

# O desafio da informação longitudinal: um “passeio guiado” sobre modelação hierárquica, “tracking” e informação omissa com um conjunto de dados do estudo de Muzambinho

CDD. 20.ed. 001.422  
796.03

José António Ribeiro MAIA\*  
Luciano BASSO\*  
Jorge Alberto de OLIVEIRA\*\*  
Cláudia Lúcia de Moraes FORJAZ\*\*  
António PRISTA\*\*\*  
Go TANI\*\*

\*Faculdade de Desporto, Universidade do Porto - Portugal.

\*\*Escola de Educação Física e Esporte, Universidade de São Paulo.

\*\*\*Faculdade de Ciências de Educação Física e Desportos, Universidade Pedagógica de Maputo - Moçambique.

## Resumo

O grande propósito deste texto é apresentar um tutorial para investigadores das Ciências do Desporto e da Educação Física acerca dos desafios que se colocam quando se analisa informação longitudinal. A partir de um exemplo com dados reais do estudo longitudinal-misto de Muzambinho percorrem-se três avenidas de preocupações: 1) a construção de um discurso desenvolvimentista com base na modelação hierárquica; 2) a apresentação de duas soluções para lidar com informação omissa; 3) a pesquisa sobre a estabilidade das diferenças interindividuais nas mudanças intraindividuais (i.e., do “tracking”). Em cada uma dessas avenidas são lançadas questões cujas soluções são sempre acompanhadas de leituras dos principais resultados dos distintos programas estatísticos utilizados.

UNITERMOS: Modelação hierárquica; Dados omissos; “Tracking”.

## Introdução e essência dos problemas

A circunscrição, clara e extensa, da ideia de “desenvolvimento” continua um desafio (CONNOLLY & PRECHTL, 1981; ELLMAN, 2003; MICHAEL & MOORE, 1995). O seu diálogo é constante com investigadores de diferentes territórios das Ciências do Desporto (CD) e Educação Física (EF). Contudo, uma das facetas que melhor ilustra a ideia de desenvolvimento - a mudança - merece um consenso generalizado.

Descrever e interpretar, de modo distinto e esclarecedor, a mudança que ocorre numa característica qualquer (representada de modo contínuo ou discreto) que inquiete pesquisadores da EF e CD continua um desafio interminável. Esse desafio é tão mais vasto quando se é aluno ou investigador principiante na matéria desenvolvimentista.

Mudança implica, necessariamente, o uso de um delineamento longitudinal para a identificar (BOYLE & WILLMS, 2001). Claro que a visão longitudinal pode, operacionalmente, ser estrita (longitudinal puro) ou mista (longitudinal-misto). Sobre esse assunto valerá

a pena uma leitura atenta de dois livros de texto que reputamos de extremamente didáticos e esclarecedores [*Longitudinal research* de MENARD (2002) e *Analyzing social and political change: a casebook of methods* editado por DALE e DAVIES (1994)]. Estamos, pois, no domínio da ilustração, tão extensa quanto possível, de fatias relevantes do curso da “história natural” de uma dada característica, ou facetas importantes de um estudo de intervenção. Nunca é demais salientar o fato de a aventura desenvolvimentista ser sempre parcelada, fatiada. É impossível, em tempo real, não obstante a presença de métodos designados de análise funcional associada a “intensive longitudinal designs” (FOK & RAMSAY, 2006) ou sistemas dinâmicos (BOKER & GRAHAM, 1998), marcar toda a complexidade do desenvolvimento. Restam-nos janelas de leitura amostradas sequencialmente. Espera-se que essa visão, necessariamente fragmentada e reduzida, ajude a perscrutar a essência desenvolvimentista.

A extensão da inquietação pelo curso desenvolvimentista de uma determinada característica é bem antiga. Relembramos algumas mais atuais no estudo da força muscular (MAIA, BEUNEN, LEFEVRE, RENSON & VANREUSEL, 2003), consumo máximo de O<sub>2</sub> (BEUNEN, BAXTER-JONES, MIRWALD, THOMIS, LEFEVRE, MALINA & BAILEY, 2002), crescimento ósseo (BAXTER-JONES, KONTULAINEN, FAULKNER & BAILEY, 2008), alteração do índice de massa corporal (DANNER, 2008), coordenação motora grossa (BASSO, SOUZA, ARAÚJO, MEIRA JÚNIOR, OLIVEIRA, PRISTA, TANI & MAIA, submetido; DEUS, BUSTAMANTE, LOPES, SEABRA, SILVA & MAIA, 2009), ou textos mais clássicos e altamente desafiadores pelo pioneirismo das suas metodologias e procedimentos de análise como são os de ESPENCHADE (1940) e de CARRON e BAILEY (1974).

A análise estatística da informação longitudinal representa um desafio constante ao investigador que pretende extrair dos seus dados toda a potencialidade de resposta às inquietações que estão na base do seu trabalho. Felizmente que dispomos, atualmente, de um vasto “arsenal” metodológico e estatístico para conduzir pesquisas longitudinais que vão, por exemplo, de livros de textos muito didáticos (BIJVELD, VAN DER KAMP, MOOIJART, VAN DER KLOOT, VAN DER LEEDEN & VAN DER BURG, 1998; DUNCAN, DUNCAN & STRYCHER, 2006; GOTTMAN, 1995), aos mais especializados (FITZMAURICE, LAIRD & WARE, 2004; FREES, 2004). Acrescenta-se a essa versatilidade a facilidade de acesso a “softwares” (SPSS, STATA, HLM, LDA) com diferentes modelos para analisar dados longitudinais cujos manuais são altamente didáticos na sua escrita e modo como guiam o utilizador (HEDEKER, GIBBONS, DU TOIT & CHENG, 2008; RABE-HESKETH & SKRONDAL, 2005). Artigos breves e de “entrada facilitada” na análise de informação longitudinal são cada vez em maior número (veja, por exemplo, ANTRETTTER, DUNKEL, OSVATH, VOROS, FEKETE & HARING, 2006; DANNER, 2008; MILLAR, BRAY, MCGOWAN, MACDONALD & MCCARTNEY, 2007). Há também algum esforço na construção de textos sobre a matéria em língua Portuguesa (FAUSTO, CARNEIRO, ANTUNES, PINTO & COLOSIMO, 2008; MAIA, GARGANTA, SEABRA, LOPES, SILVA, BUSTAMANTE, FERMINO, FREITAS, PRISTA & MEIRA JÚNIOR, 2007a; MAIA, SOUZA, FERMINO, SEABRA, SILVA, SILVA, BUSTAMANTE, LOPES, PRISTA & FREITAS, 2007b; SPYRIDES, STRUCHINER, BARBOSA & KAC, 2008).

A “urgência” do discurso desenvolvimentista condicionado à flecha temporal, ao caráter sistemático da mudança, à sua sequencialidade e (ir)reversibilidade num determinado domínio de pesquisa da EF e CD exige que não se limite o estudo da mudança a somente dois pontos no tempo (ROGOSA & SANNER, 1995; WILLETT, SINGER & MARTIN, 1998). É imperioso ter

mais do que dois pontos no tempo pelo simples fato de se poder testar, de modo competitivo e alternativo, distintas funções para melhor descrever a estrutura presente nos dados. Nunca será demais lembrar os leitores que se deve ter, sempre, uma visão integrada que associe, de forma interativa: 1) um modelo bem articulado acerca do processo da mudança; 2) um delineamento temporal suficientemente detalhado e preciso; 3) bem como um modelo estatístico suficientemente flexível e robusto quanto a análise da informação (BRYK & RAUDENBUSH, 1987; COLLINS, 2006; ROGOSA & SANER, 1995; WILLETT, SINGER & MARTIN, 1998).

No território da EF e CD dos países lusófonos não são correntes pesquisas longitudinais cujos trabalhos estejam disponíveis para consulta nacional ou internacional (BASSO et al., submetido), não obstante a sua relevância. Todo o projeto desenvolvimentista coloca o pesquisador diante de um conjunto de desafios que convém responder o mais esclarecidamente possível (BALTES & NESSELROADE JUNIOR, 1979):

1. Identificação direta da mudança intraindividual;
2. Identificação direta das diferenças (ou similaridades) interindividuais na mudança intraindividual;
3. Análise das inter-relações na mudança intraindividual;
4. Análise das “causas” (i.e., dos “determinantes”, ou variáveis correlatas) da mudança intraindividual;
5. Análise das “causas” das diferenças interindividuais na mudança intraindividual.

Face à reduzida presença de textos didáticos breves e utilitários em língua Portuguesa para ajudar alunos e investigadores iniciantes no território da EF e CD no que concerne à análise de informação longitudinal, escrevemos este tutorial com os seguintes propósitos: 1) apresentar modelos para descrever a melhor função que representa a estrutura da mudança; 2) o que fazer quando há dados omissos; 3) como interpretar a noção de estabilidade nas diferenças interindividuais na mudança intraindividual.

Esse “menu” é demasiadamente ambicioso pela vastidão de assuntos que encerra. Contudo, trataremos de manter uma forte “vigilância” na nossa abordagem, centrando no essencial dos aspectos que consideramos mais importantes - a prática guiada.

A estratégia a seguir é bem simples. Em primeiro lugar apresentaremos um conjunto de dados que serão utilizados ao longo dos vários propósitos; em seguida trataremos dos problemas a resolver em cada ponto, bem como a sua relevância; em terceiro lugar apresentaremos soluções numéricas e gráficas oriundas de vários “softwares” estatísticos que serão usados nas análises e que serão comentadas nos seus aspectos

mais importantes; finalmente, remeteremos os leitores para bibliografia mais técnica e especializada no que se

refere a aspectos computacionais e de otimização dos algoritmos implementados nos diferentes “softwares”.

## Métodos

### Os dados originais

Vamos então à nossa viagem, começando pelos dados disponíveis e que se referem a uma pequena fatia da informação da 1ª coorte do estudo longitudinal-misto intitulado **Crescimento e desenvolvimento motor de escolares de Muzambinho - MG** (uma visão mais extensa deste projeto é encontrada em BASSO, MEIRA JÚNIOR, OLIVEIRA, FORJAZ, SOUZA, PRISTA, MAIA & TANI, 2009). Muito genericamente, nesta pesquisa as avaliações de todas as variáveis foram efetuadas semestralmente nas diferentes coortes. Focalizaremos a nossa atenção no desempenho da prova da corrida da milha, a que se associou informação repetida sobre a atividade física das crianças da 1ª coorte cujas idades médias no início do

estudo foram: meninos =  $7,02 \pm 0,27$  anos; meninas =  $6,82 \pm 0,21$  anos. A corrida da milha (valores em segundos) foi realizada de acordo com o protocolo da bateria de testes Fitnessgram, e a atividade física foi avaliada com o questionário de GODIN e SHEPHARD (1985) cuja unidade de medida é METS/15min/semana. Os dados estão na TABELA 1. É evidente a presença de informação omissa (do inglês “missing values”) com frequências distintas: a mais baixa é na Milha1 e na At\_Fis 2 (3 registros), e a mais elevada na Milha 2 e At\_Fis 1 e 4 (6 registros). Mais adiante veremos como resolver essa situação.

Em seguida, a nossa tarefa centrou-se na análise exploratória da informação. Recorremos a dois tipos de gráficos: o de extremos e quartis (do inglês “box plot”), e o do comportamento das médias ( $\pm$  desvio-padrão).

TABELA 1 - Dados do desempenho na prova da milha (em segundos) e valores de atividade física (u.a.) de uma amostra de crianças (n = 24) da 1ª coorte em quatro momentos (registros semestrais).

id	Sexo	milha1	milha2	milha3	milha4	At_Fis1	At_Fis2	At_Fis3	At_Fis4
1	0	909	898	962	855		77	119	49
2	0	913	1027	748			68		
3	0	657	1008	976	951	46	90	81	119
4	0	815	857	675	681	52	37	43	53
5	0		946	978	1109	68	113	119	35
6	0	697	884	800	694	58	57	65	109
7	0	648	673	681	777		40	63	86
8	0		669	660	645	34	41	39	45
9	0	798	827	839	814	92	41	95	81
10	0	848	685	1120	863	94	119	55	40
11	1		621	536	495	80	60	95	58
12	1	837	869	929	949	40	81	31	64
13	1	812	990	825		74	48	94	
14	1	923				56			
15	1	690	548	528	525	94	96	93	69
16	1	631			565	57		72	101
17	1	668		783	608	62	55	41	62
18	1	678		763	613	101	74	89	119
19	1	690	800	650	597		89	98	87
20	1	822	668	991	668	64	119	75	
21	1	573		732		74		64	
22	1	661	700	657	913		71		79
23	1	733	761	682	596		68	43	27
24	1	682				79	119		

Gráfico da esquerda = extremos e quartis; Gráfico da direita = médias ± dp, a que se associa uma linha de tendência ("output" do SPSS 17).

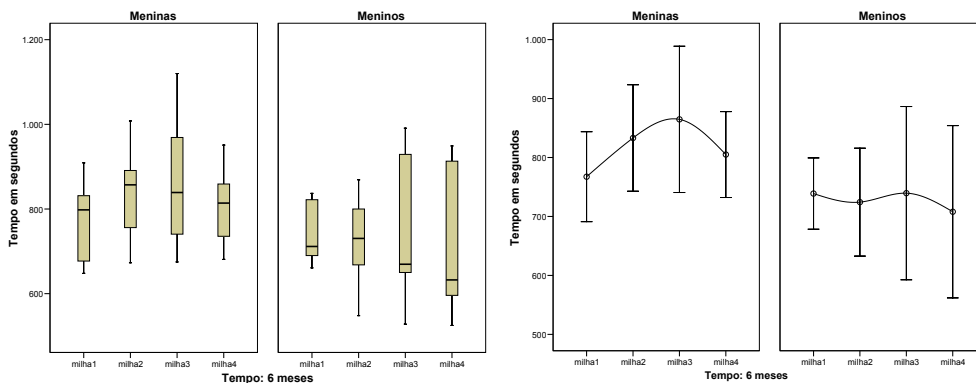


FIGURA 1 - Representação gráfica da distribuição do desempenho na prova da milha de meninos e meninas ao longo dos quatro momentos de avaliação.

Decorrem daqui os seguintes comentários: 1) não se verifica, em qualquer ponto do tempo, a presença de valores extravagantes (i.e., candidatos a "outliers"); 2) meninos e meninas têm uma variação substancial no desempenho na prova da milha, com particular destaque para os meninos; 3) a trajetória do desempenho sugere não linearidade na sua forma, sobretudo nas meninas; 4) os perfis do desempenho ao longo do tempo sugerem algum dimorfismo sexual favorecendo os meninos.

Os testes relativos à aderência de normalidade das distribuições empíricas da prova da milha<sup>1</sup> não mostraram qualquer violação desse pressuposto, conforme é evidente da FIGURA 2. Ainda que não mostremos tal informação, os valores da atividade física ao longo dos quatro pontos do tempo não mostraram, também, qualquer violação do pressuposto da normalidade da sua distribuição.

Tests of normality							
	Sex	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig	Statistic	df	Sig
milha 1	0	0,190	7	0,200*	0,920	7	0,468
	1	0,245	6	0,200*	0,865	6	0,208
milha 2	0	0,194	7	0,200*	0,925	7	0,509
	1	0,141	6	0,200*	0,985	6	0,974
milha 3	0	0,154	7	0,200*	0,939	7	0,629
	1	0,292	6	0,120	0,883	6	0,283
milha 4	0	0,161	7	0,200*	0,953	7	0,755
	1	0,255	6	0,200*	0,845	6	0,145

\* Este é o limite inferior da significância verdadeira.  
<sup>a</sup> Correção de significância de Lilliefors.

FIGURA 2 - Resultados dos testes de aderência à normalidade das distribuições de resultados da prova da milha de meninos (1) e meninas (0) nos quatro pontos do tempo ("output" do SPSS 17).

A seguir atacamos um problema mais complexo, e que pode ser colocado nestes termos (BURCHINAL & APPELBAUM, 1991): será possível modelar o desempenho da prova da milha, descrevendo a trajetória modal de meninos e meninas (se forem diferentes), mostrando trajetórias intraindividuais e diferenças entre sujeitos, ao mesmo tempo que consideramos o efeito (positivo, assim esperamos) da mudança que acontece nos valores de atividade física? Se a resposta for sim, como é que tal pode ser realizado? E como lidar com o problema da informação omissa? É evidente que essas questões

serão abordadas sequencialmente porque reclamam explicações próprias. Uma forma de resolver esses problemas é recorrer a modelos mistos que consideram, em simultâneo, efeitos fixos e aleatórios (ver, por exemplo, o tratamento dessa matéria na análise de dados antropométricos de crianças brasileiras por FAUSTO et al., 2008). Relembramos que a informação longitudinal é constituída por um tipo particular de estrutura hierárquica na sua organização: medidas repetidas "nested within subjects" (sobre esse assunto ver o exemplo didático de MAIA et al., 2007a). Os

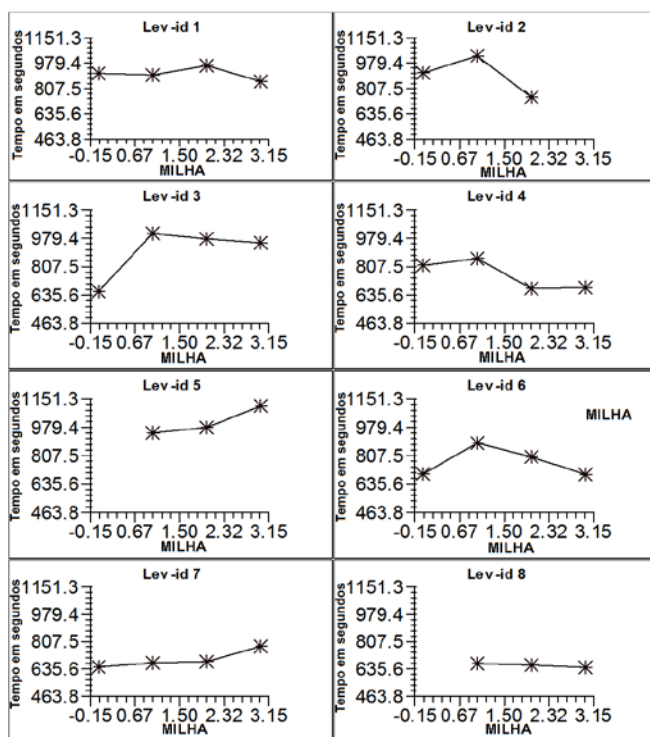
modelos mistos e a sua estrutura estatística teórica a que se adiciona os métodos de estimação de parâmetros (GOLDSTEIN, 2003; RAUDENBUSH & BRYK, 2002) e algoritmos implementados nos diferentes “softwares” disponíveis (por exemplo, SPSS, STATA, HLM, SUPERMIX, SAS) são suficientemente flexíveis para permitir a presença de informação

omissa (caso contrário da ANOVA de medidas repetidas - sobre essa matéria ver, por exemplo, KRISTJANSSON, KIRCHER & WEBB, 2007). Os resultados obtidos são precisos, uma vez que consideram toda a informação disponível para estimar, pelo método da máxima verosimilhança, os diferentes parâmetros dos vários modelos a serem testados.

## Modelação hierárquica e multinível

Antes de nos lançarmos na aventura da modelação da informação vejamos, em primeiro lugar, as FIGURAS 3 e 4. A FIGURA 3 refere-se a uma simples amostra das trajetórias do desempenho da milha de oito crianças. É bem evidente a presença de diferenças interindividuais do seu desempenho aeróbio ao longo do tempo salientando, em três crianças, informação omissa.

A FIGURA 4 ilustra, muito bem, diferentes aspectos do desempenho, que testaremos mais adiante: 1) a presença de diferenças interindividuais no início; 2) o mesmo ocorre nas trajetórias intraindividuais; 3) a eventualidade de, no tempo, a mudança no desempenho na prova da milha ser de natureza não linear na sua forma.



No programa HLM o primeiro momento assume valor 0 no início da trajetória.

FIGURA 3 - Trajetórias do desempenho na prova da milha de oito crianças nos quatro momentos.



Chamamos a atenção dos leitores que o momento inicial foi considerado como “baseline”, e portanto é o nosso ponto zero; neste sentido, a métrica temporal é 0, 1, 2 e 3.

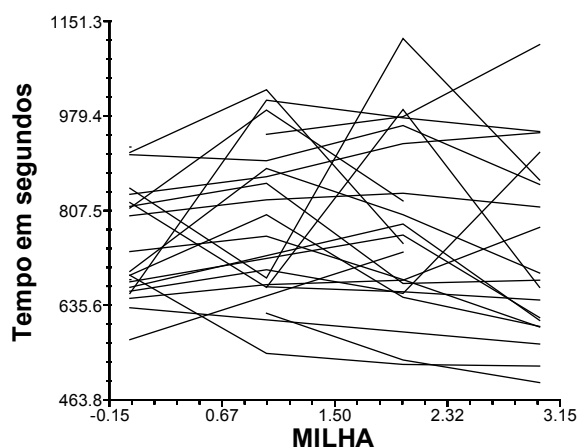


FIGURA 4 - “Spaghetti plot” da distribuição de valores da prova da milha ao longo dos quatro pontos do tempo.

O problema seguinte é modelar as trajetórias. Tal como tivemos a oportunidade de mostrar noutros textos (MAIA et al., 2007a; 2007b), a modelação hierárquica ou multinível (MHMN) apresenta uma forma de pensar e analisar a informação desenvolvimentista com uma organização estrutural precisa que se encontra distribuída por níveis (BOYLE & WILLMS, 2001). No nosso caso, no 1º nível temos informação repetida na prova da milha (bem como um preditor que também muda no tempo - a  $At\_Fi$ ), e no segundo nível temos os sujeitos com um único preditor - o gênero. Trata-se, em suma, de efetuar um esforço de pensamento e de modelação em dois níveis: no primeiro, a construção de um modelo para descrever a mudança; no segundo modelar o comportamento dos parâmetros do 1º nível em função de um conjunto variado de preditores. Esse esforço deve ser efetuado a partir de uma teoria ou corpo substantivo de hipóteses, falsificável<sup>2</sup>, para interpretar a variação interindividual nos parâmetros descritores das trajetórias individuais (BRYK & RAUDENBUSH, 1987).

Aparentemente a estrutura informacional é simples, contudo o modelo que as liga não é. Vejamos que tipo de perguntas poderemos colocar à nossa informação (BRYK & RAUDENBUSH, 1987): 1) será possível descrever a trajetória modal do desempenho da milha ao longo dos quatro pontos do tempo (visão normativista)? 2) será possível estimar a variação interindividual existente no início do estudo e na mudança (visão diferencialista)?; 3) será possível estimar a fiabilidade dos valores da trajetória modal no “baseline”?; 4) poderemos estimar a relação existente entre valores de partida e mudança?; 5) qual será o efeito do preditor que também muda no tempo ( $At\_Fis$ ) e do que permanece fixo (gênero) nas alterações individuais do desempenho na milha?

A resposta a essas perguntas implica um esforço sistemático de construção de modelos sequenciais de complexidade crescente que serão testados para verificar não só a sua qualidade, mas também o seu valor referente a explicação do comportamento da informação. Contrastam-se, pois, modelos mais parcimoniosos relativamente a mais saturados (i.e., com mais parâmetros). A estatística por excelência da qualidade de cada modelo é a “Deviance” que só tem sentido quando se contrastam dois modelos “nested within”. O melhor modelo é o que apresenta uma “Deviance” menor, e que no contraste salientou um valor de qui-quadrado significativo (sobre esse assunto consultar RAUDENBUSH & BRYK, 2002).

Testamos uma série de modelos. O primeiro tratou da descrição da trajetória modal do desempenho, ao mesmo tempo que procurou identificar a presença de heterogeneidade nas mudanças intraindividuais. Os resultados desse primeiro esforço estão nas FIGURAS 5 e 6. Interpretamos os resultados: 1) o melhor modelo é representado por um polinômio do segundo grau, implicando uma parte não linear na sua representação gráfica (foi necessário incluir o termo  $tempo^2$  para conseguir um melhor ajustamento aos dados); 2) o valor médio do desempenho da prova da milha de todas as crianças, no início do estudo (“baseline”), foi de  $743,91 \pm 23,28$  segundos,  $p < 0,001$ ; 3) há uma parte do modelo que descreve um aumento do tempo para cobrir a distância ( $65,65 \pm 27,59$ ,  $p < 0,001$ ); 4) que depois de atingir um “pico” volta a decrescer ( $-22,27 \pm 28,93$ ,  $p = 0,015$ ) (ver FIGURA 6); 5) constata-se a presença de heterogeneidade (i.e., variância) significativa em torno do valor de “baseline” ( $11088,54$ ,  $p < 0,001$ ); 6) há variância residual, cuja magnitude é de  $9415,48$ ; 7) quando se incluíram mais dois parâmetros no modelo, variâncias na parte linear e não linear do

modelo, o seu valor não foi significativo de forma que foram removidas e os resultados do modelo com essas variâncias não serão apresentados; 8) o valor da

consistência da média do “baseline” é elevado, 0,778; 9) finalmente, a “Deviance” desse modelo de descrição da mudança é de 982,42.

Fixed effect	Standard		Approx.		
	Coefficient	Error	T-ratio	df	P-value
For INTRCPT1, P0 INTRCPT2, B00	743.911805	23.278058	31.958	23	0.000
For TEMPO slope, P1 INTRCPT2, B10	65.646456	27.587467	2.380	76	0.020
For TEMPO2 slope, P2 INTRCPT2, B20	-22.266804	8.929400	-2.494	76	0.015

Final estimation of variance components:

Random effect	Standard deviation	Variance component	df	Chi-square	P-value
INTRCPT1, R0	105.30213	11088.53788	23	9415.48009	0.000
level-1, E	97.03340	9415.48009			

Statistics for current covariance components model

Deviance = 982.421197

Number of estimated parameters = 5

FIGURA 5 - Resultados do melhor modelo descritivo da mudança do tempo para completar a prova da milha ao longo dos quatro momentos (partes do “output” do HLM 6.06).

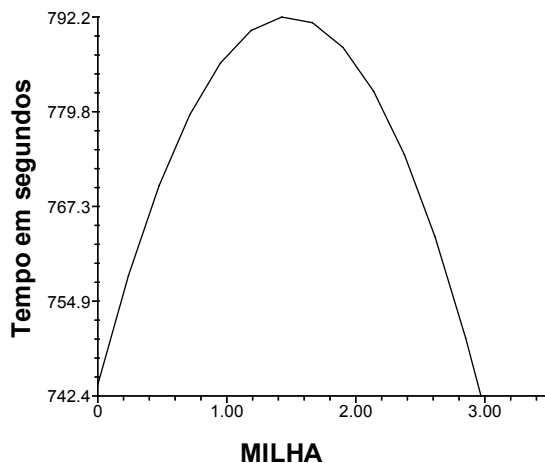


FIGURA 6 - Trajetória modal do desempenho da prova da milha das crianças ao longo dos quatro semestres.

Em seguida foi construído um modelo mais complexo com os dois preditores disponíveis: gênero (fixo, do inglês “time invariant”) e Act\_Fi (variável no tempo, do inglês “time varying”). Está representado graficamente na FIGURA 6. Os resultados numéricos encontram-se na FIGURA 7. Vejamos o seu significado. Primeiro os numéricos: 1) o valor

médio do desempenho na prova no início do estudo relativo às meninas foi de  $762,33 \pm 92,15$  segundos (essa interpretação é consistente com o fato do gênero ter sido codificado em 0 = meninas e 1 = meninos); 2) os meninos não têm um desempenho significativamente distinto ( $p = 0,526$ ); 3) tal como anteriormente, a parte linear (com declive positivo) é

evidente ( $124,17 \pm 48,58$ ,  $p = 0,013$ ), sem que haja qualquer diferença entre sexos ( $p = 0,492$ ); 4) não se constatou qualquer efeito positivo dos níveis de atividade física no “baseline” ( $p = 0,613$ ), embora se note uma tendência ligeira ( $p = 0,064$ ) de associação com as trajetórias ao longo do tempo dos meninos ( $-2,30 \pm 1,18$ ,  $p = 0,064$ ); 5) a componente não linear continua significativa ( $-40,61 \pm 12,91$ ,  $p = 0,003$ ) mostrando que a partir, sensivelmente, do segundo semestre de medição, há uma melhoria no

desempenho, sem que haja qualquer diferenciação entre meninos e meninas ( $p = 0,508$ ). A “Deviance” desse modelo é bem menor que a do modelo anterior ( $D_{M2} = 858,19$  versus  $D_{M1} = 982,42$ ), salientando a sua relevância estatística ( $\chi^2_{(7)} = 124,23$ ,  $p < 0,001$ ).

A FIGURA 8 mostra as trajetórias paralelas do desempenho na prova da milha salientando a tendência dos mais ativos que estão no percentil 75 ( $P_{75}$ ) terem tempos de prova mais baixos do que as crianças cujos níveis de atividade física estão no  $P_{25}$ .

Final estimation of fixed effects  
(with robust standard errors)

Fixed effect	Standard		Approx.		
	Coefficient	Error	T-ratio	df	P-value
For INTRCPT1, P0					
INTRCPT2, B00	762.332792	92.145763	8.273	22	0.000
SEXO, B01	70.772064	109.892716	0.644	22	0.526
For TEMPO slope, P1					
INTRCPT2, B10	124.173432	48.502841	2.560	62	0.013
SEXO, B11	-45.535092	65.871095	-0.691	62	0.492
For ACTF slope, P2					
INTRCPT2, B20	0.515756	1.007465	0.512	22	0.613
SEXO, B21	-2.295964	1.179624	-1.946	22	0.064
For TEMPO2 slope, P3					
INTRCPT2, B30	-40.605774	12.909872	-3.145	62	0.003
SEXO, B31	13.162935	19.801363	0.665	62	0.508

Final estimation of variance components:

Random effect	Standard deviation	Variance component	df	Chi-square	P-value
INTRCPT1, R0	216.73553	46974.28937	17	48.64603	0.000
ACTF slope, R2	1.89501	3.59106	17	33.42896	0.010
level-1, E	81.23855	6599.70210			

Statistics for current covariance components model

Deviance	= 858.193415
Number of estimated parameters	= 12

Model comparison test

Chi-square statistic	= 124.22778
Number of degrees of freedom	= 7
P-value	= 0.000

FIGURA 7 - Resultados do melhor modelo descritivo e explicativo da mudança na prova da milha de meninas e meninos ao longo dos quadro pontos do tempo (“output” do HLM 6.06).



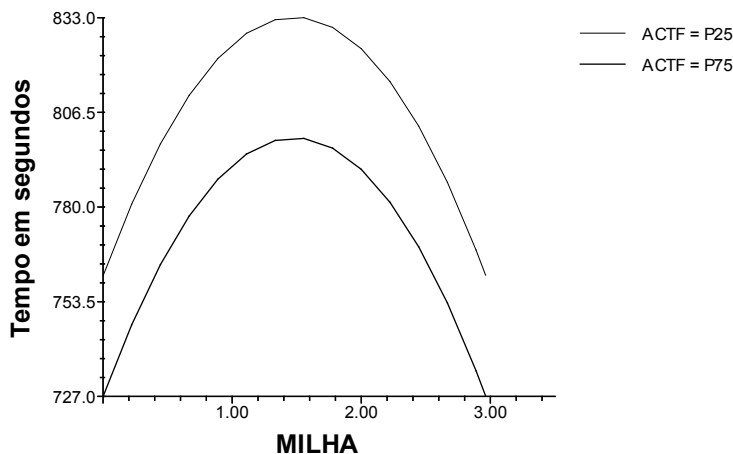


FIGURA 8- Trajetória modal do desempenho da prova da milha em função de dois percentis ( $P_{25}$  e  $P_{75}$ ) dos níveis de atividade física das crianças ao longo dos quatro semestres.

## Informação omissa e seu tratamento

Centremos a nossa atenção num novo e desafiante problema, inicialmente colocado no final da introdução. Pretendemos estudar aspectos da estabilidade e previsibilidade (do inglês “tracking”) da mudança ocorrida na prova de corrida da milha. Contudo, tal tarefa só pode ser levada a cabo com informação completa em todas as crianças. Os algoritmos implementados no “software” LDA (que utilizaremos para analisar o “tracking”) só aceitam informação completa em cada unidade de análise. Se houver informação omissa espera-se, na melhor das situações, que seja reduzida e com um padrão conhecido uma vez que tais omissões de informação numa determinada célula implica a remoção de toda a linha de dados. É desejável que a eliminação seja a mais reduzida possível. Claro que esse fato traz problemas relativamente sérios - a redução da dimensão amostral implica um abaixamento do poder dos testes estatísticos, aumenta a magnitude dos erros-padrão diminuindo a precisão das estimativas dos vários parâmetros em apreço e limita a generalização dos resultados.

Tal como informamos anteriormente os leitores, a TABELA 1 contém os valores originais. Há dados omissos em diferentes crianças dos dois sexos e em distintas ocasiões. A estratégia a seguir para ultrapassar esse problema, guiando o leitor, é a seguinte: 1) inicialmente apresentaremos, de modo muito genérico e sem qualquer formulação matemática, conceitos essenciais associados à informação omissa (do inglês “missing data”). É importante salientar que o estudo das diferentes estratégias para lidar com dados omissos

é uma necessidade e urgência de qualquer pesquisador, tendo sempre presente que não há nenhum método “mágico” que resolva, por si mesmo, todos os problemas (ver, por exemplo, COLLINS, SCHAFER & KAM, 2001; GRAHAM, 2009); 2) em seguida utilizaremos os “softwares” SYSTAT 12 (para testar formalmente o padrão de dados omissos) e NORM para remediar os problemas em apreço; 3) construiremos, no NORM, a partir de imputação múltipla (do inglês “multiple imputation”), cinco “novas” bases completas de dados (pelo processo de “data augmentation”) que serão posteriormente utilizadas quando tratarmos do problema do “tracking” da mudança na aptidão cardiorespiratória de meninos e meninas marcada pela prova da corrida da milha. Contrastaremos os resultados dessa opção com uma outra designada de imputação simples (do inglês “single imputation”) com base no algoritmo EM e no método de máxima verosimilhança (ver, por exemplo, os manuais do SYSTAT 12, SPSS 17 ou ALLISON, 2002). É importante frizar que o que segue é somente uma fatia muito fina, altamente prática, centrada no exemplo das crianças de Muzambinho e na prova da milha. Não é nosso propósito apresentar um texto extenso sobre a problemática do “missing data” mas, tão somente, uma simples proposta para sanar o problema da informação omissa. É nosso dever dirigir o leitor para o estudo sério de um livro altamente didático e nada “pesado” do ponto de vista da sua formalização matemático-estatística que é o *Missing data: a gentle introduction* (MCKNIGHT, MCKNIGHT, SIDANI & FIGUEIREDO, 2007).

O problema da informação omissa é extremamente atual pelo simples motivo que é um fato consumado em qualquer pesquisa, seja ela de natureza transversal ou longitudinal obtida em estudos observacionais ou de intervenção. É um problema que apoquentava qualquer investigador sobretudo quando a sua pesquisa é de natureza longitudinal, como é o caso de Muzambinho. Felizmente que nos últimos 10-20 anos estão disponíveis procedimentos e algoritmos altamente poderosos e flexíveis para lidar com esse problema (ver por exemplo o brilhante tratamento dessa matéria em RUBIN, 1976, 1987; LITTLE & RUBIN, 1987; SCHAFER, 1997) e cujo propósito é sempre o mesmo - obter estimativas de parâmetros dos modelos construídos com os nossos dados se a informação estivesse completa. Isto é, trata-se de “preencher” (imputar) os espaços em branco nas nossas planilhas de uma forma muito precisa, de tal modo que os valores imputados são aqueles que teriam a maior probabilidade de ser muito semelhantes aos que seriam realmente obtidos pelo investigador se tivesse informação completa de toda a gente. Essa imputação pode ser individual (i.e., imputação simples) que significa que a nossa planilha fica completa; uma outra alternativa, e a mais atual, é a que constrói novas planilhas, sendo que cada valor omissa é objeto de imputação múltipla, i.e., a cada dado em falta é atribuído diferentes valores de uma distribuição construída para o efeito.

Começemos pelo princípio. O que é informação omissa? É informação que por diferentes motivos não foi colhida pelo pesquisador. De acordo com RUBIN (1987) a informação omissa pode ser de diferentes tipos: 1) “missing completely at random” (MCAR), i.e., omissão completamente aleatória, se a probabilidade de omissão de  $y_i$  não está relacionado com o verdadeiro valor de  $y$  ou do valor de qualquer outra variável no conjunto informacional. Essa é uma assunção muito forte, para a qual se dispõe de um teste formal; 2) “missing at random” (MAR), i.e., omissão aleatória, se a probabilidade de omissão de  $y_i$  não está relacionada com o verdadeiro valor de  $y$ , depois de controlar para outras covariáveis do conjunto informacional; 3) “missing not at random” (MNAR), i.e., omissão não aleatória que depende de dados não observados. Os mecanismos de informação omissa podem ser ignorados se o seu padrão for do tipo MAR, ou não ignorados, situação muito mais complexa de resolver porque exige um conhecimento bem sustentado do mecanismo que originou tal omissão informacional (ver, por exemplo, tratamentos dessa matéria em ALLISON, 2002 e HECKMAN, 1976).

Relembramos os leitores que os procedimentos clássicos de análise de dados consideram, na sua essência, que a informação é completa para cada unidade

de análise. Por exemplo, se fosse efetuado um teste de hipóteses à diferença entre sexos na corrida da milha somente no primeiro momento de avaliação era esperado ter informação completa para todos os sujeitos. Genericamente, a maioria dos “softwares” estatísticos tem uma opção (“default”) designada de “listwise” que elimina uma unidade informacional por cada célula omissa que encontre. No nosso caso e voltando ao exemplo da diferença entre sexos no primeiro momento, só poderíamos contar com oito meninas e 13 meninos; já no segundo momento teríamos 10 meninas e oito meninos. Se porventura fosse efetuada uma ANOVA de medidas repetidas só seriam considerados 13 sujeitos, sete meninas e seis meninos. Claro que essas situações são sempre indesejáveis. Há também a opção “pairwise” normalmente utilizada em procedimentos correlacionais que implica que os valores das correlações obtidas provenham de conjuntos distintos de casos. Nenhuma dessas é uma boa solução (ALLISON, 2002). Há, também, outras opções, que não abordaremos (por exemplo, substituição dos casos omissos pela média, regressão, “hotdeck”, etc.). Um estudo guiado nessas opções pode ser encontrado em MCKINGHT et al. (2007).

O procedimento que iremos seguir é o mais atual e altamente favorecido pela maioria dos investigadores (ver, por exemplo, COLLINS, SCHAFER & KAM, 2001; ALLISON, 2002; GRAHAM, 2009; SCHAFER 1997, 2001). É designado por imputação múltipla (detalhes técnicos podem ser encontrados em LITTLE & RUBIN, 1987 e SHAFER, 1997; uma versão mais “soft” desse procedimento e estratégia pode ser consultada em ALLISON, 2002). O “software” escolhido para realizar essa imputação múltipla é o NORM desenvolvido por John Schaffer da Universidade da Pensilvânia, USA ([www.stat.psu.edu/~jls/misoftwa.html](http://www.stat.psu.edu/~jls/misoftwa.html)). É altamente versátil, de fácil utilização e manuseamento. O SPSS 17 e o SYSTAT 12, por exemplo, não têm implementado qualquer módulo de imputação múltipla, mas somente de imputação simples que faz parte de um módulo mais genérico designado “Missing Data (Value) Analysis”. Em contraposição, o “software” SOLAS ([www.statsol.ie/solas/solas.htm](http://www.statsol.ie/solas/solas.htm)) foi especialmente concebido para lidar com problemas de informação omissa. Os “softwares” SAS e STATA 11 têm módulos muito esclarecedores e fortemente flexíveis quanto a modelação para imputação múltipla, tornando mais facilitada a vida do investigador nessa matéria.

Na metodologia da imputação múltipla há que percorrer três etapas: 1) imputação: imputação dos casos omissos não uma única vez (tipo imputação

simples), mas  $m$  vezes, resultando em  $m$  conjuntos completos de dados (normalmente entre 3 a 5); 2) análise: análise de cada um dos  $m$  conjuntos completos de dados num determinado “software” estatístico (no nosso caso será o LDA para calcular uma medida de “tracking”); 3) “pooling”: integração dos resultados das  $m$  análises num resultado final (média, variância e intervalos de confiança) de acordo com as sugestões de RUBIN (1987) que estão implementadas no NORM. O STATA 11 contém esta possibilidade que está automatizada para um conjunto relativamente extenso de problemas e modelos.

Contudo, ainda temos um problema para resolver. Como a nossa informação é repetida no tempo, tem uma estrutura relacional relativamente complexa que está bem expressa na matriz de covariância que reclama um esforço de modelação. Daqui que John Schaffer tenha proposto técnicas para resolver esse problema que estão implementadas no “software” que desenvolveu e que se chama PAN. O problema desse “software” é que não tem uma versão Windows, dado que só “corre” no S-Plus. Uma forma bem expedita de sanar essa dificuldade (o da estrutura da informação repetida) é referida por ALLISON (2002) e é a seguinte: em primeiro lugar há que re-estruturar a planilha dos dados do seu formato habitual (que está expresso na TABELA 1 que é designado de “person-level”) para um novo formato designado de “person-period” (nesse caso temos somente quatro colunas: a primeira com o identificador de cada criança, a segunda relativa ao tempo, a terceira com o desempenho na prova da milha e a quarta com a atividade física; tudo isto em vez das 10 colunas no formato “person-level”) - (TABELA 2).

Em seguida constroem-se, por imputação múltipla no NORM, cinco novos arquivos de dados completos. A tarefa seguinte é voltar a re-estruturar a planilha de “person-period” para “person-level”. Foi exatamente isso que fizemos.

TABELA 2 - Planilha com uma estrutura de dados do tipo “person period” para as três primeiras crianças.

Id	Tempo	Milha	AtF
1	0	909	
1	1	898	77
1	2	962	119
1	3	855	49
2	0	913	
2	1	1027	68
2	2	748	
2	3		
3	0	657	46
3	1	1008	90
3	2	976	81

Os espaços em branco correspondem a informação omissa.

Vamos então aos preliminares da análise da informação omissa. Vejamos a primeira opção de imputação simples realizada no SYSTAT 12. A FIGURA 9 mostra os padrões de “missing” ao longo dos quatro momentos (com os dados sob a forma de “person level”). Por exemplo, há 13 casos com informação completa. Dois casos têm informação omissa no último momento de avaliação. A informação seguinte refere-se ao teste formal ao padrão de MCAR, cujo resultado não permite rejeitar a hipótese do padrão de omissão ser completamente aleatório ( $p = 0,301$ ). Finalmente temos as estimativas EM<sup>3</sup> para as médias da prova da milha de todas as crianças em cada um dos momentos de avaliação, bem como a matriz de inter-correlações entre períodos (há quem a designe por matriz de auto-correlações).

Com base nessa imputação simples, a nova planilha (ver TABELA 3) do tipo “person level” passa a ser a seguinte (chamamos a atenção dos leitores que a planilha não tem informação sobre o gênero das crianças embora a sequência das linhas esteja em consonância com o código das meninas e meninos da TABELA 1).

#### Missing value patterns

N of cases	Missing value patterns
	(X = nonmissing; . = mis.)
13	XXXX
2	XXX.
3	.XXX
2	X...
1	X..X
2	X.XX
1	X.X.

FIGURA 9 - Resultados da análise ao padrão de informação omissa (resultados do SYSTAT 12).

Continua

Little MCAR test statistic	16.201			
Df	14			
p-value	0.301			
<b>EM estimate of means</b>				
milha 1	milha 2	milha 3	milha 4	
743.376	783.511	784.673	745.551	
EM estimate correlation matrix				
	milha 1	milha 2	milha 3	milha 4
milha 1	1.000			
milha 2	0.523	1.000		
milha 3	0.565	0.431	1.000	
milha 4	0.510	0.643	0.656	1.000

FIGURA 9 - Resultados da análise ao padrão de informação omissa (resultados do SYSTAT 12) (continuação).

TABELA 3 - Nova planilha dos resultados da prova da milha construída com o método de imputação simples com recurso ao algoritmo EM (Resultados do SYSTAT 12) .

<b>Milha 1</b>	<b>Milha 2</b>	<b>Milha 3</b>	<b>Milha 4</b>
909,0	898,0	962,0	855,0
913,0	1027,0	748,0	856,1
657,0	1008,0	976,0	951,0
815,0	857,0	675,0	681,0
841,0	946,0	978,0	1109,0
697,0	884,0	800,0	694,0
648,0	673,0	681,0	777,0
681,2	669,0	660,0	645,0
798,0	827,0	939,0	814,0
848,0	685,0	1120,0	863,0
633,7	621,0	536,0	495,0
837,0	869,0	929,0	949,0
812,0	990,0	825,0	871,0
923,0	913,6	938,4	892,0
690,0	548,0	528,0	525,0
631,0	662,9	646,5	565,0
668,0	686,0	783,0	608,0
678,0	694,4	763,0	613,0
690,0	800,0	650,0	597,0
822,0	668,0	991,0	668,0
573,0	677,0	732,0	659,8
661,0	700,0	657,0	913,0
733,0	761,0	686,0	596,0
682,0	739,0	732,0	695,4

## “Tracking” da prova de milha

É bem conhecido o comportamento da aptidão aeróbia de crianças e jovens ao longo da idade, quer se utilizem testes de laboratório quer de terreno. Um resumo da literatura sobre essa matéria pode ser encontrado em MALINA, BOUCHARD e BAR-OR (2004) e ROWLAND (1996). Contudo, a informação de caráter longitudinal, expressa em termos absolutos [ $\dot{V}O_2$  máx (kg/min)] mostra incrementos notórios a partir dos oito anos de idade. O que não conseguimos localizar foi informação longitudinal acerca dos valores normativos da prova da milha. Em contrapartida, é conhecido o comportamento percentílico dos resultados nessa prova em crianças e jovens açorianos (MAIA, LOPES, MORAIS, SILVA & SEABRA, 2002) salientando uma forte variação do desempenho em cada valor discreto de idade, dos sete aos 17 anos.

As referências anteriores, não obstante a sua relevância, são sobretudo pensadas num quadro normativo relevando para a importância dos valores médios ou modais. Nada referem sobre a enorme variabilidade do desempenho intraindividual do e entre sujeitos ao longo da idade. Ora é exatamente esse problema da estabilidade ou instabilidade da mudança que interessa pesquisar porque nos informa sobre constância, variabilidade e/ou comportamentos residuais na mudança intraindividual e nas diferenças interindividuais. Esse é um problema interessante, porque nos remete para o estudo do “tracking”. Afinal o que é que se entende por “tracking”, e porque é que esta etapa da interpretação da informação contida na prova da milha é importante? Estabilidade, mudança e previsão (facetas do “tracking”) exigem, necessariamente, informação de natureza longitudinal supondo, ou não, que os pontos de observação e registro das características em estudo (neste caso a prova da milha) sejam equidistantes no tempo. “Tracking” é um termo genérico que pretende descrever um padrão regular de crescimento ou mudança numa coleção de padrões de crescimento (KOWALSKI & SCHNEIDERMAN, 1992). Não obstante a presença de divergência na sua definição e formalização matemática (ver, por exemplo, FOULKES & DAVIES 1981; WARE & WU,

1981; em língua portuguesa consultar MAIA, LOPES, SILVA & SEABRA, 2002), recorreremos à estatística  $\gamma$  desenvolvida por FOULKES e DAVIES (1981). Essa estatística pretende examinar a quantidade de indivíduos (no nosso caso crianças) que mantêm a mesma posição relativa (“rank” relativo). Isto é, calcula-se a probabilidade de duas curvas ou perfis aleatórios de performance não se inter cruzarem. O “tracking” será perfeito quando o grupo de perfis de crescimento individual da performance não se entrecruzam, i.e., quando a sua posição relativa no seio da distribuição for mantida no tempo do estudo. Se  $\gamma=0,50$  não se verifica qualquer “tracking”; se  $\gamma=1,00$  o “tracking” é perfeito, e se  $\gamma < 0,50$ , o “tracking” é irrelevante. Dado que essa estatística é não-paramétrica, não há qualquer pressuposto a cumprir acerca da forma da curva de crescimento da performance. É evidente que quanto mais simples for a curva, tanto mais elevado será o valor de  $\gamma$ . Um dos maiores problemas no que se refere a interpretação do desempenho motor é pesquisar o padrão da estabilidade<sup>4</sup> ao longo do tempo. Chamamos a atenção dos leitores para a FIGURA 4 que contém as trajetórias das crianças ao longo dos quatro pontos do tempo.

Vamos aos resultados que serão apresentados em diferentes etapas: na primeira lidaremos com os resultados da imputação simples para meninos e meninas sendo que para gênero será calculada uma medida de “tracking”; a seguir trataremos os resultados da imputação múltipla também para meninos e meninas.

A partir da base de dados construída com a imputação simples e que está na TABELA 3, o valor da estatística  $\gamma$  de FOULKES e DAVIES dos meninos foi de  $0,71 \pm 0,04$  (IC95% = 0,62; 0,79), enquanto que nas meninas foi também  $0,71 \pm 0,05$  (IC95% = 0,61; 0,82). Esses resultados mostram uma estabilidade moderada no comportamento das diferenças interindividuais na prova da milha ao longo dos quatro anos. Em contrapartida, os resultados obtidos com as cinco novas planilhas (não mostradas no texto) resultantes da imputação múltipla foram os seguintes (ver TABELA 4).

TABELA 4 - Valores da estatística  $\gamma$  de Foulkes & Davies em cada uma das cinco planilhas construídas com o procedimento da imputação múltipla no NORM .

Número de imputações	$\gamma$ de Foulkes & Davies	IC 95%
Meninas		
Imputação 1	0,36 $\pm$ 0,07	0,22; 0,49
Imputação 2	0,64 $\pm$ 0,06	0,53; 0,76
Imputação 3	0,71 $\pm$ 0,06	0,60; 0,82
Imputação 4	0,71 $\pm$ 0,06	0,60; 0,82
Imputação 5	0,69 $\pm$ 0,06	0,57; 0,81
Meninos		
Imputação 1	0,51 $\pm$ 0,06	0,39; 0,62
Imputação 2	0,51 $\pm$ 0,06	0,40; 0,63
Imputação 3	0,57 $\pm$ 0,05	0,37; 0,58
Imputação 4	0,47 $\pm$ 0,05	0,37; 0,58
Imputação 5	0,60 $\pm$ 0,05	0,52; 0,69

Finalmente, no NORM e com base nas regras de cálculo desenvolvidas por RUBIN (1987) que são relativamente simples (podem ser consultadas despidas do seu formalismo matemático em MCKINGHT et al., 2007), foi possível estimar um único valor de “tracking” para a prova da milha de meninos e meninas (ver FIGURA 10). Nos meninos o  $\gamma$  de FOUKES & DAVIES é de 0,51  $\pm$  0,08, e nas meninas é de 0,62  $\pm$  0,18. Esses valores, a que se associam os respectivos intervalos de

confiança, mostram um comportamento das diferenças interindividuais na mudança intraindividual da prova da milha completamente distintos do valor reportado pela imputação simples. Ao contrário do referido anteriormente na imputação simples, não se verifica qualquer estabilidade nas trajetórias interindividuais ao longo dos quatro anos na prova da milha de meninos e meninas (ver o comportamento interindividual das trajetórias do desempenho na FIGURA 4).

Meninos					
NORM Version 2.03 for Windows 95/98/NT					
Output from MI INFERENCE: SCALAR METHOD					
TraFD					
Thursday, 27 August 2009					
16:34:13					
Data read from file:					
C:\Documents and Settings\jmaia\Desktop\meninos.txt					
Number of estimands = 1					
Number of imputations = 5					
File format: stacked rows					
QUANTITY	ESTIMATE	STD.ERR.	T-RATIO	DF	P-VALUE
QTY_1	0.514200	0.793262E-01	6.48	13	0.0000
CONFIDENCE LEVEL FOR INTERVAL ESTIMATES (%): 95.00					
QUANTITY	LOW ENDPT.	HIGH ENDPT	%MIS.INF.		
QTY_1	0.342826	0.685574	60.6		

Continua

FIGURA 10 - “Outputs” do NORM para calcular um único valor de “tracking” a partir dos resultados apresentados na TABELA 3 .



Meninas

NORM Version 2.03 for Windows 95/98/NT

Output from MI INFERENCE: SCALAR METHOD

Untitled - Thursday, 27 August 2009 - 16:31:51

Data read from file:

C:\Documents and Settings\jmaia\Desktop\meninas.txt

Number of estimands = 1

Number of imputations = 5

File format: stacked rows

QUANTITY	ESTIMATE	STD.ERR.	T-RATIO	DF	P-VALUE
QTY_1	0.622200	0.175932	3.54	5	0.0166
CONFIDENCE LEVEL FOR INTERVAL ESTIMATES (%): 95.00					
QUANTITY	LOW ENDPT.	HIGH ENDPT.	%MIS.INF.		
QTY_1	0.169953	1.07445	91.5		

FIGURA 10 - “Outputs” do NORM para calcular um único valor de “tracking” a partir dos resultados apresentados na TABELA 3 (continuação).

Claro que esta precisão na conclusão acerca dos valores de  $\gamma$  de FOUKES e DAVIES radica no fato de ser mais plausível construir novos  $m$  conjuntos de dados e estimar o valor de “tracking” do que somente com um conjunto, tal como é sugerido, por exemplo, por SCHAFER (1999), COLLINS, SCHAFER e KAM (2001) e GRAHAM (2009). Contudo, convém ter sempre presente que a imputação múltipla não é um remédio universal para todos os problemas de informação omissa. Não é um

método “mágico”, tampouco uma espécie de alquimia estatística para inventar informação onde ela não existe (SCHAFER, 1999). O tipo de omissão, a frequência de casos omissos, a natureza da informação e dos problemas a que se adiciona o tipo de modelo a testar devem guiar o pesquisador a decidir acerca do procedimento mais adequado de imputação, não se esquecendo, nunca, de testar diferentes métodos e procedimentos para melhor ajuizar da qualidade dos seus resultados.

## Conclusões e sugestões

O desafio da pesquisa longitudinal continua bem vivo, e que genericamente pode ser referido em três pontos: 1) estabelecimento de um conjunto sólido de problemas a resolver; 2) construção de um delineamento robusto que mapeie a mudança e que reduza a frequência de “drop out”; 3) a análise da mudança intraindividual, das diferenças interindividuais e seus preditores a partir de distintos modelos estatísticos. Esses foram, e são, os grandes pilares orientadores do estudo longitudinal-misto de Muzambinho.

A modelação hierárquica ou multinível é um modo de pensar e operar de enorme valor na descrição e construção de modelos competitivos para descrever e modelar a mudança normativa e diferencial. Acresce a possibilidade de analisar, em conjunto, a presença de preditores que podem mudar, ou não, no tempo. Os dados da prova da milha mostraram

um comportamento não linear, sem diferenças entre gêneros, mas com uma influência notória dos níveis distintos ( $P_{25}$  e  $P_{75}$ ) de atividade física.

O “drop-out” (i.e., abandono definitivo ou temporário) é um problema sério em qualquer pesquisa longitudinal. Neste sentido, foi explorada, muito brevemente, a situação da informação omissa, para a qual foram apresentadas estratégias para a sua solução - concretamente a imputação simples e a imputação múltipla.

Com base nos resultados obtidos das duas estratégias de imputação foi analisada a estabilidade das diferenças interindividuais na mudança da aptidão aeróbia de meninos e meninas. Os resultados são distintos, sendo que, face às representações dos comportamentos interindividuais, é mais plausível sugerir uma instabilidade acentuada no desenvolvimento da aptidão aeróbia dos meninos e meninas.

Finalmente, para referir que este texto, não obstante o seu caráter introdutório a um assunto muito vasto, pode ser um auxiliar importante, uma

porta de entrada, para os que se iniciam na difícil tarefa de lidar com informação proveniente de uma pesquisa longitudinal.

## Abstract

The challenge of longitudinal information: a "guided tour" on hierarchical modeling, tracking and missing information with a set of data from Muzambinho study

The main aim of this study is to present a tutorial to Sport Sciences and Physical Education researchers when facing challenges emerging from longitudinal data analysis. Based on a real data set from Muzambinho mixed-longitudinal study, we shall deal with three main concerns: 1) to build a developmental view based on hierarchical or multilevel modeling; 2) to present two solutions to the missing data problem; 3) to search for stability of interindividual differences in intraindividual change (i.e., tracking). In each of these main issues questions will be asked whose answers will be presented alongside with main results coming from different statistical softwares used.

UNITERMS: Hierarchical modeling; Missing data; Tracking.

## Notas

1. Não é muito sensato realizar um teste de aderência à normalidade com distribuições de  $n$  muito reduzidas como é o nosso caso. Contudo, mantemos os resultados para fins ilustrativos.
2. Relembramos os leitores acerca do caráter circunstancial e contextual do conhecimento. Todas as hipóteses acerca da descrição e interpretação de um qualquer fenómeno são plausíveis num dado ponto do tempo. São falsificáveis, uma vez que novos conhecimentos as contradizem com frequência. Chamamos a atenção dos leitores, também, para o fato de a hipótese que se coloca à prova, em termos estatísticos, ser a hipótese nula que se assume como verdadeira e se pretende rejeitar! Essa hipótese nunca é a do investigador, que é a hipótese alternativa. Ora essa hipótese nunca é colocada à prova em termos estatísticos formais.
3. No passo E ("Expectation") as estatísticas suficientes (médias, desvios-padrão e covariâncias) da base com dados omissos são substituídas por valores esperados condicionados nos dados disponíveis utilizando, para o efeito, estimativas dos parâmetros; na fase M ("Maximization") os parâmetros são continuamente atualizados pelas suas estimativas de máxima verosimilhança a partir das estatísticas suficientes obtidas na etapa E. Esse processo iterativo pára quando se obtém convergência.
4. Relembramos o leitor que à exceção da estabilidade das médias do desenvolvimento da aptidão aeróbia, estabilidade no contexto do nosso problema não significa ausência de mudança.

## Referências

- ALLISON, P.D. **Missing data**. Thousand Oaks: Sage, 2002. (Sage University papers Series on Quantitative Applications in the Social Sciences, 07-136).
- ANTRETTTER, E.; DUNKEL, D.; OSVATH, P.; VOROS, V.; FEKETE, S.; HARING, C. Multilvel modeling was a convenient alternative to common regression designs in longitudinal suicide research. **Journal of Clinical Epidemiology**, New York, v.59, p.576-86, 2006.
- BALTES, P.B.; NESSELROADE, J.R. History and rationale of longitudinal research. In: NESSELROADE, J.R.; BALTES, P.B. (Eds.). **Longitudinal research in the study of behavior and development**. New York: Academic Press, 1979. p.1-39.
- BASSO, L.; MEIRA JÚNIOR, C.M.; OLIVEIRA, J.A.; FORJAZ, C.L.M.; SOUZA, J.A.; PRISTA, A.; MAIA, J.A.R.; TANI, G. Crescimento e desenvolvimento motor de escolares de Muzambinho: um estudo com implicações académicas, sociais e de política interinstitucional. **Revista Portuguesa de Ciência do Desporto**, Porto, v.24, n.2/3, p.247-57, 2009.

- BASSO, L.; SOUZA, C.J.F.; ARAÚJO, U.O.; MEIRA JÚNIOR, C.M.; OLIVEIRA, J.A.; PRISTA, A.; TANI, G.; MAIA, J.A.R. Olhares distintos sobre a noção de estabilidade e mudança no desempenho da coordenação motora grossa. Artigo submetido a **Revista Brasileira de Educação Física e Esporte**, São Paulo.
- BAXTER-JONES, A.D.; KONTULAINEN, S.A.; FAULKNER, R.A.; BAILEY, D.A. A longitudinal study of the relationship of physical activity to bone mineral accrual from adolescence to young adulthood. **Bone**, Elmsford, v.43, n.6, p.1101-7, 2008.
- BEUNEN, G.; BAXTER-JONES, A.D.G.; MIRWALD, R.L.; THOMIS, M.; LEFEVRE, J.; MALINA, R.M.; BAILEY, D.A. Intraindividual allometric development of aerobic power in 8- to 16-year-old boys. **Medicine and Science in Sports and Exercise**, Madison, v.34, p.503-10, 2002.
- BIJLEVELD, C.C.J.H.; van der KAMP, L.J.T.; MOOIJART, A.; van der KLOOT, W.A.; van der LEEDEN, R.; van der BURG, E. **Longitudinal data analysis: designs, models and methods**. Thousand Oaks: Sage, 1998.
- BOKER, S.M.; GRAHAM, J. A dynamical systems analysis of adolescent substance abuse. **Multivariate Behavioral Research**, Philadelphia, v.33, p.479-507, 1998.
- BOYLE, M.H.; WILLMS, J.D. Multilevel modeling of hierarchical data in developmental studies. **Journal of Child Psychology and Psychiatry**, New York, v.42, p.141-62, 2001.
- BRYK, A.S.; RAUDENBUSH, S.W. Application of hierarchical linear models to assessing change. **Psychological Bulletin**, Lancaster, v.101p.147-58, 1987.
- BURCHINAL, M.; APPELBAUM, M.I. Estimating individual developmental functions: methods and their assumptions. **Child development**, Lafayette, v.62, p.23-43, 1991.
- CARRON, A.V.; BAILEY, D.A. Strength development in boys from 10 to 16 years. **Monographs of the Society for Research in Child Development**, Chicago, v.39, n.4, 1974. (Serial n.157).
- COLLINS, L.M. Analysis of longitudinal data: the integration of theoretical model, temporal design, and statistical model. **Annual Review of Psychology**, Palo Alto, v.57, p.505-28, 2006.
- COLLINS, L.M.; SCHAFER, J.L.; KAM, C-M. A comparison of inclusive and restrictive strategies in modern missing data procedures. **Psychological Methods**, Washington, v.6, p.330-51, 2001.
- CONNOLLY, K.J.; PRECHTL, H.F.R. **Maturation and development: biological and psychological perspectives**. London: William Heinemann Medical Books, 1981. (Clinics in Developmental Medicine, n.77/78).
- DALE, A.; DAVIES, R.B. (Eds.). **Analyzing social and political change: a casebook of methods**. Thousand Oaks: Sage, 1994.
- DANNER, F.W. A national longitudinal study of the association between hours of TV viewing and the trajectory of BMI growth among US children. **Journal of Pediatric Psychology**, Washington, v.33, p.1100-7, 2008.
- DEUS, R.K.B.C.; BUSTAMANTE, A.; LOPES, V.P.; SEABRA A.F.T.; SILVA, R.M.; MAIA, J.A.R. Coordenação motora: estudo de tracking em crianças dos 6 aos 10 anos da Região Autónoma dos Açores, Portugal. **Revista Brasileira de Cineantropometria & Desempenho Humano**, Florianópolis, v.10, p.215-22, 2008.
- DUNCAN, T.E.; DUNCAN, S.C.; STRYCKER, L.A. **An introduction to latent variable growth curve modeling: concepts, issues, and applications**. 2nd. ed. Mahwah: Lawrence Erlbaum, 2006.
- ELMAN, J. Development: it's about time. **Developmental Science**, Malden, v.6, p.430-3, 2003.
- ESPENCHADE, A. Motor performance in adolescence including the study of relationships with measures of physical growth and maturity. **Monographs of the Society for Research in Child Development**, Chicago, v.5, n.1, 1940. (Serial n.24).
- FAUSTO, M.A.; CARNEIRO, M.; ANTUNES, C.M.F.; PINTO, J.A.; COLOSIMO, E.A. O modelo de regressão linear misto para dados longitudinais: uma aplicação na análise de dados antropométricos desbalanceados. **Cadernos de Saúde Pública**, Rio de Janeiro, v.24, p.513-24, 2008.
- FITZMAURICE, G.M.; LAIRD, N.M.; WARE, J.H. **Applied longitudinal analysis**. New Jersey: John Wiley, 2004.
- FOK, C.C.T.; RAMSAY, J.O. Periodic trends, non-periodic trends and their interactions in longitudinal or functional data. In: WALLS, T.A.; SHAFER, J.L. (Eds.). **Models for intensive longitudinal data**. New York: Oxford University Press, 2006. p.109-23.
- FOULKES, M.A.; DAVIES, C.E. An index of tracking for longitudinal data. **Biometrics**, Washington, v.37, p.439-46, 1981.
- FREES, E.W. **Longitudinal and panel data analysis and applications in the social sciences**. Cambridge: Cambridge University Press, 2004.
- GODIN, G.; SHEPHARD, R.J. A simple method o assess exercise behaviour in the community. **Canadian Journal of Applied Sciences**, Champaign, v.10, n.3, p.141-6, 1985
- GOLDSTEIN, H. **Multilevel statistical models**. 3rd. ed. London: Arnold, 2003.
- GOTTMAN, J.M. (Ed.). **The analysis of change**. Mahwah: Lawrence Erlbaum, 1995.

- GRAHAM, J.W. Missing data analysis: making it work in the real world. **Annual Review of Psychology**, Palo Alto, v.60, p.549-76, 2009.
- HECKMAN, J.J. The common structure of statistical models of truncated, sample selection and limited dependent variables, and a single estimator of such models. **Annals of Economic and Social Measurement**, New York, v.5, p.475-92, 1976.
- HEDEKER, D.; GIBBONS, R.; du TOIT, M.; CHENG, Y. **Manual of SuperMix: mixed effects models**. Chicago: Scientific Software International, 2008.
- KOWALSKI, C.J.; SCHNEIDERMAN, E.D. Tracking: concepts, methods and tools. **International Journal of Anthropology**, Firenze, v.7, p.33-50, 1992.
- KRISTJANSSON, S.D.; KIRCHER, J.C.; WEBB, A.K. Multilevel models for repeated measures research designs in psychophysiology: an introduction to growth curve modeling. **Psychophysiology**, Baltimore, v.44, p.728-36, 2007.
- LITTLE, R.J.A.; RUBIN, D.B. **Statistical analysis with missing data**. New York: Wiley, 1987.
- MAIA, J.A.; BEUNEN, G.; LEFEVRE, J.; CLAESSENS, A.L.; RENSON, R.; VANREUSEL, B. Modeling stability and change in strength development: a study in adolescent boys. **American Journal of Human Biology**, Malden, v.15, n.4, p.579-91, 2003.
- MAIA, J.A.R.; LOPES V.P.; MORAIS, F.P.; SILVA, R.M.G.; SEABRA, A. **Estudo do crescimento somático, aptidão física, atividade física e capacidade de coordenação corporal de crianças do 1º ciclo do ensino básico da Região Autónoma dos Açores**. Porto: DREFD/FCDEF-UP, 2002.
- MAIA, J.A.R.; LOPES, V.P.; SILVA, R.G.; SEABRA, A. A Importância do estudo do tracking (estabilidade e previsão) em delineamentos longitudinais: um estudo aplicado à epidemiologia da atividade física e à performance desportivo-motora. **Revista Portuguesa de Ciências do Desporto**, Porto, v.4, p.41-56, 2002.
- MAIA, J.A.R.; SILVA, R.M.G.; SEABRA, A.; LOPES, V.; SILVA, S.; BUSTAMANTE, A.; FERMINO, R.C.; FREITAS, D.; PRISTA, A.; MEIRA JÚNIOR, C. Uma ajuda na análise e interpretação de informação de aptidão física de crianças e jovens provenientes de amostras de grande dimensão: um tutorial centrado na modelação hierárquica ou multinível. **Revista Portuguesa de Ciências do Desporto**, Porto, v.7, n.3, p.379-90, 2007a.
- MAIA, J.A.R.; SOUSA, A.C.; FERMINO, R.; SEABRA, A.; SILVA, S.; SILVA, R.G.; BUSTAMANTE, A.; LOPES, V.P.; PRISTA, A.; FREITAS, D.L. Análise e interpretação dos de atividade física de crianças: um tutorial baseado na modelação hierárquica ou multinível. **Revista Brasileira de Cineantropometria e Desempenho Humano**, Florianópolis, v.9, n.4, p.424-45, 2007b.
- MALINA, R.M.; BOUCHARD, C.; BAR-OR, O. **Growth, maturation and physical activity**. 2nd. ed. Champaign: Human Kinetics, 2004.
- McKNIGHT, P.E.; McKNIGHT, K.M.; SIDANI, S.; FIGUEREDO, A.J. **Missing data: a gentle introduction**. New York: The Guilford Press, 2007.
- MENARD, S. **Longitudinal research**. 2nd. ed. Thousand Oaks: Sage, 2002. (Sage University papers Series on Quantitative Applications in the Social Sciences, n.7-76).
- MICHAEL, G.F.; MOORE, C.L. **Developmental psychobiology: an interdisciplinary science**. Cambridge: MIT Press, 1995.
- MILLAR, P.J.; BRAY, S.R.; McGOWAN, C.L.; MacDONALD, M.J.; McCARTNEY, N. Effects of handgrip training among people medicated for hypertension: a multilevel analysis. **Blood Pressure Monitoring**, London, v.12, p.307-14, 2007.
- RABE-HESKETH, S.; SKRONDAL, A. **Multilevel and longitudinal modeling using STATA**. 2nd. ed. College Station: StataPress, 2005.
- RAUDENBUSH, S.W.; BRYK, A.S. **Hierarchical linear models: applications and data analysis methods**. 2nd. ed. Thousand Oaks: Sage, 2002.
- ROGOSA, D.; SANER, H. Longitudinal data analysis with random coefficients models. **Journal of Educational and Behavioral Statistics**, Washington, v.20, p.149-70, 1995.
- ROWLAND, T.W. **Developmental exercise physiology**. Champaign: Human Kinetics, 1996.
- RUBIN, D.B. Inference and missing data. **Biometrika**, Cambridge, v.63, p.581-92, 1976.
- \_\_\_\_\_. **Multiple imputation for nonresponse in surveys**. London: Wiley, 1987.
- SCHAFFER, J.L. **Analysis of incomplete multivariate data**. Boca Raton: Chapman & Hall, 1997.
- \_\_\_\_\_. Multiple imputation: a primer. **Statistical Methods in Medical Research**, London, v.8, p.3-15, 1999.
- \_\_\_\_\_. Multiple imputation with PAN. In: COLLINS, L.M.; SAYER, A.G. (Eds.). **New methods for the analysis of change**. Washington: American Psychological Association, 2001. p.357-77.
- SPYRIDES, M.H.C.; STRUCHINER, C.J.; BARBOSA, M.T.; KAC, G. Efeito da duração da amamentação predominante no crescimento infantil: um estudo prospectivo com modelos não lineares de efeitos mistos. **Jornal de Pediatria**, Rio de Janeiro, v.84, p.237-43, 2008.

WARE, J.H.; WU, M.C. Tracking: prediction of future values from serial observations. **Biometrics**, Washington, v.37, p.424-37, 1981.

WILLET, J.B.; SINGER, J.D.; MARTIN, N.C. The design and analysis of longitudinal studies of development and psychopathology in context: statistical models and methodological recommendations. **Development and Psychopathology**, New York, v.10, p.395-426, 1998.

ENDEREÇO

José António Ribeiro Maia  
Laboratório de Cineantropometria  
Faculdade de Desporto  
Universidade do Porto  
R. Dr. Plácido Costa, 91  
4200-450 - Porto - PORTUGAL  
e-mail: jmaia@fade.up.pt

Recebido para publicação: 17.05.2010  
Aceito: 21.06.2010