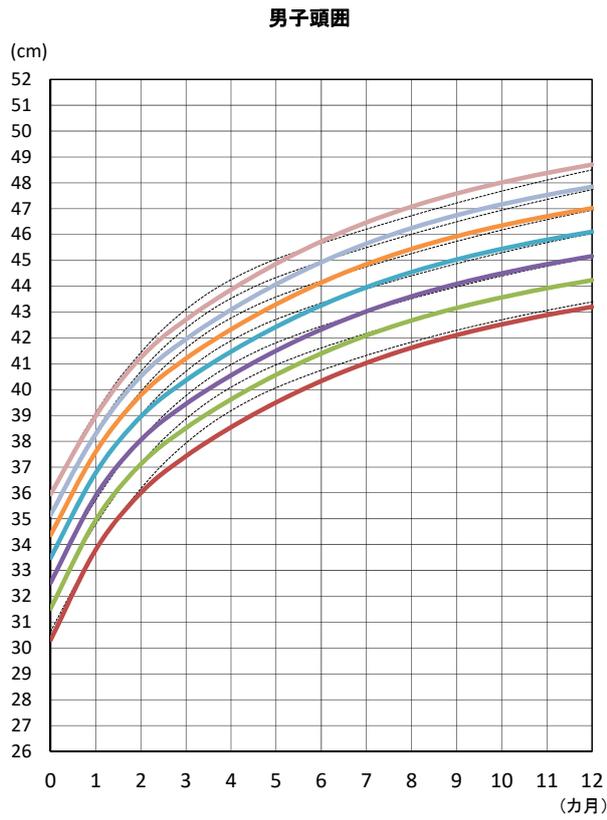
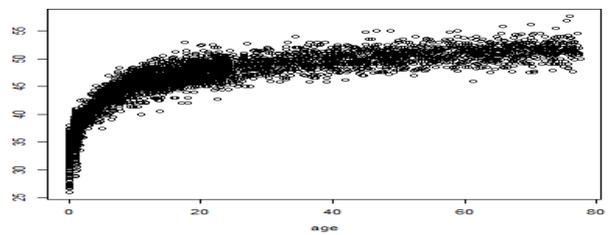
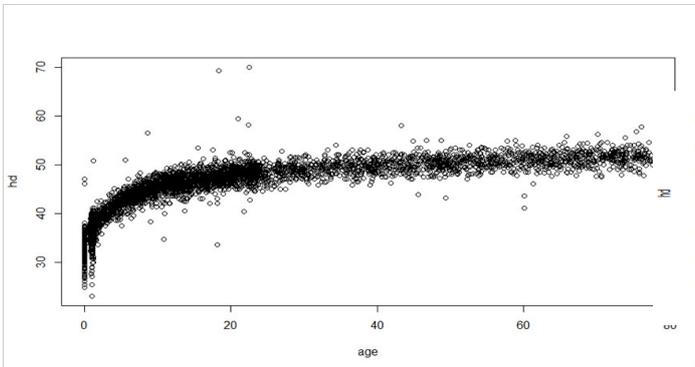


図7 男子頭囲 (生後3, 4, 5日 1カ月健診 一般調査 を使用)

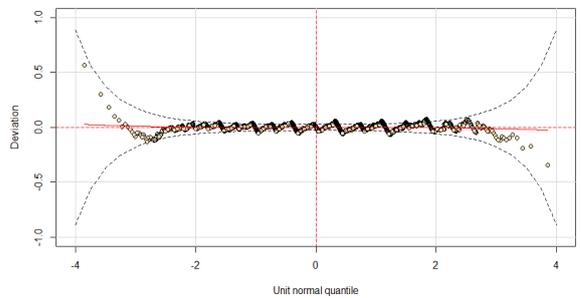
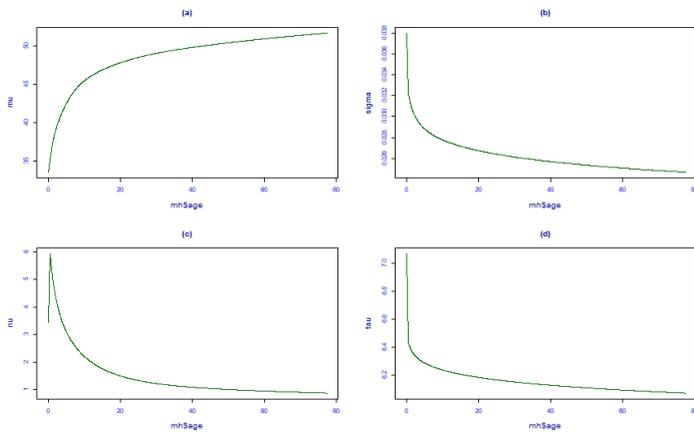


(点線 2010年公表値 実線 今回検討)

3SD 以内



3SD 以内

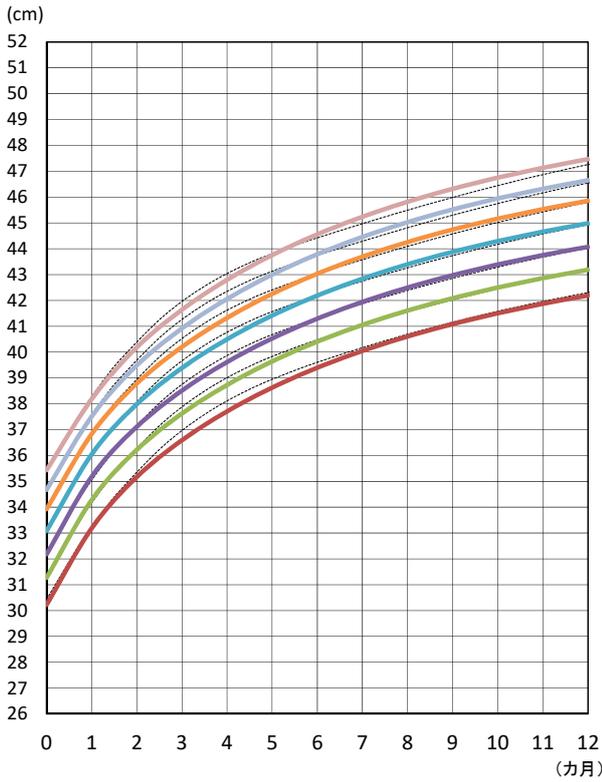


3SD 以内

3SD 以内

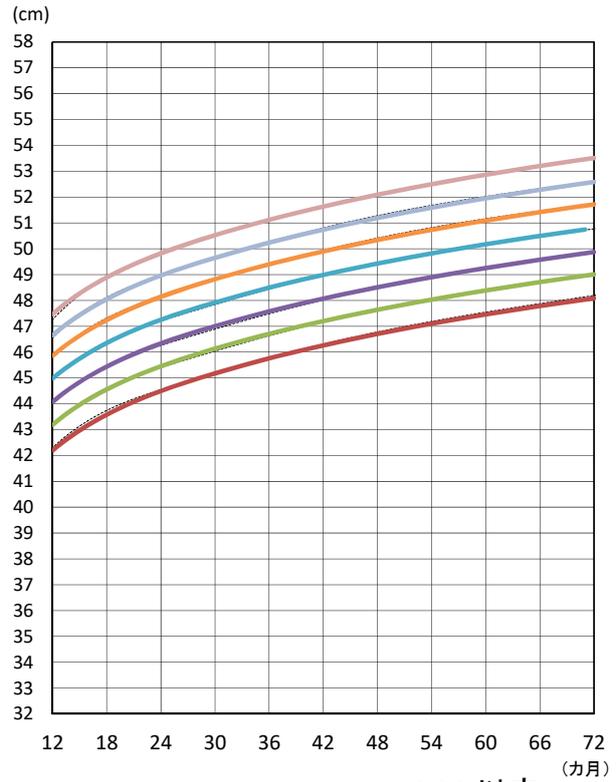
図8 女子頭囲 (生後3, 4, 5日 1カ月健診 一般調査 を使用)

女子頭囲

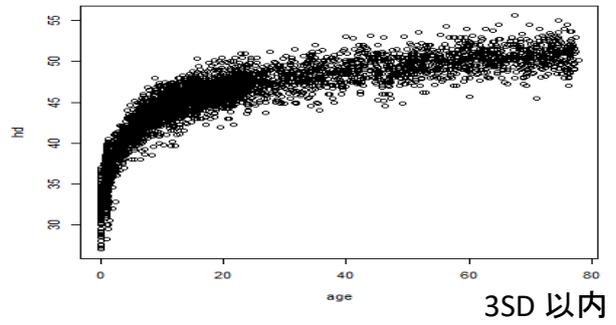
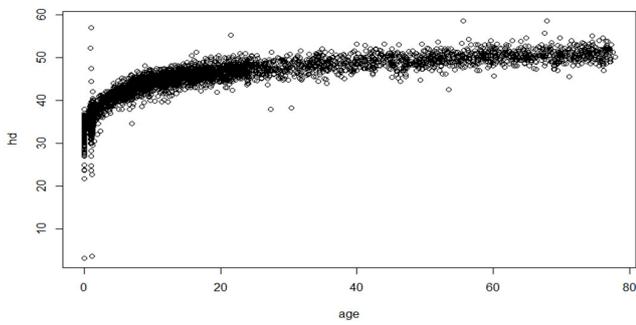


(点線 2010年公表値 実線 今回検討)

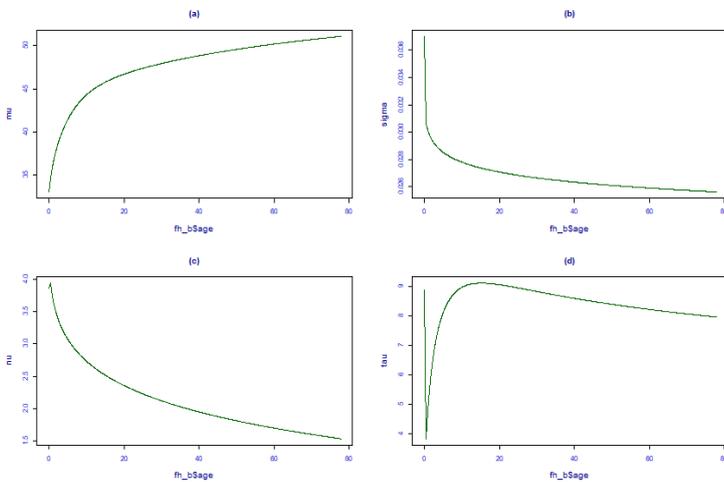
女子頭囲



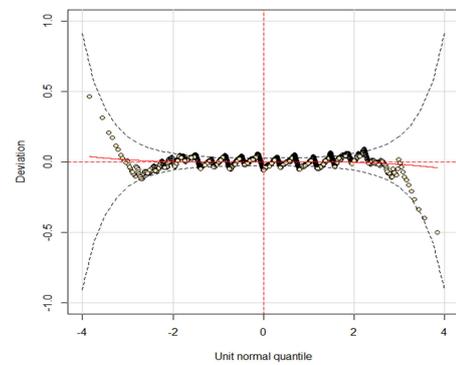
3SD 以内



3SD 以内



3SD 以内



3SD 以内