

Metodologias de identificação de padrões alimentares *a posteriori* em crianças brasileiras: revisão sistemática

Methods of *a posteriori* identification of food patterns in Brazilian children: a systematic review

Carolina Abreu de Carvalho¹

Poliana Cristina de Almeida Fonsêca²

Luciana Neri Nobre³

Silvia Eloiza Priore⁴

Sylvia do Carmo Castro Franceschini⁴

Abstract *The objective of this study is to provide guidance for identifying dietary patterns using the a posteriori approach, and analyze the methodological aspects of the studies conducted in Brazil that identified the dietary patterns of children. Articles were selected from the Latin American and Caribbean Literature on Health Sciences, Scientific Electronic Library Online and Pubmed databases. The key words were: Dietary pattern; Food pattern; Principal Components Analysis; Factor analysis; Cluster analysis; Reduced rank regression. We included studies that identified dietary patterns of children using the a posteriori approach. Seven studies published between 2007 and 2014 were selected, six of which were cross-sectional and one cohort, Five studies used the food frequency questionnaire for dietary assessment; one used a 24-hour dietary recall and the other a food list. The method of exploratory approach used in most publications was principal components factor analysis, followed by cluster analysis. The sample size of the studies ranged from 232 to 4231, the values of the Kaiser-Meyer-Olkin test from 0.524 to 0.873, and Cronbach's alpha from 0.51 to 0.69. Few Brazilian studies identified dietary patterns of children using the a posteriori approach and principal components factor analysis was the technique most used.*

Key words *Food consumption, Child, Factor analysis*

Resumo *O objetivo deste estudo é fornecer orientações para identificação de padrões alimentares por abordagem a posteriori, bem como analisar os aspectos metodológicos dos estudos realizados no Brasil que os identificaram em crianças. Os artigos foram selecionados nas bases de dados da Literatura Latino-Americana e do Caribe em Ciências da Saúde, Scientific pattern; Principal component analysis; Factor analysis; Cluster analysis; Reduced rank regression. Incluíram-se pesquisas que identificaram padrões alimentares de crianças por meio da abordagem a posteriori. Selecionou-se 7 estudos, sendo 6 transversais e 1 de coorte, publicados entre 2007 e 2014. Cinco usaram como inquérito o questionário de frequência alimentar, um o recordatório de 24h e outro uma lista de alimentos. O método de abordagem exploratória mais utilizado nas publicações foi a análise fatorial por componentes principais, seguida da análise de agrupamento. O tamanho amostral dos estudos variou de 232 a 4231, os valores do teste Kaiser-Meyer-Olkin de 0,524 a 0,873, e o alfa de Cronbach de 0,51 a 0,69. Poucos estudos brasileiros identificaram padrões alimentares de crianças utilizando abordagem a posteriori, e a análise fatorial por componentes principais foi a técnica mais usada.*

Palavras-chave *Consumo alimentar, Crianças, Análise fatorial*

¹ Programa de Pós-Graduação em Saúde Coletiva, Universidade Federal do Maranhão. R. Barão de Itapary 155, Centro. 65020-070 São Luís MA Brasil. carol.103@me.com

² Programa de Pós-Graduação em Ciência da Nutrição, Universidade Federal de Viçosa (UFV).

³ Departamento de Nutrição, Universidade Federal dos Vales do Jequitinhonha e Mucuri.

⁴ Departamento de Nutrição, UFV.

Introdução

Os hábitos alimentares são formados durante a infância e têm importantes implicações não apenas em curto prazo, mas também na vida adulta^{1,2}. Crianças que apresentam consumo alimentar inadequado podem estar expostas a déficit estatural, sobrepeso, obesidade, além de carências de micronutrientes, tais como anemia e hipovitaminose A³. Dessa forma, a associação entre dieta infantil e o desenvolvimento de doenças e desvios nutricionais justificam a avaliação do consumo alimentar nessa faixa etária.

Durante muito tempo a relação entre hábitos alimentares e saúde foi avaliada apenas levando-se em consideração o consumo de nutrientes ou alimentos isolados. Entretanto, o entendimento de que alimentos e nutrientes interagem entre si e podem apresentar efeitos sinérgicos, levou a epidemiologia nutricional ao estudo dos padrões alimentares^{4,5}.

O padrão alimentar pode ser definido como um conjunto de alimentos frequentemente consumidos por indivíduos e populações^{3,6}. Essa abordagem permite avaliar a dieta de uma perspectiva global, facilitando o estabelecimento de estratégias de promoção da alimentação saudável e prevenção de doenças e agravos nutricionais. Assim, o estudo de padrões alimentares representa um instrumento complementar para avaliação do efeito da dieta na saúde^{2,7}. Além disso, essa proposta supera algumas limitações como incapacidade de detectar pequenos efeitos de nutrientes e dificuldades da avaliação de interações entre os nutrientes⁸.

Na identificação de padrão alimentar exige-se um complexo manejo estatístico dos dados de consumo alimentar. Comumente, duas abordagens analíticas são utilizadas: *a priori* e *a posteriori*. Na abordagem *a priori*, o padrão alimentar é definido a partir de índices ou escores dietéticos previamente estabelecidos com base em evidências científicas para doenças específicas⁹. O Índice de Alimentação Saudável é um exemplo de avaliação do padrão alimentar por escore baseado em recomendações. Este índice verifica o grau em que a dieta do indivíduo está em conformidade com as recomendações do *US Department of Agriculture Food Guide Pyramid*. O Índice de Qualidade da Dieta é outro exemplo de escore que indica o grau que a dieta do indivíduo está de acordo com as recomendações do *Committee on Diet and Health*¹⁰.

Por outro lado, a abordagem *a posteriori* é um método exploratório que utiliza técnicas de aná-

lise multivariada para obter padrões alimentares. Esse tipo de análise permite agregar os alimentos consumidos pelo indivíduo e, posteriormente, reduzir essas informações a conjuntos de dados menores que representem a exposição à dieta⁶. Nessa abordagem, os padrões alimentares mais relevantes na população estudada são identificados a partir da correlação entre os dados coletados nos inquéritos alimentares, sendo possível se obter diversos padrões que, não necessariamente, refletem um padrão alimentar saudável^{11,12}. Os principais métodos utilizados na abordagem *a posteriori* são a análise fatorial (*factor analysis*), a análise de agrupamento (*cluster analysis*) e a regressão por redução de posto (*reduced rank regression*)^{9,13}.

A identificação *a posteriori* utiliza técnicas estatísticas mais robustas e permite encontrar os padrões alimentares da população em estudo, mesmo que estes não reflitam uma dieta saudável. Isto possibilita uma avaliação mais ampla da dieta e dá margem para a realização de ações de prevenção melhor direcionadas. As abordagens *a priori* e *a posteriori*, são muito diferentes do ponto de vista estatístico e, conseqüentemente, também dos resultados gerados.

No presente artigo optou-se por trabalhar apenas com a abordagem *a posteriori*, uma vez que a identificação de padrões alimentares nesta perspectiva tem sido tema de crescente interesse dentro da saúde coletiva nacional, porém os estudos nesta linha ainda são muito escassos, especialmente no público infantil. As técnicas estatísticas utilizadas *a posteriori* são complexas e as orientações para sua aplicação na extração de padrões alimentares têm sido pouco documentadas na literatura científica nacional. Até o momento, não foi encontrado outro manuscrito com esta proposta.

O interesse pelo público infantil decorre da importância da avaliação precoce do consumo alimentar como forma de possibilitar intervenções de prevenção para agravos na vida adulta. São poucos os estudos brasileiros que se dedicaram a identificação de padrões alimentares em crianças e o presente trabalho pretende colaborar para estimular e dar base teórica a futuros estudos com essa temática.

Portanto, esta revisão se propõe a verificar as publicações brasileiras sobre identificação de padrões alimentares na população infantil que utilizaram a abordagem *a posteriori*; e propõe, ainda, analisar e fornecer orientações relativas aos aspectos metodológicos para estudos que pretendem utilizar esta abordagem.

Metodologia

Trata-se de uma revisão sistemática da literatura, baseada na análise de aspectos metodológicos dos artigos sobre padrão alimentar de crianças brasileiras identificados pela abordagem *a posteriori*. Os artigos foram selecionados por meio de busca nas bases de dados PubMed, Literatura Latino-Americana e do Caribe em Ciências da Saúde (Lilacs) e *Scientific Electronic Library Online* (SciELO), entre abril e maio de 2014. Os descritores utilizados na busca dos artigos foram: *Dietary pattern*; *Food pattern*; *Principal component analysis*; *Factor analysis*; *Cluster analysis*; *Reduced rank regression*. As combinações de descritores foram: *Dietary pattern* + *Principal component analysis* ou *Factor analysis* ou *Cluster analysis* ou *Reduced rank regression*; e *Food pattern* + *Principal component analysis* ou *Factor analysis* ou *Cluster analysis* ou *Reduced rank regression*.

Nesta revisão incluíram-se somente os artigos originais publicados em português, inglês ou espanhol, que identificaram padrão alimentar de crianças brasileiras por meio de abordagem *a posteriori*. Não houve delimitação quanto ao ano de publicação dos artigos.

A busca foi realizada pelos autores, identificando-se inicialmente 637 artigos que possuíam pelo menos um dos descritores em seu título ou resumo. Posteriormente, foi realizada a avaliação dos artigos encontrados, procedendo-se a leitura dos títulos e resumos tanto para a exclusão de manuscritos repetidos, quanto para a identificação de artigos que não se encaixassem nos critérios de elegibilidade estabelecidos neste estudo. A partir desta análise foram excluídos 325 artigos por serem repetidos, 131 por não terem identificado padrão alimentar apesar de constar no título, 72 por terem utilizado a abordagem *a priori*, 52 por não terem extraído o padrão alimentar de crianças, 39 por não terem sido realizados no Brasil e 12 por serem artigos de revisão. Após estas exclusões restaram 6 trabalhos e foram realizadas buscas a partir das referências destes. A partir da busca ativa nas referências foram encontrados mais quatro artigos de potencial relevância para o objetivo deste estudo, sendo que após a análise do resumo destes, apenas um foi selecionado para compor esta revisão. Finalmente, foram selecionados sete artigos para esta revisão (Figura 1).

Os sete artigos selecionados foram lidos integralmente pelos autores. As informações de interesse foram obtidas principalmente a partir da seção de métodos, uma vez que o presente estudo

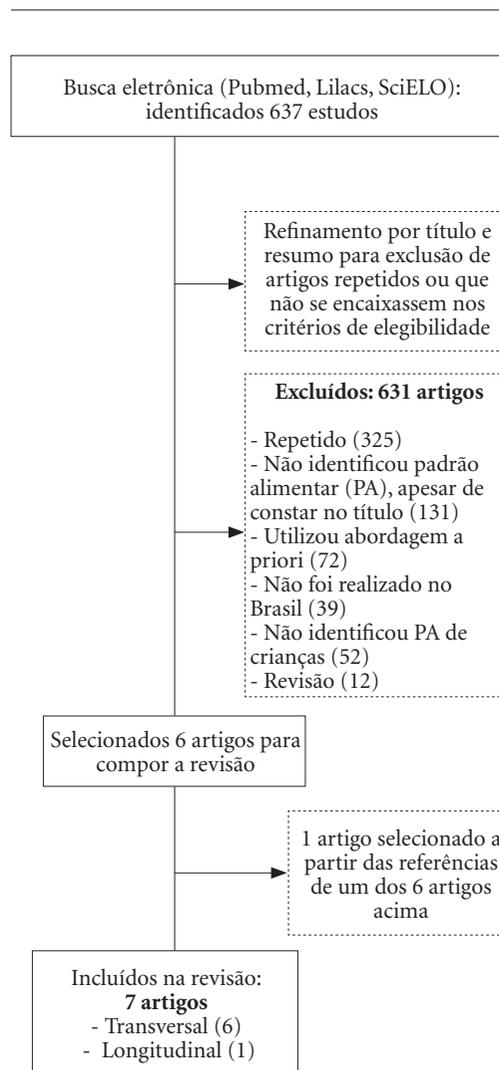


Figura 1. Fluxograma de busca de artigos.

tem como foco a abordagem metodológica dos trabalhos. Entretanto, em muitos artigos analisados as informações de interesse constavam na seção de resultados, a exemplo da nomenclatura dada aos padrões identificados. A discussão dos artigos também foi cuidadosamente analisada, pois nesta seção muitos autores destacaram os pontos fortes e as limitações das técnicas e métodos de identificação do padrão alimentar que utilizaram.

A discussão das metodologias de identificação dos padrões alimentares e orientações para a realização desta análise *a posteriori* foi embasada

em artigos internacionais, capítulos de livro sobre o tema, bem como na experiência profissional dos autores.

Resultados

No Quadro 1 são apresentados os aspectos metodológicos dos estudos brasileiros que identificaram padrões alimentares de crianças, por meio da abordagem *a posteriori*. Entre os trabalhos analisados, cinco tiveram como método de inquérito alimentar o questionário de frequência alimentar (QFA), um o recordatório de 24 horas (R24h) e outro uma lista de alimentos. O método de abordagem exploratória mais utilizado foi a análise fatorial por componentes principais (6 estudos), seguida da análise de agrupamento (1 estudo).

Na identificação de padrões alimentares utilizando a análise de componentes principais (ACP), alguns procedimentos devem ser realizados, dentre eles podem ser citados: 1) verificação do tamanho amostral; 2) Preparação de matriz de correlação, que é obtida com o teste de Kaiser-Meyer-Olkin (KMO), o qual avalia se há correlações entre os itens alimentares; 3) Extração de um conjunto de fatores da matriz de correlação, que é avaliado pelo teste de esfericidade de Bartlett, e este avalia se os dados produzem uma matriz de identidade; 4) Análise do percentual da variância de cada item explicada por todos os fatores juntos, ou seja, a comunalidade, e esta varia entre 0-1; 5) Determinação do número de fatores, geralmente com uso do teste gráfico de Cattell (*scree plot*) e a rotação dos fatores para aumentar a sua interpretabilidade; 6) A interpretação dos resultados obtidos; e, 7) denominação dos fatores, ou seja, dar títulos aos padrões alimentares identificados⁸.

Apesar de não ser condição para a ACP, recomenda-se ainda proceder a análise da consistência interna das dimensões do QFA, e neste caso é comum o uso do *alpha* de Cronbach, o qual varia de 0 a 1, e considera-se adequado quando este é superior a 0,6⁸.

Considerando os procedimentos citados acima, dos seis trabalhos que utilizaram a análise fatorial na identificação de padrões alimentares, cinco referiram ter utilizado a razão entre o número de indivíduos da amostra e o de itens alimentares para verificar a adequação da amostra para a realização desta análise. Nesta regra recomenda-se que o número de indivíduos seja igual ou superior a cinco para cada alimento/grupo de

alimentos do QFA¹⁴. O tamanho amostral dos referidos estudos variou de 232 a 4231 indivíduos. Dos cinco trabalhos que citaram terem realizado o cálculo da adequação amostral para esse teste, apenas o estudo de Matos et al.³ não identificou amostra satisfatória.

Os valores obtidos nos testes de KMO e de esfericidade de Bartlett, autovalores utilizados, significância das cargas fatoriais e alfa de Cronbach são apresentados na Tabela 1.

O uso do teste de KMO foi relatado em cinco publicações^{3,12,15-17}. Vale ressaltar que neste teste é necessário um valor superior a 0,6 para considerar modelo fatorial adequado. No estudo de Matos et al.³ não foi obtido valor adequado para realização desta análise em parte da amostra do estudo. Quatro publicações^{3,12,16,17} citaram a realização do teste de esfericidade de Bartlett, no qual o valor de p deve ser menor ou igual a 0,05 para continuidade da análise.

A rotação ortogonal Varimax, para facilitar a interpretabilidade dos dados, e o teste de *scree plot*, ou do percentual da variância acumulada, para auxiliar na decisão sobre o número de fatores a serem retidos, foram empregados em todos os trabalhos que fizeram análise fatorial por componentes principais.

Das pesquisas que citaram o percentual da variância acumulada, cinco fixaram a carga fatorial em valores acima de 0,3, enquanto um fixou acima de 0,4¹⁷.

O coeficiente alpha de Cronbach foi referido em três publicações. Silva et al.¹⁶ encontraram coeficientes superiores a 0,65 para os dois padrões alimentares identificados. Nobre et al.¹² identificaram coeficiente adequado para dois padrões alimentares e menor que 0,6 para um outro. No trabalho de Souza et al.¹⁷, dois dos cinco padrões extraídos estavam abaixo do nível aceitável para este teste.

O estudo de Gama et al.¹⁸ foi o único que identificou o padrão alimentar por meio da análise de agrupamentos, utilizando a técnica de métrica euclidiana, seguida por K-média. Essas técnicas serão detalhadas na discussão deste manuscrito. Neste estudo foram identificados seis padrões alimentares. Destes vale destacar o padrão “tradicional”, que apresenta os aspectos saudáveis da alimentação tradicional brasileira, composto por arroz, feijão, óleos, leite e pão francês. Também foi identificado um grupo dos “alimentos dietéticos”, composto por produtos dietéticos, bebidas alcoólicas e naturais, mel/melado, o qual segundo os autores “reflete a mídia, vendedora de produtos pseudo-‘saudáveis’, que podem levar

Quadro 1. Aspectos metodológicos dos estudos que identificaram padrões alimentares por abordagem *a posteriori* de crianças no Brasil.

Autor/Ano	Local	Delineamento	Idade	N	Inquérito	Análise estatística	Testes
Gama et al. 2007 ¹⁸	Rio de Janeiro (RJ)	Transversal	5 a 9 anos	356	QFA	Análise de agrupamento	Métrica euclidiana
D'Innocenzo et al., 2011 ¹⁵	Salvador (BA)	Transversal	4 a 11 anos	1260	QFA	Análise fatorial por componentes principais	KMO Adequação do tamanho da amostra VARIMAX <i>Screeplot</i>
Gatica et al., 2012 ¹⁹	Pelotas (RS)	Coorte	12, 24, 48 meses	4231	Lista de alimentos	Análise fatorial por componentes principais	VARIMAX <i>Screeplot</i>
Nobre et al. 2012 ¹²	Diamantina (MG)	Transversal	5 anos	232	QFA	Análise fatorial por componentes principais	KMO Teste de esfericidade de Bartlett Adequação do tamanho da amostra VARIMAX <i>Screeplot</i> Coeficiente alpha de Cronbach
Silva et al., 2012 ¹⁶	Salvador (BA)	Transversal	7 a 14 anos	1136	QFA	Análise fatorial por componentes principais	KMO Teste de esfericidade de Bartlett VARIMAX <i>Screeplot</i> Coeficiente alpha de Cronbach
Souza et al., 2013 ¹⁷	Pelotas (RS)	Transversal	1 a 6 anos	667	QFA	Análise fatorial por componentes principais	KMO Teste de esfericidade de Bartlett Adequação do tamanho da amostra VARIMAX <i>Screeplot</i> Coeficiente alpha de Cronbach
Matos et al., 2014 ³	Salvador, Itiruçu, Santa Inês, São Félix, Salinas das Margaridas, Acajutiba, Cipó, Gongogi, Milagres, Presidente Dutra e Serrolândia (BA)	Transversal	Menores de 5 anos	3817	R24h	Análise fatorial por componentes principais	KMO Teste de esfericidade de Bartlett VARIMAX <i>Screeplot</i>

QF A= Questionário de Frequência Alimentar; KMO = Kaiser-Meyer-Olkin; R24h = Recordatório de 24 horas.

Tabela 1. Valores obtidos nos testes de KMO e de esfericidade de Bartlett, autovalores utilizados, significância das cargas fatoriais e alfa de Cronbach, nos estudos que utilizaram análise fatorial por componentes principais na identificação dos padrões alimentares de crianças no Brasil.

Autor/ano	Kaiser-Meyer-Olkin (KMO)	Teste de esfericidade de Bartlett	Autovalores	Significância das cargas fatoriais	Alfa de Cronbach
D’Innocenzo et al., 2011 ¹⁵	0,8177	Não calculado	> 1,0	≥ 0,30	Não calculado
Gatica et al., 2012 ¹⁹	Não calculado	Não calculado	Não apresentado	≥ 0,30	Não calculado
Nobre et al. 2012 ¹²	0,697	p ≤ 0,001		≥ 0,30	0,53 - 0,63
Silva et al., 2012 ¹⁶	0,873	p ≤ 0,001	≥ 1,3	≥ 0,25	>0,65
Souza et al., 2013 ¹⁷	0,720	p ≤ 0,050	Não apresentado	≥ 0,40	0,51 – 0,69
Matos et al., 2014 ³	0,524 - 0,754	p ≤ 0,001	> 1,0	≥ 0,30	Não calculado

aos agravos nutricionais”. Outro padrão de risco identificado neste estudo foi o de “bebidas industrializadas”, composto ainda por balas, biscoitos e frutas, portanto, essencialmente formado por açúcares simples e que representa um padrão de risco para o sobrepeso. Os demais padrões extraídos não estavam muito bem definidos.

No trabalho de D’Innocenzo et al.¹⁵, foram identificados quatro padrões alimentares, os quais não foram nomeados. Os autores observaram a dependência dos padrões alimentares às condições socioeconômicas das famílias das crianças e que o consumo de alimentos considerados mais saudáveis está associado a grupos de mais alto nível socioeconômico.

Gatica et al.¹⁹, identificaram quatro padrões para a faixa etária de 12 e 24 meses (leites, alimentos básicos, carne e vegetais, lanches) e cinco aos 48 meses, a saber, os anteriormente citados e mais o “guloseimas”. Estes padrões alimentares foram fortemente associados a características sociais (escolaridade materna, posição socioeconômicas e creche) e comportamentais (duração da amamentação, a mamadeira e uso de chupeta).

Três padrões foram obtidos no trabalho de Nobre et al.¹². Os padrões foram denominados: “dieta mista”, composto por alimentos típicos da alimentação dos brasileiros, tais como arroz e raízes, feijão, leites e derivados, sucos e frutas; “lanches”, formado essencialmente por alimentos que não requerem preparo, como pães, margarinas, achocolatados, biscoitos doces e salgados; e “não saudáveis”, constituído por guloseimas gordurosas e doces, sucos artificiais, biscoitos recheados, refrigerantes, doces e sobremesas. Observou-se

que o padrão “dieta mista” associou-se à menor escolaridade materna, enquanto a maior renda *per capita* ao consumo do padrão “não saudável”.

O estudo de Silva et al.¹⁶, encontrou os padrões alimentares “obesogênico” e “tradicional” em crianças e adolescentes da rede pública de Salvador. O padrão “obesogênico”, representado por leite e derivados, frituras e sanduíches, óleos e gorduras, açúcares e doces, preparações típicas e bebidas, associou-se a menor escolaridade materna. Já o padrão “tradicional” (carnes e derivados, ovos, cereais e derivados, verduras, raízes e tubérculos, frutas) não foi influenciado por nenhum dos indicadores socioeconômicos analisados no estudo.

Em Pelotas-RS, Souza et al.¹⁷ identificaram cinco padrões alimentares: “vegetais”; “tradicional” composto por pão, margarina, arroz e massas, café e açúcar; “guloseimas e embutidos” representado por salgadinhos, embutidos, refrigerantes, chocolates; “lanches”, constituído por laticínios, achocolatados, biscoitos e sucos; e “frutas”. Os padrões “vegetais” e “frutas” se associaram a maior escolaridade materna e renda familiar, enquanto o padrão “tradicional” teve maior adesão em crianças com mães de menor escolaridade e renda familiar.

Matos et al.³ não atribuíram nome aos padrões que identificaram em crianças menores de cinco anos de Salvador e de outros municípios em áreas urbanas e rurais da Bahia. Este estudo observou que as crianças maiores de seis meses apresentaram um baixo consumo de leite materno e pouca variação na ingestão de frutas e legumes.

Discussão

A discussão deste artigo será dividida em quatro seções: inquéritos alimentares, análise de agrupamento, análise fatorial por componentes principais e nomenclatura de padrões alimentares. Nelas se discutirá os aspectos metodológicos dos artigos selecionados e serão fornecidas orientações sobre a metodologia para identificação de padrões alimentares. O método de Regressão por Redução de Posto não será discutido, uma vez que não foi contemplado em nenhum dos estudos desta revisão. Entretanto, vale destacar a crescente utilização desse tipo de análise para a extração de padrões alimentares.

Inquéritos alimentares

Na escolha do método de avaliação do consumo alimentar deve-se considerar aspectos como o objetivo do estudo, idade, sexo e concordância dos métodos quanto à sua aplicação e acurácia²⁰. No caso da população infantil, recomenda-se que o inquérito alimentar seja respondido pelos pais ou responsáveis, a fim de garantir maior fidedignidade e precisão das estimativas.

Na identificação de padrão alimentar os tipos de inquéritos alimentares mais utilizados são o QFA, Registro Alimentar (RA) e R24hs. O QFA é um método prático, informativo e de baixo custo, muito utilizado em estudos que avaliam a associação entre dieta e doenças, sobretudo as crônicas não transmissíveis. Nesse tipo de inquérito tem-se uma lista de alimentos predefinida, seguida da frequência de consumo, número de vezes que o alimento é consumido por dia, semana, mês ou ano^{21,22}.

O registro alimentar é um método prospectivo, no qual o indivíduo deve anotar todos os alimentos e bebidas consumidos ao longo de um ou mais dias, discriminando detalhadamente o tipo do alimento, quantidade, forma de preparo, adição de açúcar e sal²³. Semelhantemente ao RA, no R24h também tem-se uma descrição detalhada dos alimentos e bebidas consumidos, entretanto, é um método retrospectivo cuja referência temporal são as últimas 24 horas ou o dia anterior. Nesse método, as informações são preenchidas por um entrevistador e não pelo próprio indivíduo^{20,21}.

Após a escolha do inquérito alimentar mais apropriado para o estudo que se está conduzindo e a obtenção dos dados dietéticos, deve-se proceder a digitação dos dados. Para a identificação de padrões alimentares utilizando a ACP de dados alimentares obtidos pelo QFA, estes devem

ser digitados, preferencialmente, em uma escala *likert* de apenas uma unidade temporal, ou seja, todos com consumo diário, semanal, mensal ou anual.

Os alimentos do QFA podem ainda ser organizados em grupos²⁴. Estes devem ser formados com base nas semelhanças entre os alimentos, no que se refere ao seu uso culinário habitual, conteúdo nutritivo, composição botânica, análise de correlação de Pearson entre os alimentos, entre outros. Nas pesquisas desenvolvidas por Nobre et al.¹², Silva et al.¹⁶ e Souza et al.¹⁷, os autores fizeram o agrupamento dos alimentos de acordo com a semelhança de conteúdo nutritivo para análise dos dados. No estudo de D'Innocenzo et al.¹⁵, além das características nutricionais foram utilizadas as correlações obtidas através de análise fatorial preliminar.

Na identificação de padrões alimentares utilizando a frequência de consumo dos alimentos do QFA deve-se utilizar a escala *likert* presente neste. Souza et al.¹⁷ efetuaram a transformação das informações do QFA para frequências mensais, multiplicando o consumo diário por 30,42, o semanal por 4, e dividindo o consumo anual por 12. Essas frequências mensais representavam uma mesma unidade temporal e foram utilizadas na análise multivariada.

Quando o pesquisador transforma o QFA em grupos de alimentos deve-se obter uma medida-resumo da frequência de consumo. Entre as pesquisas citadas neste estudo e que utilizaram esse método, a medida-resumo foi obtida segundo metodologia proposta por Neumann et al.²⁵, a qual é obtida pela fórmula: medida-resumo = \sum da frequência de consumo dos alimentos contidos no grupo alimentar/nº de alimentos do grupo x frequência máxima de consumo no QFA utilizado. Posteriormente, as frequências de consumo para cada grupo de alimento são codificadas para obtenção da medida-resumo.

A soma das frequências individuais codificadas, correspondentes aos alimentos consumidos em cada grupo, é utilizada como numerador no cálculo da medida-resumo. O denominador dessa medida corresponde ao número máximo de alimentos que o indivíduo pode consumir em cada grupo, multiplicado pela frequência máxima de consumo no questionário utilizado²⁵. Assim, para cada indivíduo da amostra é calculada uma medida-resumo, a qual é utilizada na análise multivariada escolhida para identificação dos padrões alimentares.

Vale destacar que, segundo Olinto⁸, para análise, identificação e interpretação dos padrões

alimentares o ideal é que cada alimento seja apresentado na sua forma isolada no QFA e não agrupado. E se o QFA utilizado for quantitativo será necessário converter as porções dos alimentos para gramas ou mililitros, sendo que todos os itens alimentares devem estar com a mesma unidade de medida.

Quando o inquérito alimentar escolhido é o RA ou R24h, os alimentos registrados ou relatados pelos indivíduos estudados podem ser alocados em grupos padronizados, com base nas mesmas semelhanças mencionadas anteriormente para o QFA, ou utilizados isoladamente. Após o agrupamento dos alimentos, obtém-se a ingestão média absoluta em gramas/dia (g/d) para cada grupo de alimentos formado. Na fase de derivação dos padrões alimentares esses são os valores utilizados nas análises²⁶⁻²⁸.

No estudo de Gatica et al.¹⁹, os autores elaboraram uma lista de alimentos frequentemente consumidos em sua faixa-etária de estudo. Posteriormente, questionaram as mães sobre o consumo desses alimentos pelas crianças nas últimas 24 horas. A lista para crianças de 12 e 24 meses foi composta por 16 alimentos e para as de 48 meses existiam 17. A identificação dos padrões alimentares foi realizada utilizando estes alimentos isoladamente na análise de componentes principais, isto é, sem agrupá-los.

Análise de agrupamento (*cluster analysis*)

A análise de agrupamento é outra técnica utilizada para identificar padrões alimentares. A principal finalidade deste método é o agrupamento de indivíduos de acordo com regularidades no seu consumo de alimentos²⁹. Os grupos (ou *clusters*) são formados a partir de características comuns de consumo de alimentos, de forma que indivíduos com ingestão alimentar semelhante irão compor o mesmo grupo⁶. Portanto, espera-se que os indivíduos pertencentes a um determinado *cluster* possuam alta homogeneidade entre si e alta heterogeneidade quando comparados aos demais grupos formados^{30,31}.

Nesse tipo de análise obtém-se padrões alimentares mutuamente exclusivos, uma vez que cada indivíduo somente pode pertencer a um determinado grupo³². Essa característica permite identificar indivíduos com elevado risco para desenvolver doenças crônicas, por exemplo³³.

A análise de agrupamento é recomendada para situações em que a amostra não é homogênea, quando as propriedades psicométricas da

análise fatorial não forem atendidas ou quando deseja-se manter todos os itens alimentares propostos no instrumento⁸.

Existem dois tipos de conjuntos de técnicas (algoritmos) de classificação nesta análise: os métodos hierárquicos e os não hierárquicos ou de particionamento. Nos métodos hierárquicos, basicamente tem-se sucessivas aglomerações ou divisões de elementos, unindo grupos pequenos em maiores ou dividindo-se grandes grupos em menores. Por outro lado, nos métodos não hierárquicos o pesquisador define um número de grupos para a aglomeração dos elementos e estes são alocados de acordo com a medida de similaridade adotada³⁴.

Três aspectos básicos devem ser observados na condução da análise de agrupamentos: medição da similaridade, formação de agrupamentos e determinação do número destes na solução final. A similaridade refere-se às semelhanças entre os indivíduos do estudo quanto a seu consumo de alimentos e deve ser medida para que se possa proceder com a formação dos grupos³⁵.

O tipo de medida de similaridade mais utilizado em análises de agrupamento são as de distância (por exemplo, métrica euclidiana). Estas medidas baseiam-se na dissimilaridade entre os indivíduos, em que valores mais distantes apresentam menor similaridade¹⁴. As medidas de distância são muito sensíveis a diferentes escalas ou magnitudes entre as variáveis, por isso, nos casos em que as variáveis apresentarem maior dispersão é importante realizar a padronização dos dados antes de calcular a medida de similaridade³³. Geralmente, a padronização é feita por meio da conversão das variáveis em *escore-z*.

Outra medida muito importante é a exclusão de *outliers*, pois a análise de agrupamentos é muito sensível à presença de valores aberrantes. Entretanto, é necessário que o pesquisador tenha bom senso para não excluir valores que sejam representativos do consumo da sua amostra³⁶.

Para a formação dos grupos os métodos de agrupamento mais utilizados em epidemiologia nutricional são a variância mínima de Ward e K-média³³. A variância mínima de Ward é um método hierárquico que utiliza a análise de variância em cada etapa de formação de grupos, de modo que os indivíduos que promovem menor variância intragrupo são agrupados no mesmo *cluster*. A técnica K-médias é um método não hierárquico que objetiva-se particionar os indivíduos em um número predefinido de *k* grupos. Nesse método, cada indivíduo é classificado de

acordo com a sua distância do centro do grupo (centroide), o qual é definido por meio do cálculo da média de cada grupo³⁵.

Na escolha do número de agrupamentos na solução final, deve-se monitorar a similaridade intragrupos e a heterogeneidade entregrupos a medida que o número dos grupos testados diminui. Portanto, deve-se optar por um número de grupos que garanta a máxima homogeneidade intragrupo e heterogeneidade entregrupos^{14,30}.

A análise de agrupamento possui a importante limitação de ser um método que exige muitas decisões subjetivas do pesquisador, tais como: decisão do número final de grupos, características a serem usadas para formar os grupos, métodos de combinação de agrupamentos, entre outros⁶.

Dentre as publicações citados na presente pesquisa, Gama et al.¹⁸ foram os únicos que utilizaram esta análise para identificação de padrões alimentares. Foi utilizando a métrica euclidiana como medida de similaridade. O critério para agrupamento baseou-se na minimização da soma das distâncias entre cada observação e o centroide do grupo. Para determinar o número de agrupamentos escolhidos os autores realizaram a comparação do ajuste obtido para 4 a 8 grupos, encontrando o melhor modelo para 6.

Análise fatorial por componentes principais

A análise de componentes principais tem sido a técnica mais utilizada na derivação de padrões alimentares em epidemiologia nutricional. O principal objetivo dessa análise é a redução do número de informações contidas nas variáveis originais para um número menor destas, porém com perda mínima de informação³⁷. Nesta análise os alimentos são agregados em conjuntos menores de variáveis (fatores ou componentes) os quais são o mais representativo possível das originais. Para tal, identificam-se combinações lineares entre os alimentos ou grupos destes que explicam a maior parte da variação na dieta dos indivíduos¹¹.

No entanto, como já citado anteriormente, quando utilizado o QFA para identificação de padrões alimentares por meio desta análise, é necessário que o mesmo tenha número de alimentos adequado ao tamanho amostral, ou seja, para cada um é preciso de pelo menos cinco voluntários¹⁴.

Além disso, para esta análise deve-se obedecer a alguns quesitos, e o pesquisador precisa estar atento ao atendimento de todos eles. Ou seja,

no teste de adequabilidade dos dados à realização da análise fatorial o valor de KMO deve ser superior a 0,6; no teste de esfericidade de Bartlett, o valor de p deve ser menor ou igual a 0,05 e deve-se também utilizar os autovalores e o percentual da variância acumulada para definir o número de fatores (padrões alimentares) a serem retidos, sendo que um critério gráfico também pode ser empregado.

Quando a análise de autovalores é o método utilizado para definição do número de fatores (padrões alimentares) a serem retidos, geralmente, utiliza-se o critério Kaiser, o qual considera como significantes autovalores superiores a 1. Um autovalor representa a quantidade da variância que pode ser explicada por um fator.

Em relação ao critério gráfico, o de Cattel ou *scree plot* é, geralmente, o mais utilizado. Com base nesse critério faz-se um gráfico dos autovalores em função do número de fatores em sua ordem de extração³⁸. O primeiro fator extraído é o que mais explica a variância na dieta do indivíduo. O segundo é independente do primeiro, e é o que mais explica a variância remanescente e assim sucessivamente. Dessa forma, de acordo com o critério do teste *scree plot*, o ponto no qual o gráfico apresenta inclinação é considerado indicativo do número máximo de fatores a serem extraídos. Além desses critérios, é importante que o pesquisador pondere a respeito do percentual desejado de variância explicada^{14,39}.

Para simplificar a interpretação das medidas de correlação (cargas fatoriais) entre os fatores, deve-se utilizar métodos de rotação ortogonal, dentre os quais o mais comumente disponível nos programas computacionais é o Varimax. O objetivo desse método é fornecer uma estrutura mais simples, por meio do aumento da variância nos fatores mais elevados e da diminuição nos menores. Esse procedimento facilita a interpretação dos fatores obtidos pela análise de componentes principais^{8,13,40}.

Após essa rotação, deve-se manter na matriz os alimentos cujas cargas fatoriais sejam superiores a 0,3, pois neste nível considera-se que seja atendido o mínimo necessário para a interpretação da estrutura¹⁴. Os valores inferiores a esse nível devem ser desconsiderados. Finalmente, deve-se realizar a avaliação da consistência interna dos fatores, e o teste do coeficiente alfa Cronbach é utilizado com esta finalidade, sendo que valores maiores ou iguais a 0,60 indicam boa consistência do fator extraído.

Nomenclatura dos padrões alimentares

Finalmente, após identificação dos padrões alimentares estes devem ser nomeados. Na maioria dos estudos, os padrões alimentares são rotulados de acordo com o item de maior carga de saturação ou com a composição nutricional dos alimentos do fator. Para Olinto⁸, os nomes elencados aos padrões alimentares devem buscar uma explicação teórica para os agrupamentos formados. No caso de padrões alimentares obtidos através de análise de agrupamento, nomes também podem ser atribuídos de acordo com as características do consumo de alimentos dentro de cada grupo formado.

Alguns estudos optam por não atribuir nomes aos padrões identificados, apenas numerando-os. Isso reduz a subjetividade envolvida na atribuição de uma nomenclatura aos padrões alimentares, entretanto, dificulta a comparabilidade com outros estudos. D’Innocenzo et al.¹⁵, optaram por não atribuir nomes aos padrões encontrados, identificando-os apenas como “padrão 1”, “padrão 2”, “padrão 3” e “padrão 4”, semelhante a Matos et al.³.

Conclusão

São poucos os estudos brasileiros que identificaram os padrões alimentares de crianças utilizando metodologias de abordagem *a posteriori*. A análise fatorial por componentes principais foi

a técnica mais usada para a extração dos padrões alimentares, conforme observado mundialmente também. A nomenclatura de padrões alimentares é uma conduta que divide os autores, pois alguns têm preferido não os nomear devido à subjetividade que pode ser agregada com essa atitude.

A análise do padrão alimentar por meio de técnicas estatísticas exploratórias representa um passo importante na epidemiologia nutricional, pois são análises robustas e que permitem conhecer o consumo alimentar da população estudada de forma mais ampla. Ademais, possibilitam a obtenção de informações mais completas a respeito da associação entre a dieta e diversas doenças.

Ressalta-se que as metodologias de avaliação do padrão alimentar possuem limitações quanto à subjetividade das decisões a serem tomadas pelo pesquisador. Dentre elas podem ser citadas: a escolha dos critérios para agrupar os alimentos, o número de fatores a serem retidos e o tipo de rotação utilizado na análise fatorial, as características para a formação de grupos e o número de grupos formados na análise de agrupamento, o nome dado para os padrões alimentares encontrados, entre outros.

Assim, em pesquisas nas quais sejam identificados padrões alimentares é importante que o estudo seja bem delineado, garantindo a escolha do inquérito alimentar mais apropriado, com tamanho amostral suficiente e a utilização dos testes estatísticos que melhor orientem as decisões inerentes ao pesquisador.

Colaboradores

CA Carvalho e PCA Fonseca participaram das etapas de busca e análise dos artigos, redação do artigo e sua revisão crítica. LN Nobre, SE Priore e SCC Franceschini participaram da redação do artigo, revisão crítica e aprovação da versão a ser publicada.

Referências

- Santos Filha EO, Araújo JS, Barbosa JS, Gaujac DP, Santos CFS, Silva DG. Consumo dos grupos alimentares em crianças usuárias da rede pública de saúde do município de Aracaju, Sergipe. *Rev Paul Pediatr* 2012; 30(4):529-536.
- Kristiansen AL, Lande B, Sexton JA, Andersen LF. Dietary patterns among Norwegian 2-year-olds in 1999 and in 2007 and associations with child and parent characteristics. *Br J Nutr* 2013; 110(1):135-144.
- Matos SMA, Barreto ML, Rodrigues LC, Oliveira VA, Oliveira LPM, D'Innocenzo S, Teles CAS, Pereira SRS, Prado MS, Assis AMO. Padrões alimentares de crianças menores de cinco anos de idade residentes na capital e em municípios da Bahia, Brasil, 1996 e 1999/2000. *Cad Saude Publica* 2014; 30(1):44-54.
- Cutler GJ, Flood A, Hannan P, Neumark-Sztainer D. Multiple sociodemographic and socioenvironmental characteristics are correlated with major patterns of dietary intake in adolescents. *J Am Diet Assoc* 2011; 111(2):230-240.
- Cribb V, Emmett P, Northstone K. Dietary patterns throughout childhood and associations with nutrient intakes. *Public Health Nutr* 2013; 16(10):1801-1809.
- Devlin UM, McNulty BA, Nugent AP, Gibney MJ. The use of cluster analysis to derive dietary patterns: methodological considerations, reproducibility, validity and the effect of energy mis-reporting. *Proc Nutr Soc* 2012; 71(4):599-609.
- Azevedo ECC, Diniz AS, Monteiro JS, Cabral PC. Padrão alimentar de risco para as doenças crônicas não transmissíveis e sua associação com a gordura corporal – uma revisão sistemática. *Cien Saude Colet* 2014; 19(5):1447-1458.
- Olinto MT. Padrões alimentares: análise dos componentes principais. In: Kac G, Sichieri R, Gigante DP, organizadores. *Epidemiologia nutricional*. Rio de Janeiro: Editora Fiocruz/Editora Atheneu; 2007. p. 213-226.
- Kastorini CM, Papadakis G, Milionis HJ, Kalantzi K, Puddu PE, Nikolaou V, Vemmos KN, Goudevenos JA, Panagiotakos DB. Comparative analysis of a-priori and a-posteriori dietary patterns using state-of-the-art classification algorithms: a case/control study. *Artif Intell Med* 2013; 59(3):175-183.
- Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002; 13(1):3-9.
- Román-Viñas B, Barba LR, Ngo J, Martínez-González MA, Wijnhoven TMA, Serra-Majem L. Validity of dietary patterns to assess nutrient intake adequacy. *Br J Nutr* 2009; 101(Supl. 2):S12-S20.
- Nobre LN, Lamounier JA, Franceschini SCC. Preschool children dietary patterns and associated factors. *J Pediatr* 2012; 88(2):129-136.
- Panagiotakos DB, Pitsavos C, Stefanadis C. Alpha-priori and alpha-posterior dietary pattern analyses have similar estimating and discriminating ability in predicting 5-Y incidence of cardiovascular disease: methodological issues in nutrition assessment. *J Food Sci* 2009; 74(7):H218-224.
- Hair Jr. JF, Anderson RE, Tatham RL, Black WC. *Análise multivariada de dados*. 6ª ed. Porto Alegre: Editora Bookman; 2009.
- D'Innocenzo S, Marchioni DML, Prado MS, Matos SMA, Pereira SRS, Barros AP, Sampaio LR, Assis AMO, Rodrigues LC, Barreto ML. Condições socioeconômicas e padrões alimentares de crianças de 4 a 11 anos: estudo SCAALA - Salvador/Bahia. *Rev Bras Saude Matern Infant* 2011; 11(1):41-49.
- Silva RCR, Assis AMO, Szarfarc SC, Pinto EJ, Costa LCC, Rodrigues LC. Iniquidades socioeconômicas na conformação dos padrões alimentares de crianças e adolescentes. *Rev Nutr* 2012; 25(4):451-461.
- Souza RLV, Madruga SW, Gigante DP, Santos IS, Barros AJD, Assunção MCF. Padrões alimentares e fatores associados entre crianças de um a seis anos de um município do Sul do Brasil. *Cad Saude Publica* 2013; 29(12):2416-2426.
- Gama SR, Carvalho MS, Chaves CRMM. Prevalência em crianças de fatores de risco para as doenças cardiovasculares. *Cad Saude Publica* 2007; 23(9):2239-2245.
- Gatica G, Barros AJ, Madruga S, Matijasevich A, Santos IS. Food intake profiles of children aged 12, 24 and 48 months from the 2004 Pelotas (Brazil) birth cohort: an exploratory analysis using principal components. *Int J Behav Nutr Phys Act* 2012; 9:43.
- Fisberg RM, Slater B, Marchioni DML, Martini LA. *Inquéritos alimentares: métodos e bases científicas*. São Paulo: Manole; 2005.
- Fisberg RM, Marchioni DML, Colucci ACA. Avaliação do consumo alimentar e da ingestão de nutrientes na prática clínica. *Arq Bras Endocrinol Metab* 2009; 53(5):617-624.
- Costa AGV, Priore SE, Sabarense CM, Franceschini SCC. Questionário de frequência de consumo alimentar e recordatório de 24 horas: aspectos metodológicos para avaliação da ingestão de lipídeos. *Rev Nutr* 2006; 19(5):631-641.
- Holanda LB, Barros Filho AA. Métodos aplicados em inquéritos alimentares. *Rev. Paul. Pediatría* 2006; 24(1):62-70.
- Kieft-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VW, Hofman A, Raat H, Moll HA. Socio-demographic and lifestyle determinants of 'Western-like' and 'Health conscious' dietary patterns in toddlers. *Br J Nutr* 2013; 109(1):137-147.
- Neumann AI, Martins IS, Marcopito LF, Araujo EA. Dietary patterns associated with risk factors for cardiovascular disease in a Brazilian city. *Rev Panam Salud Publica* 2007; 22(5):329-339.
- Mikkilä V, Rasanen L, Raitakari OT, Pietinen P, Viikari J. Consistent dietary patterns identified from childhood to adulthood: The Cardiovascular Risk in Young Finns Study. *Br J Nutr* 2005; 93(6):923-931.
- Manios Y, Kourlaba G, Grammatikaki E, Androustos O, Ioannou E, Roma-Giannikou E. Comparison of two methods for identifying dietary patterns associated with obesity in preschool children: the GENESIS study. *Eur J Clin Nutr* 2010; 64(12):1407-1414.
- Grieger JA, Scott J, Cobiac L. Dietary patterns and breast-feeding in Australian children. *Public Health Nutr* 2011; 14(11):1939-1947.
- Cunha DB, Almeida RMVR, Pereira RA. A comparison of three statistical methods applied in the identification of eating patterns. *Cad Saude Publica* 2010; 26(11):2138-2148.

30. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* 2004; 62(5):177-203.
31. Hearty AP, Gibney MJ. Comparison of cluster and principal component analysis techniques to derive dietary patterns in Irish adults. *Br J Nutr* 2009; 101(4):598-608.
32. Smith ADAC, Emmett PM, Newby PK, Northstone K. A comparison of dietary patterns derived by cluster and principal components analysis in a UK cohort of children. *Eur J Nutr* 2011; 65(10):1102-1109.
33. Siou GL, Yasui Y, Csizmadi I, McGregor SE, Robson PJ. Exploring statistical approaches to diminish subjectivity of cluster analysis to derive dietary patterns: The Tomorrow Project. *Am J Epidemiol* 2011; 173(8):956-967.
34. Ahlert JT. *Associação entre padrões alimentares e transtornos mentais em gestantes do sul do Brasil* [dissertação]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2013.
35. Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. 6^{ed}. New Jersey: Pearson Prentice Hall; 2007.
36. Rodríguez-Ramírez S, Mundo-Rosas V, García-Guerra A, Shamah-Levy T. Dietary patterns are associated with overweight and obesity in Mexican school-age children. *Arch Latinoam Nutr* 2011; 61(3):270-278.
37. Panagiotakos D. α -priori versus α -posterior methods in dietary pattern analysis: a review in nutrition epidemiology. *Nutrition Bulletin* 2008; 33(4):311-315.
38. Raïche G, Walls TA, Magis D, Riopel M, Blais J. Non-Graphical Solutions for Cattell's Scree Test. *Methodology* 2013; 9(1):23-29.
39. Ledesma RD, Valero-Mora P. Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. *Practical assessment, research & evaluation* 2007; 12(2):1-11.
40. Marchioni DML, Latorre MRDO, Eluf-Neto J, Wünsch-Filho V, Fisberg RM. Identification of dietary patterns using factor analysis in an epidemiological study in Sao Paulo. *São Paulo Med J* 2005; 123(3):124-127.

Artigo apresentado em 03/09/2014

Aprovado em 01/12/2014

Versão final apresentada em 03/12/2014