● مجلهٔ دانشگاه علوم پزشکی بابل، سال ششم، شمارهٔ ٤ (پی‌درپی ٢٤)، صفحه ٢٣ الی ٢٩، پاییز ١٣٨٣
دریافت:٨٢/٧/١٣، پذیرش: ٨٣/٧/٢٢

تأثیر داده های گمشده در نمودارهای رشد

محمود حاجی احمدی[1*]، دکتر سیدمحمدتقی آیت اللهی[2]، دکتر جواد بهبودیان[3]

١- عضو هیأت علمی گروه پزشکی اجتماعی دانشگاه علوم پزشکی بابل ٢- استاد گروه آمار زیستی و اپیدمیولوژی دانشگاه علوم پزشکی شیراز
٣- استاد گروه آمار دانشگاه شیراز

سابقه و هدف: استفاده از منحنی رشد قدرتمندترین وسیله پایش رشد کودکان می‌باشد و از این طریق می‌توان انحرافات از الگوی رشد طبیعی را به موقع تشخیص داد. ریزش داده‌ها و مقادیر گمشده از مشکلات معمولی در تجزیه و تحلیل داده‌های طولی رشد محسوب می‌شود. لذا اهمیت دارد که با برآورد نمودن مقادیر گمشده، داده‌ها کامل شده و در مسیری مناسب و صحیح جهت تحلیل قرار داده شوند.

مواد روشها: این مطالعه طولی طی دو سال بر روی ٣١٧ نوزاد (١٥٣ پسر، ١٦٤ دختر) در شهر شیراز انجام شد. اطلاعات مربوط به رشد (وزن، قد، دور سر، بازو و دور سینه) در بدو تولد و ١١ بازدید از منزل کودکان جمع‌آوری گردید. جهت تأثیر داده‌های گمشده روی نمودارهای رشد چهار روش نادیده گرفتن مقادیر گمشده، مدلهای معمولی و فردی منحنی رشد و انتساب چند گانه مورد بررسی قرار گرفتند. میانگین، صدکهای سوم، پنجاهم و نود و هفتم خام و هموار وزن، در پسران و نمودارهای هموار رشد وزن آنها بر اساس چهار روش تعیین و مورد مقایسه قرار گرفتند.

یافته‌ها: در این مطالعه تفاوت قابل ملاحظه‌ای در میانگین وزن پسران در سنین مورد مطالعه بر اساس روشهای مدل منحنی رشد و انتساب چندگانه با حالتی که مقادیر گمشده نادیده گرفته شدند بوجود نیامد ولی نمودارهای هموار رشد نشان دادند که استفاده از مدل منحنی رشد فردی (سطح دوم) و انتساب چندگانه موجب می گردد که مقادیر صدکهای سوم و نودوهفتم اختلاف قابل توجه‌ای با تحلیل مرسوم (نادیده گرفتن مقادیر گمشده) داشته باشد.

نتیجه‌گیری: با توجه به وجود مقادیر گمشده در مطالعات رشد نادیده گرفتن آنها جهت تجزیه و تحلیل اشتباه بوده و استفاده از مدل منحنی رشد می تواند جهت مطلوب نمودن تجزیه و تحلیل و مسیر صحیح رشد با اهمیت تلقی شود.

واژه‌های کلیدی: ***داده‌های گمشده، نمودار رشد، مدل منحنی رشد، داده‌های طولی.***

مقدمه

ارزیابی الگوی رشد کودکان اطلاعات مفیدی را در باره تغذیه و بهداشت جامعه بدست داده و موجب تشخیص زودرس اختلالات رشد کودکان می‌گردد. وزن بدن که از شاخصهای مهم رشد جسمانی نیز می باشد، اندازه گیری آن بدلیل اینکه تمام افزایش در اندازه، جمع‌بندی می‌شود احتمالاً عمده‌ترین ملاک رشد و تغذیه است(٣-١). با توجه به نقش مهم زمان در بسیاری از تحقیقات علوم پزشکی، خصوصاً رشد، مطالعات طولی نقش مهم ویژه‌ای در مقایسه با مطالعات مقطعی ایفا می‌کند. در این مطالعات بعضی ویژگی‌های افراد در طول زمان چندین بار اندازه‌گیری می‌شود. هر چند فاصله‌های زمانی که هر فرد مورد مشاهده قرار می‌گیرد لزوماً

▣ هزینه انجام این پژوهش در قالب طرح تحقیقاتی شماره ١١٢٨-٨٠ از اعتبارات معاونت پژوهشی دانشگاه علوم پزشکی شیراز تأمین شده است.

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

یکسان نیستند(۱۰-۴). مهمترین مزیت این نوع مطالعات تفکیک نمودن اثر زمان از تفاوتهای فردی می‌باشد. که در مطالعات مقطعی امکان‌پذیر نیست. مطالعات طولی نیاز به روشهای آماری خاصی دارد زیرا مجموعه مشاهدات مربوط به یک فرد دارای همبستگی درونی است و جهت استنباط علمی و معتبر این همبستگی را باید در هنگام تحلیل به حساب آورد. مدلهای چند سطحی که اخیراً توسعه زیادی یافته یکی از روشهای مؤثر می‌باشد(۱۱).

در مطالعات طولی ویژگی‌های بعضی از افراد مورد مطالعه ممکن است در بعضی از دوره‌های زمانی در دسترس قرار نگیرد. وجود مقادیر گمشده ممکن است دقت شاخصهای آماری محاسبه شده را کاهش داده و بر پیچیدگی مدل و روشهای آماری بیفزاید. ریزش داده‌ها و مقادیر گمشده از مشکلات معمولی در تجزیه و تحلیل مطالعات طولی محسوب می‌شود(۴). با توجه به اینکه نتایج حاصل از تحلیل داده‌های ناقص می‌تواند به سوگرائی منجر شود لذا اهمیت دارد که تحلیل این نوع داده‌ها در مسیری مناسب و صحیح قرار داده شود(۱۱و۴). اکنون روشهای چند متغیره به داده‌های کامل نیاز دارند .متأسفانه حذف کامل داده‌ها دارای چند زیان است که عمده‌ترین آن به هدر رفتن مقداری از اطلاعات جمع‌آوری شده است(۱۲). شیوه‌های متنوعی برای جایگزین داده‌های گمشده مانند جایگزینی میانگین گروه، آخرین مقدار ثبت شده مربوط در داده‌های طولی و انواع گوناگونی از برون‌یابی و درون‌یابی مورد استفاده قرار می‌گیرند(۱۳). این شیوه‌ها نیاز به مفروضات خاصی از جمله تصادفی بودن (MACR,MAR)[1] مشاهدات گمشده در مجموعه داده‌ها دارد (۱۵و۱۴). مدل منحنی رشد (Growth Curve Model) به عنوان زیر مجموعه‌ای از مدلهای چندسطحی، مدل عمومی و مناسبی برای مطالعات رشد طولی محسوب می‌شود که در مطالعات متعددی مورد استفاده قرار گرفته است(۱۶و۱۱). مدل منحنی رشد بعنوان یک ابزار محسوس، مناسب و کارآ برای انتسابی که ارتباط اساسی میان متغیرها و همبستگی میان مشاهدات مزبور به یک فرد را حفظ و نگهداری می‌نماید مورد توجه و قبول آمارشناسان می‌باشند. هدف از این مطالعه ساختن مدلی است که در آن ساختار داده‌ها بطور دقیق و صحیح به حساب آید و بتوانیم با نادیده گرفتن مقادیر گمشده و با استفاده از مدل منحنی رشد جهت برآورد و پیش‌بینی ایجاد منحنی‌های (مقادیر) فردی استفاده نمائیم و از منحنی‌های فردی جهت برآورد مقادیر گمشده استفاده کنیم.

مواد و روشها

این مطالعه طولی بر روی رشد و تکامل کودکان طراحی شد و نوزادان متولد شده در ۱۴ زایشگاه شیراز مورد پیگیری قرار گرفتند. همگروهی از ۳۱۷ نوزاد متولد شده سالم(۱۶۴ دختر و ۱۵۳ پسر) بطور تصادفی با استفاده از شیوه نمونه‌گیری احتمالی متناسب با حجم از میان متولدین در ۱۴ زایشگاه در طی مدت ۲ هفته متوالی با یک شروع تصادفی انتخاب شدند. بازدید نوزادان در بیمارستان و در خانه‌هایشان در سنین تقریبی ۱/۵، ۳، ۴/۵، ۶، ۸، ۱۰، ۱۲، ۱۵، ۱۸، ۲۱، ۲۴ ماهگی از نظر وضعیت رشد انجام گردید. تقریباً همه زنان حامله شیراز (۹۷/۵٪) در بیمارستان زایمان می‌کنند(۱۷). جهت تأثیر داده‌های گمشده روی نمودارهای رشد چهار روش مورد بررسی قرار گرفت. در روش اول مقادیر گمشده در تحلیل نادیده گرفته شد. در روش دوم مقادیر گمشده بر اساس مدل عمومی منحنی رشد برآورد و در تحلیل، جایگزین مقادیر گمشده شدند. در روش سوم مقادیر گمشده بر اساس مدل منحنی رشد فردی برآورد و در تحلیل، جایگزین مقادیر شدند. همچنین برای غلبه نمودن بر پیچیدگی روشهای برآورد داده‌های گمشده و ناقص علاوه بر روش مدل منحنی رشد از روش انتساب چندگانه (Multiple Imputation) نیز استفاده کردیم. با استفاده از مجموعه داده‌های کامل ایجاد شده بوسیله مدل منحنی رشد (عمومی و فردی) و روش انتساب چندگانه در داده‌های طولی رشد، میانگین و صدکهای سوم، پنجاهم و نود و هفتم و نمودارهای هموار رشد ایجاد و با حالتی که مقادیر گمشده نادیده گرفته شدند مقایسه نمودیم تا بتوانیم اثر مقادیر گمشده را بر روی نمودارهای رشد مورد مطالعه قرار دهیم.

با توجه به روند تدریجی رشد، تکامل، بیماری و درمان غالباً لازم است برای بررسی سیر آنها افراد مورد مطالعه بطور مکرر در طول زمان مورد مشاهده قرار گیرند. این داده‌های مکرر که در آنها موقعیت‌های تکرار مشاهدات، نقاط زمانی هستند داده‌های طولی

1. (Missing At Completely Random, Missing At Random)

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

(Longitudinal Data) و مطالعات از این نوع را مطالعات طولی (Longitudnal Studies) می‌نامند(۱۲-۴). داده‌های مشاهده شده در علوم زیستی و انسانی دارای ساختار سلسله مراتبی یا خوشه‌ای نیز می‌باشند که این ساختار در مطالعات اندازه‌گیری تکراری یا داده‌های طولی رخ می‌دهد. در این نوع مطالعات افراد یا واحدهای مورد مطالعه در بیش از یک موقعیت اندازه‌گیری می‌شوند(۱۸). مدلهای چند سطحی یا مدلهایی با اثر ترکیبی خطی، مدلهای ضرائب تصادفی و ... به عنوان ابزارهای آماری جهت تجزیه و تحلیل داده‌هایی که ساختار سلسله مراتبی دارند بسط یافته‌اند. مدلهای سلسله مراتبی می‌تواند برآوردهای پارامترهای پایدار و معقول‌تری در مقایسه با شیوه‌های تحلیلی در مرسوم فراهم آورد(۱۱ و۱۰). در مطالعات طولی و اندازه‌گیری‌های تکراری موقعیت‌ها بصورت خوشه‌ای در درون افراد قرار دارند که سطح دوم واحدها را ارائه می‌کنند و سطح اول واحدها با اندازه‌گیری موقعیت‌ها شروع می‌شود (۱۸).

همچنین داده‌های طولی رشد بعنوان یک مدل سلسله مراتبی در سطحی که سطح اول اندازه‌گیری موقعیت‌ها در درون فرد و سطح دوم بین افراد می‌باشد در نظر گرفته می‌شود(۲۰و۱۹). در این مدل Y_{tj} وضعیت مشاهده شده در زمان t برای فرد jام بعنوان تابعی از مسیر رشد یا منحنی رشد به اضافه خطای تصادفی می‌باشد(۲۱). بنابراین سطح اول مدل رشد فردی بصورت تابعی چند جمله‌ای از درجه p بصورت زیر است.

$$y_{tj}=\pi_{oj}+\pi_{1j}Z_{tj}+\pi_{2j}Z_{tj}^2+.....+\pi_{pj}Z_{tj}^p+\ell_{tj} j=1,2,.....j,$$

Z_{tj} سن فرد jام در زمان t و π_{hj} مسیر رشد hام برای فرد jام (پارامترهای رشد فردی) با چند جمله‌ای درجه p(p h=0,1,2,……,p) می‌باشد. مقادیر π_{hj} مسیر تغییرات صحیح و واقعی فردی را در طول زمان تعیین می‌کند همچنین فرض می‌شود که ساختار خطای ساده که با e_{tj} بیان می‌شود از توزیع نرمال با میانگین صفر و واریانس σ_e^2 برخوردار است که مستقل از یکدیگرند. اغلب درک کمی از مکانیزم کنترل کننده فرآیند تغییرات وجود دارد و از یک منحنی خطی یا درجه دوم برای مسیر تقریبی رشد استفاده می‌شود(۱۱). از مفروضات کلیدی و مهم مدلسازی رشد فردی آن است که مسیر رشد برای هر فرد در جمعیت دارای یک فرم و شکل تابعی یکسان می‌باشد ولی برای افراد متفاوت ممکن است مقادیر مختلفی در پارامترهای رشد فردی وجود داشته باشند برای نشان دادن تفاوت فردی، در سطح دوم مدل از یک مدل دو سطحی، پارامترهای رشد فردی بعنوان تابعی از مشخصات فرد انتخاب شده در نظر گرفته شد. خصوصاً برای هر یک از p+1 پارامتر رشد فردی بصورت زیر داریم:

$$\pi_{jh}=\gamma_{oh}+\sum_{k=1}^{q}\gamma_{oh}W_{kj}+U_{hj}$$

که W_{kj} ویژگی اندازه‌گیری شده از سابقه فردی مانند جنس یا طبقه اجتماعی یا یک مسئله تجربی باشدγ_{0h} اثر w_k روی h امین پارامتر رشد را نشان می‌دهد و U_{hj} اثر تصادفی با میانگین صفر است. این مدلها می‌تواند برای بیان الگوی رشد در مطالعات طولی مورد استفاده قرار گیرد.

انتساب چندگانه نیز یک تکنیک شبیه‌سازی مونت کارلو بر اساس زنجیرهای مارکف جهت تجزیه و تحلیل داده‌های ناقص است که اولین بار بوسیله روبین جهت ایجاد یک مجموعه داده‌ها برای مطالعه و بررسی عمومی پیشنهاد گردید(۲۴-۲۲و۴). انتساب چندگانه بعنوان یک فرآیند سه مرحله‌ای در نظر گرفته می‌شود. در مرحله اول هر داده گمشده بوسیله m مقدار شبیه‌سازی شده جایگزین می‌شود (m>۱ و معمولاً بین ۳ تا ۱۰ می‌باشد) نتایج m مجموعه از داده‌های کامل بر اساس روشهای استاندارد مبتنی بر داده‌های کامل تجزیه و تحلیل می‌شوند و در مرحله سوم نتایج جهت ایجاد گزاره‌های استنباطی مانند برآوردهای فاصله‌ای و تعیین مقادیر p آزمون با یکدیگر ترکیب می‌شوند(۲۵و۴).

در قوانین روبین برای برآوردها فرض می شود $\hat{Q}$ برآوردی از پارامتر مورد نظر و $\hat{U}$ واریانس برآورده باشد(۴). پس از انجام تحلیل یکسان بر روی هر یک از m مجموعه داده‌های کامل شده، m برآورد Q_1 , Q_2 , …. Q_M و واریانسهای متناظر U_1 , U_2 , …. U_m بدست می‌آیند. برآورد جایگزین انتساب چندگانه یا برآورد کلی $\bar{Q}=\frac{1}{m}\sum_{i=1}^{m}\hat{Q}_i$ می شود واریانس کل برای هر برآورد دارای دو مولفه می‌باشد که تغییرات درون هر مجموعه داده‌ها و بین آنها را در نظر می‌گیرد.

واریانس درونی $\hat{U}=\frac{1}{m}\sum_{i=1}^{m}U_i$ میانگین ساده واریانسها است. واریانس بین برآوردها $B=\frac{1}{m-1}\sum(\hat{Q}-\bar{Q})^2$ که واریانس خود برآوردها $(\hat{Q}_i)$ است و واریانس کل، T برابر است با مجموع واریانسهای این

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

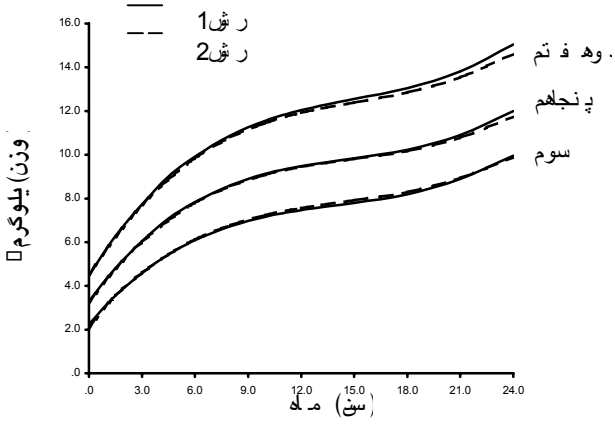
دو مولفه با لحاظ کردن یک عامل تصحیح برای خطای شبیه‌سازی در $\overline{Q}$ با $T=\hat{U}+\left(1+\frac{1}{m}\right)B$ یک فاصله اطمینان ۹۵٪ بصورت $\overline{Q}\pm 2\sqrt{T}$ و کارایی نسبی برآورد بر اساس m جایگذاری و γ (نسبت مقادیر گمشده) تقریباً $\left(1+\frac{\gamma}{m}\right)^{-1}$ می‌باشند.

فرض تصادفی بودن مقادیر گمشده و هم توزیع بودن متغیرهای پاسخ با مقادیر گمشده و بدون مقادیر گمشده (MAR) پذیرفته شد. تست لیتل در نرم‌افزار آماری SPSS فرض کاملاً تصادفی بودن مقادیر گمشده (MCAR) را رد نموده است جهت برآورد نمودن مقادیر گمشده و جایگزین نمودن آنها بر اساس مدل منحنی رشد فردی از زیربرنامه Lme (Linear Mixed Models) نرم‌افزار آماری s-plus بوسیله تابعی درجه سوم از سن کودکان استفاده گردید و مقادیر برآورد شده سطوح اول (عمومی) و دوم (فردی) جایگزین مقادیر گمشده گردید.

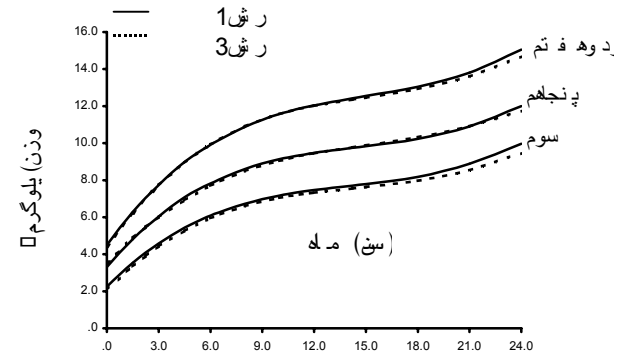
برای برآورد نمودن مقادیر گمشده و جایگزین نمودن آنها بر اساس انتساب چندگانه از برنامه NORM(۴) استفاده شد و از میانگین پنج مجموعه کامل داده‌های ایجاد شده جهت جایگزین مقادیر گمشده استفاده گردید. (۲۳و۴). هموارسازی صدکها و نمودارهای رشد در روشهای اول تا چهارم با استفاده از نرم‌افزار GROSTATII , GROSTATI و بر اساس روش HRY انجام گردید(۲۶). صدکهای هموار شده در برنامه Excel رسم و مورد مقایسه قرار گرفت.

یافته‌ها:

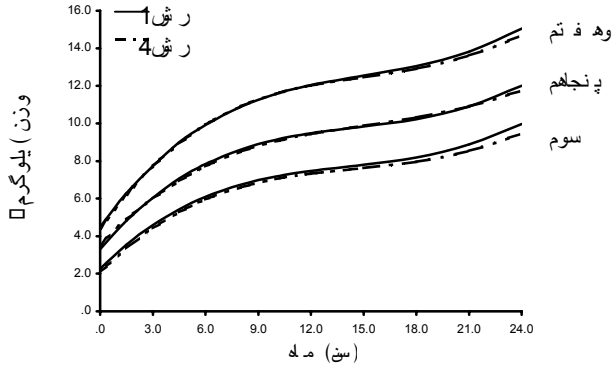
بر اساس نتایج این مطالعه میزان داده‌های گمشده برای متغیر وزن در جنس پسر حداکثر ۲۴/۲٪ بوده است (جدول ۱)، میزان کل داده‌های گمشده در تحلیل چند متغیره با حذف کامل ۳۶٪ می‌باشد که تعداد کل کودکان با داده‌های کامل به ۲۰۲ نفر کاهش می‌یابد. همبستگی وزن بازدیدهای مختلف در پسران بین ۰/۳۱ و ۰/۹۸ می‌باشند.میانگین، صدکهای خام و هموار شده سوم، پنجم، دهم، بیست و پنجم، پنجاهم، هشتاد و پنجم، نود و پنجم و نود و هفتم وزن پسران محاسبه گردید. همچنین نمودارهای هموار رشد، وزن کودکان پسر بر اساس روشهای یک تا چهار تعیین گردید(نمودارهای ۳-۱).



دارا :صدیهایهم موار وزن پسران با روشهایی و دو



ا ر۲ صدیهم وارون پسول با روشهایی و سه



ل ۳صدیهم هموار وزن پسران با روشهایی و چهار

با توجه به نتایج مطالعه و با مقایسه میانگین وزنهای بدو تولد و یازده بازدید در پسران مشاهده می‌شود که در بیشتر موارد تفاوت قابل توجه‌ای بین روشهای فوق الذکر وجود ندارد میانگین وزن پسران در ۶ ماهگی با توجه به روشهای چهارگانه بترتیب ۷/۸، ۷/۸، ۷۵/۷، ۷/۷ است و برای ۱۲ ماهگی بترتیب ۹/۴، ۹/۴، ۹/۴، ۹/۴ است(جدول ۱).

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

جدول ۱- میانگین صدکهای خام و هموار وزن (کیلوگرم) پسران سالم شیراز از بدو تولد تا ۲۴ ماهگی

سن	تعداد	درصد گمشده	میانگین				سوم (خام)				سوم(هموار)			
			روش۱	روش۲	روش۳	روش٤	روش۱	روش۲	روش۳	روش٤	روش۱	روش۲	روش۳	روش٤
بدو تولد	۱٤۸	۳/۳	۳/۲۷	۳/۲۷	۳/۲۷	۳/۲٦	۲/۲۸	۲/۳	۲/۳	۲/۳	۲/۲٥	۲/۱۱	۲/۱۲	۲/۰۷
۱/٥ ماهگی	۱٥۳	۰	٤/٤۱	٤/٤۱	٤/٤۱	٤/٤۱	۲/٦۸	۲/٦۸	۲/٦۸	۲/٦۸	۳/٥٥	۳/٥٦	۳/۳۹	۳/۳٥
۳ ماهگی	۱٤۸	۳/۳	٥/۹۱	٥/۹۰	٥/۸۸	٥/۸۷	٤/٤۳	٤/٤۳	٤/۱٦	٤/۱٤	٤/٥۸	٤/٦	٤/٤٥	٤/٤۲
٤/٥ ماهگی	۱٤٥	٥/۲	٦/۸۲	٦/۸۱	٦/۸۰	٦/۸۰	٥/۱٥	٥/۲	٥/۰۳	٥/۰۲	٥/٤۳	٥/٤۳	٥/۲۹	٥/۲٦
٦ ماهگی	۱۳۷	۱۰/٥	۷/۷۹	۷/۷۷	۷/۷٥	۷/۷٤	٥/۹۷	٥/۹۹	٥/۷۷	٥/۸٤	٦/۰۹	٦/۱۲	٥/۹٦	٥/۹٥
۸ ماهگی	۱۳۹	۹/۲	۸/٦٥	۸/٦۳	۸/٦۱	۸/٦۰	٦/٦۸	٦/۷۷	٦/٥۱	٦/٥۳	٦/۷۳	٦/۷۸	٦/٦۱	٦/٦۰
۱۰ ماهگی	۱۳۷	۱۰/٥	۹/۰٦	۹/۰٥	۹/۰٤	۹/۰۲	۷/۰۱	۷/۱۲	٦/۹٤	۷/۱۷	۷/۱۷	۷/۲٤	۷/۰٤	۷/۰٤
۱۲ ماهگی	۱۳۲	۱۳/۷	۹/٤۰	۹/۳۹	۹/٤۱	۹/۳۷	۷/۱۳	۷/۳٤	۷/۲۷	۷/۰۷	۷/٤٦	۷/٥٦	۷/۳۲	۷/۳۳
۱٥ ماهگی	۱۲۸	۱٦/۳	۹/۹۷	۹/۹٤	۹/۹٦	۹/۹٥	۷/٦۹	۷/۷٤	۷/۷٤	۷/٦٤	۷/۷۹	۷/۹۱	۷/٦۳	۷/٦۲
۱۸ ماهگی	۱۲۳	۱۹/٦	۱۰/٤٥	۱۰/۳۹	۱۰/۳۷	۱۰/۳۹	۸/۳۱	۸/۳٥	۸/۱۰	۷/۸٥	۸/۱۸	۸/۳	۷/۹۹	۷/۹٦
۲۱ ماهگی	۱۱٦	۲٤/۲	۱۱/۱۱	۱۰/۹۸	۱۰/۹۸	۱۱/۰٦	۸/۹٤	۸/۹٤	۸/٦۹	۸/٥۹	۸/۸۷	۸/۹۳	۸/٦۳	۸/٥۳
۲٤ ماهگی	۱۱۹	۲۲/۲	۱۱/٦۷	۱۱/٦۲	۱۱/٥۸	۱۱/٥٦	۹/۰٤	۹/۲۷	۸/۹٦	۸/۹٦	۹/۹٦	۹/۸۸	۹/٦٤	۹/٤٥

سن	پنجاهم(خام)				پنجاهم(هموار)				نود و هفتم(خام)				نود و هفتم(هموار)			
	روش۱	روش۲	روش۳	روش٤	روش۱	روش۲	روش۳	روش ۴	روش۱	روش۲	روش۳	روش٤	روش۱	روش۲	روش۳	روش٤
بدو تولد	۳/۲۷	۳/۲۷	۳/۲۷	۳/۲۷	۳/۳۰	۳/۲٦	۳/۳٥	۳/٥۲	٤/۳۲	٤/۳۰	٤/۳۰	٤/۳۰	٤/٤۹	٤/٥۱	٤/٦۳	٤/٤۱
۱/٥ ماهگی	٤/٤۹	٤/٤۹	٤/٤۹	٤/٤۹	٤/۸۲	٤/۷۹	٤/۸۰	٤/۸۹	٥/٦٦	٥/٦٦	٥/٦٦	٥/٦٦	٦/۲۸	٦/۲٤	٦/۲۸	٦/۲۱
۳ ماهگی	٥/۸۸	٥/۸٤	٥/۸٥	٥/۸٥	٦/۰٤	٦/۰۲	٦/۰۲	٦/۰۱	۷/٥۲	۷/٥۱	۷/٥۲	۷/٥۲	۷/۷٤	۷/۷۰	۷/٦۹	۷/۷۳
٤/٥ ماهگی	٦/۷٦	٦/۷٤	٦/۷٦	٦/۷٦	۷/۰۷	٦/۹۹	٦/۹۹	٦/۹۳	۸/٦۲	۸/٦۲	۸/٦۲	۷/٦۹	۸/۹۹	۸/۸۷	۸/۸٤	۸/۹٤
٦ ماهگی	۷/٦۱	۷/٥٦	۷/٦۱	۷/٥۸	۷/۸۳	۷/۷۹	۷/۸۰	۷/۷۰	۱۰/۳٤	۱۰/۲۳	۱۰/۲۳	۱۰/۲۳	۹/۹۱	۹/۸٤	۹/۸۱	۹/۹۲
۸ ماهگی	۸/٥۱	۸/٤٤	۸/٥۱	۸/٥۰	۸/٦	۸/٥٦	۸/٦۰	۸/٤۹	۱۱/٤٦	۱۱/٤۱	۱۱/٤۱	۱۱/٤۱	۱۰/۸۹	۱۰/۷۹	۱۰/۷۹	۱۰/۹۰
۱۰ ماهگی	۹/۰٤	۸/۹٦	۹/۰٤	۹/۰۳	۹/۱۲	۹/۰۹	۹/۱۲	۹/۰٤	۱۱/٥۹	۱۱/٤۹	۱۱/٦۷	۱۱/٤٦	۱۱/٥۷	۱۱/٤٦	۱۱/٤۸	۱۱/٥۷
۱۲ ماهگی	۹/۳۲	۹/۳٤	۹/۳٥	۹/۳۲	۹/٤۷	۹/٤٥	۹/٤۹	۹/٤٤	۱۲/٤٤	۱۲/۰۹	۱۲/۰۹	۱۲/۰۹	۱۲/۰٤	۱۱/۹۲	۱۱/۹۸	۱۲/۰۲
۱٥ ماهگی	۱۰/۰۳	۹/۷۸	۱۰/۰۲	۱۰/۰۲	۹/۸۳	۹/۸۱	۹/۸۷	۹/۸۸	۱۲/۹٥	۱۲/٦۳	۱۲/٦۳	۱۲/٦۳	۱۲/٥٤	۱۲/۳۹	۱۲/٥۰	۱۲/٤٦
۱۸ ماهگی	۱۰/٤۲	۱۰/۱۸	۱۰/٤	۱۰/٤	۱۰/۲۳	۱۰/۱۷	۱۰/۲٤	۱۰/۳۲	۱۲/۸۸	۱۲/۸۰	۱۲/۸۰	۱۲/۸۰	۱۳/۰٤	۱۲/۸٤	۱۲/۹٤	۱۲/۹۱
۲۱ ماهگی	۱۰/۹۸	۱۰/٦۰	۱۰/۹٤	۱۰/۹۷	۱۰/۹۰	۱۰/۷۸	۱۰/۸۱	۱۰/۹۰	۱۳/۹۳	۱۳/۹۱	۱۳/۹۱	۱۳/۹۱	۱۳/۸۱	۱۳/٥٤	۱۳/٥۲	۱۳/٦۰
۲٤ ماهگی	۱۱/٥٤	۱۱/٤۷	۱۱/٤۹	۱۱/٤۹	۱۲	۱۱/۷٤	۱۱/۷۰	۱۱/۷٤	۱٤/٥٦	۱٤/۱٤	۱٤/۱٤	۱٤/۱٤	۱٥/۰٤	۱٤/٦۱	۱٤/۳٦	۱٤/٦۸

همچنین تقریباً در تمام سنین صدکهای خام به مراتب کمتر از صدکهای خام هموار می‌باشند. مقدار صدک سوم در سنین مختلف در مقایسه با روشی که مقادیر گمشده نادیده گرفته شد، کاهش یافته و صدک نودوهفتم در سنین بالا بر اساس روشهای ۲ تا ۴ در مقایسه با روش ۱ کاهش پیدا کرده است و یکنوع یکنواختی در مسیر منحنی رشد بوجود می آید(نمودارهای۳ـ۱). (درصد گمشده در سنین بالا زیادتر از سنین پائین می‌باشند). در صدک پنجاهم تغییرات خیلی محسوس بوجود نیامده است.

بحث

بر اساس نتایج این مطالعه و با توجه به جدول و نمودارهای رشد تعیین شده، هر چند که تفاوت معنی‌دار آماری بین روشهای یاد شده جهت برآورد مقادیر گمشده و نادیده گرفتن مقادیر گمشده وجود ندارد ولی نادیده گرفتن مقادیر گمشده موجب می‌شود که تجزیه و تحلیل و مسیرهای منحنی رشد ناقص و نادرست باشند. با توجه به اینکه برآورد مقادیر گمشده هدف اصلی مطالعه نمی‌باشد و در واقع بدنبال ایجاد داده‌های کامل جهت تجزیه و تحلیل مقادیر بر

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

براساس روشهای آماری مختلف می‌باشیم شیوه انتساب چندگانه مناسبتر است (۲۵و۲۳و۴) و در حالتی که منحنی رشد مورد نظر باشد، استفاده از مدل و بکارگیری روشهای آماری چند سطحی مناسبتر است. اصولاً وقتی که داده‌های گمشده با درصد قابل توجه در مطالعه وجود داشته باشند سه مشکل اساسی: کاهش کارائی، پیچیدگی در تجزیه و تحلیل و دسترسی داده‌ها و سوگرایی بعلت تفاوتهای بین داده‌های مشاهده شده و نشده، در تحلیل داده‌ها می‌تواند رخ دهد. (۲۲). انتساب و جایگزینی، وضعیت داده‌های گمشده را در شروع تجزیه و تحلیل تصحیح می‌نماید و تجزیه و تحلیل کننده می‌تواند با استفاده از تکنیک‌های شناخته شده مربوط به داده‌های کامل تحلیل مناسب اعمال نماید تجزیه و تحلیل کارآ و مؤثر ممکن است تحت استفاده از یک مدل انجام گیرد. در حالیکه مدل تحلیلی ممکن است با مدل انتساب یکسان یا نابرابر باشد. انجام انتساب و جایگزینی وقتی که انتساب دهنده و تجزیه و تحلیل کننده، متفاوت هستند قابلیت اجرائی بهتری دارد. وقتی یک مدل بعنوان یک وسیله و ابزار لازم جهت انتساب مقادیر گمشده مورد استفاده قرار می‌گیرد تغییر یا مفهوم پارامترهای آن خیلی سخت نیست و مطلوبیت مدل در توانایی آن برای پیش‌بینی کردن مشاهدات گمشده می‌باشد(۲۷).

در این بررسی در سنین بالا تفاوت قابل توجه‌ای در صدکهای محاسبه شده (خام و هموار نشده) وجود دارد که می‌تواند ناشی از درصد قابل توجه مقادیر گمشده در سنین مذکور باشد. لذا با توجه به تعداد نمونه مورد مطالعه و درصد مقادیر گمشده (حداقل۶٪) پیشنهاد می‌گردد، از دو روش انتساب چندگانه و مدل منحنی رشد(فردی) جهت برآورد مقادیر گمشده استفاده گردد.

تقدیر و تشکر

بدینوسیله از معاون محترم پژوهشی و رئیس محترم تحصیلات تکمیلی دانشگاه علوم پزشکی شیراز جهت تصویب طرح پایان نامه دکتری تشکر و قدردانی می گردد.

****

References

1. Roberts SB, Dallal GE. The new childhood growth charts. Nutrition Reviews 2001; 59(2): 31-6.

۲. سلیمی اشکوری ه. گزارشی بر هنجاریابی رشد کودکان از تولد تا ٤ سالگی، تهران، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی ایران. ۱۳۷۹.

۳. خادمی ع . مدلسازی سرعت رشد کودکان ۲- ۰ ساله شیراز، پایان‌نامه کارشناسی ارشد آمار زیستی دانشگاه علوم پزشکی شیراز،۱۳۷۹.

4. Schafer JL, Olsen MK. Multiple imputation for multivariate missing data Problems: A data analyst's perspective. The Pennsylvania State University 1998; pp: 1-41.

5. Diggle PJ, Liang KY, Zeger SL. Analysis of longitudinal data. Oxford: Science Publications 1996.

6. Zeger SL, Liang KY. An overview of method for the analysis of longitudinal data. Statistics in Medicine II 1992; pp: 1825-39.

7. Rubin DB, Little RJA. Statistical analysis with missing data. NewYork: John Wiley and Sons 1987.

8. Goldstein H. The design and analysis of longitudinal studies. London, Academic Press 1994.

9. Cnann A, Laired NM, Slasor P. Tutorial in biostatistics: using the general linear mixed model to analysis unbalanced repeated measures and longitudinal data. Statistics in Medicine 1997; 16:2349-80.

10. Dwyer JH, Feinleib M, Lipert P, Hoffmeister H. Statistical models foe longitudinal studies of health. Oxford University Press 1991.

[DOR: 20.1001.1.15614107.1383.6.4.4.0]

11. Wilett JB, Singer JD, Martin NC. The design and analysis of longitudinal studies of development and psychopathology in context: statistical models and methodological recommendations. Development and Psychopathology 1998; 10: 395-426.

12. Myrtveit I, Stensurd E. Analayzing data sets with missing data: An empirical evaluation of imputation methods and likelihood based method. IEEE Transactions on Software Engineering 2001; 27:1-15.

13. Wothke W. Longitudinal and multi group modeling with missing data 2000. http:www.small waters.Com/white papers/longmiss/.

14. Rubin DB. Inference and missing data, Biometrika 1976; 63: 581-92.

15. Little RJA, Rubin DB. Statistical analysis with missing data. Wiley NewYork 1987.

16. Goldstein H. Multilevel statistical models. London: Edward Arnold & Halstead press 1995.

17. Ayatollahi SMT. Infants development milestone pattern of Shiraz (Iran) in relation to Denver chart (in press).

18. Ohtaki M. An extension of multivariate growth models with linear structure for parameters. Technical report 1994; 94: 1-18.

19. Ayatollahi SMT. Multilevel Modeling: A powerful tool for handling longitudinal growth data analysis. International congress on computation methods in engineering, Shiraz University 1993.

20. Hoseini M, Carpenter RG, Mohammad K. Growth of childern in Iran. Annals of Human Biology 1998; 25(3): 249-61.

21. Bryk AS, Raudenbush SW. Hierarchical linear models: Applications and data analysis methilds. Sage Publications, Newbury Park, CA, 1992.

22. Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of sofware packages for regression models with missing variables. The American statistican 2001; 55(3): 244-54.

23. Schafer JL. Imputation procedures for missing data. Continuing education session prepared for the INCLEN XVI Global meeting, March 1-4, 1999, Bangkok. Software for multiple imputation available from: http:/ww.stat.psu.edu/~jls/misostwa.html.

24. Rubin DB. Multiple imputation after 18^{+} years (with discussion). Journal of the American Statistical Association, 1996; 91: 473-89.

25. Schafer JL. Imputation of missing covariate under a general linear mixed model. Technical Report, Dept of statistics, Penn State University 1997.

26. Rasbash J, Healy M. Grandstat: A statistical package for the acquisition and analysis of data relate to growth and development: London School of Hygiene and Tropical Medicine 1998.

27. Schafer JL, Yucel RM. Multivariate linear mixed- effect models with missing values. Journal of the American Statistical Association 1999; 102: 379-94.

* آدرس نویسنده مسئول: بابل، دانشگاه علوم پزشکی، بخش پزشکی اجتماعی، تلفن: ٤-۲۲۲۹۵۹۱-۰۱۱۱.

[DOR: 20.1001.1.15614107.1383.6.4.4.0]