

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE MEDICINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA**



**TESE DE DOUTORADO**

**Influência da suposição sobre a frequência do consumo, diário ou episódico, e do uso de covariáveis na distribuição do consumo alimentar usual nos métodos MSM, NCI e SPADE**

Greice Helen da Costa Laureano

Orientador: Prof. Dra. Suzi Alves Camey

Co-orientador: Prof. Dra. Vanessa Bielefeldt Leotti

Porto Alegre, Setembro de 2018.

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE MEDICINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA**

**TESE DE DOUTORADO**

**Influência da suposição sobre a frequência do consumo, diário ou episódico, e do uso de covariáveis na distribuição do consumo alimentar usual nos métodos MSM, NCI e SPADE**

Greice Helen da Costa Laureano

**Orientador: Prof. Dra. Suzi Alves Camey**

**Co-orientador: Prof. Dra. Vanessa Bielefeldt Leotti**

A apresentação desta tese é exigência do Programa de Pós-graduação em Epidemiologia, da Universidade Federal do Rio Grande do Sul, para obtenção do título de Doutor.

Porto Alegre, Brasil  
2018

## CIP - Catalogação na Publicação

Laureano, Greice Helen da Costa

Influência da suposição sobre a frequência do consumo, diário ou episódico, e do uso de covariáveis na distribuição do consumo alimentar usual nos métodos MSM, NCI e SPADE / Greice Helen da Costa Laureano. -- 2019.

117 f.

Orientadora: Suzi Alves Camey.

Coorientadora: Vanessa Bielefeldt Leotti.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, Programa de Pós-Graduação em Epidemiologia, Porto Alegre, BR-RS, 2019.

1. Estimação de consumo alimentar usual. 2. Comparação de métodos de estimação. 3. NCI. 4. SPADE. 5. MSM. I. Camey, Suzi Alves, orient. II. Leotti, Vanessa Bielefeldt, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

## **BANCA EXAMINADORA**

Prof. Dr. Alvaro Vigo, Programa de Pós-graduação em Epidemiologia, UFRGS.

Prof. Dr. Eliseu Verly Junior, Departamento de Epidemiologia do Instituto de Medicina Social  
da UERJ.

Prof. Dra. Sandra Patrícia Crispim, Departamento de Nutrição, UFPR.

*Dedico essa tese aos meus pais Sirlei Laureano e Santo Laureano, aos meus amores Rodrigo Coster, Zelda e Appa que me acompanharam com compreensão e carinho nesta jornada.*

## **AGRADECIMENTOS**

Agradeço a minha família, aos meus amigos – em especial à Josiene S. Niesciur e ao Ueldo M. P. Rodrigues – e as minhas orientadoras pelo apoio e incentivo em todos esses anos.

## SUMÁRIO

ABREVIATURAS E SIGLAS	9
RESUMO	10
1. APRESENTAÇÃO	14
2. INTRODUÇÃO	15
3. REVISÃO DE LITERATURA	17
3.1. ESTIMAÇÃO DO CONSUMO ALIMENTAR	17
3.1.1. TIPOS DE ERROS	17
3.1.2. DIÁRIO VS EPISÓDICO	21
3.1.3. USO DE COVARIÁVEIS	22
3.1.4. MÉTODOS DE ESTIMAÇÃO DO CONSUMO ALIMENTAR USUAL	25
3.1.4.1. NATIONAL CANCER INSTITUTE (NCI)	29
3.1.4.2. MULTIPLE SOURCE METHOD (MSM)	35
3.1.4.3. STATISTICAL PROGRAM TO ASSESS DIETARY EXPOSURE (SPADE)	40
3.2. JUSTIFICATIVA DO PRESENTE TRABALHO	44
4. OBJETIVOS	45
4.1. OBJETIVO GERAL	45
4.2. OBJETIVOS ESPECÍFICOS	45
5. ARTIGO 1	50
6. ARTIGO 2	67
7. CONCLUSÕES E CONSIDERAÇÕES FINAIS	80
APÊNDICE A: Artigo 1 - Tabela S1: Valores verdadeiros para alimentos em nutrientes com base na média dos 20 R24h.	83
APÊNDICE B: Artigo 1 - Figura S1: Componentes alimentares excluídos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.	85
APÊNDICE C: Artigo 1 - Figura S2: Componentes alimentares mantidos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.	86
APÊNDICE D: Artigo 1 - Figura S3: Componentes alimentares mantidos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.	87
APÊNDICE E: Artigo 1 - Tabela S2: Vícios relativos para a suposição de consumo episódico.	88
APÊNDICE F: Artigo 1 - Tabela S3: Vícios relativos para a suposição de consumo diário.	93

APÊNDICE G: Artigo 1 - Tabela S4: Vícios relativos para a suposição de consumo episódico sem uso de QFA.	97
APÊNDICE I: Artigo 1 - Quadro S1: Detalhamento dos modelos com suposição de consumo diário.	100
APÊNDICE J: Artigo 1 - Quadro S2: Detalhamento dos modelos com suposição de consumo episódico.	102
APÊNDICE K: Artigo 2 - Tabela S1: Vícios relativos com e sem a utilização das covariáveis.	104
APÊNDICE L: Artigo 2 - Tabela S2: Diferença do vício relativo (DVR) com e sem covariáveis para componentes alimentares de frequência de consumo diário	113
APÊNDICE M: Artigo 2 - Figura S1: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo diário, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].	115
APÊNDICE N: Artigo 2 - Figura S2: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo episódico de nível intermediário, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].	116
APÊNDICE O: Artigo 2 - Figura S3: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo episódico de nível baixo, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].	117



## ABREVIATURAS E SIGLAS

AGE MODE	<i>Age-dependent dietary assessment model</i>
AIC	Critério de informação de Akaike
ANOVA	<i>Analysis of Variance</i>
BBN	<i>BetaBinomial-Normal Model</i>
DVR	Diferença entre os vícios relativos
EFCOVAL	<i>European Food Consumption Validation</i>
FFQ	<i>Food Frequency Questionnaire</i>
GEE	<i>Generalized Estimating Equations</i>
ISU	<i>Iowa State University</i>
LNN	<i>LogisticNormal - Normal model with correlation</i>
LNN0	<i>LogisticNormal - Normal model without correlation</i>
MM	Modelo Misto
MMNL	Modelo misto não linear
MSM	<i>Multiple Source Method</i>
NCI	<i>National Cancer Institute</i>
QFA	Questionário de frequência alimentar
R24h	Recordatório alimentar de 24 horas
SAS	<i>Statistical Analysis System</i>
SPADE	<i>Statistical Program to Assess Dietary Exposure</i>
VR	Vício relativo

## RESUMO

Este trabalho foi realizado com o intuito de preencher lacunas da literatura de epidemiologia nutricional no que se refere a avaliação do impacto sobre a utilização das suposições de frequência do consumo alimentar, diário ou episódico, e o uso de covariáveis (Questionário de Frequência Alimentar (QFA), sexo e idade) na estimação da distribuição do consumo alimentar usual para populações dos métodos *National Cancer Institute* (NCI), *Multiple Source Method* (MSM) e o *Statistical Program to Assess Dietary Exposure* (SPADE). Para tanto, foi elaborado um artigo que comparou o impacto da escolha da suposição de consumo diário ou episódico, assim como a influência da utilização do QFA na acurácia dos métodos. Ainda, no segundo artigo, foi verificado se o uso das covariáveis sexo e idade influenciaram a acurácia dos métodos. Em ambos artigos o vício relativo foi a medida de acurácia empregada e os dados de consumo usados foram os de uma pesquisa que coletou 20 Recordatórios Alimentares de 24 horas (R24h) de 302 pessoas residentes na cidade do Rio de Janeiro – Brasil.

Os resultados mostraram que:

- Para componentes alimentares de frequência de consumo diária (até 0,1% de R24h iguais a zero) as estimativas mais acuradas foram as com uso da suposição de consumo diário em todos os métodos.
- Em componentes alimentares com frequência de consumo episódica de nível intermediário (entre 20% e 75% de R24h iguais a zero) as estimativas mais acuradas ocorreram com o MSM fazendo o uso da suposição de consumo episódico.
- Já em componentes alimentares com frequência de consumo episódica de nível baixo (no mínimo 76% de consumos iguais a zero nos R24h), as estimativas mais acuradas utilizando a suposição de consumo episódico foram:
  - No método MSM na estimação dos percentis menores ou iguais a 25 e nos métodos SPADE e NCI na estimação da média e os percentis maiores ou iguais a 50.
- Em relação ao uso do QFA no ajuste dos métodos:
  - Discreto aumento de acurácia dos métodos MSM (principalmente para a média e percentis acima de 75) e SPADE (sobretudo para percentis abaixo de 25).
- Quanto a utilização das covariáveis sexo e idade no ajuste dos métodos:
  - Os métodos não foram impactados para componentes alimentares de frequência de consumo diário;
  - Para as frequências de consumo episódica de nível intermediário:
    - O NCI teve resultados, em geral, piores quando comparado aos sem ajuste;
    - Os métodos MSM e o SPADE não foram impactados.
  - Para as frequências de consumo episódica de nível baixo:
    - O NCI foi impactado, em geral, com piora nos resultados;
    - O MSM sofreu impacto, em geral, com melhora nos resultados;
    - O SPADE não foi impactado.

Com esses achados, aconselha-se:

- Utilizar em todos os métodos a suposição de frequência de consumo diário para componentes alimentares de frequência de consumo diária.
- Utilizar a suposição de frequência de consumo episódico em componentes de frequência de consumo episódico de nível intermediário:
  - Emprego do método MSM na estimação dos percentis 5 e 10;
  - Uso do NCI, MSM ou SPADE na estimação dos demais parâmetros da distribuição.
- Utilizar da suposição de frequência de consumo episódico em componentes de frequência de consumo episódico de nível baixo:
  - Emprego do método MSM na estimação dos percentis menores ou iguais a 25;
  - Uso dos métodos SPADE e NCI para estimativa da média e dos percentis acima de 25.
- Utilizar o QFA no SPADE e MSM quando se desejar o máximo de acurácia nas estimações e não houver restrições orçamentárias.
- Fazer o uso de sexo e idade como covariável no método MSM quando se tratarem de componentes alimentares com frequência de consumo episódico de nível baixo.

**Palavras-chave:** NCI. MSM. SPADE. Consumo episódico. Consumo diário. QFA. R24h.

## ABSTRACT

This work was carried out to fill gaps in nutritional epidemiology literature regarding the evaluation of the impact on the use of food frequency assumptions, daily or episodic, and the use of covariates (Food Frequency Questionnaire (FFQ), sex and age) in the estimation of the usual food intake distribution for the methods National Cancer Institute (NCI), Multiple Source Method (MSM) and the Statistical Program to Assess Dietary Exposure (SPADE). To do so, we developed one article with the impact comparison of the choice of the daily or episodic intake assumption, as well as the influence of the QFA use on the accuracy of the methods. On the other hand, in the second article, was verified if the sex and age used as covariable has influenced in the accuracy of the methods. In both articles, the relative bias was the measure of accuracy used and the intake data used data from a survey that collected 20 24-hour dietary recall (24HR) from 302 people living in the city of Rio de Janeiro - Brazil.

The results showed that:

- For daily intake (up to 0.1% 24HR equal to zero) the most accurate estimates was using the daily intake assumption in all methods.
- In food components with intermediate level of episodic intake (between 20% and 75% of 24HR equal to zero), the most accurate estimates occurred with the MSM using the of episodic intake supposition.
- In food components with low level of episodic intake(at least 76% of intakes equal to zero in 24HR) the most accurate estimates using the episodic intake assumption was:
  - MSM method in the estimation of percentiles smaller or equal to 25 and in the SPADE and NCI methods in the estimation of average and percentiles greater or equal to 50.
- About the usage of FFQ:
  - Slight increase in accuracy of MSM (mainly for the average and percentiles above 75) and SPADE (especially for percentiles below 25).
- About the usage of sex and age as covariates:
  - The methods were not impacted for food components of high intake frequency;
  - For intermediate level of episodic intake:
    - The NCI had, in general, worse results when compared to those without adjustment;
    - Both MSM and SPADE were not impacted.
  - For low level of episodic intake:
    - The NCI was generally impacted by worsening results;
    - The MSM had an impact by improving results;
    - SPADE was not impacted.

With these findings, we advised that:

- Use daily assumption in the estimation for food components of high intake frequency, regardless of method;
- Use episodic assumption in intermediate intake frequency components:
  - MSM method in the estimation of the 5th and 10th percentiles;

- NCI, MSM or SPADE in the estimation of the other parameters of the distribution.
- Use episodic intake frequency assumption in low intake frequency components:
  - MSM in the estimation of percentiles less than or equal to 25;
  - SPADE and NCI to estimate average and percentiles above 25.
- Use the FFQ in SPADE and MSM when you want maximum accuracy in estimates and there are no budget constraints.
- Make use of sex and age as covariate in the MSM method when dealing with food components with low intake frequency.

**Keywords:** NCI. MSM. SPADE. Episodic assumption. Daily assumption. FFQ. 24HR.

## 1. APRESENTAÇÃO

Este trabalho consiste na tese de doutorado intitulada “**Influência da suposição sobre a frequência do consumo, diário ou episódico, e do uso de covariáveis na distribuição do consumo alimentar usual nos métodos MSM, NCI e SPADE**”, apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em 28 de Setembro de 2018. O trabalho é apresentado em três partes, na ordem que segue:

- I. Introdução, Revisão da Literatura e Objetivos;
- II. Artigo(s);
- III. Conclusões e Considerações Finais.

Documentos de apoio estão dispostos nos apêndices.

## 2. INTRODUÇÃO

A distribuição do consumo de componentes alimentares tem sido objeto de diversos estudos (1–3). O conhecimento da distribuição é importante para identificar o comportamento alimentar da população em análise e com isso, por exemplo, avaliar se ela tem consumo adequado dos componentes alimentares (4,5).

Nesse contexto, tem-se que para estimar essas distribuições foram desenvolvidos diversos métodos com diferentes tipos de modelagem. Alguns estão implementados em aplicativos computacionais, tais como: STEM.I (6,7) e STEM.II (7); Wallace (8); ISU (9–11); Gay (12); S-Nusser (13); MSM (14,15); SPADE (16–18); NCI (19–22) e AGE MODE (23). Em geral, os métodos diferem de acordo com a frequência de consumo do componente alimentar cuja distribuição pretende-se estimar e com o uso ou não de covariáveis na estimação.

Com essa variedade de métodos disponíveis, há necessidade de pesquisas que os comparem, para que se possa determinar se eles são equivalentes ou quais métodos são mais precisos ou acurados. Vários estudos dessa natureza já foram conduzidos (9,10,13,16,24–26), mas ainda há poucos estudos explorando a influência nos métodos quanto a escolha da suposição de frequência de consumo (diário ou episódico) e a utilização de covariáveis (20,27,16,24,28,29) na estimação da distribuição de consumo alimentar usual. Salienta-se que em nossa experiência empírica de atendimento a pesquisadores verificamos que a questão da escolha do tipo de suposição de frequência de consumo, assim como a utilização de covariáveis, geram dúvidas nos usuários dos métodos, fazendo com que haja dificuldade para o pesquisador fazer essas definições.

Em vista da falta de referências na literatura que mostrem a influência da escolha da suposição de consumo e do uso de covariáveis nos métodos, este trabalho realizou a comparação da acurácia das

estimações para a distribuição de consumo alimentar usual para populações dos métodos *National Cancer Institute* (NCI), *Multiple Source Method* (MSM) e o *Statistical Program to Assess Dietary Exposure* (SPADE). Essas comparações foram realizadas com base em dados de 302 residentes do Rio de Janeiro – Brasil, com 20 Recordatórios Alimentares de 24 horas (R24h) por indivíduo, onde foram mensurados o consumo de alguns componentes alimentares, ao passo que nesses dados foi avaliando o impacto da escolha da suposição de frequência de consumo (diário ou episódico) e da utilização de covariáveis (Questionário de frequência alimentar (QFA), sexo e idade).



### **3. REVISÃO DE LITERATURA**

Nessa revisão de literatura foram apresentados os conceitos básicos para estimação do consumo alimentar, passando pelos tipos de erros inerentes à mensuração do consumo, assim como as técnicas disponíveis para a correção desses erros, além do panorama atual da utilização de covariáveis nos métodos de estimação da distribuição de consumo alimentar e do detalhamento técnico dos métodos NCI, MSM e SPADE.

#### **3.1. ESTIMAÇÃO DO CONSUMO ALIMENTAR**

A seguir foram apresentadas algumas definições relacionadas à estimação do consumo alimentar que serão necessárias para o seguimento do trabalho.

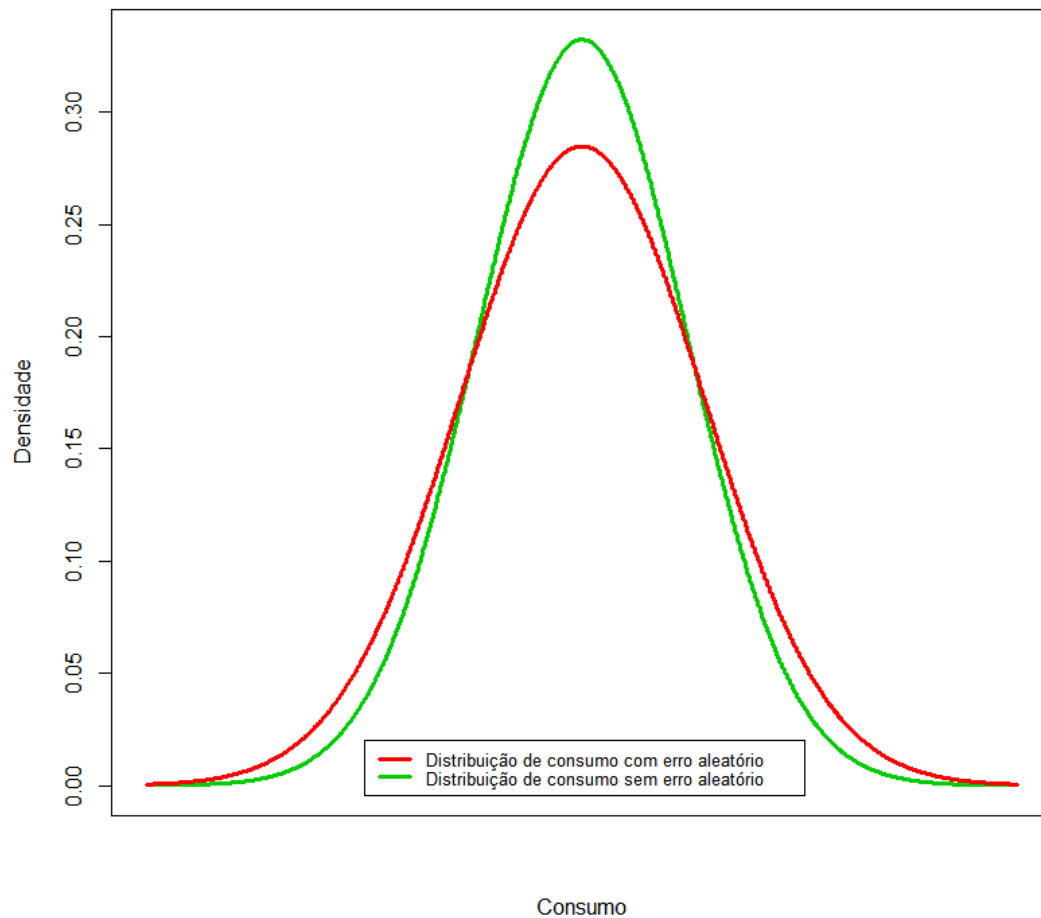
##### **3.1.1. TIPOS DE ERROS**

A mensuração de dados de consumo alimentar pode gerar dois tipos de erros: o erro sistemático e o erro aleatório (4,30). Todavia, ao se trabalhar com esses dados, é necessário que esses erros sejam minimizados, para que os resultados provenientes não sejam prejudicados.

O *National Research Council (US) Subcommittee on Criteria for Dietary Evaluation* (4) define que o erro aleatório ocorre quando o entrevistado reporta consumos maiores e/ou menores de forma aleatória. Essa variação pode ocorrer nos consumos diários dos sujeito (variação intraindividual) e nos consumos entre os sujeitos (variação interindividual) (30). Já o erro sistemático envolve grupos das populações, por exemplo, quando todos ou quase todos participantes em estudo que têm baixo nível socioeconômico reportam consumo maior de componentes alimentares mais caros (como por exemplo a carne vermelha).

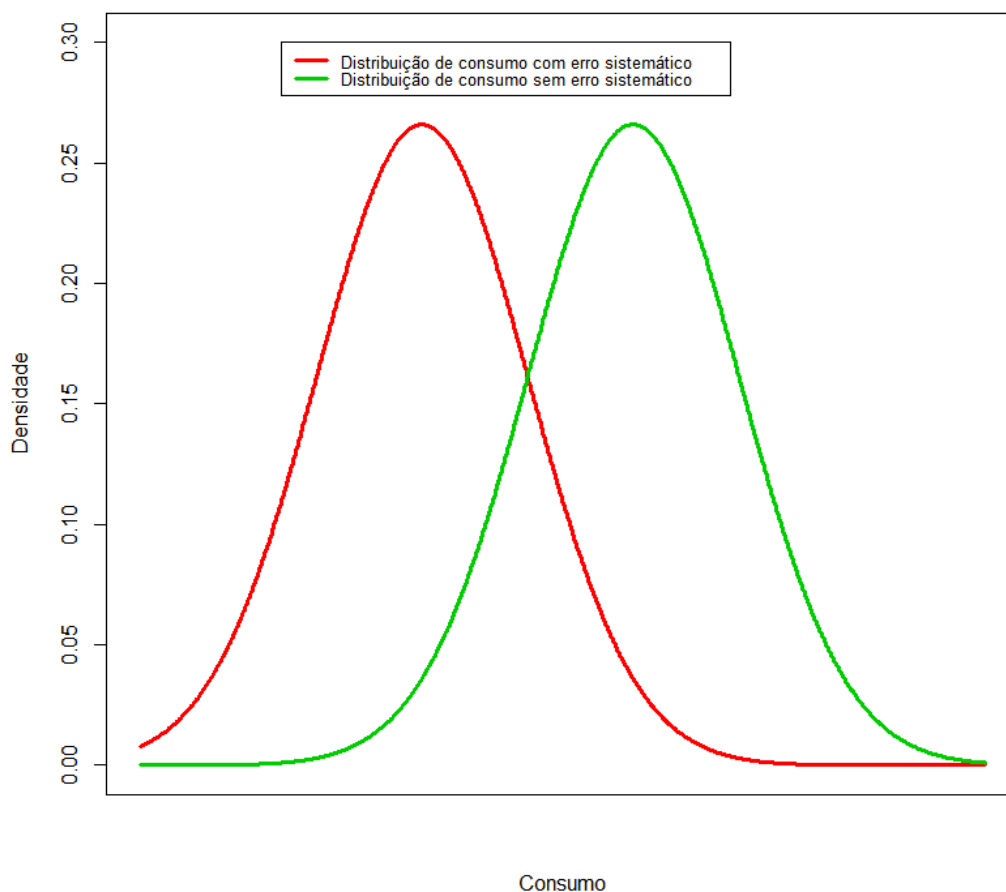
A seguir foi disposto o detalhamento dos tipos de erro, com base no trabalho do *Webinar Series* (31). Na Figura 1 se tem um exemplo do impacto do erro aleatório na distribuição de consumo alimentar

usual, nela é possível observar que um instrumento com erro aleatório não é viciado, mas pode ser impreciso.



**Figura 1:** Distribuições de consumo alimentar com e sem erro aleatório.

Na Figura 2 foi disposto o impacto do erro sistemático na distribuição de consumo alimentar usual. Nessa figura é ilustrado que, quando esse tipo de erro está presente, as medições de consumo se afastam do valor verdadeiro na mesma direção, gerando estimativas com viés.



**Figura 2:** Distribuições de consumo alimentar usual com e sem erro sistemático.

Quando se trata de erros aleatórios, tem-se que a média de várias medidas repetidas aproxima-se do verdadeiro valor do consumo alimentar usual, enquanto que para os erros sistemáticos essa relação assintótica não ocorre (4,5). Com isso, para corrigir o erro sistemático é necessário a utilização de um instrumento de referência que permita estimar a distribuição verdadeira e, com isso, corrigir esse viés. Já o erro aleatório pode ser minimizado pela média de múltiplas medidas repetidas de consumo.

Quando se trata da mensuração do consumo alimentar os instrumentos de coleta mais utilizados são o R24h e o QFA (5). No QFA os indivíduos respondem a respeito da frequência (e, às vezes, da quantidade) em que consomem determinados alimentos em um período. Já no R24h, o indivíduo relata

a frequência e a quantidade de tudo que foi consumido durante as últimas 24 horas.

Apesar do QFA requerer investimento financeiro inferior ao R24h, ele é limitado a uma lista finita de alimentos e é prejudicado pela dificuldade que os indivíduos têm para relatar seu consumo alimentar de um longo período. Em contraste, o R24h fornece detalhes ricos sobre os tipos e quantidades de alimentos consumidos.

Ambas as metodologias têm erros de mensuração atribuídos a elas, sendo que, no caso do QFA os erros sistemáticos são maiores e os aleatórios menores, já no R24h os erros sistemáticos são menores e os aleatórios são maiores (5,31). Apesar dos R24h fornecerem estimativas com menor erro sistemático, a medição de um único dia de R24h é insuficiente para se estimar o consumo de longo prazo, uma vez que geralmente existe grande variabilidade do consumo individual nas populações (5).

A estrutura do erro de medição para o consumo alimentar foi definida a seguir

$$R_{ij} = \beta_0 + \beta_1 T_i + u_i + e_{ij} \quad (1)$$

onde  $R_{ij}$  é o consumo reportado do indivíduo  $i$  no dia  $j$ ,  $\beta_0$  é o erro constante aditivo,  $\beta_1$  é o erro relacionado ao consumo,  $T_i$  é o consumo verdadeiro do indivíduo  $i$ ,  $u_i$  erro relacionado ao indivíduo  $i$  e  $e_{ij}$  é o erro aleatório.

Dos erros contidos nessa equação, um é referente ao erro aleatório e três são referentes aos erros sistemáticos. Dentre esses erros sistemáticos, tem-se:

- Erro relacionado ao consumo: referente a um valor que é multiplicado pelo consumo verdadeiro dos participantes.
- Erro constante aditivo: erro sistemático que ocorre igualmente para todos os participantes.
- Erro relacionado ao indivíduo: erro sistemático que surge como resultado de características pessoais, como sexo, idade ou peso corporal que levam os sujeitos a relatarem consumos maiores ou menores do que realmente foi ingerido.

Para fazer a correção do erro sistemático tem se utilizado resultados de biomarcadores para se obter a distribuição de consumo verdadeira, e, por conseguinte, remover esse erro. Entretanto, essa solução está disponível apenas para alguns componentes alimentares (5,31).

Com base na estrutura de erros de medição o R24h é a melhor aproximação existente de um instrumento de referência para a aferição do consumo, exceto nos casos para os quais se tem um biomarcador para a atenuação dos erros sistemáticos. Além de haver erro sistemático menor, também é possível corrigir o erro aleatório por meio de modelos estatísticos quando há múltiplas medidas dos R24h por sujeito (5). Salienta-se que além da utilização dos R24h também é possível utilizar o registro alimentar como instrumento de medição de consumo (30).

Por fim, evidencia-se que os erros de medição podem afetar de forma distinta as estimações a nível individual (quando se deseja estimar o valor de consumo de um indivíduo) e a nível populacional (quando se deseja estimar a distribuição de consumo de uma população com base em uma amostra de indivíduos) fazendo com que, por exemplo, sejam necessários um número maior de R24h individuais para estimar a distribuição de consumo individual, quando comparado ao número de R24h individuais para estimar o consumo populacional (30).

### **3.1.2. DIÁRIO VS EPISÓDICO**

Hoffman et al. (13) define o consumo diário quando os componentes alimentares são consumidos diariamente ou quase diariamente por todos indivíduos, por outro lado o consumo episódico é definido quando os componentes alimentares são consumidos raramente ou apenas por uma parte da população. Já no Webinar Series (31) o consumo diário é definido pelo consumo quase diário por quase todas as pessoas e o consumo episódico quando o componente alimentar é consumido episodicamente pela maioria das pessoas. Em Souverein et al. (16) o consumo diário é definido quando os componentes

alimentares são consumidos frequentemente (quase todos os dias) e os de consumo episódico não são consumidos todos os dias.

Apesar de haver as definições já citadas, apenas duas recomendações existentes na literatura determinam de forma quantitativa o consumo diário ou episódico. Em Guenther et al. (32) o consumo diário foi definido quando menos de 10% dos R24h têm consumos iguais a zero. Já no site do NCI (33), tem-se que quando menos de 5% dos R24h têm consumos iguais a zero considera-se consumo diário e consumo episódico quando este percentual for de 10% ou mais. Quando se tratar de percentuais entre 5% a 10% de consumos iguais a zero, a recomendação é ajustar o modelo considerando ora consumo diário e ora consumo episódico e decidir como mais adequado o que obtiver o melhor ajuste aos dados.

### **3.1.3. USO DE COVARIÁVEIS**

Ao se tratar de consumo alimentar, é intuitivo assumir que há diferença no comportamento do consumo entre grupos populacionais, motivo pelo qual as pesquisas que avaliam esta questão, geralmente, mostram os resultados com estratificações destes grupos (3,34–36).

Em Rossum et al. (3), por exemplo, foi avaliada a dieta de crianças e adultos com idades entre 7 e 69 anos através da investigação do consumo alimentar dos holandeses de 2007 a 2010. Nele, os dados dos R24h foram analisados com base em estratificações por sexo e por idade, onde se observou que, para a faixa etária de 19 a 30 anos, o consumo de frutas, nozes e azeitonas de homens ocorreu em 54% dos dias apreciados, à medida que o consumo mediano desses homens foi de 51 gramas/dia, enquanto as mulheres tiveram consumo mediano de 75 gramas/dia em 60% dos dias. Já ao se tratar da faixa etária de 51 a 69 anos, os homens apresentaram consumo em 66% dos dias, com consumo mediano de 102 gramas/dia. Alternativamente, as mulheres obtiveram consumo em 77% dos dias e consumo mediano de 132 gramas/dia. Com isso, ao comparar os sujeitos mais velhos com os mais jovens, verificou-se que houve aumento da quantidade de consumo de 100% para homens e de 78% para

mulheres. Por sua vez, a frequência teve um aumento de 12% para homens e de 17% para mulheres. Contudo, quando se compararam os sexos, verificou-se que os homens consomem em torno de 32% a menos destes alimentos por dia em comparação com as mulheres em ambas as faixas etárias, ao mesmo tempo que a frequência de dias de consumo dos homens variou de 6% a 11% a menos que as mulheres. Evidenciando que nessa população existe diferença no consumo alimentar de frutas, nozes e azeitonas não só entre homens e mulheres, mas também entre faixas etárias. Também em Herrick et al. (36) foi utilizada a estratificação por sexo e por idade, já em Crispim et al. (34) há estratificação por sexo e nacionalidade. Outro exemplo é em Yang et al. (35), que mostra a estratificação de sexo, idade, etnia, grau de escolaridade e status de tabagismo.

Quando se trata da estimação da distribuição de consumo alimentar usual, os métodos de estimação, em geral, permitem a utilização de covariáveis, ao passo que esse uso tem o objetivo de mapear as diferenças entre os grupos populacionais e, conseqüentemente, melhorar a estimação (15,16,18,22,36). Apesar de ser possível usar covariável de qualquer natureza, só o impacto da utilização do QFA já foi avaliado em seis estudos. Em Subar et al. (27) foram analisados componentes alimentares consumidos episodicamente por 965 sujeitos que responderam a quatro R24h e um QFA no estudo *Eating at America's Table* (37). Nos resultados foi observado a existência de correlação positiva entre a probabilidade de consumo (obtida pela estimação da distribuição de consumo alimentar individual do método NCI) e a frequência de dias de consumo relatadas no QFA, assim como houve correlação entre a quantidade média consumida e a frequência de consumo. Com esses achados, os autores recomendaram a utilização dos dados do QFA como covariável na estimação do consumo usual com o objetivo de reduzir erros.

Adicionalmente, Tooze et al. (20) utilizaram simulação de dois R24h por sujeito, com base nas estimativas do total de grãos do estudo *Eating at America's Table* (37), enquanto que a verdadeira

distribuição de consumo foi aproximada com base na média de 365 dias de consumo simulado. Com base nessa verdadeira distribuição, foram comparados dois tipos de suposição do método NCI (consumo episódico – com e sem correlação entre quantidade e frequência de consumo) e a utilização do QFA, chegando-se à conclusão de que o uso do QFA não causou alterações no resultado dos métodos, mas que a suposição de consumo episódico com correlação teve melhores resultados que a sem correlação.

Souverein et al. (16) utilizaram os dados de 293 homens que responderam a dois R24h no estudo EFCOVAL (34), com o objetivo de examinar a influência da inserção do QFA como covariável na qualidade dos métodos NCI e MSM na estimação do consumo de vegetais, frutas e peixe. Nos resultados foi constatado que a inserção do QFA não alterou de forma contundente os métodos.

Por outro lado, em Goedhart et al. (24) utilizaram simulação de dados para analisar o impacto do emprego das informações do QFA na estimação da distribuição de consumo usual. Foi avaliado o impacto do emprego do QFA como covariável e da informação de pessoas que nunca consomem na estimação dos métodos *BetaBinomial-Normal model* (BBN), *LogisticNormal-Normal model without correlation* (LNN0), *LogisticNormal-Normal model with correlation* (LNN e NCI), SPADE e MSM. Chegou-se à conclusão que a utilização do QFA como covariável não influenciou os métodos, já a informação de quem nunca consome traz benefícios na estimação de percentis mais baixos da distribuição nos métodos.

Em Verly et al. (28) foi utilizado o critério de informação de Akaike (AIC, do inglês *Akaike information criterion*) como medida de qualidade do ajuste dos métodos NCI e MSM, para avaliar a influência de dois tipos de covariável: QFA e uma variável binária indicadora de consumo (ela recebeu valor igual a zero quando não houve consumo em nenhum dos R24h e a frequência de consumo no QFA foi igual a zero; ou recebeu o valor igual a 1 quando houve consumo positivo em qualquer dia dos R24h e/ou a frequência de consumo no QFA não foi nula). Com isso, chegaram à conclusão de que o emprego



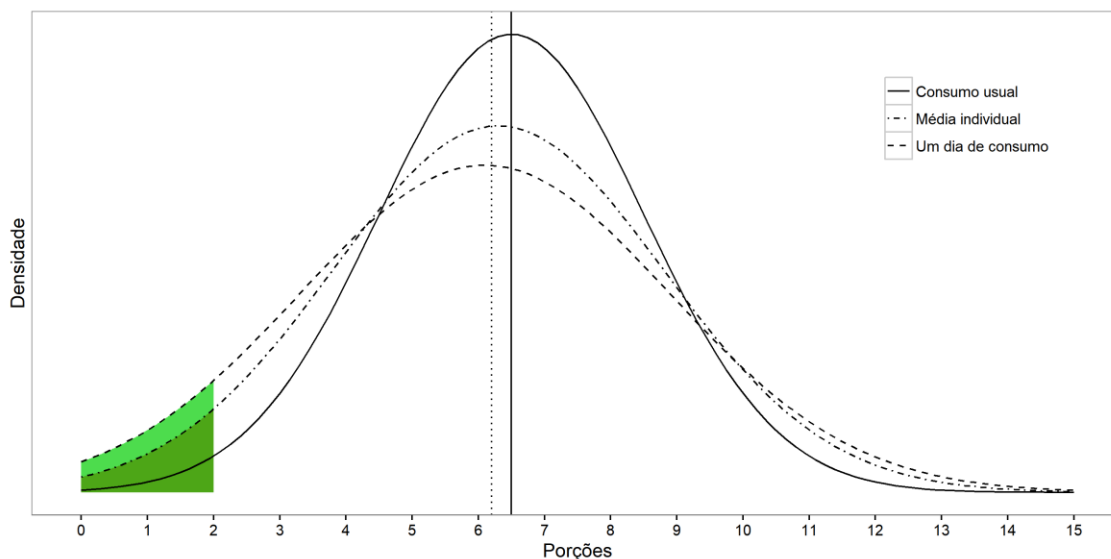
do QFA como covariável não melhorou a qualidade do ajuste dos modelos. Todavia, a variável indicadora trouxe melhora na qualidade do ajuste dos métodos (principalmente para componentes de consumo episódico).

Em Ost et al. (29) foi avaliado se há ganho em se utilizar as informações do QFA para identificar quem consome e quem nunca consome o componente alimentar na estimação da distribuição de consumo alimentar usual. Para isso, o método SPADE foi ajustado aos dados simulados de três componentes alimentares com proporções diferentes de não consumidores. Os resultados mostram que a indicação dos sujeitos que não consomem não melhorou as estimativas dos percentis mais altos e que, em contrapartida, houve melhora na estimação dos mais baixos.

#### **3.1.4. MÉTODOS DE ESTIMAÇÃO DO CONSUMO ALIMENTAR USUAL**

No trabalho de Dodd et al. (5) foi realizada uma revisão geral a respeito da estimação de consumo alimentar usual. Nesse trabalho foi relatado que o consumo usual não é bem retratado com a aferição de um único dia de consumo, e, por isso, começou-se a se utilizar mais de uma medição de R24h por sujeito, ao passo que a estimação da distribuição de consumo usual passou a ser feita pela distribuição empírica das médias individuais de consumo. Contudo, para muitos componentes alimentares, a média dos R24h ainda contém considerável variação intraindividual, fazendo com que a distribuição das médias individuais possua variância maior do que a verdadeira distribuição de consumo usual, o que leva a uma estimativa imprecisa da distribuição.

Esses potenciais vieses foram ilustrados na Figura 3 que mostra as distribuições de consumo estimadas de consumos gerados aleatoriamente e que representam um R24h por sujeito (linha tracejada), a média de dois R24h por indivíduo (linha tracejada e pontilhada) e a distribuição de consumo usual estimada (linha sólida). Já a linha vertical sólida é referente a média de dois R24h por indivíduo e a vertical tracejada é referente a média das outras duas distribuições.



**Figura 3:** Distribuições de consumo estimadas.

Já as áreas sob as curvas à esquerda das retas verticais são referentes a porcentagem da população com consumo de menos de sete porções por dia (que é aproximadamente 50% para todas as três distribuições). No entanto, quando se avalia se o consumo alimentar está adequado às recomendações, muitas vezes esses valores estão localizados nas caudas da distribuição de consumo, onde os vieses podem ser maiores. Por exemplo, a porcentagem da população com consumo usual de menos de duas porções por dia é estimada em 5,7%, 3,7% e 1,1%, respectivamente, para um R24h por entrevistado, para a média de dois R24h por indivíduo e para a distribuição de consumo usual estimada. Com base em um exemplo similar ao da Figura 3, Dood et al. (5) relataram que a utilização de uma observação de R24h ou da média de múltiplos R24h é suficiente para estimar o consumo usual médio. Entretanto, são insuficientes para estimar os percentis da distribuição de consumo alimentar usual, sendo necessário a utilização de métodos estatísticos que atenuem a variabilidade intraindividual da distribuição de consumo usual.

Em vista disso, nessa sessão foram apresentados alguns métodos para estimação da distribuição do consumo alimentar usual que fazem essa atenuação da variabilidade intraindividual de consumo,

conforme o Quadro 1. Tais métodos foram propostos com a finalidade de fazer a estimação para populações em relação ao consumo de componentes alimentares se valendo dos R24h.

**Quadro 1:** Relação de métodos para a estimação da distribuição do consumo alimentar usual.

Método	Ano de publicação	Software	Licença	Permite uso de covariáveis?	Modela o consumo episódico?	Permite a modelagem de mais de um item alimentar?	Permite a definição de pontos de corte para valores recomendados de consumo?	Permite planos amostrais complexos?
Gay (12)	1994	Programa que tenha a ANOVA* e permita simulação.	-	Não	Não	Não	Não	Não
ISU (9–11)**	1996	Versão SIDE no SAS, Versão C-SIDE e PCSIDE.	Pago	Sim	Não	Não	Não	Sim
STEM.I e II (6,7)	1996	Programa que tenha a ANOVA*	-	Sim	Sim	Não	Não	Não
Wallace (8)	1999	Programa <i>Small</i>	-	-	Não	Não	Não	Não
S-Nusser (13)	2002	SAS	Pago	Sim	Não	Não	Não	Sim
AGE MODE (23)	2006	S-PLUS	Pago	Sim	Não	Não	Não	Não
NCI (19–22)	2009	SAS	Pago	Sim	Sim	Sim	Sim	Sim
SPADE (16–18)	2010	R	Livre	Sim	Sim	Sim	Sim	Sim
MSM (14–16)	2011	Web	Livre	Sim	Sim	Não	Não	Não

\*Análise de Variância (ANOVA, do inglês *Analysis of Variance*).

\*\* Primeiro método com software desenvolvido para estimação da distribuição de consumo usual.

Em geral, os dados de consumo alimentar têm distribuições assimétricas, o que impossibilita a utilização direta de métodos que pressuponham normalidade para estimar a distribuição de consumo de componentes alimentares (5). Para contornar esse problema a maioria dos métodos utiliza a transformação dos dados dos R24h para fazer a melhor aproximação possível da distribuição normal.

Esses métodos de estimação da distribuição de consumo usual geralmente têm dois tipos de suposição sobre a frequência de consumo: diário e episódico.

Quando se trata da estimação com suposição de consumo diário, as metodologias utilizam apenas a informação da quantidade consumida pelos indivíduos nos R24h e, após transformar esses dados para a escala normal, realizam a estimação da média e das variâncias interindividual e intraindividual. Posteriormente, faz-se a atenuação do efeito do erro intraindividual por meio da remoção da variância intrarindividual. Assim, após a obtenção das estimativas atenuadas na escala normal, é feita a transformação inversa dos resultados para a escala original e, então, se obtém a distribuição de consumo alimentar usual.

Já quando se trata da suposição de consumo episódico, os métodos fazem uso da informação da quantidade e da frequência de consumo. Em vista disso, além deles realizarem a estimação considerando a quantidade de consumo (da mesma forma que no consumo usual), eles também estimam a probabilidade de consumo com base na informação de frequência. Após a estimação da quantidade e da probabilidade de consumo, faz-se a multiplicação das estimativas e se obtém a distribuição de consumo alimentar usual.

Os métodos se diferenciam em relação à forma de estimação dos parâmetros (média e variâncias) e da transformação dos dados, mas – em geral – eles têm as seguintes pressuposições (4-12,16,18-22,32):

- O R24h fornece mensurações não viciadas para o consumo usual;
- Os dados transformados podem ser decompostos em dois componentes independentes (variação interindividual e intraindividual);
- A variância intraindividual não influencia no consumo em longo prazo.

Neste trabalho, foram abordados três métodos que permitem escolher se o consumo alimentar é diário ou episódico: NCI, MSM e SPADE. O detalhamento de cada um deles foram dispostos logo a seguir.

#### **3.1.4.1. NATIONAL CANCER INSTITUTE (NCI)**

O NCI (20–22) foi proposto para estimar a distribuição do consumo alimentar diário ou episódico de componentes alimentares. Ele foi adaptado ao passar dos anos e neste trabalho foi abordada a versão mais atual (Versão 2.1, de 2015) que está implementada em duas macros do SAS (39):

- **MIXTRAN**: utiliza modelos mistos para ajustar os dados de consumo alimentar, com o intuito de estimar os parâmetros necessários para a segunda macro;
- **DISTRIB**: utiliza o método de simulação Monte Carlo para estimar a distribuição de consumo usual da população (21,22).

Ambas as macros podem ser encontradas no site: <http://riskfactor.cancer.gov/diet/usualintakes/>.

##### **3.1.4.1.1. Macro MIXTRAN**

A estimação dos parâmetros depende da suposição sobre o consumo do componente alimentar, que pode ser:

- episódico, sem correlação entre frequência e quantidade de consumo;
- episódico, com a frequência de consumo correlacionada com a quantidade de consumo;
- diário.

Para qualquer uma dessas situações, nos dados referentes a quantidade de consumo, a macro ajusta o modelo de equações de estimação generalizadas (GEE do inglês *Generalized Estimating Equations*), com o objetivo de estimar os valores iniciais para os parâmetros. Logo após, esses valores

são utilizados como valores iniciais para o ajuste do Modelo Misto Não Linear (MMNL), que faz uso da otimização Quasi-Newton na estimação dos parâmetros, à medida que essa otimização utiliza a verossimilhança aproximada pelo método de Quadratura Gaussiana Adaptativa (40). Apesar das questões em comum, cada suposição citada tem suas peculiaridades na estimação dos parâmetros:

- Consumo episódico sem correlação: faz-se a estimação dos parâmetros de forma separada, uma vez que um modelo é utilizado para as informações de quantidade consumida e outro para frequência de consumo;
- Consumo episódico com correlação: faz-se a estimação dos parâmetros em um modelo que considera as informações de quantidade consumida, de frequência de consumo e correlação entre quantidade e frequência;
- Consumo diário: faz-se a estimação dos parâmetros em apenas um modelo que considera apenas a informação de quantidade consumida.

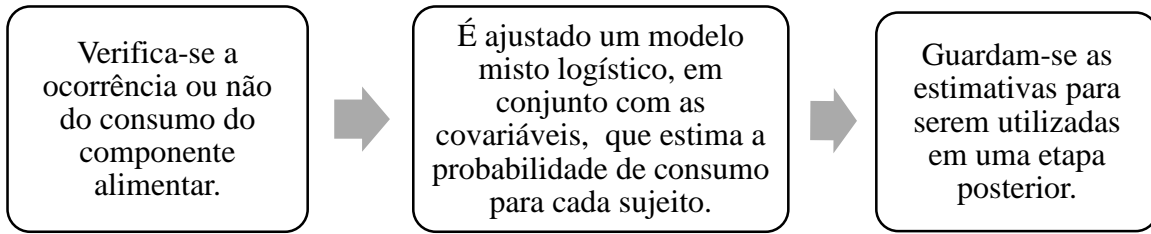
O detalhamento do processo de estimação do método NCI foi disposto a seguir. Entretanto, como neste trabalho não vai ser abordada a suposição de existência de correlação entre a frequência e a quantidade, esse processo de estimação não será detalhado.

#### Suposição de consumo episódico

Para a suposição de consumo episódico sem correlação o método é executado em duas partes: a primeira parte é referente a estimação com base na probabilidade de consumo, e a segunda parte é referente a estimação com base na quantidade de consumo.

#### ***Primeira parte:***

Na Figura 4 foi disposto um fluxograma que mostra de forma simplificada o processo de estimação dos parâmetros para a probabilidade de consumo.



**Figura 4:** Processo de estimação dos parâmetros para estimar a probabilidade de consumo.

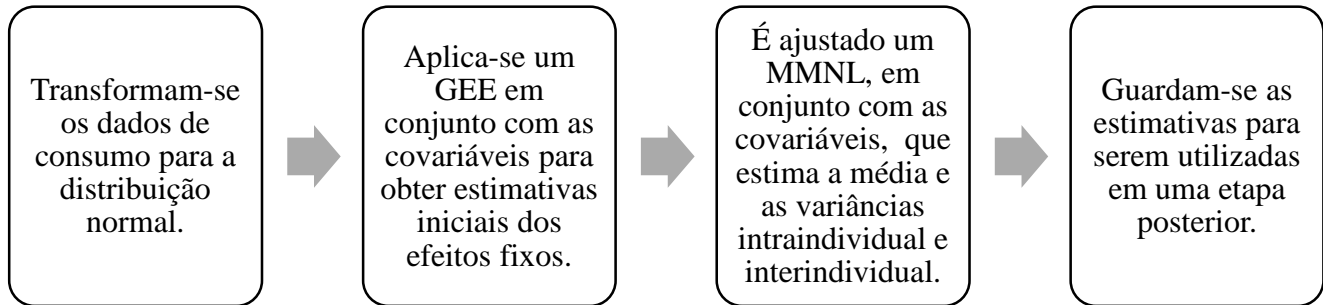
Considerando a variável aleatória  $Y_{ij}$ , que representa a quantidade consumida do componente alimentar pelo sujeito  $i$  no tempo  $j$ . Tem-se que a probabilidade de consumo do componente alimentar é estimada por um modelo misto logístico definido como

$$P(Y_{ij} > 0 | X_{1i}) = p_i = \frac{1}{1 + \exp(-(\beta_{10} + \beta_{X1}^t X_{1i} + v_i))} \quad (2)$$

onde  $p_i$  é a probabilidade de consumo do indivíduo  $i$ ,  $\exp$  é a função exponencial,  $\beta_1 = (\beta_{10}, \beta_{X1}^t)^t$  são os parâmetros de efeito fixo;  $X_{1i}$  é o vetor de covariáveis do indivíduo  $i$  e  $v_i$  é o erro aleatório da probabilidade de consumo que possui distribuição  $N(0, \sigma_v^2)$ .

**Segunda parte:**

Na Figura 5 foi disposto um fluxograma que mostra o processo de estimação para a quantidade de consumo.



**Figura 5:** Processo de estimação dos parâmetros para a quantidade de consumo.

Considerando a variável aleatória  $Y_{ij}$ , aplica-se a transformação Box-Cox de um parâmetro (41) que é definida como

$$Y_{ij}^* = \begin{cases} \frac{Y_{ij}^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0 \\ \log(Y_{ij}) & \text{se } \lambda = 0 \end{cases} \quad (3)$$

onde  $Y_{ij}^*$  são os dados de quantidade de consumo transformados e  $\lambda$  é uma constante que melhor aproxima a distribuição de  $Y_{ij}^*$  da distribuição normal (esse valor pode ser estimado pela macro ou fixado arbitrariamente) – salienta-se que não necessariamente os dados transformados terão distribuição normal, uma vez que nem sempre é possível encontrar o valor de  $\lambda$  que faça essa transformação. Nesses dados transformados é ajustado um GEE e se estima os valores iniciais dos efeitos fixo e aleatório para a quantidade de consumo em função das covariáveis. Assim, tem-se que esses valores são utilizados como valores iniciais no ajuste de um MMNL conforme a equação abaixo:

$$g(Y_{ij}, \lambda) = \beta_{20} + \beta_{X2}^t X_{2i} + u_i + \varepsilon_{ij} \quad (4)$$



onde  $g(Y_{ij}, \lambda)$  é a função que transforma os consumos originais para escala normal,  $\beta_2 = (\beta_{20}, \beta_{X2}^t)^t$  são os parâmetros de efeito fixo,  $X_{2i}$  é o vetor de covariáveis do indivíduo  $i$ ,  $u_i$  é o erro aleatório interindividual que possui distribuição  $N(0, \sigma_u^2)$  e  $\varepsilon_{ij}$  é o erro aleatório intraindividual que possui distribuição  $N(0, \sigma_\varepsilon^2)$  - ambos os erros são para a quantidade de consumo.

Salienta-se que os vetores de covariáveis da primeira ( $X_{1i}$ ) e segunda ( $X_{2i}$ ) parte podem conter covariáveis distintas.

#### Suposição de consumo diário

Para a suposição de consumo diário, o método utiliza apenas a segunda parte de modelagem, que é referente a estimação da quantidade de consumo.

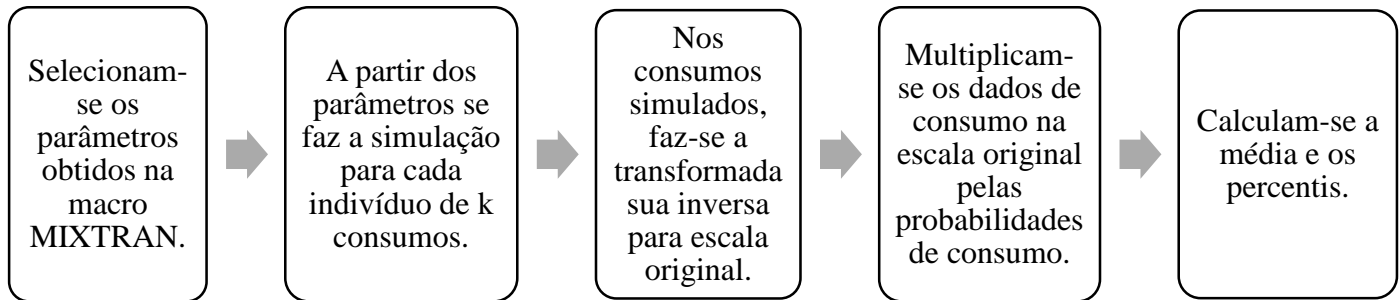
Por fim, para ambas as suposições, após a estimação dos parâmetros a macro guarda os valores para posteriormente serem utilizadas pela macro DISTRIB.

#### **3.1.4.1.2. Macro DISTRIB**

A macro DISTRIB utiliza simulação de Monte Carlo para simular  $k$  consumos, por sujeito, fazendo o uso das estimativas obtidas pela macro MIXTRAN, com o objetivo de estimar a média e os percentis da distribuição do consumo usual. Nessa etapa também há diferença na forma de estimação quando se supõe consumo diário ou episódico.

#### Suposição de consumo episódico

O processo simplificado de estimação de quando há a suposição de consumo episódico sem correlação foi apresentado na Figura 6 a seguir



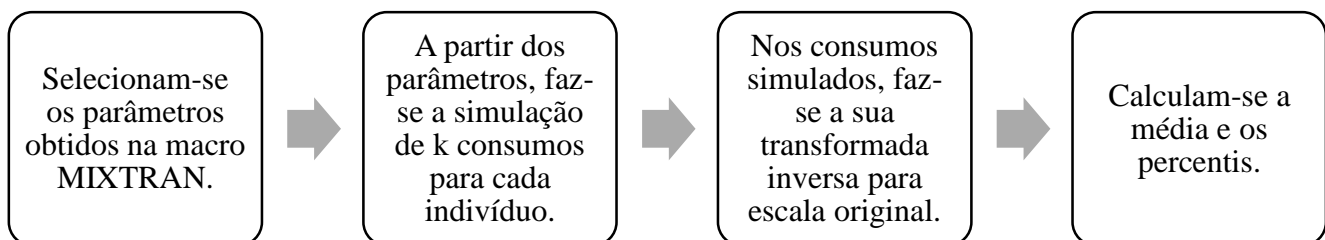
**Figura 6:** Processo de estimação da distribuição do consumo alimentar usual para quando se supõe consumo episódico.

Primeiramente, é gerado o consumo usual na escala transformada com base na simulação de  $k$  consumos por indivíduo. Para isso, primeiramente, com base na variância interindividual ( $\sigma_u^2$ ), faz-se a simulação de  $k$  realizações do erro aleatório interindividual ( $u_l$ ), onde  $l = 1, \dots, kn$  e  $n$  é o número de sujeitos. Posteriormente, os valores simulados de  $u_l$  são aplicados na fórmula  $\mu_l^* = \beta_{20} + \beta_{X2}X_{2l} + u_l$ , onde  $X_{2l} = X_{2i}$ , e se obtém  $l$  consumos simulados na escala transformada.

Após a geração dos consumos simulados, é realizada a transformação inversa para escala original desses consumos. Por fim, multiplicam-se o consumo na escala original e a probabilidade de consumo individual obtida pela equação (2) e, em seguida, calculam-se os percentis e a média dos dados obtidos.

#### Suposição de consumo diário

Foi disposto na Figura 7 o fluxograma do processo de estimação para a suposição de consumo diário.



**Figura 7:** Processo de estimação da distribuição do consumo alimentar usual para quando se supõe consumo diário.

Primeiramente, faz-se a geração dos consumos simulados (conforme descrito na suposição de consumo episódico). Após a criação do banco de dados simulados, eles são transformados para a escala original e se calcula a média e os percentis de forma empírica (42).

#### **3.1.4.2. MULTIPLE SOURCE METHOD (MSM)**

O Método MSM foi proposto por Haubrock *et al.* (14) e Harttig *et al.* (15) e está implementado no site: <https://msm.dife.de/>. Para estimar a distribuição de consumo usual esse método permite escolher entre duas suposições sobre o consumo do componente alimentar: diário ou episódico.

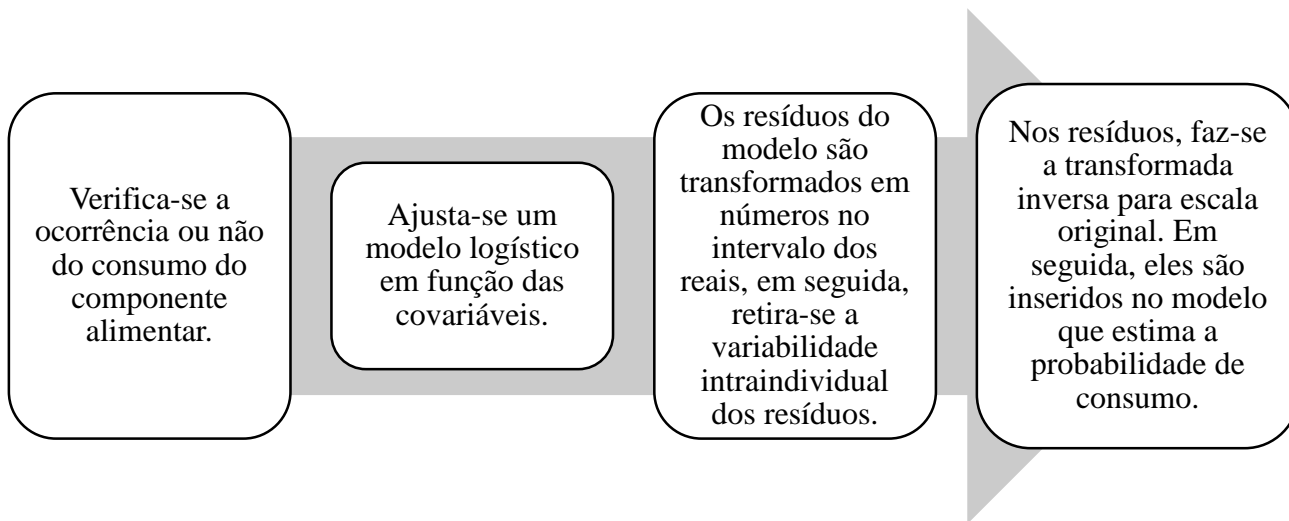
Como no NCI, a estimação para cada suposição tem suas peculiaridades, uma vez que na suposição de consumo episódico a estimação é realizada em duas etapas (uma que utiliza as informações de quantidade consumida e outra que usa a probabilidade de consumo) e, na de consumo diário, a estimação é feita com base na informação de quantidade consumida. Os detalhamentos do processo de estimação dos dois tipos de suposição foram mostrados a seguir.

##### *Suposição de consumo episódico*

Para a suposição de consumo episódico, o método divide-se em duas partes. Em cada uma é possível utilizar covariáveis, mas – diferentemente do NCI – elas têm que ser as mesmas em ambas as partes da modelagem.

##### ***Primeira parte:***

Na Figura 8 foi disposto um fluxograma para o processo de estimação da probabilidade de consumo.



**Figura 8:** Processo de estimação para a probabilidade de consumo.

Considerando  $Y_{ij}$  que representa a quantidade consumida do sujeito  $i$  e no dia  $j$ , tem-se que  $p_{ij}$ , que representa a ocorrência de consumo do sujeito  $i$  e no dia  $j$ , é definida por

$$p_{ij} = \begin{cases} 0, & \text{se } Y_{ij} = 0 \\ 1, & \text{se } Y_{ij} \neq 0 \end{cases} \quad (5)$$

Com isso, a probabilidade de consumo de um alimento é estimada com um modelo de regressão logística em função do vetor de covariáveis ( $X_{1i}$ )

$$p_{ij} = \frac{1}{1 + \exp(-(\beta_{10} + \beta_{X1}^t X_{1i} + r_{ij}))} \quad (6)$$

onde  $\exp$  é a função exponencial,  $\beta_1 = (\beta_{10}, \beta_{X1}^t)^t$  são parâmetros de efeito fixo,  $X_{1i}$  é o vetor de covariáveis do indivíduo  $i$  e  $r_{ij}$  é o resíduo do modelo definido no intervalo  $[-1; 1]$

Esse modelo foi definido de forma abreviada pelos autores conforme a equação

$$p_{ij} = m_{i/X_{1i}} + r_{ij} \quad (7)$$

onde  $m_{i/X_{1i}}$  é a previsão do modelo logístico para o indivíduo  $i$ . Posteriormente, os resíduos do modelo são transformados para números no intervalo dos reais utilizando a função  $g(r_{ij}) = \log((1 + r_{ij})/(1 - r_{ij}))$ .

Após a transformação dos resíduos é utilizada a decomposição

$$z_{ij} = g(r_{ij}) = t_i + \varepsilon_{ij} \quad (8)$$

onde  $t_i$  é uma variável estocástica com esperança  $E_{t_i}$  (que é o valor esperado do resíduo do modelo de regressão logística para o indivíduo  $i$ ), já  $\varepsilon_{ij}$  é a variação diária do resíduo para o sujeito  $i$ . Esses dois termos ( $t_i$  e  $\varepsilon_{ij}$ ) são supostos independentes entre si e têm variâncias interindividuais ( $\sigma_t^2$ ) e intraindividuais ( $\sigma_\varepsilon^2$ ), respectivamente.

Para atenuar o efeito da variabilidade intraindividual nos resíduos do modelo de regressão logística o método primeiramente estima a variância interindividual do resíduo

$$\hat{\sigma}_t^2 = \hat{\sigma}_z^2 - \frac{\hat{\sigma}_\varepsilon^2}{k} \quad (9)$$

onde  $\hat{\sigma}_z^2$  é a variância da média individual do resíduo na escala transformada ( $\bar{z}_i$ ),  $\hat{\sigma}_\varepsilon^2$  é a variância intraindividual estimada do resíduo e  $k$  é o número de dias em que os indivíduos foram avaliados.

Posteriormente, faz-se a redução da variância intraindividual nos resíduos do modelo de regressão logística

$$\hat{t}_i = \frac{\hat{\sigma}_t}{\hat{\sigma}_z} [\bar{z}_i - \bar{z}_..] + \bar{z}_.. \quad (10)$$

onde  $\hat{\sigma}_t$  é o desvio padrão interindividual do resíduo,  $\hat{\sigma}_z$  é o desvio padrão da média individual do resíduo e  $\bar{z}_..$  é a média geral dos resíduos.

Assim, tem-se que a probabilidade de consumo para o sujeito  $i$  é definida por

$$p_i^* = m_{i/X_{1i}} + g_{retrans}(\hat{t}_i) \quad (11)$$

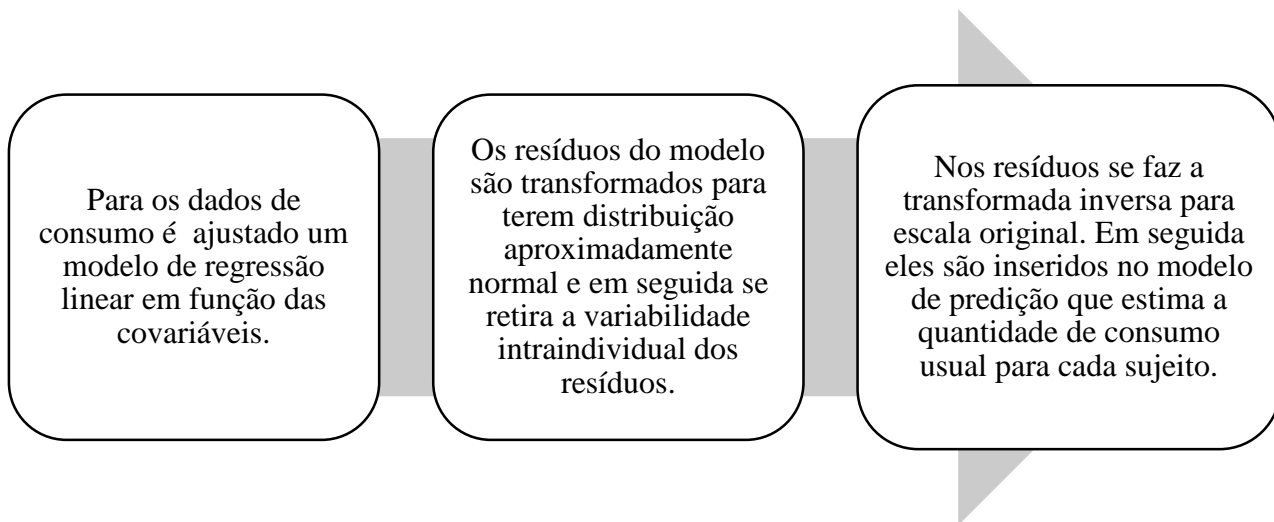
onde  $g_{retrans}$  é a transformação inversa dos resíduos que é definida por

$$g_{retrans}(\hat{t}_i) = \frac{\exp(\hat{t}_i) - 1}{\exp(\hat{t}_i) + 1} - \frac{\exp(\varpi \bar{t}) - 1}{\exp(\varpi \bar{t}) + 1} \quad (12)$$

onde  $\bar{t}$  é a média geral dos resíduos estimados ( $\hat{t}_i$ ) do modelo de regressão logística e  $\varpi$  é raiz quadrada de [1- coeficiente de correlação intra classe].

**Segunda parte:**

Essa etapa é realizada para os casos de consumo positivo, ou seja,  $Y_{ij} > 0$ . Antes de detalhar as etapas do método, foi apresentada a Figura 9, na qual foi disposto um fluxograma do processo de estimação.



**Figura 9:** Processo de estimação para a quantidade de consumo.

Nessa etapa os dados de consumo alimentar são ajustados por uma regressão linear em função do conjunto de covariáveis ( $X_{1i}$ )

$$Y_{ij} = \beta_{20} + \beta_{X1}^t X_{1i} + R_{ij} \quad (13)$$

onde  $\beta_2 = (\beta_{20}, \beta_{X1}^t)^t$  são parâmetros de efeito fixo e  $R_{ij}$  são os resíduos do modelo. Esse modelo de regressão foi definido de forma abreviada pelos autores conforme a equação abaixo

$$Y_{ij} = M_{i/X_{1i}} + R_{ij} \quad (14)$$

onde  $M_{i/X_{1i}}$  é a previsão do modelo para o consumo do indivíduo  $i$ . Logo em seguida, os resíduos do modelo são transformados para ter uma distribuição aproximadamente normal pela transformada de Box-Cox de dois parâmetros (41):

$$f(R_{ij}) = \begin{cases} ((R_{ij} + \omega)^\tau - 1)\tau^{-1} & \text{se } \tau \neq 0 \\ \ln(R_{ij} + \omega) & \text{se } \tau = 0 \end{cases} \quad (15)$$

onde os valores de  $\omega$  e  $\tau$  são determinados por maximização de Shapiro-Wilk, onde  $\tau$  varia ao longo do conjunto de valores  $\left[1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{10}, 0\right]$  e  $\omega$  varia ao longo do mesmo conjunto multiplicado pela média dos dados originais. Destaca-se que não necessariamente os dados transformados terão distribuição normal, uma vez que nem sempre é possível encontrar os valores de  $\omega$  e  $\tau$  que transformem os dados para a distribuição normal.

Após a transformação dos resíduos é utilizada a decomposição

$$Z_{ij} = f(R_{ij}) = T_i + e_{ij} \quad (14)$$

onde  $T_i$  é a variável aleatória com esperança  $E_{T_i}$  (que é o valor esperado do resíduo do modelo de regressão linear para o indivíduo  $i$ ) e  $e_{ij}$  é a variação diária dos resíduos para o sujeito  $i$ . Esses dois termos ( $T_i$  e  $e_{ij}$ ) são supostos independentes entre si com variâncias interindividual ( $\sigma_T^2$ ) e intraindividual ( $\sigma_e^2$ ), respectivamente.

Para atenuar o efeito da variabilidade intraindividual nos resíduos do modelo de regressão linear o método primeiramente estima a variância interindividual

$$\hat{\sigma}_T^2 = \hat{\sigma}_{\bar{Z}}^2 - \frac{\hat{\sigma}_e^2}{k} \quad (16)$$

onde  $\hat{\sigma}_T^2$  é a variância interindividual estimada,  $\hat{\sigma}_{\bar{Z}}^2$  é a variância da média individual do resíduo na escala transformada ( $\bar{Z}_i$ ),  $\hat{\sigma}_e^2$  é a variância intraindividual estimada e  $k$  é o número de dias em que os indivíduos consumiram ( $k > 0$ ).

Em seguida, faz-se a redução da variância intraindividual nos resíduos do modelo de regressão linear

$$\hat{T}_i = \frac{\hat{\sigma}_T}{\hat{\sigma}_{\bar{Z}}} [\bar{Z}_i - \bar{Z}_{..}] + \bar{Z}_{..} \quad (17)$$

onde  $\hat{\sigma}_T$  é o desvio padrão interindividual do resíduo,  $\hat{\sigma}_{\bar{Z}}$  é o desvio padrão da média individual do resíduo e  $\bar{Z}_{..}$  é a média geral dos resíduos.

Por fim, adiciona-se ao modelo de estimação do consumo os resíduos transformados para a escala original.

$$Y_i^* = M_{i/X_{1i}} + f_{retrans}(\hat{T}_i) \quad (18)$$

onde  $f_{retrans}$  é a transformação inversa dos resíduos, que é definida conforme a equação a seguir

$$f_{retrans}(\hat{T}_i) = \lambda^{-\lambda} \sum_{s=0}^{int(\lambda/2)} \binom{\lambda}{2s} (\hat{T}_i + \lambda)^{\lambda-2s} \sigma_{\varepsilon}^{2s} (2s-1)!! - \omega \quad (19)$$

onde  $\lambda = \tau^{-1}$ , e  $x!!$  é denotado como o produto de todos números ímpares de 1 a  $x$ .

Após o consumo ser estimado, ele é multiplicado pela probabilidade de consumo estimada e, por fim, são calculados os percentis, a média, o desvio padrão, a curtose e a assimetria.

### Suposição de consumo diário

Já na suposição de consumo diário é utilizada apenas a segunda parte da metodologia recém descrita.

#### **3.1.4.3. STATISTICAL PROGRAM TO ASSESS DIETARY EXPOSURE (SPADE)**

O método SPADE (16–18) realiza a estimação da distribuição do consumo usual de componentes alimentares utilizando a informação da variável idade, assim como seu antecessor, o método AGEMODE (23). Na estimação da distribuição de consumo alimentar usual o método tem



características adicionais as apresentadas do Quadro 1:

- O padrão do método é utilizar a idade como covariável na modelagem, entretanto é possível realizar a estimação sem essa informação desde que se faça alguns ajustes,
- Além da idade é possível utilizar outras covariáveis,
- Os resultados podem ser apresentados por unidade de idade (ex. anos ou meses) ou para grupos de idade,
- Podem ser apresentados os intervalos de confiança calculados por *Bootstrap*.

O método tem dois tipos suposição quanto a frequência de consumo na estimação: consumo diário ou consumo episódico. A seguir, foram apresentados os detalhamentos técnicos da estimação para cada uma das suposições.

O SPADE está implementado no R e utiliza os seguintes pacotes: *mpf* (43), *amer* (44), *lme4* (45), *gamlss* (46).

#### Suposição de consumo episódico

Para esse tipo de suposição é possível utilizar o modelo que permite identificar nos dados quem são os indivíduos que realmente não consomem o componente alimentar. A identificação desses indivíduos possibilita que o modelo atribua probabilidade de consumo igual a zero para os que não consomem e probabilidade de consumo positiva para os que são consumidores em potencial.

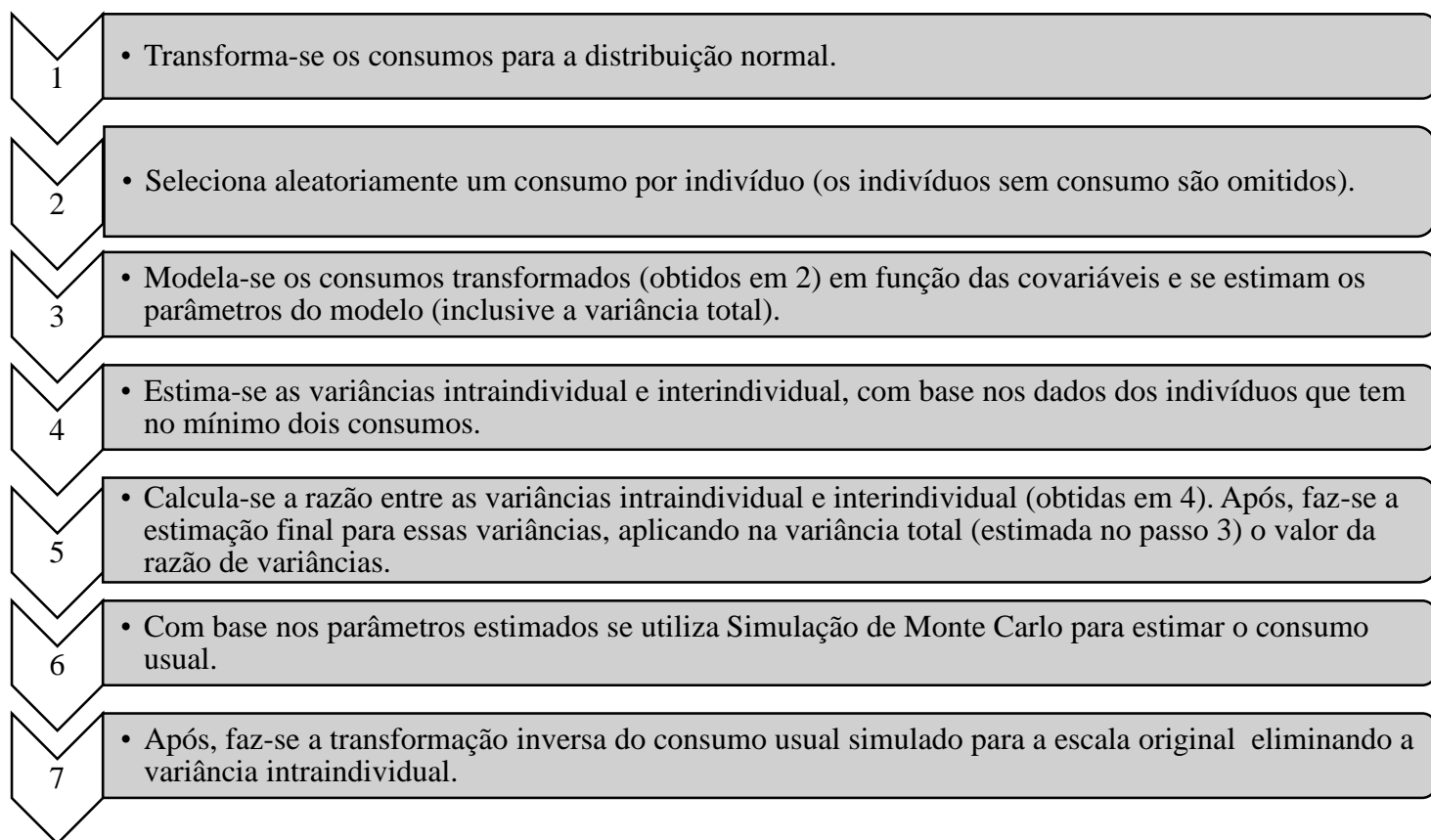
Assim como no NCI e no MSM para esse tipo de suposição o método é separado em duas partes: uma que estima a probabilidade de consumo e outra que estima a quantidade de consumo.

#### ***Primeira parte:***

Nessa etapa a probabilidade de consumo é estimada em função das covariáveis com um modelo Beta-Binomial.

#### ***Segunda parte:***

O processo de estimação da quantidade de consumo considerando a suposição de consumo episódico foi detalhado a seguir na Figura 10.

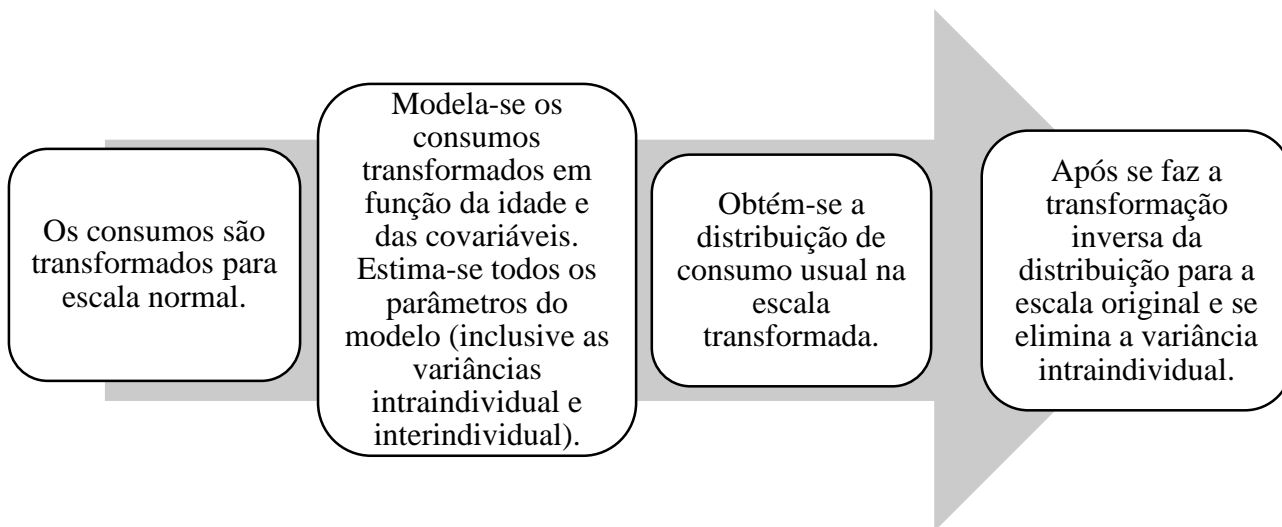


**Figura 10:** Segunda parte do processo de estimação para a suposição de consumo episódico.

Por fim, a distribuição do consumo alimentar usual é obtida pela multiplicação da probabilidade de consumo pelos dados simulados na escala original.

#### Suposição de consumo diário

Essa metodologia tem o objetivo de estimar a distribuição de consumo usual de componentes alimentares que são consumidos diariamente. Antes de detalhar as etapas do método, foi apresentada a Figura 11, na qual foi disposto um fluxograma que mostra o processo de estimação.



**Figura 11:** Processo de estimação para a suposição de consumo diário.

Primeiramente os consumos positivos são transformados para a distribuição normal, para isso o modelo utiliza a transformação Box-Cox de um parâmetro (conforme realizado no método NCI). Assim como nos demais métodos essa transformação dos dados não é garantia de que os mesmos tenham distribuição normal. Entretanto, diferentemente dos métodos NCI e MSM, o SPADE possibilita verificar se os dados foram transformados satisfatoriamente para a distribuição normal com base na disponibilização de um gráfico comparativo.

Após a transformação dos dados, o consumo médio individual é modelado com um Modelo Misto (MM) que permite a estimação da variância intraindividual ( $\hat{\sigma}_\varepsilon^2$ ) e da variância interindividual ( $\hat{\sigma}_u^2$ ) conforme a equação a seguir

$$Y_{ij}^* = a + u_i + b(x_i)^p + c(x_i)^q + X_{2i} + \varepsilon_{ij} \quad (20)$$

onde  $Y_{ij}^*$  é o consumo transformado do indivíduo  $i$  no dia  $j$ ,  $a$  é o intercepto,  $u_i$  é o erro aleatório interindividual,  $x_i$  é a idade do indivíduo  $i$ ;  $b$  e  $c$  são os coeficientes do modelo para a idade;  $q$  e  $p$  são as potências do polinômio (que podem assumir valores no conjunto  $\{-2; -1; -0,5; 0; 0,5; 1; 2\}$ ),  $X_{2i}$  é o vetor de covariáveis e  $\varepsilon_{ij}$  é o erro aleatório intraindividual.

Em seguida o método de Quadratura Gaussiana (17) é utilizado para realizar a transformação inversa da distribuição (por unidade de idade) eliminando a variância intraindividual.

### **3.2. JUSTIFICATIVA DO PRESENTE TRABALHO**

Como não há literatura que trate do impacto da escolha das suposições de frequência de consumo, nem do uso de covariáveis sexo e idade na estimação da distribuição do consumo alimentar usual, assim como há poucos estudos que realizaram a avaliação do impacto do QFA como covariável (6 estudos ao total), o presente trabalho avaliou o impacto de diferentes suposições sobre a frequência do consumo (diário ou episódico) e a utilização de covariáveis (QFA, sexo e idade) na estimação da distribuição de consumo alimentar usual.

Essa análise foi realizada para os métodos NCI, MSM e SPADE com base nos dados de Verly et al. (26), que consiste em um conjunto de dados com 20 R24h por indivíduo onde foram mensurados o consumo de diversos componentes alimentares em uma população de residentes na cidade do Rio de Janeiro – Brasil. Para fazer as comparações dos métodos foi utilizada como verdadeira distribuição de consumo a distribuição da média dos 20 R24h individuais dos 302 sujeitos disponíveis.

## **4. OBJETIVOS**

A seguir estão dispostos os objetivos do presente estudo.

### **4.1. OBJETIVO GERAL**

Comparar a acurácia dos métodos de estimação da distribuição do consumo alimentar usual sob diferentes suposições.

### **4.2. OBJETIVOS ESPECÍFICOS**

- Avaliar o impacto da escolha da suposição de frequência alimentar na acurácia dos métodos de estimação da distribuição do consumo alimentar usual;
- Estudar o impacto do uso de covariáveis (QFA, idade e sexo) na acurácia dos métodos de estimação da distribuição do consumo alimentar usual.

## REFERÊNCIAS

1. on behalf of the EFCOVAL Consortium, de Boer EJ, Slimani N, van 't Veer P, Boeing H, Feinberg M, et al. The European Food Consumption Validation Project: conclusions and recommendations. *Eur J Clin Nutr.* julho de 2011;65(S1):S102–7.
2. U.S. Department of Agriculture and U.S. Department of Health and Human Services. Dietary Guidelines for Americans, 2010 [Internet]. 7th Edition, Washington, DC: U.S. Government Printing Office; 2010. Disponível em: <https://health.gov/dietaryguidelines/dga2010/DietaryGuidelines2010.pdf#page=27>
3. Rossum, CTM, Fransen, HP, Verkaik-Kloosterma, J, Buurma-Rethans, EJM, Ocké, MC. Dutch National Food Consumption Survey 2007-2010 [Internet]. National Institute for Public Health and the Environment; 2011 p. 33–7. (Diet of children and adults aged 7 to 69 years). Report No.: 350050006/2011. Disponível em: <https://www.rivm.nl/bibliotheek/rapporten/350050006.pdf>
4. National Research Council (US) Subcommittee on Criteria for Dietary Evaluation. Nutrient Adequacy: Assessment Using Food Consumption Surveys [Internet]. Washington (DC): National Academies Press (US); 1986 [citado 18 de agosto de 2018]. Disponível em: <http://www.ncbi.nlm.nih.gov/books/NBK217533/>
5. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc.* outubro de 2006;106(10):1640–50.
6. Slob W. Modeling long-term exposure of the whole population to chemicals in food. *Risk Anal Off Publ Soc Risk Anal.* outubro de 1993;13(5):525–30.
7. Slob W. Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc.* julho de 2006;44(7):933–51.
8. Wallace LA, Duan N, Ziegenfus R. Can Long-Term Exposure Distributions Be Predicted from Short-Term Measurements? *Risk Anal.* fevereiro de 1994;14(1):75–85.
9. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *J Am Stat Assoc.* dezembro de 1996;91(436):1440–9.
10. Guenther PM, Kott PS, Carriquiry AL. Development of an approach for estimating usual nutrient intake distributions at the population level. *J Nutr.* junho de 1997;127(6):1106–12.
11. Nusser SM, Fuller WA, Guenther PM. Estimating Usual Dietary Intake Distributions: Adjusting for Measurement Error and Nonnormality in 24-Hour Food Intake Data. In: Lyberg L, Biemer P, Collins M, De Leeuw E, Dippo C, Schwarz N, et al., organizadores. *Survey Measurement and*

Process Quality [Internet]. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2012. p. 689–709. Disponível em: <http://doi.wiley.com/10.1002/9781118490013.ch30>

12. Gay C. Estimation of population distributions of habitual nutrient intake based on a short-run weighed food diary. *British Journal of Nutrition*. 83° ed 200DC;287–93.
13. Hoffmann K, Boeing H, Dufour A, Volatier JL, Telman J, Virtanen M, et al. Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr*. maio de 2002;56 Suppl 2:S53-62.
14. Haubrock J, Nothlings U, Volatier J-L, Dekkers A, Ocke M, Harttig U, et al. Estimating Usual Food Intake Distributions by Using the Multiple Source Method in the EPIC-Potsdam Calibration Study. *J Nutr*. 23 de março de 2011;141(5):914–20.
15. Harttig U, Haubrock J, Knüppel S, Boeing H. The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *Eur J Clin Nutr*. julho de 2011;65 Suppl 1:S87-91.
16. Souverein OW, Dekkers AL, Geelen A, Haubrock J, de Vries JH, Ocké MC, et al. Comparing four methods to estimate usual intake distributions. *Eur J Clin Nutr*. julho de 2011;65 Suppl 1:S92-101.
17. Dekkers ALM, Slob W. Gaussian Quadrature is an efficient method for the back-transformation in estimating the usual intake distribution when assessing dietary exposure. *Food Chem Toxicol*. outubro de 2012;50(10):3853–61.
18. Dekkers AL, Verkaik-Kloosterman J, van Rossum CT, Ocke MC. SPADE, a New Statistical Program to Estimate Habitual Dietary Intake from Multiple Food Sources and Dietary Supplements. *J Nutr*. 1° de dezembro de 2014;144(12):2083–91.
19. Tooze JA, Grunwald GK, Jones RH. Analysis of repeated measures data with clumping at zero. *Stat Methods Med Res*. agosto de 2002;11(4):341–55.
20. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc*. outubro de 2006;106(10):1575–87.
21. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics*. dezembro de 2009;65(4):1003–10.
22. Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. *Stat Med*. 30 de novembro de 2010;29(27):2857–68.
23. Waijers PMCM, Dekkers ALM, Boer JMA, Boshuizen HC, van Rossum CTM. The potential of AGE MODE, an age-dependent model, to estimate usual intakes and prevalences of inadequate intakes in a population. *J Nutr*. novembro de 2006;136(11):2916–20.

24. Goedhart PW, van der Voet H, Knüppel S, Dekkers ALM, Dodd KW, Boeing H, et al. A comparison by simulation of different methods to estimate the usual intake distribution for episodically consumed foods. EFSA - Supporting publications [Internet]. EN299 ed 15 de junho de 2012 [citado 28 de maio de 2018]; Disponível em: <http://doi.wiley.com/10.2903/sp.efsa.2012.EN-299>
25. Laureano GHC, Torman VBL, Crispim SP, Dekkers ALM, Camey SA. Comparison of the ISU, NCI, MSM, and SPADE Methods for Estimating Usual Intake: A Simulation Study of Nutrients Consumed Daily. *Nutrients*. 15 de março de 2016;8(3):166.
26. Verly-Jr E, Oliveira DCRS, Fisberg RM, Marchioni DML. Performance of statistical methods to correct food intake distribution: comparison between observed and estimated usual intake. *Br J Nutr*. setembro de 2016;116(5):897–903.
27. Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M, et al. The Food Propensity Questionnaire: Concept, Development, and Validation for Use as a Covariate in a Model to Estimate Usual Food Intake. *J Am Diet Assoc*. outubro de 2006;106(10):1556–63.
28. Verly- E, Fisberg RM, Marchioni DML. Is the food frequency consumption essential as covariate to estimate usual intake of episodically consumed foods? *Eur J Clin Nutr*. novembro de 2012;66(11):1254–8.
29. Ost C, De Ridder KAA, Tafforeau J, Van Oyen H. The added value of food frequency questionnaire (FFQ) information to estimate the usual food intake based on repeated 24-hour recalls. *Arch Public Health* [Internet]. dezembro de 2017 [citado 22 de julho de 2018];75(1). Disponível em: <http://archpublichealth.biomedcentral.com/articles/10.1186/s13690-017-0214-8>
30. Willett W. *Nutritional epidemiology*. Third edition. Oxford ; New York: Oxford University Press; 2013. 529 p. (Monographs in epidemiology and biostatistics).
31. Applied Research: Cancer Control and Population Sciences.Measurement Error [Internet]. Webinar Series. [citado 20 de agosto de 2018]. Disponível em: <http://appliedresearch.cancer.gov/measurementerror/>
32. Guenther PM, Kirkpatrick SI, Reedy J, Krebs-Smith SM, Buckman DW, Dodd KW, et al. The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. *J Nutr*. março de 2014;144(3):399–407.
33. Usual Dietary Intakes: Food Intakes, U.S. Population, 2007-10 [Internet]. Disponível em: <https://epi.grants.cancer.gov/diet/usualintakes/pop/2007-10/>
34. Crispim SP, de Vries JHM, Geelen A, Souverein OW, Hulshof PJM, Lafay L, et al. Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres – results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr*. fevereiro de 2011;105(03):447–58.
35. Yang Q, Zhang Z, Gregg EW, Flanders WD, Merritt R, Hu FB. Added Sugar Intake and Cardiovascular Diseases Mortality Among US Adults. *JAMA Intern Med*. 1º de abril de 2014;174(4):516.



36. Herrick, KA R LM, Parsons, R D KW. Estimating usual dietary intake from National Health and Nutrition Examination Survey data using the National Cancer Institute method. National Center for Health Statistics. 2018. Report No.: Vital Health Stat 2(178).
37. Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires : the Eating at America's Table Study. *Am J Epidemiol*. 15 de dezembro de 2001;154(12):1089–99.
38. Buck RJ, Hammerstrom KA, Ryan PB. Estimating long-term exposures from short-term measurements. *J Expo Anal Environ Epidemiol*. setembro de 1995;5(3):359–73.
39. SAS Institute. Base SAS 9.3 procedures guide [Internet]. Cary, N.C.: SAS Institute; 2011 [citado 29 de maio de 2018]. Disponível em: <http://www.books24x7.com/marc.asp?bookid=44483>
40. Pinheiro JC, Bates DM. Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model. *J Comput Graph Stat*. 1º de março de 1995;4(1):12–35.
41. Box GEP, Cox DR. An Analysis of Transformations. *Series B (Methodological)*. 1964;211–52.
42. Usual Dietary Intakes: Details of the NCI Method [Internet]. Disponível em: <https://epi.grants.cancer.gov/diet/usualintakes/details.html>
43. Ambler G, Benner A. MFP: Multivariable Fractional Polynomials. 2010; Disponível em: <http://cran.r-project.org/web/packages/mfp/>
44. Scheipl F. AMER: Additive mixed models with lme4. 2011; Disponível em: <http://cran.r-project.org/web/packages/amer/amer.pdf>
45. Bates DM, Maechler M, Bolker B. LME4: Linear mixed-effects models using S4 classes. setembro de 2013; Disponível em: <http://cran.r-project.org/web/packages/lme4/>
46. Stasinopoulos DM, Rigby RA. Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*. 23º ed dezembro de 2007;

## 5. ARTIGO 1

### **Impacto da suposição sobre a frequência do consumo, diário ou episódico, e do uso do questionário de frequência alimentar na estimação da distribuição de consumo alimentar usual nos métodos NCI, MSM e SPADE**

Impact of daily or episodic intake frequency supposition and FFQ use in usual dietary intake distribution estimation for NCI, MSM and SPADE methods

Greice H. C. Laureano, Vanessa B. Leotti e Suzi A. Camey

**Artigo a ser submetido ao jornal *British Journal of Nutrition***

## Resumo

Para estimar distribuições de consumo alimentar usual em populações é necessário que se defina a suposição sobre a frequência de consumo (diário ou episódico) do componente alimentar de interesse. Entretanto, não há estudos na literatura que se refiram à influência dessa seleção na qualidade das estimativas dos métodos. O presente trabalho avaliou o impacto da escolha da suposição sobre a frequência de consumo alimentar e a utilização das informações do Questionário de Frequência Alimentar (QFA) na estimação da distribuição de consumo alimentar usual. Para isto, os métodos *National Cancer Institute* (NCI), *Multiple Source Method* (MSM) e *Statistical Program to Assess Dietary Exposure* (SPADE) foram aplicados nos dados de uma pesquisa que coletou 20 recordatórios alimentares de 24 horas (R24h) de 302 pessoas residentes na cidade do Rio de Janeiro, Brasil.

Na análise dos dados, a frequência de consumo dos componentes alimentares foi classificada da seguinte forma: diário (no máximo 0,1% de consumos iguais a zero nos R24h), episódico de nível intermediário (quando a proporção de R24h iguais a zero esteve entre 20% e 75%) e episódico de nível baixo (no mínimo 76% dos R24h com medições iguais a zero). Já a mensuração da qualidade de estimação dos métodos foi realizada pelo vício relativo, ao passo que os resultados mostraram que, quando o QFA não é considerado nos componentes alimentares de frequência de consumo diário, as medidas mais acuradas foram obtidas com a suposição de consumo diário e que, neste caso, não há diferença entre os três métodos. No entanto, em componentes alimentares com frequência de consumo episódico de nível intermediário, em geral, obteve-se maior acurácia no MSM com a suposição de consumo episódico. Já nos componentes com frequência de consumo episódico de nível baixo, os resultados mais acurados foram obtidos considerando a suposição de consumo episódico nos métodos SPADE e NCI (na mensuração da média e dos percentis superiores a 25) e no MSM (em percentis menores ou iguais a 25). A utilização do QFA, mesmo tendo apresentado um discreto aumento de acurácia nos métodos SPADE (sobretudo nos percentis abaixo de 25) e MSM (especialmente na média e nos percentis acima de 75), de forma geral não trouxe benefícios expressivos nos resultados dos métodos.

Portanto, sugere-se escolher a suposição de frequência de consumo diário em componentes alimentares de frequência de consumo diário, independentemente do método utilizado. Por outro lado, em componentes de frequência de consumo episódico de nível intermediário, recomenda-se o emprego do método MSM com suposição de consumo episódico. Por fim, em se tratando de componentes com frequência de consumo episódico de nível baixo, indica-se o uso da suposição de consumo episódico e

do método MSM para estimar percentis menores ou iguais a 25, e dos métodos SPADE e NCI para estimar a média e dos percentis acima de 25. Quanto à utilização do QFA, sugere-se que seja avaliado o custo-benefício da coleta desses dados, uma vez que a elaboração e aplicação do QFA demanda investimentos que podem não compensar as melhoras obtidas nos métodos MSM e SPADE.

**Palavras-chave:** MSM. NCI. SPADE. Consumo diário. Consumo episódico. R24h. QFA.

## Introdução

Com o advento da epidemia de obesidade e das morbidades correlacionadas, há um crescente interesse no conhecimento do consumo alimentar das populações pela saúde pública, o que vem fomentando a investigação dos componentes alimentares que as populações consomem usualmente <sup>(1-3)</sup>. Um dos métodos utilizados para mensurar o consumo alimentar é o recordatório alimentar de 24 horas (R24h), mas além deles também é possível utilizar os relatos alimentares (que são de uso comum no Brasil). Os métodos de mensuração do consumo são aplicados em amostras representativas da população e, a partir desses resultados, estima-se a distribuição de consumo alimentar usual <sup>(4)</sup>. Essa distribuição é reflexo da frequência e da quantidade consumida pelos indivíduos. Com ela, é possível identificar, por exemplo, se há carência ou excesso de consumo e, com isso, pode-se traçar estratégias de saúde pública que propiciem a recomendação de níveis considerados adequados <sup>(5)</sup>.

Existem diferentes métodos que realizam a estimação da distribuição de consumo alimentar usual <sup>(6-26)</sup>. Dentre eles, encontram-se o *National Cancer Institute* (NCI) <sup>(14,15,19,20)</sup>, o *Multiple Source Method* (MSM) <sup>(21,27)</sup> e o *Statistical Program to Assess Dietary Exposure* (SPADE) <sup>(22,24,25)</sup>. Os três exigem que se faça uma suposição da frequência de consumo do componente alimentar objeto de análise. Há duas suposições <sup>(14,19,25,27)</sup>: uma considera que quase todos os indivíduos o consomem quase diariamente (consumo diário) e outra que nem todos os indivíduos o consomem diariamente (consumo episódico).

Em Guenther et al. <sup>(28)</sup> há o entendimento de que o consumo diário é definido quando menos de 10% dos R24h têm consumos iguais a zero. Já na página *web* do *National Cancer Institute* <sup>(29)</sup> há a recomendação de se usar a suposição de consumo diário no caso de até 5% dos R24h possuírem consumos iguais a zero, como também de se fazer o uso da suposição de consumo episódico quando este percentual seja de 10% ou mais. Não obstante, quando se trata de percentuais entre 5% a 10% de consumos iguais a zero, a indicação é utilizar os dois tipos de suposição e decidir como mais adequada aquela que conseguir o melhor ajuste aos dados. Embora existam estas orientações, não há pesquisas que

mostrem a repercussão da escolha da suposição sobre a frequência do consumo na acurácia da estimação da distribuição.

Outra questão relacionada à estimação do consumo alimentar usual é a inclusão do Questionário de Frequência Alimentar (QFA) como informação adicional nos modelos dos métodos para reduzir erros de estimação. O uso do QFA como covariável na estimação da distribuição de consumo alimentar vem sendo discutido por alguns autores. Em Subar et al. <sup>(30)</sup> foi observado a existência de correlação positiva entre a probabilidade de consumo (obtida pela estimação da distribuição de consumo alimentar do método NCI) e a frequência de dias de consumo relatadas no QFA. Já Tooze et al. <sup>(15)</sup> chegaram à conclusão de que o QFA não trouxe mudanças nos resultados do método NCI. Em Goedhart et al. <sup>(31)</sup> foi avaliado o impacto do emprego do QFA como covariável e da informação de pessoas que nunca consomem na estimação dos métodos *BetaBinomial-Normal model* (BBN), *LogisticNormal-Normal model without correlation* (LNN0), *LogisticNormal-Normal model with correlation* (LNN e NCI), SPADE e MSM, chegando à conclusão que a utilização do QFA como covariável não influenciou os métodos, enquanto que a informação de quem nunca consome traz benefícios na estimação de percentis inferiores da distribuição nos métodos. Adicionalmente, Souverein et al. <sup>(22)</sup> verificaram a influência da inserção do QFA como covariável na qualidade dos métodos NCI e MSM e constataram que não houve alteração expressiva nos resultados dos métodos. Já em Verly et al. <sup>(32)</sup> foi utilizada a qualidade do ajuste dos métodos NCI e MSM como técnica de comparação a respeito da inserção de dois tipos de covariável (QFA e/ou uma variável indicadora de consumo baseada nos R24h), chegando-se à conclusão de que a inclusão do QFA como covariável não alterou a qualidade do ajuste dos métodos. Em contrapartida, a variável indicadora de consumo trouxe melhorias na qualidade do ajuste dos métodos. Em Ost et al. <sup>(33)</sup> foi avaliado se há ganho em se utilizar as informações do QFA para identificar quem consome e quem nunca consome o componente alimentar na estimação do método SPADE, chegando-se à conclusão que não houve melhora nas estimativas dos percentis mais altos. Em contrapartida, houve melhora na estimação dos mais baixos.

Em estudos que comparam métodos para estimação da distribuição de consumo alimentar usual, é necessário definir qual a verdadeira distribuição. Há divergência entre as pesquisas da área, uma vez que, para se definir a verdadeira distribuição, se faz uso de simulação de dados <sup>(10,15,20,22,24,31,34)</sup>, média de múltiplos R24h <sup>(11,13,35)</sup> ou biomarcadores <sup>(5)</sup>. À medida que se emprega a simulação de dados, a construção da distribuição verdadeira é realizada computacionalmente e, apesar dessa geração normalmente ser baseada em dados reais, há críticas <sup>(31)</sup> quanto à reprodutibilidade dos resultados em

outros conjuntos de dados. Na utilização de diversos R24h, se considera que a média dessas múltiplas medidas representam o comportamento usual de consumo da população em questão. Essa técnica também recebe críticas <sup>(5)</sup> pelos altos custos na mensuração de múltiplos recordatórios e pela baixa reprodutibilidade dos resultados em outros conjuntos de dados. Por fim, também há críticas <sup>(5)</sup> no cálculo dos valores verdadeiros via biomarcadores, uma vez que essa medida está disponível apenas para alguns componentes alimentares, também possuindo alto custo.

Diante desta situação, o objetivo deste trabalho é mensurar o impacto da suposição sobre a frequência do consumo, diário ou episódico, na estimação da distribuição de consumo alimentar usual, além de verificar se a utilização das informações do QFA no ajuste dos métodos corrobora com os achados atuais da literatura. Todas as análises foram comparadas com a distribuição de consumo usual verdadeira baseada na distribuição da média de múltiplos recordatórios. Esta análise foi realizada nos métodos NCI, MSM e SPADE e foram empregados os dados da pesquisa realizada por Verly et al. <sup>(35)</sup>.

## **Método**

Verly et al. <sup>(35)</sup> apresentaram um estudo longitudinal com 302 pessoas residentes na cidade do Rio de Janeiro, Brasil, que responderam a 20 dias não consecutivos de R24h, onde se considerou a ingestão dos componentes alimentares a seguir:

- Alimentos: arroz, carnes em geral, vegetais totais, pães (excluindo os recheados), feijão, café, derivados do leite, refrigerante, frutas, carne de gado, doces e chocolates, carne de frango, biscoitos, suco natural, bolos, ovos, pizza e chá.
- Macronutrientes: carboidratos, lipídios totais, proteínas, e gorduras saturadas.
- Micronutrientes: zinco, cobre, cálcio, fósforo, potássio, sódio, folato diário, magnésio, vitamina A, vitamina B1, vitamina B2, vitamina B6 e vitamina C.

Adicionalmente, foi aplicado um QFA em 285 participantes. Apenas sete alimentos do QFA tinham correspondência com aqueles mensurados no R24h: carne de gado, bolo, frango, ovo, pizza, refrigerante e suco natural. Na elaboração do banco de dados dos R24h foi utilizado o software Brasil Nutri que contém diversos itens alimentares, formas de preparo e unidades de medidas caseiras pré-definidas. Ele foi desenvolvido pelo Ministério da Saúde brasileiro em parceria com o Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro.

Na análise dos dados, foi usada a seguinte classificação da frequência de consumo dos componentes alimentares:

- Episódico de nível baixo – no mínimo 76% de consumos iguais a zero nos R24h;
- Episódico de nível intermediário – entre 20% e 75% de consumos iguais a zero nos R24h;
- Diário – no máximo 0,1% de consumos iguais a zero nos R24h.

Nessas classificações apenas os nutrientes ficaram na faixa de consumo diário e não há componentes alimentares com consumo entre 0,1% e 20%.

Os métodos empregados para a estimação da distribuição de consumo alimentar usual foram o NCI, MSM e o SPADE. Apesar de terem diferenças metodológicas entre si, um ponto em comum é que o usuário necessita definir qual a suposição sobre a frequência do consumo que se pretende fazer uso. Assim, ao se definir a suposição de consumo diário, os métodos realizam a estimação da distribuição utilizando apenas a informação da quantidade consumida pelos indivíduos, enquanto que, ao se escolher a suposição de consumo episódico, além de usar a informação da quantidade, também se usa a da frequência de consumo. Salienta-se que o método SPADE é o único que possui um modelo de consumo episódico que permite identificar se o sujeito nunca consome o componente alimentar, ao passo que, nos demais métodos, é possível aproximar essa informação com base no uso do QFA como covariável. No material suplementar estão dispostos resumos comparativos dos métodos, considerando a suposição de consumo diário (Quadro S1) e a suposição de consumo episódico (Quadro S2).

Os métodos NCI, MSM e SPADE foram utilizados no ajuste dos dados dos 20 R24h por pessoa, fazendo-se em cada método o uso das suposições de consumo diário ou episódico. Já quando se tratou da inserção da informação do QFA, foram ajustados os métodos NCI e MSM com a quantidade de consumo informada no QFA como covariável e o método SPADE constando a informação de quando o sujeito não consumia o componente alimentar (essa informação foi obtida do QFA, criando-se uma variável que indicava se o sujeito não consumia o componente alimentar). Nesses ajustes, foi usada apenas a suposição de consumo episódico, uma vez que todos os componentes alimentares mensurados pelo R24h e QFA, simultaneamente, tiveram frequência de consumo episódica de nível baixo ou intermediário.

O vício relativo foi a medida de acurácia empregada na aferição da influência da suposição sobre a frequência do consumo, diário ou episódico, e do uso do QFA na modelagem.

O vício relativo foi definido por

$$VR = \begin{cases} \frac{\hat{\theta} - \theta}{\theta}, & \text{se } \theta > 0 \\ 0, & \text{se } \theta = 0 \text{ e } \hat{\theta} = 0 \\ \infty, & \text{se } \theta = 0 \text{ e } \hat{\theta} > 0 \end{cases} \quad (1)$$

onde  $\hat{\theta}$  é o valor estimado do parâmetro (média ou percentis da distribuição do consumo) e  $\theta$  é o valor verdadeiro do parâmetro. Caso o valor verdadeiro do parâmetro seja zero, tem-se duas possíveis situações: vício relativo igual a zero (se o valor estimado é igual a zero) e vício relativo igual a infinito (quando o valor estimado é diferente de zero), sendo que neste último caso os métodos não foram capazes de estimar consumo igual a zero. Os verdadeiros valores dos parâmetros foram calculados com base na distribuição empírica das médias dos 20 consumos dos indivíduos e podem ser vistos na Tabela S1 do material suplementar.

Foram excluídos deste estudo os componentes alimentares referentes a carnes processadas, peixe, carne de porco, cerveja, niacina, ferro, fibra total e vitamina B12; uma vez que foi observado que a utilização dos 20 R24h não foram suficientes para estabilizar a estimativa para o valor verdadeiro. No material suplementar está disposta a metodologia utilizada para a seleção e no gráfico S1 estão dispostos os componentes alimentares excluídos das análises, já nos gráficos S2 e S3 estão as razões calculadas dos componentes alimentares que permaneceram.

A fim de facilitar a visualização dos resultados, os vícios maiores do que 2,5 foram excluídos das figuras, lembrando que um vício maior do que 2,5 significa que o estimador superestimou em 250% ou mais o verdadeiro valor do parâmetro. Por fim, na avaliação dos resultados dos nutrientes, em alguns casos, os métodos NCI e SPADE não conseguiram produzir estimativas dos parâmetros, impossibilitando a obtenção dos valores dos vícios relativos.

Foi usado no ajuste do método NCI o *software* SAS <sup>(36)</sup> versão 9.3. No ajuste do método SPADE e análise dos resultados, foi utilizado o *software* R <sup>(37)</sup> versão 3.3.3. No método MSM, foi utilizado a sua interface na *web*: <https://msm.dife.de/>.

## Resultados

Na Tabela 1 está disposto o percentual mensurado de consumos iguais a zero, considerando os 20 R24h por participante ou considerando o QFA. Observa-se que, nos componentes alimentares avaliados a frequência de consumo sempre foi superior no QFA.

**Tabela 1:** Percentual estimado de consumos iguais a zero nos componentes alimentares, considerando-se os 20 R24h ou o QFA.



<b>Componente alimentar</b>	<b>Percentual de consumos iguais a zero nos R24h's</b>	<b>Percentual de consumos iguais a zero no QFA</b>
<b>NUTRIENTES</b>		
Vitamina C	0,1	-
Demais nutrientes*	0	-
<b>ALIMENTOS</b>		
Arroz	21	-
Carnes em geral	27	-
Vegetais totais	29	-
Pães (excluindo os recheados)	34	-
Feijão	44	-
Café	49	-
Derivados do leite	51	-
Refrigerante	56	-
Frutas	62	-
Carne de gado	64	2
Doces e chocolates	66	-
Carne de frango	67	2
Biscoitos	68	-
Suco natural	81	8
Bolos	87	6
Ovos	89	5
Pizza	94	6
Chá	95	-

\*cálcio, carboidratos, fósforo, lipídios totais, potássio, proteínas, sódio, zinco, cobre, folato diário, magnésio, gordura saturada, vitamina A, vitamina B1, vitamina B2 e vitamina B6.

Na Figura 1 está a comparação dos vícios dos métodos NCI, MSM e SPADE, considerando as duas suposições de frequência de consumo alimentar e as categorias de frequência de consumo do componente alimentar de acordo com os R24h. Constatou-se que, em componentes alimentares com frequência de consumo diário, obtiveram-se resultados semelhantes em todos os métodos, com vícios

relativos menores (variando entre -0,56 e 0,37) na suposição de consumo diário quando comparado com a suposição de consumo episódico. Em contrapartida, em componentes alimentares com frequência episódico de nível intermediário de consumo, em geral, obteve-se estimativas mais acuradas usando o MSM com a suposição de consumo episódico, apresentando vícios relativos inferiores a 0,55 em valores absolutos. Por fim, nos componentes alimentares com frequência de consumo episódico de nível baixo, as medidas com maior acurácia foram obtidas com suposição de consumo episódico, sendo que o SPADE e o NCI tiveram menores vícios relativos para a estimativa da média e dos percentis maiores que 25, já o método MSM apresentou os menores vícios (entre 0 e -1) para os percentis de 5 a 25, enquanto os demais métodos obtiveram a maioria dos vícios iguais a valores infinitos – esses casos podem ser verificados na Tabela S2 do material suplementar.

Ainda sobre a comparação das suposições de frequência de consumo, foi identificado que alguns dos métodos não conseguiram estimar consumo igual a zero nos casos a seguir:

- Em ambas as suposições nos métodos NCI e SPADE: biscoitos e frutas (percentil 5); bolos, café, e ovos (percentis 5 e 10); pizza e suco natural (do percentil 5 ao percentil 25); chá (do percentil 5 ao percentil 75);
- Na suposição de consumo diário: café (percentis 5 e 10) no método MSM.

Pode-se notar que a escolha da suposição de frequência de consumo não alterou a ocorrência dessa situação nos métodos NCI e SPADE.

Observou-se também que ocorreram vícios relativos superiores a 2,5 nos seguintes casos:

- Em ambas as suposições: refrigerante (percentil 5) no método NCI e frutas (percentil 10) nos métodos NCI e SPADE;
- Na suposição de consumo diário: refrigerante (percentil 5) no método SPADE, biscoitos (percentil 10) nos métodos NCI e SPADE, café (média, percentis 90 e 95) nos métodos NCI e SPADE e frutas (percentil 95) no método SPADE.

Por fim, assumindo a suposição de consumo episódico, os métodos NCI e SPADE não conseguiram realizar as estimações dos parâmetros para quase todos os nutrientes, com exceção da vitamina C. Nas Tabelas S2 e S3 do material suplementar estão dispostos os resultados completos dos vícios relativos para as duas suposições.

Na Figura 2 está disposta a comparação dos vícios relativos dos métodos NCI, MSM e SPADE com e sem a utilização da informação do QFA. Esses resultados mostraram que, para componentes alimentares de frequência de consumo episódico de nível intermediário, não houve impacto expressivo

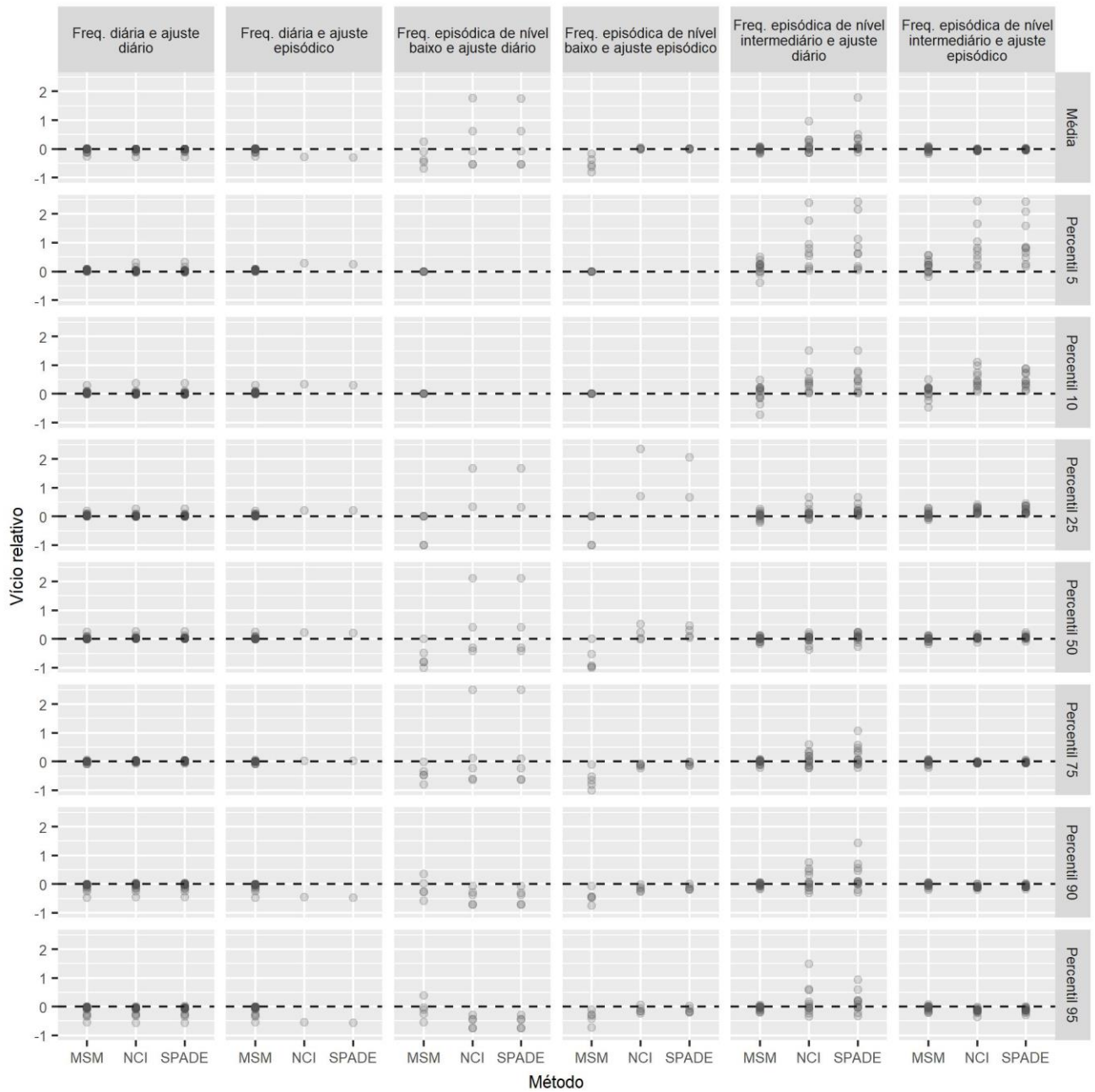
nas estimações dos métodos NCI (para todos os parâmetros), MSM (para a média e percentis maiores ou iguais a 10) e SPADE (para a média e percentis maiores ou iguais a 50). Entretanto, observou-se impacto com a utilização do QFA nos métodos MSM (para o percentil 5) e SPADE (para percentis 5, 10 e 25).

Já para componentes alimentares de frequência de consumo episódico de nível baixo, não houve impacto do QFA nas estimações dos métodos NCI (para todos os parâmetros), MSM (nos percentis 5 e 10) e SPADE (para todos os parâmetros com exceção do percentil 25). Por outro lado, observou-se impacto com a utilização do QFA nos métodos MSM (para o percentil maiores ou iguais a 50) e SPADE (para o percentil 25). Adicionalmente, nas Tabelas S4 e S5 do material suplementar estão organizados os resultados completos dos vícios relativos com e sem aplicação do QFA como informação adicional. Nesses resultados, verificou-se que, de um total de 56 estimativas por método, houve redução do vício relativo ao se usar o QFA: no MSM (29 reduções – principalmente na média e nos percentis acima de 75); no NCI (16 reduções) e no SPADE (21 reduções – basicamente nos percentis abaixo de 25). Ainda para esses resultados, observou-se que os valores mínimos e máximos dos vícios relativos foram: NCI sem QFA  $[-0,23; \infty)$ , NCI com QFA  $[-0,23; \infty)$ , MSM sem QFA  $[-1,00; 0,55]$ , MSM com QFA  $[-1,00; 0,24]$ , SPADE sem QFA  $[-0,18; \infty)$  e SPADE com QFA  $[-1,00; \infty)$ . Chama atenção também que o método NCI obteve vícios relativos superiores a 2,5 – independentemente do uso do QFA – na mensuração do percentil 5 do refrigerante.

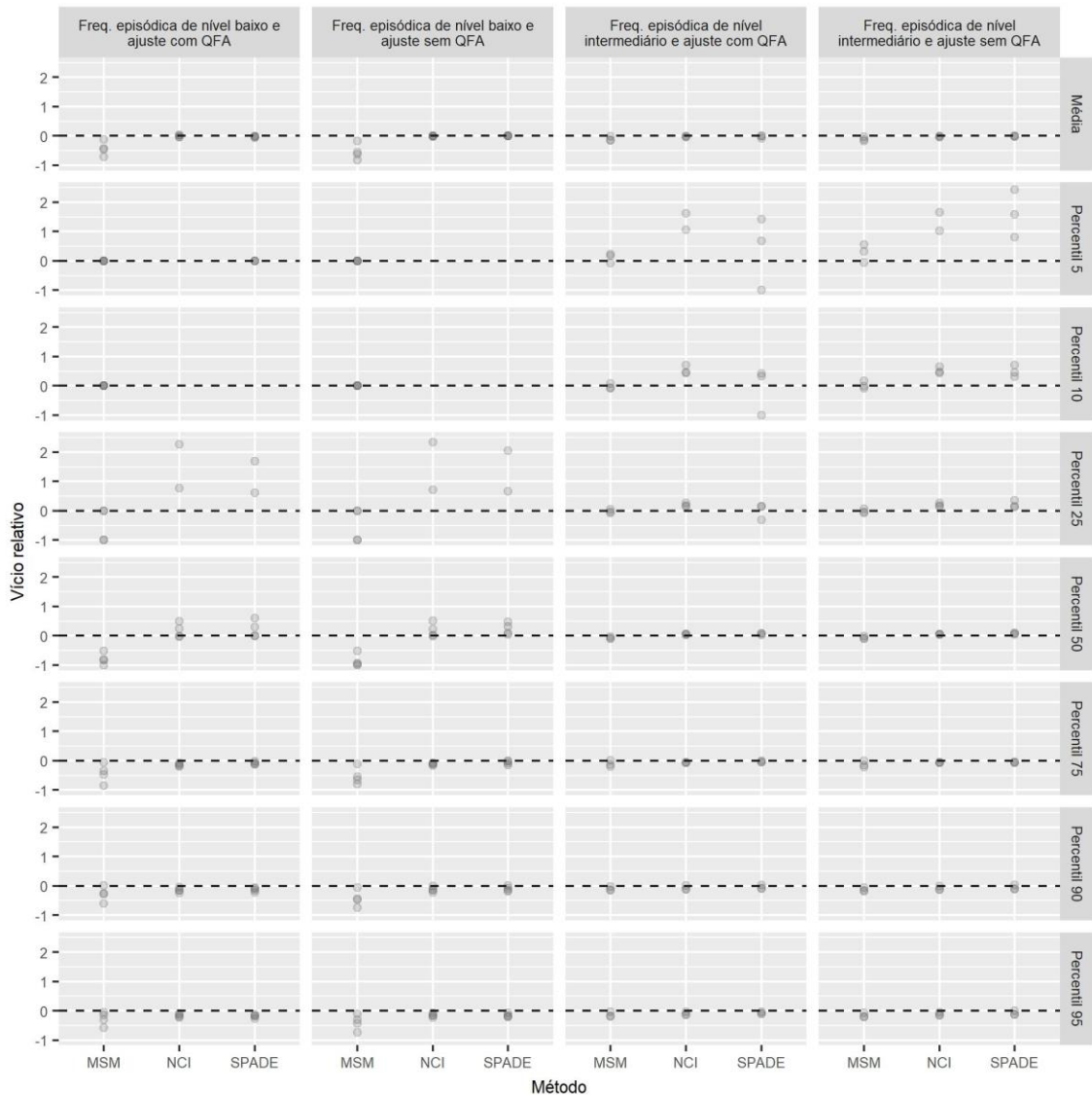
Em suma, o MSM foi o método em que o vício relativo teve o maior número de reduções com a incorporação do QFA. Entretanto, realizando a comparação dos três métodos, verificou-se que o NCI e o SPADE sem o uso do QFA apresentaram valores mais acurados na estimação da média e nos percentis superiores a 25, independentemente da frequência de consumo dos componentes alimentares. Por sua vez, o MSM sem QFA obteve estimativas mais acuradas nos percentis menores ou iguais a 25, independentemente da frequência de consumo dos componentes alimentares.

Ao se fazer a comparação da utilização ou não do QFA, foi identificado que, em alguns casos, alguns dos métodos não conseguiram estimar consumo igual a zero:

- Independentemente do uso do QFA: bolos (percentil 10), ovos (percentis 5 e 10) e pizza e suco natural (percentis 10 e 25) nos métodos NCI e SPADE; bolos (percentil 5), pizza e suco natural (percentil 5) no método NCI.
- Sem emprego do QFA: bolos (percentil 25) no método NCI; bolos, pizza e suco natural (percentil 5).



**Figura 1:** Vícios relativos menores ou iguais a 2,5 para cada método, considerando a frequência de consumo do componente alimentar e tipo de suposição de ajuste.



**Figura 2:** Vícios relativos menores ou iguais a 2,5 para cada método, considerando a frequência de consumo do componente alimentar e o ajuste com ou sem informação do QFA.

## Discussão

A respeito da escolha das suposições sobre a frequência de consumo alimentar, evidenciou-se que é possível considerar a utilização da suposição de consumo diário quando se tem, no máximo, 0,1% de ausência de consumos nos R24h. Esse achado não permite corroborar ou refutar as duas referências existentes na literatura que indicam o uso da suposição de consumo diário em casos em que a frequência de consumos iguais a zero nos R24h seja menor do que 5% <sup>(29)</sup> e do que 10% <sup>(28)</sup>. No entanto, pode-se

afirmar que o uso da suposição de consumo episódico em geral produz estimativas mais acuradas quando mais do que 20% de consumos forem iguais a zero nos R24h.

No que se refere a comparação dos métodos em componentes alimentares de frequência de consumo diário e utilização de suposição de consumo diário, os resultados mostraram que todos os métodos realizaram estimações parecidas entre si com acurácia alta. Além disso, em componentes alimentares com frequência de consumo episódico de nível intermediário e utilização de suposição de consumo episódico, o método MSM foi, em geral, o mais acurado na estimação de todos os parâmetros. Por outro lado, quando se tratou de componentes alimentares de frequência de consumo episódico de nível baixo e utilização de suposição de consumo episódico, os métodos NCI e SPADE foram mais acurados na estimação da média e dos percentis superiores a 25, em contrapartida o método MSM teve acurácia maior na estimação de percentis menores ou iguais a 25.

O emprego do QFA como informação adicional nos modelos, em alguns casos, mostrou uma leve redução dos valores do vício relativo nos métodos SPADE e MSM. Adicionalmente, foi observado que, em algumas estimações dos parâmetros, a utilização do QFA auxiliou o método SPADE na estimação do consumo verdadeiro igual a zero, verificando-se que esse uso ajudou nessa estimação em 5 casos.

Esses resultados corroboram os achados de Goedhart et al. <sup>(31)</sup> e Ost et al. <sup>(33)</sup> que identificaram melhora na estimação de percentis inferiores para o método SPADE quando utilizado a informação de quem nunca consome o componente alimentar, como também dos trabalhos de Tooze et al. <sup>(15)</sup>, Souverein et al. <sup>(22)</sup>, Goedhart et al. <sup>(31)</sup> e Verly et al. <sup>(32)</sup>, que não observaram impacto da utilização do QFA na estimação do método NCI. Apesar dessa consonância de resultados, o mesmo não foi observado para o MSM, uma vez que no presente estudo foi verificado melhora na estimação dos percentis superiores da distribuição e na literatura essa situação não foi relatada. Apesar dessa diferença, os resultados, de forma geral, estão em concordância com os já observados na literatura atual para os métodos SPADE e NCI. Apesar dos métodos MSM e SPADE terem apresentado leve redução dos vícios relativos estimados o custo-benefício da coleta dessas informações deve ser avaliado, visto que a elaboração e aplicação do QFA demanda investimentos que podem não compensar as melhoras obtidas.

Esse estudo possui duas limitações. Uma delas é que não foram mensurados pelo QFA componentes alimentares de frequência de consumo diário, portanto, não foi possível verificar a influência da utilização do QFA na estimação da distribuição de consumo destes componentes. A outra é que não havia alimentos com frequência de consumo na faixa entre 0,1% e 20%, portanto não se pode

concluir a respeito dos pontos de corte sugeridos na literatura (5% e 10% de R24h com consumos iguais a zero).

## Conclusão

Em suma, aconselha-se que se use quaisquer um dos três métodos – com a suposição de consumo diário – na estimação da distribuição de consumo usual de componentes com frequência de consumo diário. Sugere-se também a utilização do método MSM – com suposição de consumo episódico – na estimação para componentes alimentares com frequência de consumo episódico de nível intermediário. Além disso, nas frequências de consumo episódico de nível baixo, o método MSM é indicado na estimação dos percentis menores ou iguais a 25 e os métodos SPADE e NCI para estimar a média e os percentis maiores ou iguais a 50. Por fim, foi verificado que a adição das informações do QFA, apesar de causar aumento de acurácia nas estimativas dos métodos MSM (principalmente para a média e percentis acima de 75) e SPADE (sobretudo para percentis abaixo de 25), não trouxe melhora expressiva nos resultados dos métodos.

## Referências

1. Rossum, CTM, Fransen, HP, Verkaik-Kloosterman, J, et al. (2011) *Dutch National Food Consumption Survey 2007-2010*. 33–37. National Institute for Public Health and the Environment.
2. U.S. Department of Agriculture and U.S. Department of Health and Human Services (2010) *Dietary Guidelines for Americans, 2010*. 7th Edition, Washington, DC: U.S. Government Printing Office.
3. on behalf of the EFCOVAL Consortium, de Boer EJ, Slimani N, et al. (2011) The European Food Consumption Validation Project: conclusions and recommendations. *Eur. J. Clin. Nutr.* **65**, S102–S107.
4. Willett W (2013) *Nutritional epidemiology*. Third edition. Oxford ; New York: Oxford University Press.
5. Dodd KW, Guenther PM, Freedman LS, et al. (2006) Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J. Am. Diet. Assoc.* **106**, 1640–1650.
6. Slob W (1993) Modeling Long-Term Exposure of the Whole Population to Chemicals in Food. *Risk Anal.* **13**, 525–530.
7. Gay C (2000) Estimation of population distributions of habitual nutrient intake based on a short-run weighed food diary. *Br. J. Nutr.* **83**, 287–293.

8. Wallace LA, Duan N & Ziegenfus R (1994) Can Long-Term Exposure Distributions Be Predicted from Short-Term Measurements? *Risk Anal.* **14**, 75–85.
9. Buck RJ, Hammerstrom KA & Ryan PB (1995) Estimating long-term exposures from short-term measurements. *J. Expo. Anal. Environ. Epidemiol.* **5**, 359–373.
10. Nusser SM, Carriquiry AL, Dodd KW, et al. (1996) A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *J. Am. Stat. Assoc.* **91**, 1440–1449.
11. Guenther PM, Kott PS & Carriquiry AL (1997) Development of an Approach for Estimating Usual Nutrient Intake Distributions at the Population Level. *J. Nutr.* **127**, 1106–1112.
12. Chang H-Y, Suchindran CM & Pan W-H (2001) Using the overdispersed exponential family to estimate the distribution of usual daily intakes of people aged between 18 and 28 in Taiwan. *Stat. Med.* **20**, 2337–2350.
13. EFCOSUM Group, Hoffmann K, Boeing H, et al. (2002) Estimating the distribution of usual dietary intake by short-term measurements. *Eur. J. Clin. Nutr.* **56**, S53–S62.
14. Tooze JA, Grunwald GK & Jones RH (2002) Analysis of repeated measures data with clumping at zero. *Stat. Methods Med. Res.* **11**, 341–355.
15. Tooze JA, Midthune D, Dodd KW, et al. (2006) A New Statistical Method for Estimating the Usual Intake of Episodically Consumed Foods with Application to Their Distribution. *J. Am. Diet. Assoc.* **106**, 1575–1587.
16. Slob W (2006) Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. *Food Chem. Toxicol.* **44**, 933–951.
17. Waijers PMCM, Dekkers ALM, Boer JMA, et al. (2006) The Potential of AGE MODE, an Age-Dependent Model, to Estimate Usual Intakes and Prevalences of Inadequate Intakes in a Population. *J. Nutr.* **136**, 2916–2920.
18. Staudenmayer J, Ruppert D & Buonaccorsi JP (2008) Density Estimation in the Presence of Heteroscedastic Measurement Error. *J. Am. Stat. Assoc.* **103**, 726–736.
19. Kipnis V, Midthune D, Buckman DW, et al. (2009) Modeling Data with Excess Zeros and Measurement Error: Application to Evaluating Relationships between Episodically Consumed Foods and Health Outcomes. *Biometrics* **65**, 1003–1010.
20. Tooze JA, Kipnis V, Buckman DW, et al. (2010) A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. *Stat. Med.* **29**, 2857–2868.
21. Haubrock J, Nöthlings U, Volatier J-L, et al. (2011) Estimating Usual Food Intake Distributions by Using the Multiple Source Method in the EPIC-Potsdam Calibration Study. *J. Nutr.* **141**, 914–920.
22. Souverein, OW, Dekkers AL, Haubrock J, et al. (2011) Comparing four methods to estimate usual intake distributions. *Eur. J. Clin. Nutr.* **65**, S92–S101.



23. Nusser SM, Fuller WA & Guenther PM (2012) Estimating Usual Dietary Intake Distributions: Adjusting for Measurement Error and Nonnormality in 24-Hour Food Intake Data. In *Surv. Meas. Process Qual.*, pp. 689–709 [Lyberg L, Biemer P, Collins M, et al., editors]. Hoboken, NJ, USA: John Wiley & Sons, Inc.
24. Dekkers ALM & Slob W (2012) Gaussian Quadrature is an efficient method for the back-transformation in estimating the usual intake distribution when assessing dietary exposure. *Food Chem. Toxicol.* **50**, 3853–3861.
25. Dekkers AL, Verkaik-Kloosterman J, van Rossum CT, et al. (2014) SPADE, a New Statistical Program to Estimate Habitual Dietary Intake from Multiple Food Sources and Dietary Supplements. *J. Nutr.* **144**, 2083–2091.
26. Chernova J & Solis-Trapala I (2016) A simplified approach to estimating the distribution of occasionally-consumed dietary components, applied to alcohol intake. *BMC Med. Res. Methodol.* **16**.
27. Harttig U, Haubrock J, Knüppel S, et al. (2011) The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *Eur. J. Clin. Nutr.* **65**, S87–S91.
28. Guenther PM, Kirkpatrick SI, Reedy J, et al. (2014) The Healthy Eating Index-2010 Is a Valid and Reliable Measure of Diet Quality According to the 2010 Dietary Guidelines for Americans. *J. Nutr.* **144**, 399–407.
29. Usual Dietary Intakes: Food Intakes, U.S. Population, 2007-10. <https://epi.grants.cancer.gov/diet/usualintakes/pop/2007-10/>.
30. Subar AF, Dodd KW, Guenther PM, et al. (2006) The Food Propensity Questionnaire: Concept, Development, and Validation for Use as a Covariate in a Model to Estimate Usual Food Intake. *J. Am. Diet. Assoc.* **106**, 1556–1563.
31. Goedhart PW, van der Voet H, Knüppel S, et al. (2012) A comparison by simulation of different methods to estimate the usual intake distribution for episodically consumed foods. *EFSA - Support. Publ.*, EN299 ed.
32. Verly- E, Fisberg RM & Marchioni DML (2012) Is the food frequency consumption essential as covariate to estimate usual intake of episodically consumed foods? *Eur. J. Clin. Nutr.* **66**, 1254–1258.
33. Ost C, De Ridder KAA, Tafforeau J, et al. (2017) The added value of food frequency questionnaire (FFQ) information to estimate the usual food intake based on repeated 24-hour recalls. *Arch. Public Health* **75**.
34. Laureano GHC, Torman VBL, Crispim SP, et al. (2016) Comparison of the ISU, NCI, MSM, and SPADE Methods for Estimating Usual Intake: A Simulation Study of Nutrients Consumed Daily. *Nutrients* **8**, 166.

35. Verly-Jr E, Oliveira DCRS, Fisberg RM, et al. (2016) Performance of statistical methods to correct food intake distribution: comparison between observed and estimated usual intake. *Br. J. Nutr.* **116**, 897–903.
36. SAS Institute (2011) *Base SAS 9.3 procedures guide*. Cary, N.C.: SAS Institute.
37. R Core Team (2015) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

## **6. ARTIGO 2**

### **Impacto da utilização de covariáveis na estimação da distribuição de consumo alimentar usual nos métodos NCI, MSM e SPADE**

Impact of covariable use in usual dietary intake distribution estimation for NCI, MSM and SPADE methods

Greice H. C. Laureano, Vanessa B. Leotti e Suzi A. Camey

**Artigo a ser submetido ao jornal *British Journal of Nutrition***

## Resumo

Quando se estimam distribuições de consumo alimentar usual de populações, é intuitivo que se pense na utilização de informações adicionais ao próprio consumo alimentar (covariáveis), como, por exemplo, sexo e idade, para que as estimações se aproximem o máximo possível da realidade. Apesar de ser uma prática comum na literatura, há poucos estudos a respeito da influência do emprego de covariáveis nos modelos de estimação das distribuições de consumo alimentar usual. Em vista disso, este trabalho visou mensurar a influência da incorporação das covariáveis sexo e idade na estimação da distribuição de consumo alimentar dos métodos *National Cancer Institute* (NCI), *Multiple Source Method* (MSM) e *Statistical Program to Assess Dietary Exposure* (SPADE), através de dados de uma pesquisa com múltiplos recordatórios alimentares de 24 horas (R24h) aplicados em residentes na cidade do Rio de Janeiro, Brasil.

Na análise dos dados, a frequência de consumo dos componentes alimentares foi classificada dessa forma: diário (no máximo 0,1% de consumos iguais a zero nos R24h), episódico de nível intermediário (quando a proporção de R24h iguais a zero estiver entre 20% e 75%) e episódico de nível baixo (no mínimo 76% dos R24h com medições iguais a zero). Na averiguação do impacto da utilização de sexo e idade como covariáveis nos métodos, verificou-se que os métodos não foram impactados na estimação de componentes alimentares de frequência de consumo diário. Por outro lado, o MSM obteve maior redução do vício relativo na estimação de componentes alimentares de frequência de consumo episódico de nível baixo, enquanto o NCI teve com uma maior frequência um aumento do vício relativo quando utilizou covariáveis para componentes alimentares com as frequências de consumo episódico de nível intermediário e baixo. Já o SPADE, em geral, não teve modificações nos vícios relativos quando se utilizou covariáveis.

Frente ao exposto, recomenda-se o emprego dos métodos sem ajuste por sexo e idade na estimação de componentes alimentares de frequência de consumo diário e episódico de nível intermediário. Já para componentes alimentares com frequência de consumo episódico de nível baixo, indica-se o método MSM com ajuste por sexo e idade e os métodos NCI e SPADE sem ajuste.

**Palavras-chave:** NCI. MSM. SPADE. Covariáveis. Sexo. Idade. R24h.

## Introdução

Quando se trata do consumo alimentar, é esperado que haja diferenças entre grupos populacionais<sup>(1,2)</sup>, e assim, para se realizar pesquisas que avaliam tais diferenças, é necessário aferir quais situações podem influenciar nesse consumo, além de mensurar a quantidade e a frequência de consumo dos componentes alimentares analisados. Uma ferramenta que permite medir tanto a frequência quanto a quantidade de consumo é o recordatório alimentar de 24 horas (R24h), enquanto que, para se avaliar as demais situações que podem influenciar no consumo, usualmente são utilizadas covariáveis como sexo, etnia, idade, dentre outras.

Existem diferentes métodos que se propõem a realizar a estimação das distribuições de consumo alimentar usual<sup>(3-23)</sup>. A maioria desses métodos permite o uso de covariáveis na modelagem para que se possa reduzir erros de estimação das distribuições, sendo que dois deles foram criados especificamente com a ideia de se empregar a idade como informação adicional na modelagem: o *Age-dependent Dietary Assessment Model* (AGE MODE)<sup>(14)</sup> e o *Statistical Program to Assess Dietary Exposure* (SPADE)<sup>(21,22,24)</sup> que foi derivado do AGE MODE.

Embora a utilização de covariáveis seja comum nessa área de pesquisa, há escassez de análises que compararam a influência dessa utilização nos resultados dos métodos. Ao passo que nos estudos disponíveis foi realizada a comparação do uso de covariáveis com foco na influência da utilização do Questionário de Frequência Alimentar (QFA) como covariável<sup>(25,26,24,27-30)</sup>. A utilização de estratificações por sexo e idade nos estudos que se referem ao consumo alimentar de populações é comum<sup>(1,31,32)</sup>, de maneira que esses resultados segmentados podem ser obtidos com o uso de covariáveis nos modelos de estimação ou com o ajuste dos modelos em cada estrato de interesse. Em Tooze et. al<sup>(17)</sup> foi analisado se utilizar o NCI com sexo e idade no ajuste era mais apropriado do que usar o NCI em grupos estratificados dos dados. Para isso se utilizou dados dos nutrientes cálcio, ferro e vitamina A do estudo *Eating at America's Table*<sup>(33)</sup> e o teste de razão de verossimilhanças, chegando-se à conclusão que os resultados com o ajuste das informações de sexo e idade tem melhores resultados que o estratificado. Apesar da utilização das informações de sexo e idade serem comuns, não há estudos que meçam a influência do uso dessas covariáveis no ajuste dos métodos de estimação da distribuição de consumo usual.

Diante da falta de literatura a respeito do uso de covariáveis que não sejam o QFA, este trabalho realizou a comparação dos métodos *National Cancer Institute* (NCI)<sup>(11,12,16,17)</sup>, *Multiple Source Method* (MSM)<sup>(18,19)</sup> e SPADE quando utilizadas as covariáveis sexo e idade na modelagem com o objetivo de

mensurar o impacto delas na acurácia dos métodos de estimação da distribuição de consumo alimentar usual. Fez-se uma comparação empírica dos métodos a partir de dados de Verly et al.<sup>(34)</sup>.

## **Método**

Nesse estudo foram empregados os dados da pesquisa de Verly et al.<sup>(34)</sup>, referentes ao R24h de 20 dias não consecutivos de 302 pessoas residentes na cidade do Rio de Janeiro, Brasil. Além do emprego dos R24h, também foram coletadas informações de sexo e idade dos participantes. Os componentes alimentares selecionados para a análise são os dispostos a seguir:

- Alimentos: arroz, carnes em geral, vegetais totais, pães (excluindo os recheados), feijão, café, derivados do leite, refrigerante, frutas, carne de gado, doces e chocolates, carne de frango, biscoitos, suco natural, bolos, ovos, pizza e chá.
- Macronutrientes: carboidratos, lipídios totais, proteínas, e gorduras saturadas.
- Micronutrientes: zinco, cobre, cálcio, fósforo, potássio, sódio, folato diário, magnésio, vitamina A, vitamina B1, vitamina B2, vitamina B6 e vitamina C.

Mais detalhes a respeito destes dados estão dispostos em Verly et al.<sup>(34)</sup>.

Na análise dos dados, foi utilizada a seguinte classificação da frequência de consumo dos componentes alimentares:

- Episódico de nível baixo – no mínimo 76% de consumos iguais a zero nos R24h;
- Episódico de nível intermediário – entre 20% e 75% de consumos iguais a zero nos R24h;
- Diário (nutrientes) – no máximo 0,1% de consumos iguais a zero nos R24h.

Nessas classificações apenas os nutrientes ficaram na faixa de consumo diário e não há componentes alimentares com proporção de consumos iguais a zero entre 0,1% e 20%. Detalhes adicionais a respeito das frequências de consumo dos componentes alimentares e os critérios de escolha para os mesmos foram dispostos em Laureano et al.<sup>(35)</sup>.

Neste estudo foram utilizados os métodos de estimação da distribuição de consumo alimentar usual NCI, MSM e SPADE. Salienta-se que, apesar de terem diferenças metodológicas entre si, todos incorporam as covariáveis de forma similar: com a suposição de consumo diário, os métodos usam apenas a informação de quantidade consumida na estimação e, em vista disso, empregam as covariáveis apenas nessa parte do modelo. Porém, quando se trata da suposição de consumo episódico, os métodos fazem uso tanto da informação da quantidade quanto da frequência consumida, fazendo com que as informações das covariáveis, por sua vez, possam ser utilizadas em ambas as partes do modelo. Neste

caso, o método MSM sempre utiliza as mesmas covariáveis nas duas partes do modelo, enquanto que o NCI e o SPADE permitem a escolha das covariáveis para as duas partes de forma distinta <sup>(26,16,19,22)</sup>.

Para realizar os ajustes dos métodos, foram usados os 20 R24h dos 285 sujeitos que tinham informação de sexo e idade. Esses ajustes foram feitos para cada método utilizando a suposição de consumo diário (em componentes alimentares com frequência de consumo diário, ou seja, os nutrientes) ou a suposição de consumo episódico (em componentes alimentares com frequência de consumo episódico de nível intermediário ou baixo). Por outro lado, com o intuito de aproveitar toda a informação disponível, o verdadeiro valor dos parâmetros foi baseado na distribuição empírica das médias dos 20 R24h dos 302 indivíduos, e estão dispostos no material suplementar de Laureano et al. <sup>(35)</sup>. Tanto as estimativas quanto os valores verdadeiros foram calculados para a população geral, isto é, sem estratificação por sexo ou idade.

O Vício Relativo (VR) foi a medida de acurácia escolhida para a aferição da influência do uso de covariáveis na estimação da distribuição de consumo alimentar usual, sendo definido por

$$VR = \begin{cases} \frac{\hat{\theta} - \theta}{\theta}, & \text{se } \theta > 0 \\ 0, & \text{se } \theta = 0 \text{ e } \hat{\theta} = 0 \\ \infty, & \text{se } \theta = 0 \text{ e } \hat{\theta} > 0 \end{cases} \quad (1)$$

onde  $\hat{\theta}$  é o valor estimado do parâmetro e  $\theta$  é o valor verdadeiro do parâmetro. Quando o valor verdadeiro do parâmetro for igual a zero, há duas possíveis circunstâncias: VR igual a infinito (cada vez que o valor estimado é diferente de zero – quando o método não foi capaz de estimar consumo igual a zero) e VR igual a zero (se o valor estimado é igual a zero).

Também foi utilizada a diferença entre os VR's (DVR) para avaliar o impacto do emprego das covariáveis na estimação da distribuição de consumo alimentar usual, essa diferença foi definida por:

$$DVR = |VR_{com\ covar}| - |VR_{sem\ covar}| \quad (2)$$

onde  $|VR_{com\ covar}|$  é o valor absoluto do VR com covariáveis e  $|VR_{sem\ covar}|$  é o valor absoluto do VR sem covariáveis. Na análise dos resultados do DVR foi considerada a faixa de tolerância de [-0,05; 0,05] e classificadas as seguintes situações na estimação:

- Piorou:  $DVR > 0,05$ , quando o VR aumentou com o uso das covariáveis.

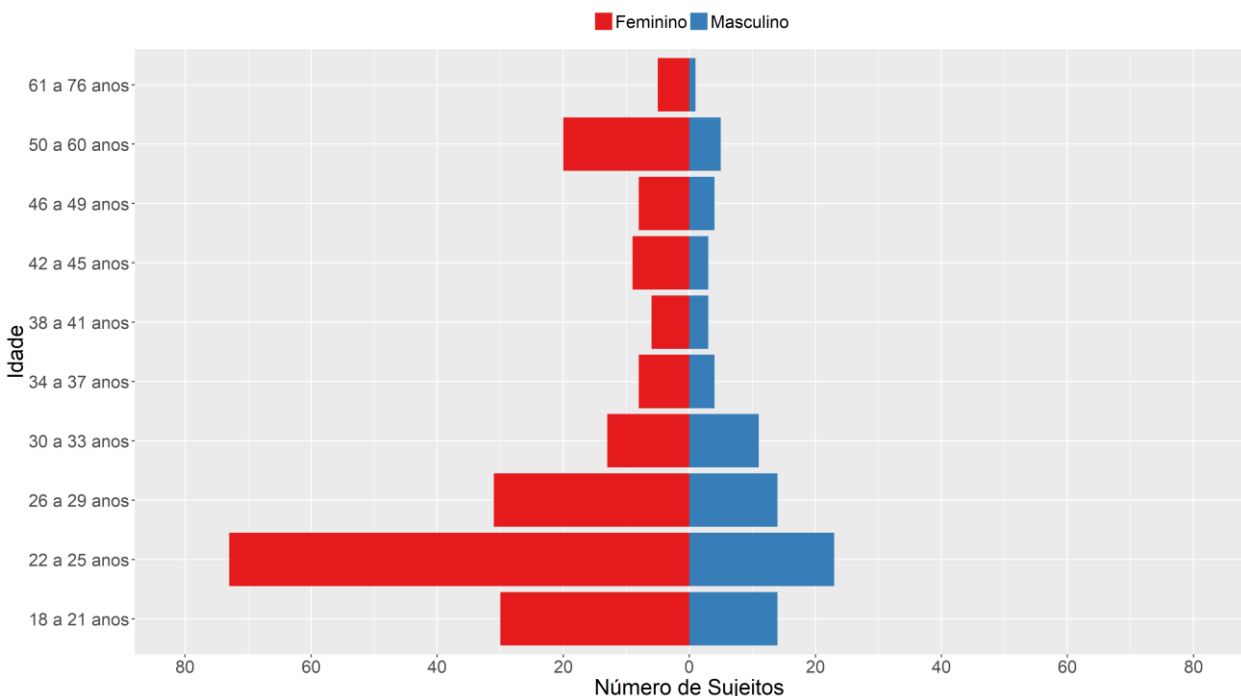
- Não alterou: DVR dentro do intervalo  $[-0,05; 0,05]$ , quando a diferença dos VR's for irrelevante (até 5% em termos absolutos).
- Melhorou:  $DVR < -0,05$ , quando o VR diminuiu com o uso das covariáveis.

Para facilitar a visualização dos resultados, em algumas análises, os VR's maiores do que 2,5 foram excluídos. Evidencia-se que um VR maior do que 2,5 é referente a um valor que indicaria uma superestimação do valor verdadeiro de no mínimo 250%.

Para o ajuste do método NCI foi empregado o *software Statistical Analysis System (SAS)* <sup>(36)</sup> versão 9.3, no ajuste do método SPADE e diagnóstico dos resultados foi utilizado o *software R* <sup>(37)</sup> versão 3.3.3., já no método MSM foi usada a sua interface na *web*: <https://msm.dife.de/>.

## Resultados

Na Figura 1 está disposta a pirâmide etária dos sujeitos utilizados no ajuste dos métodos. Nesse estudo há uma maior proporção de mulheres (71%) quando comparada a de homens, assim como a maioria dos sujeitos tem até 29 anos de idade (65%).



**Figura 1:** Pirâmide etária dos participantes do estudo

Na Figura 2 está a comparação dos vícios relativos menores ou iguais a 2,5 dos métodos NCI, MSM e SPADE considerando a utilização ou não de sexo e idade como covariáveis. Nela se verificou que, para componentes alimentares de frequência de consumo diário, não há impacto do ajuste por sexo

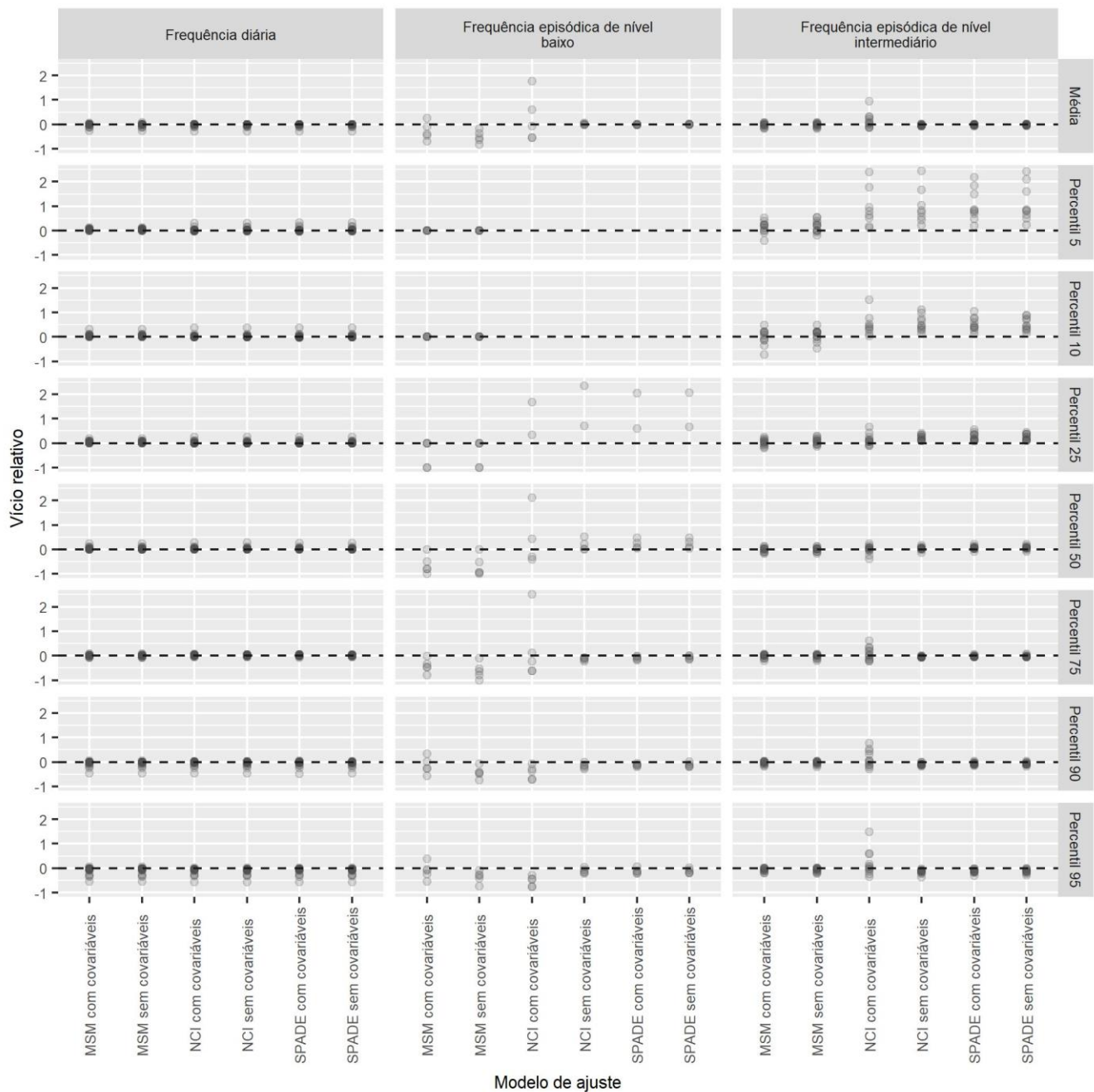


e idade. Por outro lado, quando são considerados alimentos com frequência de consumo episódico de nível intermediário, o maior impacto do ajuste por sexo e idade ocorreu no método NCI, com aumento do intervalo interquartil do vício relativo na estimação da média, do percentil 5 e dos percentis maiores ou iguais a 75. Todavia, quando foram considerados alimentos com frequência de consumo episódico de nível baixo, verificou-se impacto do ajuste por sexo e idade em todos os métodos, à medida que no MSM houve redução do intervalo interquartil quando utilizado o ajuste na estimação da média e dos percentis maiores ou iguais a 50. Já no NCI houve aumento do intervalo interquartil na estimação da média e dos percentis maiores ou iguais a 25. Por fim, o SPADE sofreu impacto do ajuste por sexo e idade na estimação do percentil 25. Na Tabela S1 do material suplementar estão dispostos os resultados completos dos vícios relativos com e sem ajuste por sexo e idade. Nesses resultados se observou que eles variaram da seguinte forma na estimação para componentes alimentares de frequência de consumo episódico de nível intermediário e baixo:

- Frequência de consumo episódico de nível intermediário :
  - Sem ajuste: NCI  $[-0,36; \infty)$ , MSM  $[-0,47; 0,55]$  e SPADE  $[-0,29; \infty)$ ;
  - Com ajuste: NCI  $[-0,38; \infty)$ , MSM  $[-0,73; \infty)$  e SPADE  $[-0,30; \infty)$ .
- Frequência de consumo episódico de nível baixo:
  - Sem ajuste: NCI  $[-0,28; \infty)$ , MSM  $[-1,00; 0,00]$  e SPADE  $[-0,21; \infty)$ ;
  - Com ajuste: NCI  $[-0,75; \infty)$ , MSM  $[-1,00; 0,38]$  e SPADE  $[-0,22; \infty)$ .

Tem-se que os métodos NCI e SPADE não conseguiram estimar consumo igual a zero, independentemente do uso de covariáveis, nos casos a seguir: biscoitos e frutas (percentil 5); bolos, café, e ovos (percentil 5 e 10); pizza e suco natural (do percentil 5 ao percentil 25); e chá (do percentil 5 ao percentil 50). Já o método MSM não conseguiu estimar consumo igual a zero no percentil 5 e 10 quando se utilizou covariáveis na estimação para o café.

Adicionalmente, ocorreram vícios relativos superiores a 2,5, independentemente do emprego de covariáveis nos métodos NCI e SPADE na estimação do percentil 10 para as frutas. No método NCI houve essa situação, independentemente do uso de covariáveis para percentil 5 do refrigerante, mas ocorreram esses valores de vício relativo apenas quando se utilizou covariáveis para o biscoito (percentil 10) e para o café (média, percentil 90 e 95). Com isso, pode-se verificar que a inserção de covariáveis aumentou a ocorrência de vícios relativos superiores a 2,5 no método NCI.



**Figura 2:** Vícios relativos menores ou iguais a 2,5 para cada componente alimentar, o método com ou sem utilização de covariáveis.

Na Tabela 1 estão dispostos os DVR's por método, para componentes alimentares com frequência de consumo episódico de nível intermediário ou baixo. Nela se constatou que em componentes alimentares com frequência de consumo episódico de nível intermediário os DVR's dos métodos variaram de -2,13 a  $\infty$  no NCI, de -0,28 a  $\infty$  no MSM e de -0,26 a  $\infty$  no SPADE. Nesses resultados, o

ajuste por sexo e idade provocou a maior redução do vício relativo no percentil 10 das frutas com o uso do NCI (DVR = -2,13) e o maior aumento do vício relativo no percentil 95 do café usando o NCI (DVR = 7,1). Adicionalmente, em componentes alimentares com frequência de consumo episódico de nível baixo, os DVR's dos métodos variaram de -0,68 a  $\infty$  no NCI, de 0,53 a 0,11 no MSM e de -0,1807 a  $\infty$  no SPADE. À medida que o ajuste por sexo e por idade provocou a maior redução do vício relativo nos bolos para o percentil 10 usando o NCI (DVR = -0,68) e o maior aumento do vício relativo no percentil 75 do chá utilizando o NCI (DVR = 2,26). Com esses resultados, foi possível verificar que a utilização das covariáveis causaram impacto na estimação dos métodos, principalmente no método NCI, exceto para a frequência de consumo diário (resultados disponíveis na Tabela S2 do material suplementar) onde todos os métodos obtiveram a maioria dos DVR's iguais a zero e com valor máximo de 0,02 (em termos absolutos).

Na Figura 3 foi apresentado o impacto do uso das covariáveis sexo e idade na acurácia das estimativas para os componentes alimentares de frequências de consumo episódico de nível intermediário e baixo. Nota-se que, entre os componentes alimentares de consumo intermediário, o maior impacto do ajuste por sexo e por idade ocorreu nas estimativas obtidas pelo NCI, pois ele apresentou maior proporção de pioras (para todos os parâmetros) e de melhoras (percentis 5, 10, 25, 50, 90 e 95) quando comparado aos demais métodos. Para os métodos MSM e SPADE, o impacto da inclusão das covariáveis foi semelhante, sendo que não houve alteração relevante da acurácia na estimação da média e dos percentis maiores ou iguais a 50. Em contraposição, em componentes alimentares frequência de consumo episódico de nível baixo, os maiores impactos do ajuste por sexo e por idade ocorreram nos métodos NCI e MSM, ao passo que no NCI houve predominância de piora (na estimação da média e do percentil 25 ao 95) e no MSM houve predominância de melhora (média e dos percentis 50 a 95). Já, para o método SPADE, na maioria dos casos, não houve alteração nos resultados das estimativas.

Ao se analisar individualmente os itens alimentar que tiverem melhora ou piora com o ajuste por sexo e idade verificou-se:

- **Melhora no MSM:** bolos, ovos (média e percentis 50, 75, 90 e 95); chá (média, percentis 75 e 90); pizza (média e percentis 90 e 95); suco natural (média e percentis 75 e 95); refrigerante (percentis 5 e 10); frutas (percentil 10), café (percentil 25); carne de gado (percentil 5).
- **Piora no MSM:** doces e chocolates (percentis 5 e 10); carne de gado (percentil 10); biscoitos (percentis 10 e 25); frutas (percentil 25); chá (percentil 95).

- **Melhora no NCI:** arroz (percentis 5, 90 e 95), derivados do leite (percentis 5, 10, 25 e 95), doces e chocolates, frango (percentis 5, 10 e 25), vegetais totais (percentis 5 e 10), frutas (percentis 10, 25 e 50), bolos, café, ovos (percentil 25); pães (excluindo os recheados) percentis 25, 90 e 95; pizza (percentil 75); carnes em geral (percentis 90 e 95); chá e refrigerante (percentil 90).
- **Piora no NCI:** carne de gado (percentis 5, 10, 25, 90 e 95); feijão (média, percentis 5, 50, 75, 90 e 95); biscoitos (média, percentis 10, 25, 75, 90 e 95); bolo, café, ovos, vegetais totais (média, percentis 50, 75, 90 e 95); chá (média, percentis 75 e 95); derivados do leite (média, percentis 75 e 90); doces e chocolates (média, percentis 50, 75 e 90); carne de frango, frutas (média, percentis 75, 90 e 95); pizza (média, percentis 50, 90 e 95); suco natural (percentis 50, 75, 90 e 95); refrigerante (percentis 5, 10 e 25); pães (excluindo os recheados) percentil 5; arroz (percentis 50 e 75).
- **Melhora no SPADE:** carne de gado (percentil 5), doces e chocolates (percentis 5, 10 e 25), refrigerante (percentis 5 e 10) e ovos (percentis 25 e 95).

Ao se fazer a comparação da utilização das covariáveis, verificou-se piora nos seguintes casos:

- **Piora no SPADE:** carne de frango (percentil 5), biscoitos (percentis 10 e 25), frutas (percentil 10) e café (percentil 25).

Nas Figuras S1, S2 e S3 do material suplementar foram disponibilizadas a proporção de casos que houve piora, melhora ou não houve modificação das estimativas ao se utilizar sexo e idade no ajuste, considerando DVR's sem faixa de tolerância e com diferentes faixas de tolerância: [-0,01; 0,01], [-0,025; 0,025], [-0,05; 0,05] para todas as classificações de consumo. Nesses gráficos é possível observar que as faixas de tolerância não causaram impacto expressivo nos métodos NCI e MSM. Entretanto, para o método SPADE, houve redução nos casos de piora e melhora ao se aumentar a faixa de tolerância.

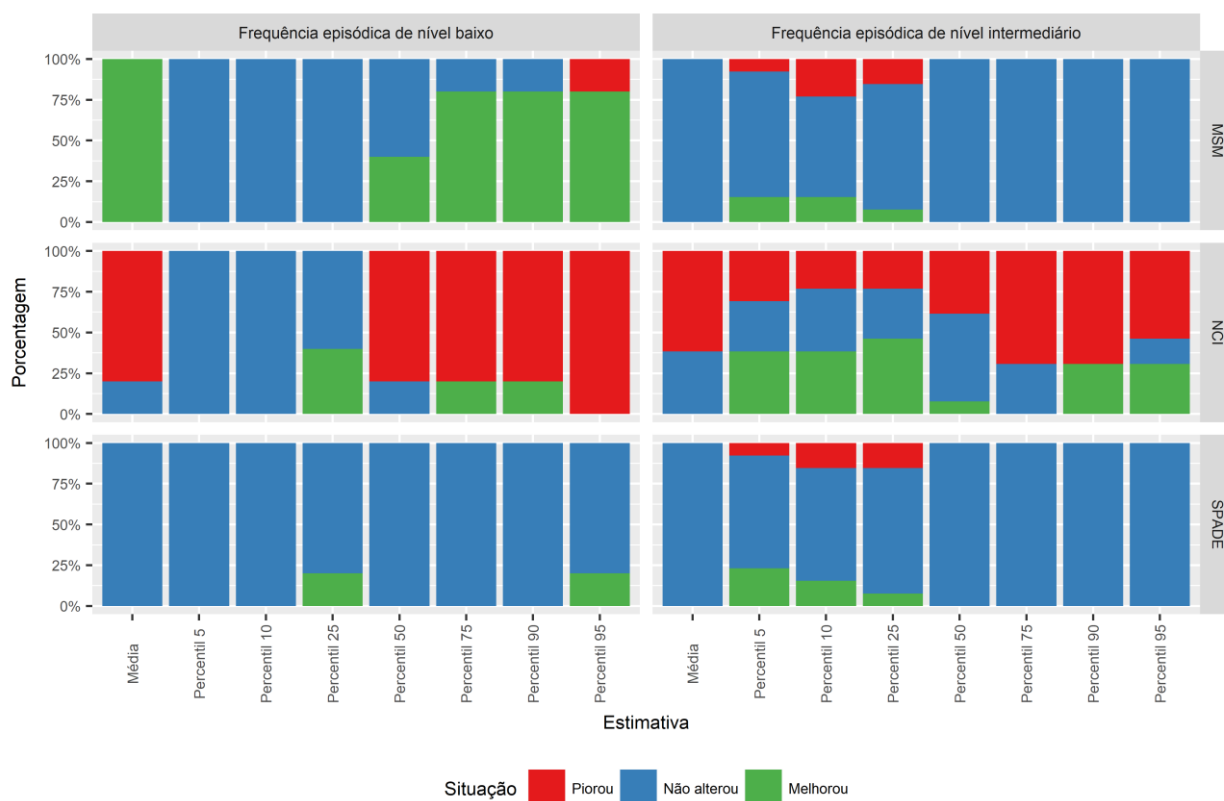
**Tabela 1:** Diferença do vício relativo (DVR) com e sem covariáveis para componentes alimentares de frequência episódico de nível intermediário ou baixo.

Componente alimentar	Frequência de consumo	MSM								NCI								SPADE							
		Média	P5	P10	P25	P50	P75	P90	P95	Média	P5	P10	P25	P50	P75	P90	P95	Média	P5	P10	P25	P50	P75	P90	P95
Arroz	Episódico de nível Intermediário	-0,01	0,00	0,00	-0,01	0,00	-0,01	0,00	-0,01	-0,03	-0,12	-0,05	0,02	0,07	0,08	-0,07	-0,06	-0,03	0,02	0,02	0,03	0,03	0,03	-0,03	-0,03
Biscoitos		0,00	0,00	0,26	0,08	0,00	0,00	0,00	-0,03	0,08	I	<b>1,65</b>	0,38	0,01	0,15	0,21	0,21	0,00	I	0,17	0,08	0,01	0,00	0,02	0,04
Carne de gado		0,02	-0,28	0,06	0,00	0,00	-0,01	-0,02	-0,01	0,02	<b>0,73</b>	0,29	0,11	0,01	0,05	0,09	0,10	0,00	-0,10	-0,05	-0,03	-0,01	0,00	-0,01	-0,02
Café		-0,01	I+	I+	-0,3	0,00	0,02	0,01	-0,01	<b>6,12</b>	I	I	-0,24	0,26	0,54	<b>3,0</b>	<b>7,1</b>	0,01	I	I	0,18	0,01	0,02	0,02	0,02
Carnes em geral		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	-0,04	-0,02	0,00	0,05	0,00	-0,09	-0,09	0,01	-0,05	-0,03	-0,02	-0,01	0,01	0,01	0,00
Derivados do leite		0,00	-0,04	-0,01	0,00	0,00	0,00	0,00	0,00	0,17	-0,29	-0,30	-0,20	-0,04	0,17	0,21	-0,27	0,00	-0,01	-0,01	-0,02	-0,02	-0,02	0,02	0,02
Doces e chocolates		0,00	0,2	0,12	0,00	0,03	0,00	0,00	-0,01	0,11	-0,68	-0,45	-0,09	0,21	0,21	0,12	0,02	-0,02	-0,26	-0,09	-0,06	-0,04	0,02	-0,01	-0,01
Feijão		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,24	0,06	0,02	0,03	0,13	0,29	0,36	0,51	-0,01	-0,03	0,02	0,00	0,01	0,01	0,00	0,00
Carne de frango		0,00	0,02	0,05	0,03	0,02	0,00	-0,01	-0,01	0,10	-0,08	-0,09	-0,11	0,01	0,11	0,11	0,08	-0,01	0,06	0,05	0,03	0,01	-0,01	0,00	0,01
Frutas		0,00	0,00	-0,07	0,12	0,00	0,00	-0,01	0,00	0,87	I	<b>-2,13</b>	-0,37	-0,14	0,14	0,57	<b>1,26</b>	-0,01	I	0,11	-0,01	-0,01	0,01	-0,02	-0,04
Pães (excluindo os recheados)		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	-0,03	-0,06	0,01	-0,03	-0,06	-0,17	-0,01	0,05	0,02	0,02	0,01	-0,01	-0,01	-0,01
Refrigerante		0,00	-0,21	-0,09	-0,03	0,01	0,00	0,00	0,00	0,05	<b>3,30</b>	0,85	0,17	-0,03	-0,02	-0,07	0,02	-0,02	-0,23	-0,11	-0,02	0,00	-0,02	-0,03	-0,02
Vegetais totais		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,30	-0,33	-0,22	-0,05	0,17	0,33	0,47	0,43	-0,01	-0,01	0,01	0,01	0,01	-0,01	-0,01	0,00
Bolos	Episódico de nível baixo	-0,16	0,00	0,00	0,00	-0,11	-0,18	-0,18	-0,17	0,54	I	I	-0,68	0,38	0,50	0,59	0,58	0,02	I	I	-0,02	-0,01	0,03	0,04	0,04
Chá		-0,13	0,00	0,00	0,00	0,00	-0,53	-0,09	0,11	<b>1,71</b>	I	I	I	I	<b>2,26</b>	-0,21	0,24	-0,03	I	I	I	I	0,05	-0,03	0,03
Ovos		-0,16	0,00	0,00	0,00	-0,16	-0,19	-0,21	-0,21	0,54	I	I	-0,38	0,09	0,53	0,56	0,59	0,01	I	I	-0,06	-0,04	0,00	-0,03	-0,07
Pizza		-0,13	0,00	0,00	0,00	0,00	-0,01	-0,16	-0,20	0,60	I	I	I	<b>1,60</b>	-0,06	0,07	0,21	0,00	I	I	I	0,01	0,01	0,02	0,02
Suco natural		-0,07	0,00	0,00	0,00	-0,03	-0,09	-0,05	-0,06	0,05	I	I	I	0,41	0,15	0,37	0,38	0,01	I	I	I	-0,01	0,04	0,04	0,05

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

I – quando os VR's com e sem covariáveis foram iguais a infinito

I+ – quando o VR sem covariáveis foi finito e o VR com covariáveis foi infinito



**Figura 3:** Proporção de casos em que o ajuste por sexo e por idade melhorou ( $DVR < -0,05$ ), piorou ( $DVR > 0,05$ ) ou não alterou ( $-0,05 < DVR < 0,05$ ) a acurácia das estimativas, de acordo com as frequências de consumo episódico de nível intermediário ou baixo nos R24h.

## Discussão

Com este trabalho foi possível verificar que, para componentes alimentares com frequência de consumo diário (isto é, nutrientes), o ajuste por sexo e idade não causou alterações substanciais nas estimativas de todos os métodos. Em contrapartida, para componentes alimentares de frequência episódico de nível Intermediário (arroz, biscoitos, carne de gado, café, carnes em geral, derivados do leite, doces e chocolates, feijão, carne de frango, frutas, pães, refrigerantes e vegetais totais) os métodos MSM e SPADE não mostraram alteração de desempenho na maioria dos casos, enquanto o NCI obteve maior proporção de casos com piora na estimação da média e dos percentis 50 a 95 e com melhora na estimação dos percentis 5, 10 e 25. Já quando se tratou da frequência de consumo episódico de nível baixo (bolos, carnes processadas, cerveja, chá, ovos, peixe, pizza, carne de porco e o suco natural) o método que obteve mais casos de melhora nas estimativas com o ajuste por sexo e idade foi o MSM, enquanto o NCI teve a maior

quantidade de piora quando ajustado por sexo e idade, todavia o SPADE não evidenciou melhora ou piora das estimativas na maioria dos casos.

Com isso, pode-se observar que o uso das covariáveis sexo e idade no ajuste dos métodos não causou impacto na estimação da distribuição para componentes alimentares de frequência de consumo diário. Por outro lado, na estimação de componentes alimentares de frequência de consumo episódico de nível intermediário, houve impacto no método NCI, com piora nas estimativas na maioria dos casos. Também ocorreu impacto da utilização dessas covariáveis na estimação de componentes alimentares de frequência de consumo episódico de nível baixo nos métodos NCI e MSM, à medida que no NCI houve piora na maioria das estimativas e no MSM houve melhora.

No presente trabalho as informações de sexo e idade foram utilizadas no ajuste dos métodos, fazendo com que a estimação dos parâmetros fosse também baseada nessas informações adicionais. Também é possível estratificar os dados e fazer as estimações separadamente para cada grupo de idade e sexo, entretanto essa estratégia pode causar prejuízo aos resultados, uma vez que se reduz o tamanho de amostra <sup>(17,26,38)</sup>. Estudos futuros poderiam comparar impacto da estratificação e do ajuste por sexo e idade em componentes alimentares de frequência de consumo episódico de nível intermediário e episódico de nível baixo, podendo assim confirmar se os achados de Tooze et al. <sup>(17)</sup> permaneceriam os mesmos.

Uma limitação deste estudo é que os dados de sexo e idade não são igualmente distribuídos entre homens e mulheres, assim como entre as faixas etárias. Em vista disso, há uma maior proporção de mulheres quando comparada a de homens, assim como a maioria dos sujeitos têm até 29 anos de idade. Com isso, para que se possa extrapolar os achados do presente trabalho para outros conjuntos de dados, é necessário que sejam realizados mais testes em outras bases de dados ou simulação.

## **Conclusão**

Chegou-se à conclusão de que, para componentes alimentares de frequência de consumo diário nenhum método foi impactado pelo emprego das covariáveis sexo e idade. Em contrapartida, nas demais situações, o NCI foi impactado pelo uso dessas covariáveis, com resultados, em geral, piores quando comparado aos sem ajuste. Já o MSM não foi impactado quando a frequência de consumo foi episódico de nível intermediário, mas sofreu impacto quando a frequência de consumo foi episódico de nível baixo, com melhora nos resultados quando se utilizou o ajuste por sexo e idade. Por fim,

o SPADE não evidenciou melhora ou piora das estimativas na maioria dos casos quando a frequência de consumo foi episódico de nível intermediário ou baixo.

Portanto, recomenda-se o uso de todos os métodos sem covariáveis na estimação de componentes alimentares de frequência de consumo diário e episódico de nível intermediário. Todavia, nos componentes alimentares com frequência de consumo episódico de nível baixo, recomenda-se a utilização do método MSM com ajuste por sexo e idade e dos métodos NCI e SPADE sem esse ajuste.

## Referências

1. Rossum, CTM, Fransen, HP, Verkaik-Kloosterman, J, et al. (2011) *Dutch National Food Consumption Survey 2007-2010*. 33–37. National Institute for Public Health and the Environment.
2. on behalf of the EFCOVAL Consortium, de Boer EJ, Slimani N, et al. (2011) The European Food Consumption Validation Project: conclusions and recommendations. *Eur. J. Clin. Nutr.* **65**, S102–S107.
3. Slob W (1993) Modeling long-term exposure of the whole population to chemicals in food. *Risk Anal. Off. Publ. Soc. Risk Anal.* **13**, 525–530.
4. Gay C (2004) Estimation of population distributions of habitual nutrient intake based on a short-run weighed food diary. *Br. J. Nutr.*, 83rd ed., 287–293.
5. Wallace LA, Duan N & Ziegenfus R (1994) Can Long-Term Exposure Distributions Be Predicted from Short-Term Measurements? *Risk Anal.* **14**, 75–85.
6. Buck RJ, Hammerstrom KA & Ryan PB (1995) Estimating long-term exposures from short-term measurements. *J. Expo. Anal. Environ. Epidemiol.* **5**, 359–373.
7. Nusser SM, Carriquiry AL, Dodd KW, et al. (1996) A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *J. Am. Stat. Assoc.* **91**, 1440–1449.
8. Guenther PM, Kott PS & Carriquiry AL (1997) Development of an approach for estimating usual nutrient intake distributions at the population level. *J. Nutr.* **127**, 1106–1112.
9. Chang H-Y, Suchindran CM & Pan W-H (2001) Using the overdispersed exponential family to estimate the distribution of usual daily intakes of people aged between 18 and 28 in Taiwan. *Stat. Med.* **20**, 2337–2350.
10. Hoffmann K, Boeing H, Dufour A, et al. (2002) Estimating the distribution of usual dietary intake by short-term measurements. *Eur. J. Clin. Nutr.* **56 Suppl 2**, S53-62.
11. Tooze JA, Grunwald GK & Jones RH (2002) Analysis of repeated measures data with clumping at zero. *Stat. Methods Med. Res.* **11**, 341–355.



12. Tooze JA, Midthune D, Dodd KW, et al. (2006) A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J. Am. Diet. Assoc.* **106**, 1575–1587.
13. Slob W (2006) Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* **44**, 933–951.
14. Waijers PMCM, Dekkers ALM, Boer JMA, et al. (2006) The potential of AGE MODE, an age-dependent model, to estimate usual intakes and prevalences of inadequate intakes in a population. *J. Nutr.* **136**, 2916–2920.
15. Staudenmayer J, Ruppert D & Buonaccorsi JP (2008) Density Estimation in the Presence of Heteroscedastic Measurement Error. *J. Am. Stat. Assoc.* **103**, 726–736.
16. Kipnis V, Midthune D, Buckman DW, et al. (2009) Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics* **65**, 1003–1010.
17. Tooze JA, Kipnis V, Buckman DW, et al. (2010) A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. *Stat. Med.* **29**, 2857–2868.
18. Haubrock J, Nöthlings U, Volatier J-L, et al. (2011) Estimating usual food intake distributions by using the multiple source method in the EPIC-Potsdam Calibration Study. *J. Nutr.* **141**, 914–920.
19. Harttig U, Haubrock J, Knüppel S, et al. (2011) The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *Eur. J. Clin. Nutr.* **65 Suppl 1**, S87-91.
20. Nusser SM, Fuller WA & Guenther PM (2012) Estimating Usual Dietary Intake Distributions: Adjusting for Measurement Error and Nonnormality in 24-Hour Food Intake Data. In *Surv. Meas. Process Qual.*, pp. 689–709 [Lyberg L, Biemer P, Collins M, et al., editors]. Hoboken, NJ, USA: John Wiley & Sons, Inc.
21. Dekkers ALM & Slob W (2012) Gaussian Quadrature is an efficient method for the back-transformation in estimating the usual intake distribution when assessing dietary exposure. *Food Chem. Toxicol.* **50**, 3853–3861.
22. Dekkers AL, Verkaik-Kloosterman J, van Rossum CT, et al. (2014) SPADE, a New Statistical Program to Estimate Habitual Dietary Intake from Multiple Food Sources and Dietary Supplements. *J. Nutr.* **144**, 2083–2091.
23. Chernova J & Solis-Trapala I (2016) A simplified approach to estimating the distribution of occasionally-consumed dietary components, applied to alcohol intake. *BMC Med. Res. Methodol.* **16**.
24. Souverein OW, Dekkers AL, Geelen A, et al. (2011) Comparing four methods to estimate usual intake distributions. *Eur. J. Clin. Nutr.* **65 Suppl 1**, S92-101.

25. Subar AF, Dodd KW, Guenther PM, et al. (2006) The Food Propensity Questionnaire: Concept, Development, and Validation for Use as a Covariate in a Model to Estimate Usual Food Intake. *J. Am. Diet. Assoc.* **106**, 1556–1563.
26. Tooze JA, Midthune D, Dodd KW, et al. (2006) A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J. Am. Diet. Assoc.* **106**, 1575–1587.
27. Goedhart PW, van der Voet H, Knüppel S, et al. (2012) A comparison by simulation of different methods to estimate the usual intake distribution for episodically consumed foods. *EFSA - Support. Publ.*, EN299 ed.
28. Verly- E, Fisberg RM & Marchioni DML (2012) Is the food frequency consumption essential as covariate to estimate usual intake of episodically consumed foods? *Eur. J. Clin. Nutr.* **66**, 1254–1258.
29. Ost C, De Ridder KAA, Tafforeau J, et al. (2017) The added value of food frequency questionnaire (FFQ) information to estimate the usual food intake based on repeated 24-hour recalls. *Arch. Public Health* **75**.
30. Freedman LS, Midthune D, Arab L, et al. (2018) Combining a Food Frequency Questionnaire With 24-Hour Recalls to Increase the Precision of Estimating Usual Dietary Intakes – Evidence From the Validation Studies Pooling Project. *Am. J. Epidemiol.*
31. Crispim SP, de Vries JHM, Geelen A, et al. (2011) Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres – results from the European Food Consumption Validation (EFCOVAL) study. *Br. J. Nutr.* **105**, 447–458.
32. Herrick, KA R LM & Parsons, R D KW (2018) *Estimating usual dietary intake from National Health and Nutrition Examination Survey data using the National Cancer Institute method. National Center for Health Statistics.* .
33. Guenther PM, Kirkpatrick SI, Reedy J, et al. (2014) The Healthy Eating Index-2010 Is a Valid and Reliable Measure of Diet Quality According to the 2010 Dietary Guidelines for Americans. *J. Nutr.* **144**, 399–407.
34. Verly-Jr E, Oliveira DCRS, Fisberg RM, et al. (2016) Performance of statistical methods to correct food intake distribution: comparison between observed and estimated usual intake. *Br. J. Nutr.* **116**, 897–903.
35. Laureano GHC, Leotti VB & Camey S Impacto da suposição sobre a frequência do consumo, diário ou episódico, e do uso do questionário de frequência alimentar na estimação da distribuição de consumo alimentar usual nos métodos NCI, MSM e SPADE. Tese de doutorado, UFRGS.
36. SAS | Business Analytics and Business Intelligence Software. <http://www.sas.com/>.
37. R Core Team (2015) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

38. Dodd KW, Guenther PM, Freedman LS, et al. (2006) Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J. Am. Diet. Assoc.* **106**, 1640–1650.

## 7. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Esse trabalho teve como objetivo preencher lacunas existentes na literatura de nutrição epidemiológica no que se refere ao impacto da escolha da suposição de frequência de consumo alimentar (diário ou episódico), assim como da utilização de covariáveis (QFA, sexo e idade) na estimação da distribuição de consumo alimentar usual, chegando-se às seguintes recomendações:

- Utilizar a suposição de frequência de consumo diário na estimação da distribuição de consumo alimentar de componentes alimentares de frequência de consumo diária, independentemente do método;
- Utilizar a suposição de frequência de consumo episódico em componentes de frequência de consumo episódico de nível intermediário:
  - Emprego do método MSM na estimação dos percentis 5 e 10;
  - Uso do NCI, MSM ou SPADE na estimação dos demais parâmetros da distribuição.
- Utilizar da suposição de frequência de consumo episódico em componentes de frequência de consumo episódico de nível baixo:
  - Emprego do método MSM na estimação dos percentis menores ou iguais a 25;
  - Uso dos métodos SPADE e NCI para estimativa da média e dos percentis acima de 25.

Conclui-se que, em relação ao uso do QFA, apesar de ter sido apurado discreto aumento de acurácia em algumas situações, principalmente nos métodos MSM e SPADE, o custo-benefício da coleta dessas informações deve ser avaliado, visto que a elaboração e aplicação do QFA demanda investimentos que podem não compensar as melhoras

obtidas.

Já quando se trata da utilização das covariáveis sexo e idade, recomenda-se o uso de todos os métodos sem covariáveis na estimação de componentes alimentares de frequência de consumo diário e episódico de nível intermediário. Todavia, nos componentes alimentares com frequência de consumo episódico de nível baixo, recomenda-se a utilização do método MSM com ajuste por sexo e idade e dos métodos NCI e SPADE sem esse ajuste.

Os achados desse estudo não corroboraram em totalidade os da literatura atual no que se refere a escolha das suposições, uma vez que nos dados analisados não houveram componentes alimentares com percentual de zeros nos R24h entre 5% e 10%, entretanto, foi possível verificar que para componentes alimentares com até 0,1% de R24h iguais a zero, a suposição com maior acurácia é a de consumo diário e, quando no mínimo 20% de R24h são iguais a zero, a utilização da suposição de consumo episódico é mais adequada.

Já para o uso do QFA como informação adicional na modelagem os resultados confirmam os achados atuais da literatura no que se refere ao melhor desempenho do método SPADE - quando utilizado a informação de quem nunca consome o componente alimentar - na estimação de percentis inferiores, como também, da falta de impacto da utilização do QFA como covariável no NCI. Além disso, o presente trabalho trouxe novos resultados a respeito do impacto do uso do QFA como covariável no método MSM.

Embora tenha sido possível realizar uma série de recomendações, ainda são necessárias mais investigações a respeito da suposição de frequência de consumo alimentar e do uso de covariáveis na estimação dos métodos, para que se possa ter maior confiança na escolha dos métodos, assim como saber se é imperativo que se utilize covariáveis na estimação da distribuição de consumo alimentar usual. Em vista disso,

estudos futuros poderiam investigar a influência da suposição de frequência de consumo em componentes alimentares com consumos iguais a zero entre 0,1% e 20% nos R24h, como também a análise do impacto da utilização de outras covariáveis na estimação dos métodos, além de repetir as análises realizadas em outros conjuntos de dados.

**APÊNDICE A: Artigo 1 - Tabela S1: Valores verdadeiros para alimentos em nutrientes com base na média dos 20 R24h.**

<i>Componente alimentar</i>	<i>Média</i>	<i>P5</i>	<i>P10</i>	<i>P25</i>	<i>P50</i>	<i>P75</i>	<i>P90</i>	<i>P95</i>
Arroz	99,47	35,21	44,98	59,94	86,48	118,25	164,20	201,13
Biscoitos	16,01	0,00	1,50	4,81	12,13	22,08	35,03	45,67
Carne de gado	48,95	7,93	17,05	29,26	42,63	65,19	86,97	101,74
Bolos	12,21	0,00	0,00	1,50	8,88	18,44	28,95	39,89
Café	90,19	0,00	0,00	10,00	74,65	138,45	219,30	257,38
Cálcio	714,14	364,77	409,49	541,92	682,07	857,12	1048,85	1179,77
Carnes em geral	142,59	59,69	73,85	95,28	129,69	174,71	228,29	264,69
Chá	11,18	0,00	0,00	0,00	0,00	9,95	42,43	59,95
Carboidratos	246,00	149,90	166,95	195,42	238,25	275,13	325,85	368,31
Cobre	1,35	0,71	0,77	0,92	1,14	1,52	2,09	2,72
Derivados do leite	45,53	4,51	6,89	15,82	32,75	56,81	92,00	158,26
Folato diário	297,42	172,68	200,56	235,15	288,61	340,84	417,27	472,96
Doces e chocolates	38,03	2,00	4,98	14,50	30,40	53,37	77,40	97,69
Feijão	108,43	13,10	23,55	49,00	87,50	136,45	202,78	244,80
Fósforo	1184,45	725,59	812,26	944,10	1113,04	1358,46	1587,90	1845,87
Carne de frango	57,63	10,00	17,50	30,13	48,25	75,94	96,36	120,50
Frutas	77,47	0,00	1,50	16,56	47,66	110,47	190,65	238,07
Lipídios totais	70,93	44,48	48,80	57,24	67,17	79,98	98,92	112,75
Magnésio	238,43	145,94	157,88	191,21	226,79	266,59	320,41	367,65
Ovos	7,97	0,00	0,00	2,06	5,00	11,19	19,13	24,50
Pães (excluindo-se os recheados)	49,75	9,33	15,53	28,14	43,70	64,69	87,45	109,19

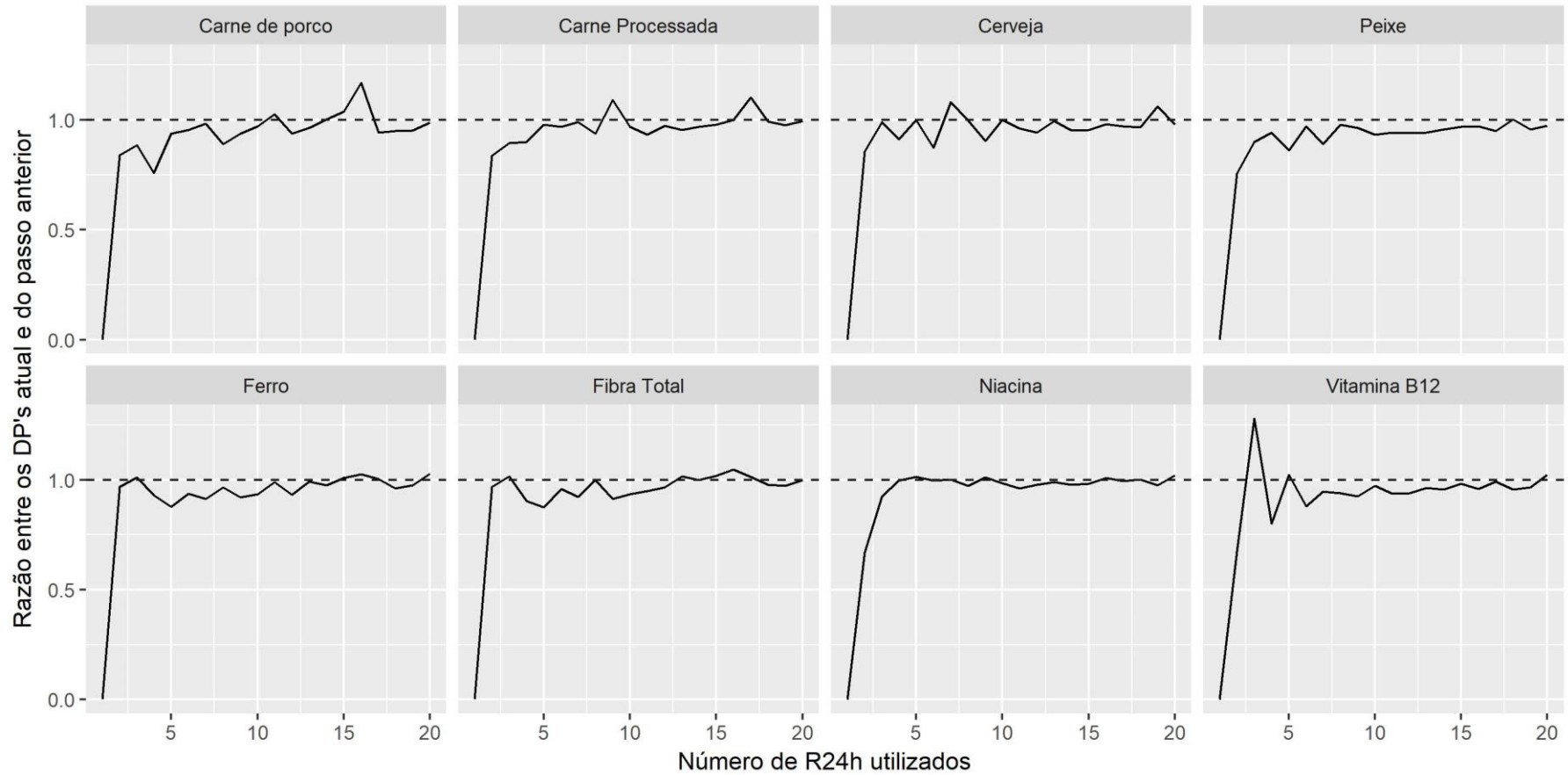
<i>Componente alimentar</i>	<i>Média</i>	<i>P5</i>	<i>P10</i>	<i>P25</i>	<i>P50</i>	<i>P75</i>	<i>P90</i>	<i>P95</i>
Pizza	19,41	0,00	0,00	0,00	10,00	30,00	50,23	65,00
Carne de porco	6,82	0,00	0,00	0,00	0,00	9,50	19,00	28,50
Potássio	2268,24	1378,35	1542,04	1837,27	2211,86	2589,96	3002,93	3450,91
Proteínas	89,70	57,37	62,95	72,31	84,76	100,55	120,07	137,85
Refrigerante	209,73	6,25	24,00	68,50	167,50	299,48	445,15	537,77
Gorduras saturadas	24,79	14,10	16,28	19,19	23,68	27,99	34,77	39,32
Sódio	2130,44	1148,77	1354,76	1624,91	1991,00	2537,09	3098,25	3616,74
Suco natural	63,62	0,00	0,00	0,00	35,50	95,81	164,93	227,25
Vegetais totais	80,20	20,86	30,35	48,28	69,76	101,15	133,74	164,97
Vitamina A	642,84	246,69	297,42	383,14	494,17	763,98	1184,32	1531,66
Vitamina B1	1,29	0,80	0,90	1,05	1,22	1,48	1,80	2,00
Vitamina B2	1,94	1,18	1,30	1,51	1,80	2,20	2,83	3,21
Vitamina B6	1,73	1,04	1,17	1,34	1,58	1,90	2,33	2,74
Vitamina C	120,15	16,66	20,84	35,21	55,69	106,14	301,87	463,86
Zinco	15,06	7,89	8,46	9,88	12,01	16,06	21,64	28,48

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

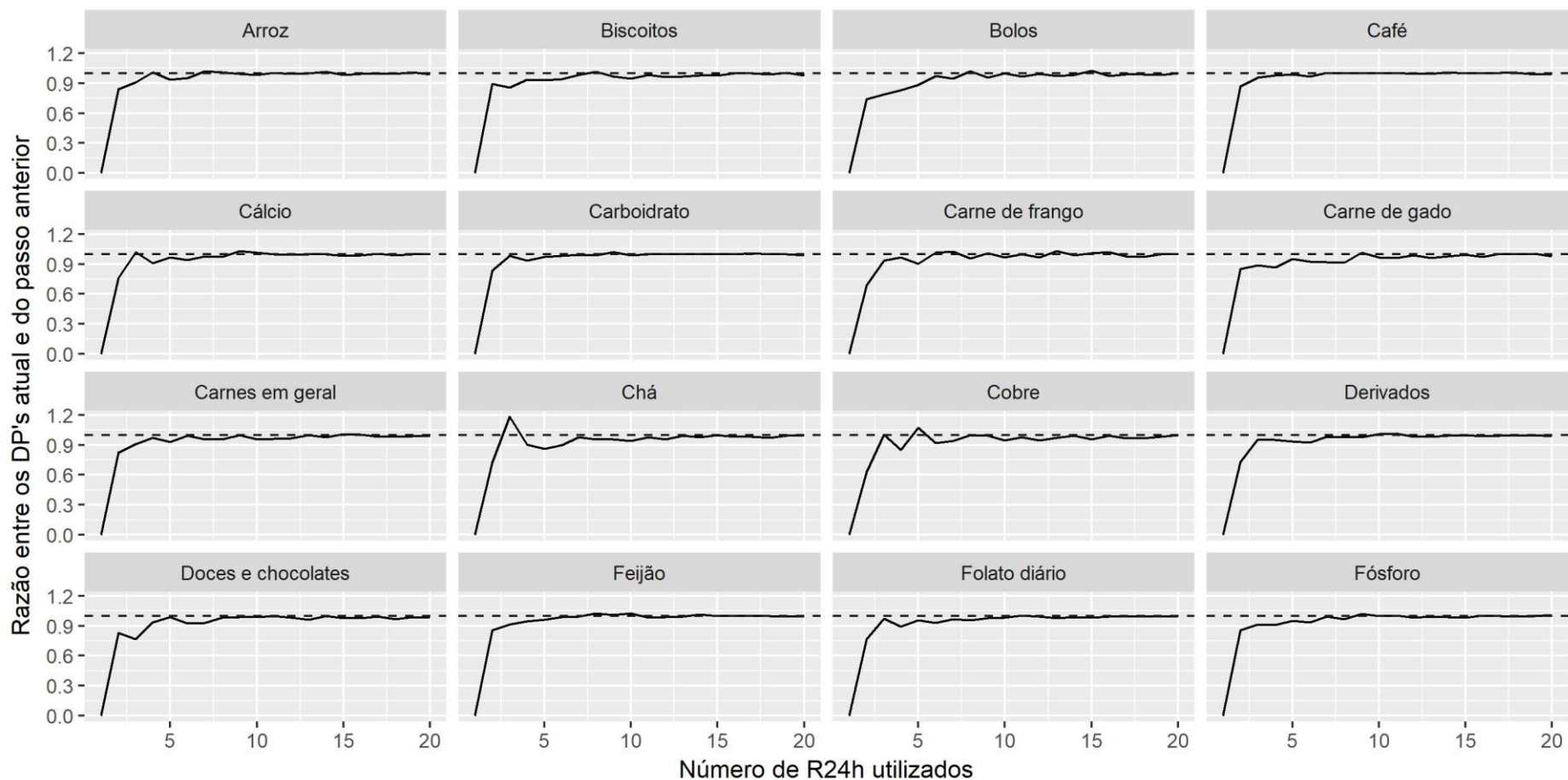
A seleção dos componentes alimentares foi realizada pela observação visual da razão entre o desvio padrão (DP) utilizando um número N R24h (atual) e o desvio padrão utilizando N-1 R24h (passo anterior), por exemplo a razão para 20 R24h foi calculada por  $\frac{DP_{20\ R24h}}{DP_{19\ R24h}}$ . Ao passo que, não foi excluído o componente alimentar quando ao aumentar o número de R24h, a razão entre os desvios padrões foi próxima de 1.



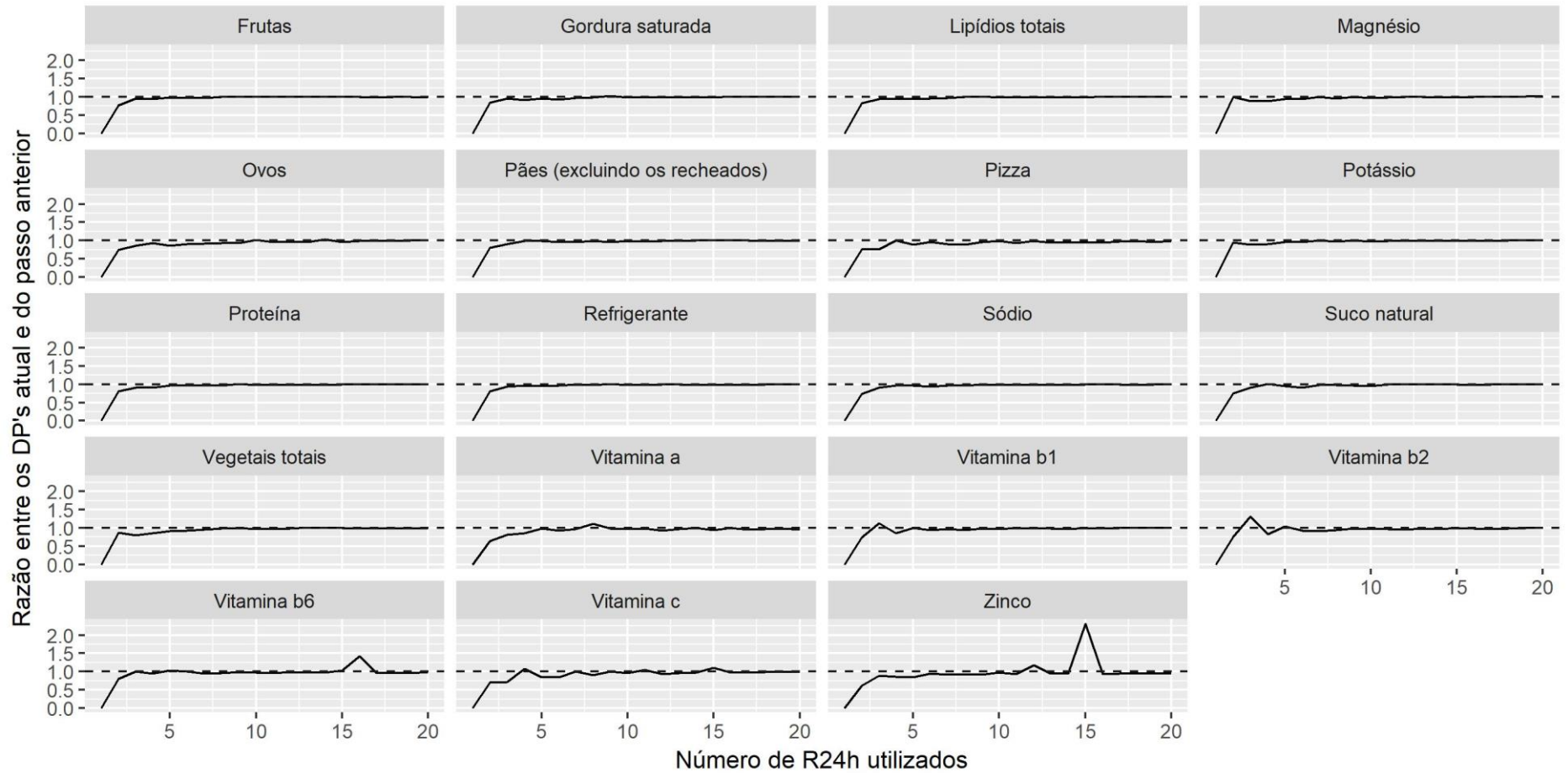
**APÊNDICE B: Artigo 1 - Figura S1: Componentes alimentares excluídos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.**



**APÊNDICE C: Artigo 1 - Figura S2: Componentes alimentares mantidos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.**



**APÊNDICE D: Artigo 1 - Figura S3: Componentes alimentares mantidos após o cálculo da razão entre os desvios padrões (atual e do passo anterior) considerando diferentes números de dias consecutivos de R24h por sujeito.**



Nas Tabelas a seguir (S2, S3, S4 e S5) estão dispostos os resultados das comparações entre os métodos. Quando a estimativa do vício relativo for igual a infinito, tem-se que o valor verdadeiro de parâmetro foi igual a zero e a estimativa do método foi diferente de zero.

**APÊNDICE E: Artigo 1 - Tabela S2: Vícios relativos para a suposição de consumo episódico.**

Método	Componente alimentar	Média	P5	P10	P25	P50	P75	P90	P95	Frequência de consumo
MSM	Cálcio	0,00	0,05	0,04	0,00	0,01	0,00	-0,01	-0,04	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Carboidratos	0,00	0,02	0,02	0,01	0,00	0,00	-0,01	-0,01	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Cobre	-0,11	-0,01	0,03	0,09	0,05	-0,08	-0,18	-0,3	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Folato diário	0,00	0,05	0,02	0,03	0,01	-0,01	-0,02	-0,02	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Fósforo	0,00	0,03	0,03	0,02	0,01	0,00	-0,02	-0,03	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Gorduras saturadas	0,01	0,06	-0,02	0,04	0,01	0,00	-0,02	-0,01	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	
SPADE		NA	NA	NA	NA	NA	NA	NA	NA	
MSM	Lipídios totais	0,00	0,01	0,00	0,01	0,01	0,00	-0,03	-0,02	Diário
NCI		NA	NA	NA	NA	NA	NA	NA	NA	

<b>Método</b>	<b>Componente alimentar</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>	<b>Frequência de consumo</b>
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		0,00	0,03	0,05	0,01	0,02	0,00	-0,03	0,00	
<b>NCI</b>	Magnésio	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		0,00	0,00	0,00	0,00	-0,01	-0,01	0,01	-0,03	
<b>NCI</b>	Potássio	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		0,00	0,06	0,02	0,02	0,00	-0,01	-0,02	-0,03	
<b>NCI</b>	Proteínas	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		0,00	0,09	0,04	0,01	0,00	-0,01	-0,02	-0,06	
<b>NCI</b>	Sódio	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		-0,09	0,07	0,10	0,09	0,09	-0,08	-0,25	-0,34	
<b>NCI</b>	Vitamina A	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		0,01	0,00	0,00	0,05	-0,01	0,02	-0,05	-0,05	
<b>NCI</b>	Vitamina B1	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		-0,02	0,02	0,08	0,06	0,00	0,00	-0,08	-0,1	
<b>NCI</b>	Vitamina B2	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>		-0,02	0,05	0,03	0,04	0,01	0	-0,01	-0,09	
<b>NCI</b>	Vitamina B6	NA	NA	NA	NA	NA	NA	NA	NA	Diário
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>	Zinco	-0,14	0,05	0,09	0,05	0,08	-0,07	-0,12	-0,26	Diário

<b>Método</b>	<b>Componente alimentar</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>	<b>Frequência de consumo</b>
<b>NCI</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>SPADE</b>		NA	NA	NA	NA	NA	NA	NA	NA	
<b>MSM</b>	Vitamina C	-0,27	0,08	0,3	0,19	0,24	0,06	-0,47	-0,55	Diário
<b>NCI</b>		-0,29	0,28	0,33	0,21	0,23	0,02	-0,46	-0,55	
<b>SPADE</b>		-0,30	0,25	0,3	0,21	0,21	0,02	-0,47	-0,56	
<b>MSM</b>	Arroz	0,08	0,14	0,11	0,08	0,10	0,08	0,05	0,07	Episódico de nível Intermediário
<b>NCI</b>		-0,04	0,14	0,07	0,08	0,02	0,00	-0,07	-0,12	
<b>SPADE</b>		-0,03	0,17	0,1	0,09	0,02	0,00	-0,06	-0,1	
<b>MSM</b>	Carnes em geral	0,09	0,22	0,21	0,19	0,13	0,06	0,00	-0,01	Episódico de nível Intermediário
<b>NCI</b>		-0,01	0,21	0,13	0,11	0,03	-0,03	-0,1	-0,12	
<b>SPADE</b>		-0,01	0,24	0,14	0,11	0,03	-0,03	-0,09	-0,13	
<b>MSM</b>	Vegetais totais	0,05	0,39	0,22	0,14	0,1	0,05	0,03	-0,02	Episódico de nível Intermediário
<b>NCI</b>		-0,02	0,44	0,26	0,11	0,06	-0,03	-0,07	-0,14	
<b>SPADE</b>		-0,02	0,49	0,26	0,1	0,05	-0,03	-0,08	-0,13	
<b>MSM</b>	Pães (excluindo-se os recheados)	0,03	0,21	0,16	0,07	0,03	0,02	-0,01	0,00	Episódico de nível Intermediário
<b>NCI</b>		-0,05	0,57	0,3	0,11	0,03	-0,06	-0,12	-0,2	
<b>SPADE</b>		-0,03	0,65	0,35	0,11	0,04	-0,05	-0,1	-0,17	
<b>MSM</b>	Feijão	0,02	0,22	0,19	0,1	0,05	0,02	0,02	0,00	Episódico de nível Intermediário
<b>NCI</b>		-0,06	0,73	0,37	0,11	0,02	-0,01	-0,08	-0,09	
<b>SPADE</b>		-0,02	0,86	0,44	0,17	0,05	0,02	-0,02	-0,03	
<b>MSM</b>	Café	0,01	0,00	0,00	0,3	0,02	0,01	-0,01	0,02	Episódico de nível Intermediário
<b>NCI</b>		-0,08	Infinito	Infinito	0,33	-0,12	-0,06	-0,13	-0,1	
<b>SPADE</b>		-0,05	Infinito	Infinito	0,37	-0,09	-0,03	-0,1	-0,07	
<b>MSM</b>	Derivados do leite	-0,01	0,55	0,49	0,26	0,01	-0,01	0,03	-0,14	Episódico de nível Intermediário
<b>NCI</b>		-0,05	0,82	0,74	0,34	0,11	0,01	-0,1	-0,36	
<b>SPADE</b>		0,01	0,79	0,74	0,37	0,15	0,06	-0,03	-0,29	

<b>Método</b>	<b>Componente alimentar</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>	<b>Frequência de consumo</b>
<b>MSM</b>	Refrigerante	-0,03	0,31	0,17	0,07	-0,01	00,00	-0,05	-0,08	Episódico de nível Intermediário
<b>NCI</b>		-0,06	2,85	0,66	0,26	0,04	-0,05	-0,13	-0,15	
<b>SPADE</b>		-0,02	2,42	0,7	0,36	0,10	-0,03	-0,11	-0,13	
<b>MSM</b>	Frutas	-0,04	0,00	0,2	-0,03	0,01	0,00	-0,08	-0,05	Episódico de nível Intermediário
<b>NCI</b>		-0,08	Infinito	5,19	0,41	0,17	-0,05	-0,19	-0,22	
<b>SPADE</b>		-0,07	Infinito	4,17	0,46	0,22	-0,04	-0,18	-0,22	
<b>MSM</b>	Carne de gado	-0,12	0,55	0	-0,04	-0,09	-0,14	-0,20	-0,20	Episódico de nível Intermediário
<b>NCI</b>		0,00	1,66	0,48	0,14	0,06	-0,08	-0,12	-0,14	
<b>SPADE</b>		0,00	1,59	0,46	0,15	0,07	-0,07	-0,1	-0,13	
<b>MSM</b>	Doces e chocolates	-0,08	-0,2	-0,24	0,03	-0,08	-0,1	-0,04	-0,07	Episódico de nível Intermediário
<b>NCI</b>		0,02	2,45	0,98	0,21	0,03	-0,03	0,00	-0,02	
<b>SPADE</b>		0,03	2,09	0,88	0,24	0,07	-0,01	-0,02	-0,04	
<b>MSM</b>	Carne de frango	-0,17	-0,06	-0,09	-0,07	-0,11	-0,21	-0,15	-0,21	Episódico de nível Intermediário
<b>NCI</b>		-0,01	1,03	0,43	0,17	0,05	-0,06	0,00	-0,06	
<b>SPADE</b>		-0,01	0,81	0,32	0,12	0,04	-0,05	0,03	-0,01	
<b>MSM</b>	Biscoitos	-0,13	0,00	-0,47	-0,13	-0,18	-0,09	-0,09	-0,15	Episódico de nível Intermediário
<b>NCI</b>		-0,05	Infinito	1,11	0,28	-0,02	-0,06	-0,09	-0,13	
<b>SPADE</b>		-0,05	Infinito	0,86	0,26	-0,02	-0,06	-0,10	-0,12	
<b>MSM</b>	Suco natural	-0,17	0,00	0,00	0,00	-0,52	-0,11	-0,07	-0,10	Episódico de nível baixo
<b>NCI</b>		-0,03	Infinito	Infinito	Infinito	-0,01	-0,09	-0,01	-0,07	
<b>SPADE</b>		0,01	Infinito	Infinito	Infinito	0,10	0,00	0,01	-0,07	
<b>MSM</b>	Bolos	-0,55	0,00	0,00	-1,00	-0,92	-0,53	-0,45	-0,42	Episódico de nível baixo
<b>NCI</b>		0,00	Infinito	Infinito	2,35	0,03	-0,13	-0,11	-0,17	
<b>SPADE</b>		0,01	Infinito	Infinito	2,05	0,04	-0,08	-0,09	-0,16	
<b>MSM</b>	Ovos	-0,62	0,00	0,00	-1,00	-0,96	-0,65	-0,48	-0,31	Episódico de nível baixo
<b>NCI</b>		0,00	Infinito	Infinito	0,71	0,22	-0,08	-0,16	-0,16	

<b>Método</b>	<b>Componente alimentar</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>	<b>Frequência de consumo</b>
<b>SPADE</b>		0,00	Infinito	Infinito	0,67	0,31	-0,02	-0,15	-0,21	
<b>MSM</b>		-0,82	0,00	0,00	0,00	-1,00	-0,8	-0,74	-0,74	Episódico de nível
<b>NCI</b>	Pizza	0,01	Infinito	Infinito	Infinito	0,52	-0,17	-0,23	-0,23	baixo
<b>SPADE</b>		0,00	Infinito	Infinito	Infinito	0,47	-0,14	-0,19	-0,2	
<b>MSM</b>		-0,38	0,00	0,00	0,00	0,00	-1,00	-0,43	-0,27	Episódico de nível
<b>NCI</b>	Chá	0,05	Infinito	Infinito	Infinito	Infinito	-0,24	-0,28	0,05	baixo
<b>SPADE</b>		-0,03	Infinito	Infinito	Infinito	Infinito	-0,14	-0,2	0,02	

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

NA: método não conseguiu estimar os parâmetros.



**APÊNDICE F: Artigo 1 - Tabela S3: Vícios relativos para a suposição de consumo diário.**

Método	Componente alimentar	Média	P5	P10	P25	P50	P75	P90	P95	Frequência de consumo
MSM	Cálcio	0,00	0,05	0,04	0,00	0,01	0,00	-0,01	-0,04	Diário
NCI		0,00	0,03	0,06	0,00	0,00	0,00	-0,01	-0,02	
SPADE		0,00	0,03	0,06	0,00	0,00	0,00	-0,02	-0,02	
MSM	Carboidratos	0,00	0,02	0,02	0,01	0,00	0,00	-0,01	-0,01	Diário
NCI		0,00	-0,03	-0,02	0,00	0,00	0,04	0,04	0,01	
SPADE		0,00	-0,03	-0,02	0,00	0,00	0,04	0,04	0,01	
MSM	Cobre	-0,11	-0,01	0,03	0,09	0,05	-0,08	-0,18	-0,30	Diário
NCI		-0,08	0,02	0,03	0,06	0,05	-0,04	-0,16	-0,28	
SPADE		-0,08	0,02	0,04	0,05	0,05	-0,04	-0,15	-0,28	
MSM	Folato diário	0,00	0,05	0,02	0,03	0,01	-0,01	-0,02	-0,02	Diário
NCI		0,00	0,04	0,00	0,02	0,00	0,02	-0,03	-0,06	
SPADE		0,00	0,04	0,00	0,02	0,00	0,02	-0,03	-0,07	
MSM	Fósforo	0,00	0,03	0,03	0,02	0,01	0,00	-0,02	-0,03	Diário
NCI		0,00	0,01	0,00	0,01	0,03	0,01	0,01	-0,05	
SPADE		0,00	0,01	0,00	0,01	0,03	0,01	0,01	-0,05	
MSM	Gorduras saturadas	0,01	0,06	-0,02	0,04	0,01	0,00	-0,02	-0,01	Diário
NCI		0,00	0,03	0,01	0,03	0,01	0,03	-0,02	-0,04	
SPADE		0,00	0,03	0,00	0,03	0,01	0,04	-0,02	-0,04	
MSM	Lipídios totais	0,00	0,01	0,00	0,01	0,01	0,00	-0,03	-0,02	Diário
NCI		0,00	-0,04	-0,02	0,00	0,03	0,03	-0,03	-0,06	
SPADE		0,00	-0,04	-0,02	0,00	0,03	0,03	-0,02	-0,06	
MSM	Magnésio	0,00	0,03	0,05	0,01	0,02	0,00	-0,03	0,00	Diário
NCI		0,00	0,00	0,03	0,01	0,02	0,04	0,01	-0,04	

SPADE		0,00	0,00	0,03	0	0,02	0,04	0,01	-0,03	
MSM		0,00	0,00	0,00	0,00	-0,01	-0,01	0,01	-0,03	Diário
NCI	Potássio	0,00	0,01	0,00	0,00	0,00	0,02	0,02	-0,03	
SPADE		0,00	0,01	0,00	0,00	0,00	0,02	0,02	-0,03	
MSM		0,00	0,06	0,02	0,02	0,00	-0,01	-0,02	-0,03	Diário
NCI	Proteínas	0,00	-0,01	-0,01	0,01	0,03	0,02	0,00	-0,05	
SPADE		0,00	-0,01	-0,01	0,01	0,03	0,03	0,00	-0,05	
MSM		0,00	0,09	0,04	0,01	0,00	-0,01	-0,02	-0,06	Diário
NCI	Sódio	0,00	0,04	0,00	0,02	0,03	-0,01	-0,03	-0,07	
SPADE		0,00	0,04	0,00	0,02	0,03	-0,01	-0,03	-0,08	
MSM		-0,09	0,07	0,10	0,09	0,09	-0,08	-0,25	-0,34	Diário
NCI	Vitamina A	-0,09	0,15	0,11	0,1	0,11	-0,07	-0,25	-0,33	
SPADE		-0,09	0,15	0,11	0,1	0,11	-0,07	-0,25	-0,34	
MSM		0,01	0,00	0,00	0,05	-0,01	0,02	-0,05	-0,05	Diário
NCI	Vitamina B1	0,00	0,01	0,00	0,00	0,03	0,01	-0,04	-0,05	
SPADE		0,00	0,01	-0,01	0,00	0,03	0,01	-0,04	-0,05	
MSM		-0,02	0,02	0,08	0,06	0,00	0,00	-0,08	-0,1	Diário
NCI	Vitamina B2	-0,01	0,00	0,00	0,01	0,02	0,00	-0,08	-0,1	
SPADE		-0,02	0,00	0,00	0,02	0,02	0,00	-0,08	-0,11	
MSM		-0,02	0,05	0,03	0,04	0,01	0,00	-0,01	-0,09	Diário
NCI	Vitamina B6	-0,02	0,01	-0,01	0,02	0,04	0,02	-0,02	-0,08	
SPADE		-0,02	0,01	-0,01	0,02	0,04	0,03	-0,01	-0,08	
MSM		-0,14	0,05	0,09	0,05	0,08	-0,07	-0,12	-0,26	Diário
NCI	Zinco	-0,11	0,00	0,04	0,06	0,07	-0,03	-0,14	-0,27	
SPADE		-0,11	0,00	0,04	0,06	0,06	-0,03	-0,14	-0,27	
MSM		-0,27	0,08	0,3	0,19	0,24	0,06	-0,47	-0,55	Diário
NCI	Vitamina C	-0,28	0,31	0,38	0,26	0,27	0,04	-0,46	-0,56	
SPADE		-0,28	0,32	0,38	0,26	0,27	0,04	-0,46	-0,56	

MSM		0,07	0,14	0,11	0,07	0,10	0,07	0,05	0,06	Episódico de nível Intermediário
NCI	Arroz	0,01	0,02	0,02	0,1	0,09	0,08	0	-0,06	
SPADE		0,02	0,05	0,03	0,1	0,08	0,08	0,01	-0,04	
MSM		0,09	0,22	0,21	0,19	0,13	0,06	0	-0,01	Episódico de nível Intermediário
NCI	Carnes em geral	0,04	0,17	0,11	0,11	0,08	0,03	-0,01	-0,03	
SPADE		0,04	0,17	0,11	0,11	0,08	0,04	0	-0,02	
MSM		0,05	0,39	0,22	0,14	0,1	0,05	0,03	-0,02	Episódico de nível Intermediário
NCI	Vegetais totais	0,32	0,11	0,04	0,06	0,23	0,36	0,54	0,57	
SPADE		0,33	0,11	0,04	0,07	0,23	0,37	0,55	0,58	
MSM		0,03	0,21	0,16	0,07	0,03	0,02	-0,01	0,00	Episódico de nível Intermediário
NCI	Pães (excluindo -se os recheados)	0,05	0,63	0,27	0,05	0,04	0,03	0,06	0,03	
SPADE		0,07	0,62	0,26	0,05	0,04	0,05	0,1	0,07	
MSM		0,02	0,22	0,19	0,1	0,05	0,02	0,02	0,00	Episódico de nível Intermediário
NCI	Feijão	0,3	0,79	0,39	0,14	0,15	0,3	0,44	0,6	
SPADE		0,5	0,84	0,44	0,21	0,25	0,47	0,7	0,94	
MSM		0,00	Infinito	Infinito	0	0,02	0,03	-0,02	0,01	Episódico de nível Intermediário
NCI	Café	6,2	Infinito	Infinito	-0,09	-0,38	0,6	3,13	7,2	
SPADE		11,32	Infinito	Infinito	0,05	-0,26	1,07	4,8	11,05	
MSM		-0,01	0,51	0,48	0,26	0,01	-0,01	0,03	-0,14	Episódico de nível Intermediário
NCI	Derivados do leite	0,22	0,53	0,44	0,14	0,07	0,18	0,31	0,09	
SPADE		0,35	0,62	0,52	0,21	0,14	0,29	0,46	0,22	
MSM		-0,03	0,1	0,08	0,04	-0,02	0	-0,05	-0,08	Episódico de nível
NCI	Refrigerant e	0,11	6,15	1,51	0,43	0,01	-0,03	0,06	0,17	

SPADE		0,13	6,19	1,51	0,44	0,02	-0,01	0,09	0,21	Intermediário
MSM		-0,04	0,00	-0,13	-0,15	0,01	0	-0,07	-0,05	Episódico
NCI	Frutas	0,95	Infinito	3,06	-0,04	-0,03	0,19	0,76	1,48	de nível
SPADE		1,79	Infinito	3,95	0,2	0,23	0,58	1,44	2,51	Intermediário
MSM		-0,14	0,27	-0,06	-0,04	-0,09	-0,13	-0,18	-0,19	Episódico
NCI	Carne de gado	-0,02	2,39	0,77	0,25	0,07	-0,13	-0,21	-0,24	de nível
SPADE		0	2,43	0,8	0,27	0,09	-0,11	-0,19	-0,22	Intermediário
MSM		-0,08	-0,4	-0,36	0,03	-0,11	-0,1	-0,04	-0,06	Episódico
NCI	Doces e chocolates	-0,13	1,77	0,53	-0,12	-0,24	-0,24	-0,12	-0,04	de nível
SPADE		0,04	2,16	0,75	0,02	-0,12	-0,09	0,07	0,17	Intermediário
MSM		-0,17	-0,08	-0,14	-0,1	-0,13	-0,21	-0,14	-0,2	Episódico
NCI	Carne de frango	-0,11	0,95	0,34	0,06	-0,06	-0,17	-0,11	-0,14	de nível
SPADE		-0,01	1,13	0,47	0,17	0,04	-0,07	0,01	-0,03	Intermediário
MSM		-0,13	0,00	-0,73	-0,21	-0,18	-0,09	-0,09	-0,12	Episódico
NCI	Biscoitos	-0,13	Infinito	2,76	0,66	-0,03	-0,21	-0,3	-0,34	de nível
SPADE		-0,12	Infinito	2,77	0,67	-0,02	-0,21	-0,29	-0,33	Intermediário
MSM		-0,1	0,00	0,00	0,00	-0,49	-0,02	0,02	-0,04	Episódico
NCI	Suco natural	-0,08	Infinito	Infinito	Infinito	0,42	-0,24	-0,38	-0,45	de nível
SPADE		-0,08	Infinito	Infinito	Infinito	0,41	-0,24	-0,38	-0,45	baixo
MSM		-0,39	0,00	0,00	-1	-0,81	-0,35	-0,27	-0,25	Episódico
NCI	Bolos	-0,54	Infinito	Infinito	1,67	-0,41	-0,63	-0,7	-0,75	de nível
SPADE		-0,54	Infinito	Infinito	1,67	-0,41	-0,63	-0,7	-0,75	baixo
MSM		-0,46	0,00	0,00	-1	-0,8	-0,46	-0,27	-0,1	
NCI	Ovos	-0,54	Infinito	Infinito	0,33	-0,31	-0,61	-0,72	-0,75	

SPADE		-0,54	Infinito	Infinito	0,33	-0,31	-0,61	-0,72	-0,75	Episódico de nível baixo
MSM		-0,69	0,00	0,00	0,00	-1	-0,79	-0,58	-0,54	Episódico de nível baixo
NCI	Pizza	0,61	Infinito	Infinito	Infinito	2,12	0,11	-0,3	-0,44	Episódico de nível baixo
SPADE		0,61	Infinito	Infinito	Infinito	2,11	0,11	-0,3	-0,44	Episódico de nível baixo
MSM		0,25	0,00	0,00	0,00	0,00	-0,47	0,34	0,38	Episódico de nível baixo
NCI	Chá	1,76	Infinito	Infinito	Infinito	Infinito	2,5	-0,07	-0,29	Episódico de nível baixo
SPADE		1,75	Infinito	Infinito	Infinito	Infinito	2,49	-0,07	-0,29	Episódico de nível baixo

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

#### APÊNDICE G: Artigo 1 - Tabela S4: Vícios relativos para a suposição de consumo episódico sem uso de QFA.

Método	Componente alimentar	Média	P5	P10	P25	P50	P75	P90	P95
MSM	Carne de gado	-0,12	0,55	0,00	-0,04	-0,09	-0,14	-0,20	-0,20
NCI		0,00	1,66	0,48	0,14	0,06	-0,08	-0,12	-0,14
SPADE		0,00	1,59	0,46	0,15	0,07	-0,07	-0,10	-0,13
MSM	Bolos	-0,55	0,00	0,00	-1,00	-0,92	-0,53	-0,45	-0,42
NCI		0,00	Infinito	Infinito	2,35	0,03	-0,13	-0,11	-0,17
SPADE		0,01	Infinito	Infinito	2,05	0,04	-0,08	-0,09	-0,16
MSM	Carne de frango	-0,17	-0,06	-0,09	-0,07	-0,11	-0,21	-0,15	-0,21
NCI		-0,01	1,03	0,43	0,17	0,05	-0,06	0,00	-0,06
SPADE		-0,01	0,81	0,32	0,12	0,04	-0,05	0,03	-0,01
MSM	Ovos	-0,62	0,00	0,00	-1,00	-0,96	-0,65	-0,48	-0,31
NCI		0,00	Infinito	Infinito	0,71	0,22	-0,08	-0,16	-0,16
SPADE		0,00	Infinito	Infinito	0,67	0,31	-0,02	-0,15	-0,21
MSM	Pizza	-0,82	0,00	0,00	0,00	-1,00	-0,80	-0,74	-0,74

NCI		0,01	Infinito	Infinito	Infinito	0,52	-0,17	-0,23	-0,23
SPADE		0,00	Infinito	Infinito	Infinito	0,47	-0,14	-0,19	-0,20
MSM	Refrigerante	-0,03	0,31	0,17	0,07	-0,01	0,00	-0,05	-0,08
NCI		-0,06	2,85	0,66	0,26	0,04	-0,05	-0,13	-0,15
SPADE		-0,02	2,42	0,70	0,36	0,10	-0,03	-0,11	-0,13
MSM	Suco natural	-0,17	0,00	0,00	0,00	-0,52	-0,11	-0,07	-0,10
NCI		-0,03	Infinito	Infinito	Infinito	-0,01	-0,09	-0,01	-0,07
SPADE		0,01	Infinito	Infinito	Infinito	0,10	0,00	0,01	-0,07

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

**APÊNDICE H: Artigo 1 - Tabela S5: Vícios relativos para a suposição de consumo episódico com uso de QFA**

Método	Componente alimentar	Média	P5	P10	P25	P50	P75	P90	P95
MSM	Carne de gado	-0,14	0,24	-0,06	-0,04	-0,09	-0,13	-0,16	-0,19
NCI		-0,01	1,62	0,46	0,13	0,06	-0,08	-0,11	-0,13
SPADE		0,00	1,42	0,43	0,14	0,08	-0,06	-0,10	-0,11
MSM	Bolos	-0,43	0,00	0,00	-1,00	-0,82	-0,35	-0,27	-0,30
NCI		-0,03	Infinito	Infinito	2,27	-0,01	-0,17	-0,15	-0,20
SPADE		-0,03	0,00	Infinito	1,69	0,02	-0,10	-0,11	-0,18
MSM	Carne de frango	-0,15	-0,08	-0,09	-0,07	-0,11	-0,20	-0,13	-0,18
NCI		-0,01	1,06	0,45	0,18	0,06	-0,05	0,01	-0,04
SPADE		-0,01	0,68	0,34	0,17	0,07	-0,03	0,02	-0,03
MSM	Ovos	-0,46	0,00	0,00	-1,00	-0,82	-0,47	-0,27	-0,14
NCI		0,02	Infinito	Infinito	0,78	0,25	-0,07	-0,15	-0,15
SPADE		-0,01	Infinito	Infinito	0,60	0,29	-0,02	-0,15	-0,17
MSM	Pizza	-0,72	0,00	0,00	0,00	-1,00	-0,84	-0,60	-0,57
NCI		0,02	Infinito	Infinito	Infinito	0,50	-0,19	-0,25	-0,23
SPADE		-0,02	0,00	Infinito	Infinito	0,60	-0,12	-0,23	-0,27
MSM	Refrigerante	-0,02	0,17	0,08	0,05	-0,03	0,02	-0,03	-0,03

NCI		-0,06	3,05	0,70	0,26	0,02	-0,06	-0,12	-0,15
SPADE		-0,09	-1,00	-1,00	-0,32	0,03	0,00	-0,07	-0,09
MSM		-0,12	0,00	0,00	0,00	-0,52	-0,05	0,01	-0,05
NCI	Suco natural	-0,05	Infinito	Infinito	Infinito	-0,02	-0,13	-0,05	-0,09
SPADE		-0,07	0,00	Infinito	Infinito	-0,01	-0,07	-0,05	-0,12

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

**APÊNDICE I: Artigo 1 - Quadro S1: Detalhamento dos modelos com suposição de consumo diário.**

Etapa	Método		
	MSM	NCI	SPADE
<b>Definição da suposição sobre o consumo</b>	O usuário define que o componente alimentar é considerado de consumo diário.		
<b>Transformação de dados</b>	Aplica-se uma transformação Box-Cox de dois parâmetros nos resíduos do modelo de regressão linear para se aproximar de uma distribuição normal.	Aplica-se uma transformação Box-Cox de um parâmetro nos consumos observados para se aproximar de uma distribuição normal.	
<b>Modelo de ajuste</b>	Regressão linear nos consumos observados.	Modelo misto nos consumos transformados.	
<b>Estimação das variâncias intraindividual e interindividual</b>	Obtidas a partir dos resíduos transformados.	Obtidas a partir do modelo misto.	
<b>Simulação</b>	Não tem etapa de simulação.	Simula k (valor a ser definido) pseudo consumos por pessoa com uma distribuição normal com desvio padrão igual o desvio padrão interindividual.	Não utilizado para consumo diário.
<b>Transformação inversa</b>	Realizada nos resíduos do modelo de regressão.	Realizado no conjunto de dados simulados, por uma aproximação de Taylor de dois termos.	A média na escala transformada é diretamente transformada para escala original pela Quadratura Gaussiana, usando a variância total



Etapa	Método		
	MSM	NCI	SPADE
			e o parâmetro de transformação Box-Cox $\lambda$ . Os percentis na escala transformada correspondem exatamente com os percentis na escala original, e sua transformação inversa utiliza a quadratura gaussiana com a variância interindividual e $\lambda$ .
<b>Estimação da média e dos percentis</b>	Após a transformação dos resíduos para escala original, a média e os percentis são estimados via o modelo de regressão.	A partir dos dados simulados, transformados para a escala original.	Obtidos na etapa de transformação inversa.
<p><b>NCI</b>, <i>National Cancer Institute</i> <sup>(14, 15, 19, 20)</sup>;</p> <p><b>MSM</b>, <i>Multiple Source Method</i> <sup>(21, 27)</sup>;</p> <p><b>SPADE</b>, <i>Statistical Program to Assess Dietary Exposure</i> <sup>(22, 24, 25)</sup>.</p>			

**APÊNDICE J: Artigo 1 - Quadro S2: Detalhamento dos modelos com suposição de consumo episódico.**

Etapa		Método		
		MSM	NCI	SPADE
<b>Definição da suposição sobre o consumo</b>		O usuário define que o componente alimentar é considerado de consumo episódico.		
<b>Primeira parte</b>	Transformação de dados	Transforma os resíduos ajustados do modelo para uma distribuição normal, utilizando-se uma transformação Box-Cox de dois parâmetros.	-	-
	Modelo de ajuste	A probabilidade de consumo individual de um alimento é estimada via resíduos de um modelo de regressão logística.	Estima-se a probabilidade de um indivíduo consumir usando um modelo misto logístico.	Todas as observações dos R24h são utilizadas para estimar a distribuição da probabilidade de consumo por meio de um modelo beta binomial em função da idade.
	Simulação	Não tem etapa de simulação.	Não tem etapa de simulação	Não tem etapa de simulação.
	Transformação inversa	Realizada nos resíduos do modelo de regressão logística.	-	-
<b>Segunda Parte</b>	Transformação de dados	Idem Quadro S1.	Idem Quadro S1.	Para a estimativa da distribuição de consumo usual, somente os sujeitos com pelo menos 1 dia de consumo são utilizados. Para os indivíduos com um consumo positivo, seleciona-se aleatoriamente um consumo positivo dos indivíduos que têm dois consumos positivos. Após esse ajuste, os consumos positivos são transformados por meio de uma transformação Box-Cox de um parâmetro.

Etapa	Método		
	MSM	NCI	SPADE
Modelo de ajuste	Idem Quadro S1	Idem Quadro S1	
Estimação das variâncias intraindividual e interindividual	Idem Quadro S1	Idem Quadro S1	Idem Quadro S1
Simulação	Idem Quadro S1	Idem Quadro S1	Simulação de Monte Carlo para estimar o valor de ingestão habitual em um dia de consumo.
Transformação inversa	Idem Quadro S1	Idem Quadro S1	Idem Quadro S1
Incorporação da probabilidade de consumo	Multiplicam-se os consumos estimados pelas probabilidades estimadas e se obtém a distribuição de consumo episódico.	Multiplicam-se os consumos simulados na escala original pelas probabilidades estimadas.	Os consumos usuais derivados são multiplicados pelas probabilidades estimadas.
Estimação da média e dos percentis	Cálculo das médias e dos percentis nos dados do passo anterior.	Cálculo das médias e dos percentis nos dados do passo anterior.	

**NCI, National Cancer Institute** <sup>(14, 15, 19, 20)</sup>;

**MSM, Multiple Source Method** <sup>(21, 27)</sup>;

**SPADE, Statistical Program to Assess Dietary Exposure** <sup>(22, 24, 25)</sup>.

Nas Tabela S1 e Tabela S2, assim como nas Figuras (S1, S2 e S3) a seguir, estão dispostos os resultados do impacto do uso das covariáveis sexo e idade no ajuste dos métodos. Quando a estimativa do vício relativo for igual a infinito, tem-se que o valor verdadeiro de parâmetro foi igual a zero e a estimativa do método foi diferente de zero.

**APÊNDICE K: Artigo 2 - Tabela S1: Vícios relativos com e sem a utilização das covariáveis.**

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Cálcio</b>	Diário	MSM	Com covariável	0,00	0,05	0,04	0,00	0,01	0,00	-0,01	-0,04
		MSM	Sem covariável	0,00	0,05	0,04	0,00	0,01	0,00	-0,01	-0,04
		NCI	Com covariável	0,00	0,03	0,06	0,00	0,00	0,00	-0,01	-0,02
		NCI	Sem covariável	0,00	0,03	0,06	0,00	0,00	0,00	-0,01	-0,02
		SPADE	Com covariável	-0,01	0,02	0,05	-0,01	0,00	-0,01	-0,02	-0,03
		SPADE	Sem covariável	0,00	0,03	0,06	0,00	0,00	0,00	-0,02	-0,02
<b>Carboidratos</b>	Diário	MSM	Com covariável	0,00	0,02	0,02	0,01	0,00	0,00	-0,01	-0,01
		MSM	Sem covariável	0,00	0,02	0,02	0,01	0,00	0,00	-0,01	-0,01
		NCI	Com covariável	0,00	-0,03	-0,02	0,00	0,00	0,04	0,04	0,01
		NCI	Sem covariável	0,00	-0,03	-0,02	0,00	0,00	0,04	0,04	0,01
		SPADE	Com covariável	0,00	-0,03	-0,03	0,00	0,00	0,05	0,05	0,02
		SPADE	Sem covariável	0,00	-0,03	-0,02	0,00	0,00	0,04	0,04	0,01
<b>Cobre</b>	Diário	MSM	Com covariável	-0,11	-0,01	0,03	0,09	0,05	-0,08	-0,18	-0,30
		MSM	Sem covariável	-0,11	-0,01	0,03	0,09	0,05	-0,08	-0,18	-0,30
		NCI	Com covariável	-0,08	0,02	0,03	0,06	0,05	-0,04	-0,16	-0,28
		NCI	Sem covariável	-0,08	0,02	0,03	0,06	0,05	-0,04	-0,16	-0,28
		SPADE	Com covariável	-0,08	0,01	0,03	0,05	0,05	-0,03	-0,15	-0,27

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Folato diário</b>	Diário	SPADE	Sem covariável	-0,08	0,02	0,04	0,05	0,05	-0,04	-0,15	-0,28
		MSM	Com covariável	0,00	0,05	0,02	0,03	0,01	-0,01	-0,02	-0,02
		MSM	Sem covariável	0,00	0,05	0,02	0,03	0,01	-0,01	-0,02	-0,02
		NCI	Com covariável	0,00	0,04	0,00	0,02	0,00	0,02	-0,03	-0,06
		NCI	Sem covariável	0,00	0,04	0,00	0,02	0,00	0,02	-0,03	-0,06
<b>Fósforo</b>	Diário	SPADE	Com covariável	0,00	0,04	0,00	0,02	0,01	0,02	-0,03	-0,06
		SPADE	Sem covariável	0,00	0,04	0,00	0,02	0,00	0,02	-0,03	-0,07
		MSM	Com covariável	0,00	0,03	0,03	0,02	0,01	0,00	-0,02	-0,03
		MSM	Sem covariável	0,00	0,03	0,03	0,02	0,01	0,00	-0,02	-0,03
		NCI	Com covariável	0,00	0,01	0,00	0,01	0,03	0,01	0,01	-0,05
<b>Gorduras saturadas</b>	Diário	NCI	Sem covariável	0,00	0,01	0,00	0,01	0,03	0,01	0,01	-0,05
		SPADE	Com covariável	0,00	0,00	-0,01	0,01	0,03	0,01	0,01	-0,05
		SPADE	Sem covariável	0,00	0,01	0,00	0,01	0,03	0,01	0,01	-0,05
		MSM	Com covariável	0,01	0,06	-0,02	0,04	0,01	0,00	-0,02	-0,01
		MSM	Sem covariável	0,01	0,06	-0,02	0,04	0,01	0,00	-0,02	-0,01
<b>Lipídios totais</b>	Diário	NCI	Com covariável	0,00	0,03	0,01	0,03	0,01	0,03	-0,02	-0,04
		NCI	Sem covariável	0,00	0,03	0,01	0,03	0,01	0,03	-0,02	-0,04
		SPADE	Com covariável	0,00	0,02	-0,01	0,02	0,01	0,04	-0,01	-0,03
		SPADE	Sem covariável	0,00	0,03	0,00	0,03	0,01	0,04	-0,02	-0,04
		MSM	Com covariável	0,00	0,01	0,00	0,01	0,01	0,00	-0,03	-0,02
		MSM	Sem covariável	0,00	0,01	0,00	0,01	0,01	0,00	-0,03	-0,02
		NCI	Com covariável	0,00	-0,04	-0,02	0,00	0,03	0,03	-0,03	-0,06
		NCI	Sem covariável	0,00	-0,04	-0,02	0,00	0,03	0,03	-0,03	-0,06
		SPADE	Com covariável	0,00	-0,05	-0,03	-0,01	0,03	0,04	-0,01	-0,05
		SPADE	Sem covariável	0,00	-0,04	-0,02	0,00	0,03	0,03	-0,02	-0,06

<b>Componente alimentar</b>	<b>Frequência de consumo</b>	<b>Método</b>	<b>Situação</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>
<b>Magnésio</b>	Diário	MSM	Com covariável	0,00	0,03	0,05	0,01	0,02	0,00	-0,03	0,00
		MSM	Sem covariável	0,00	0,03	0,05	0,01	0,02	0,00	-0,03	0,00
		NCI	Com covariável	0,00	0,00	0,03	0,01	0,02	0,04	0,01	-0,04
		NCI	Sem covariável	0,00	0,00	0,03	0,01	0,02	0,04	0,01	-0,04
		SPADE	Com covariável	0,00	-0,01	0,02	0,00	0,02	0,04	0,02	-0,03
		SPADE	Sem covariável	0,00	0,00	0,03	0,00	0,02	0,04	0,01	-0,03
<b>Potássio</b>	Diário	MSM	Com covariável	0,00	0,00	0,00	0,00	-0,01	-0,01	0,01	-0,03
		MSM	Sem covariável	0,00	0,00	0,00	0,00	-0,01	-0,01	0,01	-0,03
		NCI	Com covariável	0,00	0,01	0,00	0,00	0,00	0,02	0,02	-0,03
		NCI	Sem covariável	0,00	0,01	0,00	0,00	0,00	0,02	0,02	-0,03
		SPADE	Com covariável	0,00	0,00	0,00	0,00	0,00	0,02	0,02	-0,03
		SPADE	Sem covariável	0,00	0,01	0,00	0,00	0,00	0,02	0,02	-0,03
<b>Proteínas</b>	Diário	MSM	Com covariável	0,00	0,06	0,02	0,02	0,00	-0,01	-0,02	-0,03
		MSM	Sem covariável	0,00	0,06	0,02	0,02	0,00	-0,01	-0,02	-0,03
		NCI	Com covariável	0,00	-0,01	-0,01	0,01	0,03	0,02	0,00	-0,05
		NCI	Sem covariável	0,00	-0,01	-0,01	0,01	0,03	0,02	0,00	-0,05
		SPADE	Com covariável	0,00	-0,02	-0,02	0,01	0,03	0,03	0,01	-0,04
		SPADE	Sem covariável	0,00	-0,01	-0,01	0,01	0,03	0,03	0,00	-0,05
<b>Sódio</b>	Diário	MSM	Com covariável	0,00	0,09	0,04	0,01	0,00	-0,01	-0,02	-0,06
		MSM	Sem covariável	0,00	0,09	0,04	0,01	0,00	-0,01	-0,02	-0,06
		NCI	Com covariável	0,00	0,04	0,00	0,02	0,03	-0,01	-0,03	-0,07
		NCI	Sem covariável	0,00	0,04	0,00	0,02	0,03	-0,01	-0,03	-0,07
		SPADE	Com covariável	0,00	0,03	-0,01	0,01	0,02	-0,01	-0,03	-0,07
		SPADE	Sem covariável	0,00	0,04	0,00	0,02	0,03	-0,01	-0,03	-0,08
<b>Vitamina A</b>	Diário	MSM	Com covariável	-0,09	0,07	0,10	0,09	0,09	-0,08	-0,25	-0,34

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95		
<b>Vitamina B1</b>	Diário	MSM	Sem covariável	-0,09	0,07	0,10	0,09	0,09	-0,08	-0,25	-0,34		
		NCI	Com covariável	-0,09	0,15	0,11	0,10	0,11	-0,07	-0,25	-0,33		
		NCI	Sem covariável	-0,09	0,15	0,11	0,10	0,11	-0,07	-0,25	-0,33		
		SPADE	Com covariável	-0,10	0,14	0,10	0,09	0,10	-0,08	-0,25	-0,34		
		SPADE	Sem covariável	-0,09	0,15	0,11	0,10	0,11	-0,07	-0,25	-0,34		
		MSM	Com covariável	0,01	0,00	0,00	0,05	-0,01	0,02	-0,05	-0,05		
		MSM	Sem covariável	0,01	0,00	0,00	0,05	-0,01	0,02	-0,05	-0,05		
		NCI	Com covariável	0,00	0,01	0,00	0,00	0,03	0,01	-0,04	-0,05		
		NCI	Sem covariável	0,00	0,01	0,00	0,00	0,03	0,01	-0,04	-0,05		
		SPADE	Com covariável	-0,01	0,00	-0,02	0,00	0,03	0,01	-0,03	-0,05		
<b>Vitamina B2</b>	Diário	SPADE	Sem covariável	0,00	0,01	-0,01	0,00	0,03	0,01	-0,04	-0,05		
		MSM	Com covariável	-0,02	0,02	0,08	0,06	0,00	0,00	-0,08	-0,10		
		MSM	Sem covariável	-0,02	0,02	0,08	0,06	0,00	0,00	-0,08	-0,10		
		NCI	Com covariável	-0,01	0,00	0,00	0,01	0,02	0,00	-0,08	-0,11		
		NCI	Sem covariável	-0,01	0,00	0,00	0,01	0,02	0,00	-0,08	-0,11		
		SPADE	Com covariável	-0,02	-0,01	-0,01	0,01	0,02	0,00	-0,08	-0,11		
		SPADE	Sem covariável	-0,02	0,00	0,00	0,02	0,02	0,00	-0,08	-0,11		
		<b>Vitamina B6</b>	Diário	MSM	Com covariável	-0,02	0,05	0,03	0,04	0,01	0,00	-0,01	-0,09
				MSM	Sem covariável	-0,02	0,05	0,03	0,04	0,01	0,00	-0,01	-0,09
				NCI	Com covariável	-0,02	0,01	-0,01	0,02	0,04	0,02	-0,02	-0,08
NCI	Sem covariável			-0,02	0,01	-0,01	0,02	0,04	0,02	-0,02	-0,08		
SPADE	Com covariável			-0,02	-0,01	-0,02	0,01	0,03	0,03	-0,01	-0,07		
SPADE	Sem covariável			-0,02	0,01	-0,01	0,02	0,04	0,03	-0,01	-0,08		
<b>Vitamina C</b>	Diário			MSM	Com covariável	-0,27	0,08	0,30	0,19	0,24	0,06	-0,47	-0,55
				MSM	Sem covariável	-0,27	0,08	0,30	0,19	0,24	0,06	-0,47	-0,55

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Zinco</b>	Diário	NCI	Com covariável	-0,28	0,31	0,38	0,26	0,27	0,04	-0,46	-0,56
		NCI	Sem covariável	-0,28	0,31	0,38	0,26	0,27	0,04	-0,46	-0,56
		SPADE	Com covariável	-0,30	0,33	0,38	0,26	0,25	0,02	-0,48	-0,57
		SPADE	Sem covariável	-0,28	0,32	0,38	0,26	0,27	0,04	-0,46	-0,56
		MSM	Com covariável	-0,14	0,05	0,09	0,05	0,08	-0,07	-0,12	-0,26
		MSM	Sem covariável	-0,14	0,05	0,09	0,05	0,08	-0,07	-0,12	-0,26
		NCI	Com covariável	-0,11	0,00	0,04	0,06	0,07	-0,03	-0,14	-0,27
		NCI	Sem covariável	-0,11	0,00	0,04	0,06	0,07	-0,03	-0,14	-0,27
<b>Arroz</b>	Episódico de nível intermediário	SPADE	Com covariável	-0,11	-0,01	0,03	0,06	0,07	-0,02	-0,13	-0,26
		SPADE	Sem covariável	-0,11	0,00	0,04	0,06	0,06	-0,03	-0,14	-0,27
		MSM	Com covariável	0,07	0,14	0,11	0,07	0,10	0,07	0,05	0,06
		MSM	Sem covariável	0,08	0,14	0,11	0,08	0,10	0,08	0,05	0,07
		NCI	Com covariável	0,01	0,02	0,02	0,10	0,09	0,08	0,00	-0,06
		NCI	Sem covariável	-0,04	0,14	0,07	0,08	0,02	0,00	-0,07	-0,12
		SPADE	Com covariável	0,00	0,20	0,12	0,11	0,05	0,03	-0,03	-0,07
		SPADE	Sem covariável	-0,03	0,17	0,10	0,09	0,02	0,00	-0,06	-0,10
<b>Biscoitos</b>	Episódico de nível intermediário	MSM	Com covariável	-0,13	0,00	-0,73	-0,21	-0,18	-0,09	-0,09	-0,12
		MSM	Sem covariável	-0,13	0,00	-0,47	-0,13	-0,18	-0,09	-0,09	-0,15
		NCI	Com covariável	-0,13	infinito	2,76	0,66	-0,03	-0,21	-0,30	-0,34
		NCI	Sem covariável	-0,05	infinito	1,11	0,28	-0,02	-0,06	-0,09	-0,13
		SPADE	Com covariável	-0,05	infinito	1,03	0,34	0,03	-0,05	-0,12	-0,17
		SPADE	Sem covariável	-0,05	infinito	0,86	0,26	-0,02	-0,06	-0,10	-0,12
<b>Café</b>	Episódico de nível intermediário	MSM	Com covariável	0,00	infinito	infinito	0,00	0,02	0,03	-0,02	0,01
		MSM	Sem covariável	0,01	0,00	0,00	0,30	0,02	0,01	-0,01	0,02
		NCI	Com covariável	6,20	infinito	infinito	-0,09	-0,38	0,60	3,13	7,20



Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Carne de frango</b>	Episódico de nível intermediário	NCI	Sem covariável	-0,08	infinito	infinito	0,33	-0,12	-0,06	-0,13	-0,10
		SPADE	Com covariável	-0,06	infinito	infinito	0,55	-0,10	-0,05	-0,12	-0,09
		SPADE	Sem covariável	-0,05	infinito	infinito	0,37	-0,09	-0,03	-0,10	-0,07
		MSM	Com covariável	-0,17	-0,08	-0,14	-0,10	-0,13	-0,21	-0,14	-0,20
		MSM	Sem covariável	-0,17	-0,06	-0,09	-0,07	-0,11	-0,21	-0,15	-0,21
		NCI	Com covariável	-0,11	0,95	0,34	0,06	-0,06	-0,17	-0,11	-0,14
		NCI	Sem covariável	-0,01	1,03	0,43	0,17	0,05	-0,06	0,00	-0,06
		SPADE	Com covariável	0,00	0,87	0,36	0,15	0,06	-0,04	0,03	-0,02
<b>Carne de gado</b>	Episódico de nível intermediário	SPADE	Sem covariável	-0,01	0,81	0,32	0,12	0,04	-0,05	0,03	-0,01
		MSM	Com covariável	-0,14	0,27	-0,06	-0,04	-0,09	-0,13	-0,18	-0,19
		MSM	Sem covariável	-0,12	0,55	0,00	-0,04	-0,09	-0,14	-0,20	-0,20
		NCI	Com covariável	-0,02	2,39	0,77	0,25	0,07	-0,13	-0,21	-0,24
		NCI	Sem covariável	0,00	1,66	0,48	0,14	0,06	-0,08	-0,12	-0,14
		SPADE	Com covariável	0,00	1,48	0,41	0,12	0,06	-0,06	-0,10	-0,11
		SPADE	Sem covariável	0,00	1,59	0,46	0,15	0,07	-0,07	-0,10	-0,13
		MSM	Com covariável	0,09	0,22	0,21	0,19	0,13	0,06	0,00	-0,01
<b>Carnes em geral</b>	Episódico de nível intermediário	MSM	Sem covariável	0,09	0,22	0,21	0,19	0,13	0,06	0,00	-0,01
		NCI	Com covariável	0,04	0,17	0,11	0,11	0,08	0,03	-0,01	-0,03
		NCI	Sem covariável	-0,01	0,21	0,13	0,11	0,03	-0,03	-0,10	-0,12
		SPADE	Com covariável	-0,03	0,19	0,12	0,09	0,02	-0,05	-0,10	-0,13
		SPADE	Sem covariável	-0,01	0,24	0,14	0,11	0,03	-0,03	-0,09	-0,13
		MSM	Com covariável	-0,01	0,51	0,48	0,26	0,01	-0,01	0,03	-0,14
		MSM	Sem covariável	-0,01	0,55	0,49	0,26	0,01	-0,01	0,03	-0,14
		NCI	Com covariável	0,22	0,53	0,44	0,14	0,07	0,18	0,31	0,09
<b>Derivados do leite</b>	Episódico de nível intermediário	NCI	Sem covariável	-0,05	0,82	0,74	0,34	0,11	0,01	-0,10	-0,36

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Doces e chocolates</b>	Episódico de nível intermediário	SPADE	Com covariável	-0,01	0,78	0,74	0,35	0,12	0,05	-0,04	-0,31
		SPADE	Sem covariável	0,01	0,79	0,74	0,37	0,15	0,06	-0,03	-0,29
		MSM	Com covariável	-0,08	-0,40	-0,36	0,03	-0,11	-0,10	-0,04	-0,06
		MSM	Sem covariável	-0,08	-0,20	-0,24	0,03	-0,08	-0,10	-0,04	-0,07
		NCI	Com covariável	-0,13	1,77	0,53	-0,12	-0,24	-0,24	-0,12	-0,04
		NCI	Sem covariável	0,02	2,45	0,98	0,21	0,03	-0,03	0,00	-0,02
<b>Feijão</b>	Episódico de nível intermediário	SPADE	Com covariável	0,01	1,83	0,79	0,18	0,03	-0,03	-0,01	-0,03
		SPADE	Sem covariável	0,03	2,09	0,88	0,24	0,07	-0,01	-0,02	-0,04
		MSM	Com covariável	0,02	0,22	0,19	0,10	0,05	0,02	0,02	0,00
		MSM	Sem covariável	0,02	0,22	0,19	0,10	0,05	0,02	0,02	0,00
		NCI	Com covariável	0,30	0,79	0,39	0,14	0,15	0,30	0,44	0,60
		NCI	Sem covariável	-0,06	0,73	0,37	0,11	0,02	-0,01	-0,08	-0,09
<b>Frutas</b>	Episódico de nível intermediário	SPADE	Com covariável	-0,01	0,83	0,46	0,18	0,07	0,04	-0,02	-0,03
		SPADE	Sem covariável	-0,02	0,86	0,44	0,17	0,05	0,02	-0,02	-0,03
		MSM	Com covariável	-0,04	0,00	-0,13	-0,15	0,01	0,00	-0,07	-0,05
		MSM	Sem covariável	-0,04	0,00	0,20	-0,03	0,01	0,00	-0,08	-0,05
		NCI	Com covariável	0,95	infinito	3,06	-0,04	-0,03	0,19	0,76	1,48
		NCI	Sem covariável	-0,08	infinito	5,19	0,41	0,17	-0,05	-0,19	-0,22
<b>Pães (excluindo-se os recheados)</b>	Episódico de nível intermediário	SPADE	Com covariável	-0,06	infinito	4,28	0,44	0,21	-0,05	-0,17	-0,18
		SPADE	Sem covariável	-0,07	infinito	4,17	0,46	0,22	-0,04	-0,18	-0,22
		MSM	Com covariável	0,03	0,21	0,16	0,07	0,03	0,02	-0,01	0,00
		MSM	Sem covariável	0,03	0,21	0,16	0,07	0,03	0,02	-0,01	0,00
		NCI	Com covariável	0,05	0,63	0,27	0,05	0,04	0,03	0,06	0,03
		NCI	Sem covariável	-0,05	0,57	0,30	0,11	0,03	-0,06	-0,12	-0,20
		SPADE	Com covariável	-0,02	0,70	0,37	0,13	0,05	-0,04	-0,09	-0,16

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Refrigerante</b>	Episódico de nível intermediário	SPADE	Sem covariável	-0,03	0,65	0,35	0,11	0,04	-0,05	-0,10	-0,17
		MSM	Com covariável	-0,03	0,10	0,08	0,04	-0,02	0,00	-0,05	-0,08
		MSM	Sem covariável	-0,03	0,31	0,17	0,07	-0,01	0,00	-0,05	-0,08
		NCI	Com covariável	0,11	6,15	1,51	0,43	0,01	-0,03	0,06	0,17
		NCI	Sem covariável	-0,06	2,85	0,66	0,26	0,04	-0,05	-0,13	-0,15
<b>Vegetais totais</b>	Episódico de nível intermediário	SPADE	Com covariável	-0,01	2,19	0,60	0,34	0,10	-0,01	-0,08	-0,11
		SPADE	Sem covariável	-0,02	2,42	0,70	0,36	0,10	-0,03	-0,11	-0,13
		MSM	Com covariável	0,05	0,39	0,22	0,14	0,10	0,05	0,03	-0,02
		MSM	Sem covariável	0,05	0,39	0,22	0,14	0,10	0,05	0,03	-0,02
		NCI	Com covariável	0,32	0,11	0,04	0,06	0,23	0,36	0,54	0,57
		NCI	Sem covariável	-0,02	0,44	0,26	0,11	0,06	-0,03	-0,07	-0,14
		SPADE	Com covariável	-0,02	0,48	0,27	0,11	0,06	-0,02	-0,06	-0,13
<b>Bolos</b>	Episódico de nível baixo	SPADE	Sem covariável	-0,02	0,49	0,26	0,10	0,05	-0,03	-0,08	-0,13
		MSM	Com covariável	-0,39	0,00	0,00	-1,00	-0,81	-0,35	-0,27	-0,25
		MSM	Sem covariável	-0,55	0,00	0,00	-1,00	-0,92	-0,53	-0,45	-0,42
		NCI	Com covariável	-0,54	infinito	infinito	1,67	-0,41	-0,63	-0,70	-0,75
		NCI	Sem covariável	0,00	infinito	infinito	2,35	0,03	-0,13	-0,11	-0,17
		SPADE	Com covariável	-0,02	infinito	infinito	2,03	0,03	-0,11	-0,13	-0,20
<b>Chá</b>	Episódico de nível baixo	SPADE	Sem covariável	0,01	infinito	infinito	2,05	0,04	-0,08	-0,09	-0,16
		MSM	Com covariável	0,25	0,00	0,00	0,00	0,00	-0,47	0,34	0,38
		MSM	Sem covariável	-0,38	0,00	0,00	0,00	0,00	-1,00	-0,43	-0,27
		NCI	Com covariável	1,76	infinito	infinito	infinito	infinito	2,50	-0,07	-0,29
		NCI	Sem covariável	0,05	infinito	infinito	infinito	infinito	-0,24	-0,28	0,05
		SPADE	Com covariável	0,00	infinito	infinito	infinito	infinito	-0,19	-0,17	0,05
		SPADE	Sem covariável	-0,03	infinito	infinito	infinito	infinito	-0,14	-0,20	0,02

Componente alimentar	Frequência de consumo	Método	Situação	Média	P5	P10	P25	P50	P75	P90	P95
<b>Ovos</b>	Episódico de nível baixo	MSM	Com covariável	-0,46	0,00	0,00	-1,00	-0,80	-0,46	-0,27	-0,10
		MSM	Sem covariável	-0,62	0,00	0,00	-1,00	-0,96	-0,65	-0,48	-0,31
		NCI	Com covariável	-0,54	infinito	infinito	0,33	-0,31	-0,61	-0,72	-0,75
		NCI	Sem covariável	0,00	infinito	infinito	0,71	0,22	-0,08	-0,16	-0,16
		SPADE	Com covariável	0,02	infinito	infinito	0,61	0,27	-0,02	-0,12	-0,13
		SPADE	Sem covariável	0,00	infinito	infinito	0,67	0,31	-0,02	-0,15	-0,21
<b>Pizza</b>	Episódico de nível baixo	MSM	Com covariável	-0,69	0,00	0,00	0,00	-1,00	-0,79	-0,58	-0,54
		MSM	Sem covariável	-0,82	0,00	0,00	0,00	-1,00	-0,80	-0,74	-0,74
		NCI	Com covariável	0,61	infinito	infinito	infinito	2,12	0,11	-0,30	-0,44
		NCI	Sem covariável	0,01	infinito	infinito	infinito	0,52	-0,17	-0,23	-0,23
		SPADE	Com covariável	-0,01	infinito	infinito	infinito	0,48	-0,15	-0,21	-0,22
		SPADE	Sem covariável	0,00	infinito	infinito	infinito	0,47	-0,14	-0,19	-0,20
<b>Suco natural</b>	Episódico de nível baixo	MSM	Com covariável	-0,10	0,00	0,00	0,00	-0,49	-0,02	0,02	-0,04
		MSM	Sem covariável	-0,17	0,00	0,00	0,00	-0,52	-0,11	-0,07	-0,10
		NCI	Com covariável	-0,08	infinito	infinito	infinito	0,42	-0,24	-0,38	-0,45
		NCI	Sem covariável	-0,03	infinito	infinito	infinito	-0,01	-0,09	-0,01	-0,07
		SPADE	Com covariável	-0,03	infinito	infinito	infinito	0,09	-0,04	-0,05	-0,12
		SPADE	Sem covariável	0,01	infinito	infinito	infinito	0,10	0,00	0,01	-0,07

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

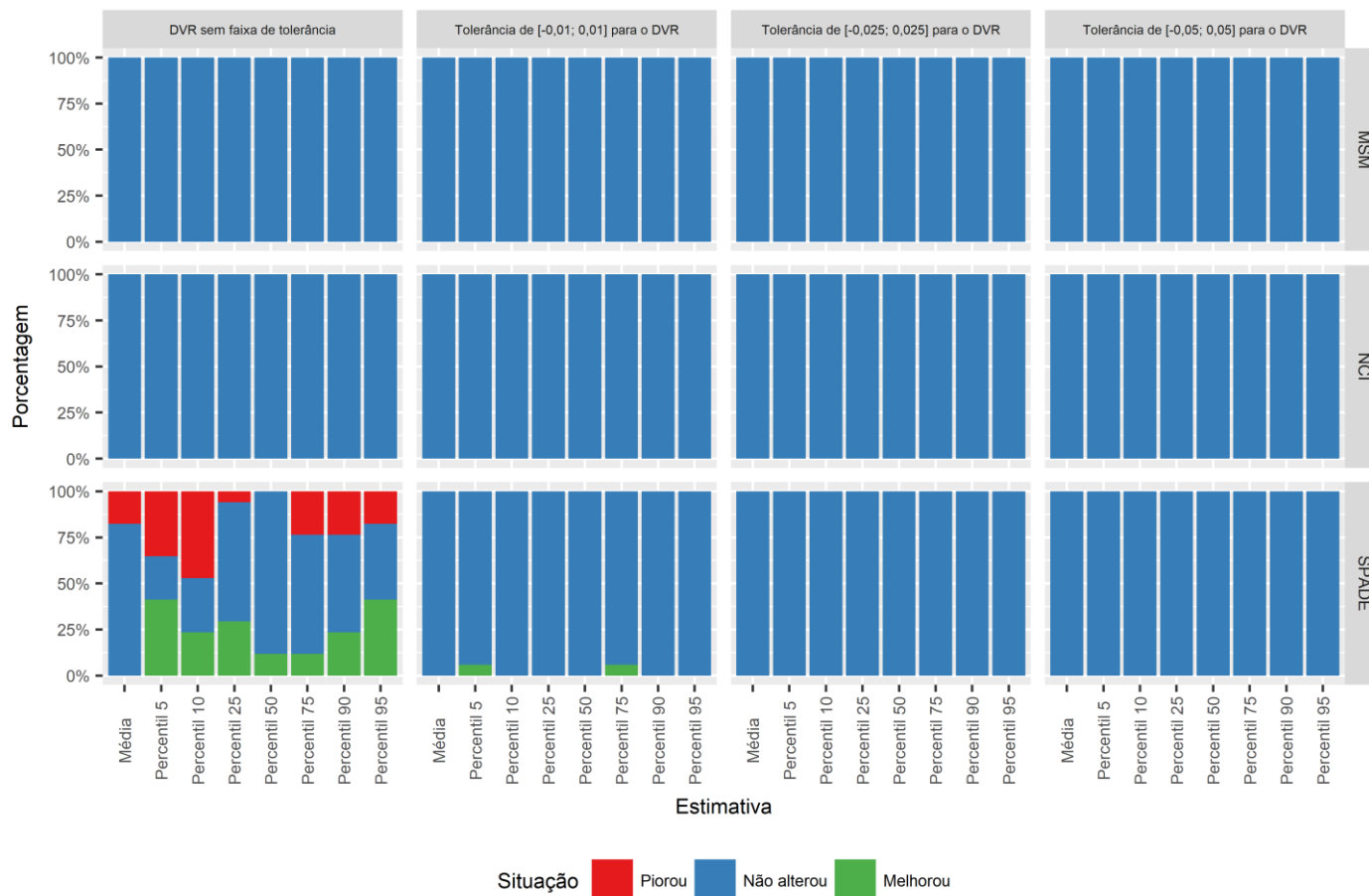
**APÊNDICE L: Artigo 2 - Tabela S2: Diferença do vício relativo (DVR) com e sem covariáveis para componentes alimentares de frequência de consumo diário**

<b>Componente alimentar</b>	<b>Método</b>	<b>Média</b>	<b>P5</b>	<b>P10</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>P90</b>	<b>P95</b>
<b>Cálcio</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,01	-0,01	-0,01	0,01	0,00	0,00	0,00	0,00
<b>Carboidratos</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,01	0,00	0,00	0,00	0,01	0,01	0,01
<b>Cobre</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	-0,01	-0,01	0,00	0,00	0,00	0,00	0,00
<b>Folato diário</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Fósforo</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,00	0,01	-0,01	0,00	0,00	0,00	0,00
<b>Gorduras saturadas</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	-0,01	0,00	0,00	0,00	0,01	-0,01	-0,01
<b>Lipídios totais</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,01	0,01	0,00	0,00	0,01	-0,01	-0,01
<b>Magnésio</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,01	0,00	0,00	0,00	0,00	0,01	-0,01

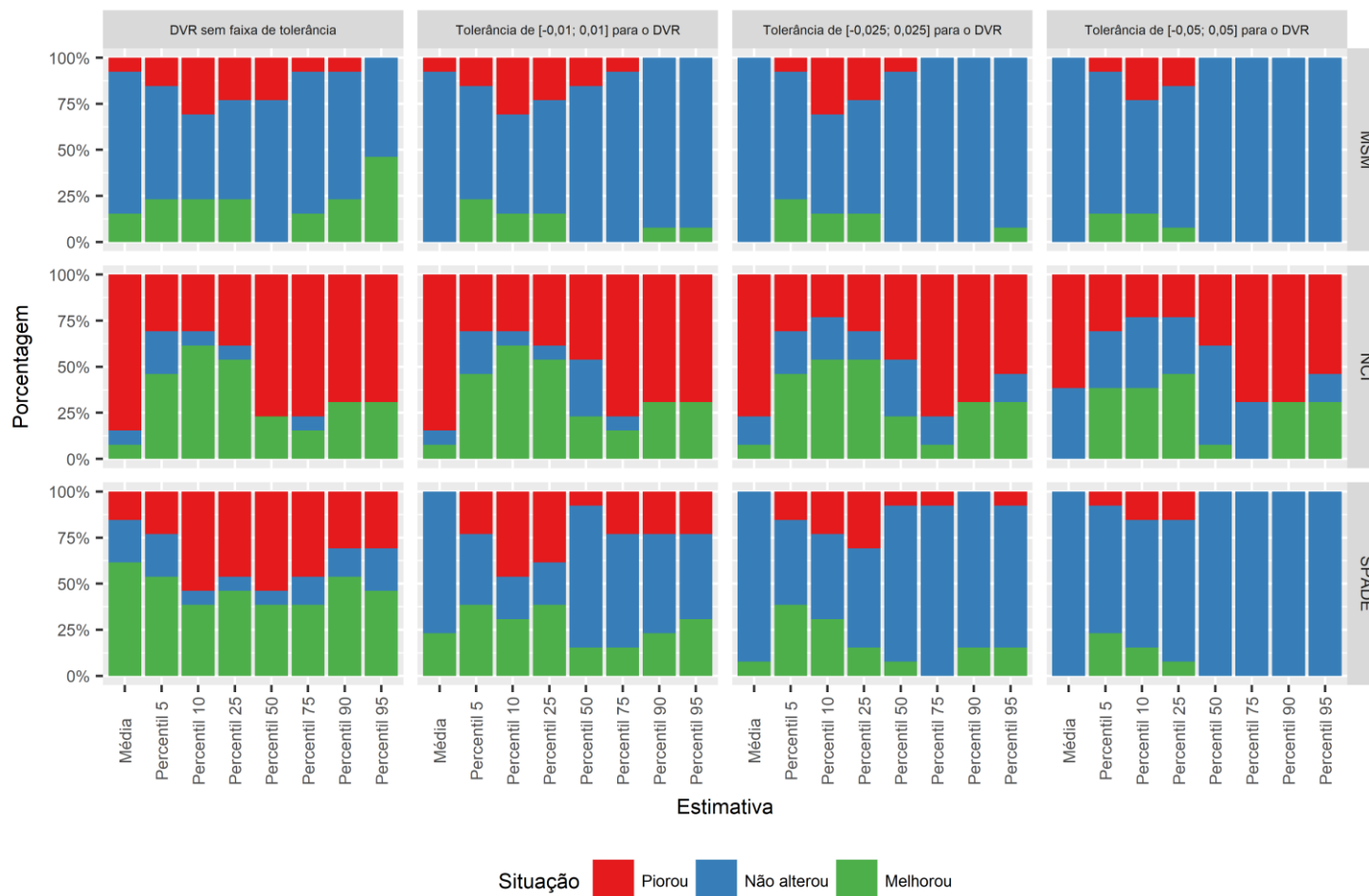
<b>Potássio</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	-0,01	0,00	0,00	0,00	0,00	0,00	0,00
<b>Proteínas</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,01	0,01	0,00	0,00	0,00	0,00	-0,01
<b>Sódio</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	-0,02	0,01	-0,01	0,00	0,00	0,00	-0,01
<b>Vitamina A</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,01	-0,01	-0,01	-0,01	-0,01	0,01	0,01	0,01
<b>Vitamina B1</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	-0,01	0,01	0,00	0,00	0,00	0,00	0,00
<b>Vitamina B2</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,00	0,01	-0,01	0,00	0,00	0,00	0,00
<b>Vitamina B6</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,00	0,01	-0,01	0,00	0,00	-0,01	-0,01
<b>Vitamina C</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,01	0,01	0,01	0,00	-0,01	-0,02	0,01	0,01
<b>Zinco</b>	MSM	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NCI	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SPADE	0,00	0,01	-0,01	0,00	0,00	-0,01	-0,01	-0,01

P5 – percentil 5, P10 – percentil 10, P25 – percentil 25, P50 – percentil 50, P75 – percentil 75, P90 – percentil 90 e P95 – percentil 95.

**APÊNDICE M: Artigo 2 - Figura S1: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo diário, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].**



**APÊNDICE N: Artigo 2 - Figura S2: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo episódico de nível intermediário, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].**





**APÊNDICE O: Artigo 2 - Figura S3: Proporção de casos em que o ajuste por sexo e por idade melhorou, piorou ou não alterou a acurácia das estimativas para componentes alimentares de frequência de consumo episódico de nível baixo, considerando o DVR sem faixa de tolerância e com as faixas de tolerância [-0,01; 0,01], [-0,025; 0,025] e [-0,05; 0,05].**

