

Universidade Federal do Rio Grande do Sul - UFRGS

Faculdade de Medicina

Programa de Pós-Graduação em Epidemiologia



## **DISSERTAÇÃO DE MESTRADO**

# **Formas de correção de consumo de produtos ultraprocessados mensurados a partir de Recordatórios Alimentares de 24 horas**

Luana Giongo Pedrotti

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Vanessa Bielefeldt Leotti

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Michele Drehmer

Porto Alegre, 28 de setembro de 2020.

Universidade Federal do Rio Grande do Sul - UFRGS

Faculdade de Medicina

Programa de Pós-Graduação em Epidemiologia



## **DISSERTAÇÃO DE MESTRADO**

# **Formas de correção de consumo de produtos ultraprocessados mensurados a partir de Recordatórios Alimentares de 24 horas**

Luana Giongo Pedrotti

Orientadora: Prof<sup>ª</sup> Dr<sup>ª</sup> Vanessa Bielefeldt Leotti

Coorientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Michele Drehmer

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

Porto Alegre

28 de setembro de 2020

## Banca Examinadora

Profa. Dra. Eurídice Martínez Steele, Universidade de São Paulo.

Profa. Dra. Sandra Patrícia Crispim, Universidade Federal do Paraná.

Prof. Dr. Rodrigo Citton Padilha dos Reis, Universidade Federal do Rio Grande do Sul.

# Agradecimentos

Agradeço a todas as pessoas que me acompanharam e me incentivaram nesta caminhada, foram dois anos de muitas mudanças e escolhas difíceis na minha vida, mas que com certeza contribuíram com meu crescimento. Obrigada aos que estiveram e estão ao meu lado!

Um agradecimento especial aos meus pais, Neiva e Maurício, que sempre estiveram presentes mesmo distantes e por todo apoio e suporte de sempre.

Ao Matias, pelo inestimável incentivo.

As minhas colegas de mestrado, Letícia e Luísa, e Alice, por compartilhar momentos de estudos, descontração, cafés, pipocas, além de angústias e sofrimentos.

À professora Vanessa, agradeço por todo aprendizado, paciência e incentivo ao longo de todo trabalho. Saiba que além de uma grande cientista, é uma grande mulher!

À professora Michele, obrigada pelo auxílio e conhecimento transmitido.

Ao grupo de pesquisa, Vanessa, Michele, Bruna e Betina, agradeço pela parceria e oportunidade de me aproximar da Nutrição e participar de um projeto tão interessante.

Um agradecimento especial para a equipe da USP, em especial Daniela Neri, Eurídice Steele e Carlos Monteiro, pela disponibilidade em receber a Bruna e eu, além de todo acolhimento, a oportunidade de compartilhar conhecimento.

*Quando penso que cheguei ao meu limite descobro que tenho forças para ir além.*

Ayrton Senna

# Resumo

**OBJETIVO:** Comparar empiricamente, utilizando dados de um estudo realizado em crianças, o efeito das abordagens *first-add-then-shrink* e *first-shrink-then-add* aplicadas na estimação do consumo de produtos ultraprocessados (PUP). **MÉTODOS:** Ensaio clínico randomizado com dados coletados nos anos de 2012 e 2013. A amostra final envolveu 155 crianças de 4 a 7 anos de idade. O consumo alimentar foi avaliado por meio de recordatórios alimentares de 24h e os alimentos consumidos foram agrupados segundo a classificação NOVA. O consumo habitual foi estimado por meio do método SPADE (*Statistical Program to Assess Dietary Exposure*), em calorias e percentual do valor energético total (VET), utilizando duas abordagens distintas, *first-add-then-shrink* e *first-shrink-then-add*. Também, foi utilizado o método de correção pela média para comparação com as abordagens. **RESULTADOS:** Observou-se um aumento do consumo médio de PUP com a idade partindo de 31,57% e 32,63% aos 4 anos para 46,68% e 45,42% aos 7 anos para as abordagens *first-add-then-shrink* e *first-shrink-then-add* respectivamente. Para percentual do VET e para calorias respectivamente, os intervalos interquartis da abordagem *first-shrink-then-add* são mais amplos se comparado à *first-add-then-shrink*: [30,90; 45,28] e [528,16; 826,20] contra [34,18; 44,63] e [587,01; 820,72]. Porém, são menos amplos quando comparados às estimativas corrigida pela média, implicando maior variabilidade na estimação com este método. **CONCLUSÕES:** A abordagem *first-shrink-then-add* foi menos precisa, apresentando maior amplitude nas estimativas quando comparada à *first-add-then-shrink*, em todos os casos. Ambas as abordagens apresentaram distribuições estimadas com menor variabilidade quando comparadas à utilização da correção pela média. Reforça-se a necessidade de utilizar métodos que diminuam o impacto da variabilidade intraindividual para obter estimativas mais precisas e recomendações adequadas de consumo de PUP tendo em vista seus possíveis efeitos na saúde. Ainda se faz necessária a utilização de um estudo de simulação para poder comparar as abordagens em relação a acurácia e precisão, tendo assim um consumo habitual verdadeiro para utilizar como parâmetro de comparação.

**Palavras-chave:** *First-add-then-shrink*. *First-shrink-then-add*. Produtos Ultraprocessados. Classificação NOVA. Epidemiologia Nutricional.

# Abstract

**OBJECTIVE:** Empirically compare the effect of the first-add-then-shrink and first-shrink-then-add approaches on the ultra-processed products (UPP) intake estimation using data from a study carried out in children. **METHODS:** Randomized clinical trial with data collected in the years of 2012 and 2013. The final sample enrolled 155 children aged 4 to 7 years. The intake was assessed by 24-hour dietary recalls (R24h) and the consumed foods were grouped using the NOVA classification. Usual intake was estimated through the SPADE (Statistical Program to Assess Dietary Exposure) method, in calories and percentage contribution to the total energy value (TEV), using two different approaches, first-add-then-shrink and first-shrink-then-add. Also, the mean-corrected estimate was compared with the approaches. **RESULTS:** There is an increase in UPP average intake by age, starting at 31,57% and 32,63% at 4 years and getting to 46,68% and 45,42% at 7 years, for the first-add-then-shrink and first-shrink-then-add approaches respectively. For both percentage of TEV and calories, the interquartile ranges for the first-shrink-then-add approach are wider than for the first-add-then-shrink: [30,90; 45,28] and [528,16; 826,20] against [34,18; 44,63] and [587,01; 820,72]. However, they are narrower when compared to the mean-corrected estimate, implying that this method estimates intake with larger variability. **CONCLUSIONS:** The first-shrink-then-add approach was less precise, with wider range estimates when compared to the first-add-then-shrink. Both approaches resulted in estimated distributions with less variability when compared to the mean-corrected method. We reinforce the need of using statistical methods to reduce the within-person variability impact to obtain more accurate estimates and adequate UPP intake recommendations given their possible effects on health outcomes. The need of a simulation study still remains, aiming to obtain a real usual intake as a parameter of comparison for the approaches in terms of accuracy and precision.

**Keywords:** First-add-then-shrink. First-shrink-then-add. Ultra-processed products. NOVA classification. Nutritional Epidemiology.

# Lista de ilustrações

Figura 1 – Distribuição de consumo alimentar com erro aleatório. . . . .	14
Figura 2 – Distribuição de consumo alimentar com erro sistemático. . . . .	15
Figura 3 – Distribuições de consumo alimentar sem a presença de erro sistemático.	20
Figura 4 – Distribuição de consumo alimentar com a presença de erro sistemático.	21
Figura 5 – Efeito do erro de medida em modelagem de relações entre dieta e desfechos em saúde. . . . .	22
Figura 6 – Esquema de funcionamento do SPADE. . . . .	25

# Lista de abreviaturas e siglas

ECR	Ensaio Clínico Randomizado
ISU	<i>Iowa State University Method</i>
MSM	<i>Multiple Source Method</i>
NCI	<i>National Cancer Institute Method</i>
POF	Pesquisa de Orçamento Familiar
PUP	Produtos Ultraprocessados
QFA	Questionário de Frequência Alimentar
R24h	Recordatório Alimentar de 24 horas
RA	Registro Alimentar
SPADE	<i>Statistical Program to Assess Dietary Exposure</i>
VET	Valor energético total

# Sumário

<b>1</b>	<b>APRESENTAÇÃO</b>	<b>10</b>
<b>2</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>3</b>	<b>REVISÃO DE LITERATURA</b>	<b>13</b>
<b>3.1</b>	<b>Erros de medida</b>	<b>13</b>
<b>3.2</b>	<b>Definição da equação do consumo habitual</b>	<b>17</b>
3.2.1	Erro aditivo constante	17
3.2.2	Erro multiplicativo	17
3.2.3	Erro específico do indivíduo	18
3.2.4	Erro aleatório intraindividual	18
<b>3.3</b>	<b>Efeitos dos erros de medida</b>	<b>19</b>
3.3.1	Estimação da distribuição	19
3.3.2	Em modelos de regressão	21
<b>3.4</b>	<b>Método SPADE</b>	<b>23</b>
3.4.1	Um componente alimentar	24
3.4.1.1	Modelo de uma parte	24
3.4.1.2	Modelo de duas partes	26
3.4.2	Múltiplos componentes alimentares	29
3.4.2.1	Modelo de três partes	29
3.4.2.2	Modelo multiparte de duas ou mais fontes	30
<b>3.5</b>	<b>Produtos Ultraprocessados</b>	<b>31</b>
<b>4</b>	<b>OBJETIVOS</b>	<b>33</b>
	<b>REFERÊNCIAS</b>	<b>34</b>
<b>5</b>	<b>ARTIGO</b>	<b>39</b>
<b>6</b>	<b>CONCLUSÕES E CONSIDERAÇÕES FINAIS</b>	<b>57</b>
	<b>Anexos</b>	<b>58</b>
	<b>Anexo A - Aprovação do Comitê de Ética</b>	<b>58</b>
	<b>Anexo B - Termo de Consentimento Livre e Esclarecido</b>	<b>59</b>
	<b>Anexo C - Recordatório Alimentar de 24 horas</b>	<b>60</b>

# 1 Apresentação

Este trabalho consiste na dissertação de mestrado intitulada “Formas de correção de consumo de produtos ultraprocessados mensurados a partir de Recordatórios Alimentares de 24 horas”, apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em 29 de junho de 2020. O trabalho é apresentado em três partes, na ordem que segue:

1. Introdução, Revisão da Literatura e Objetivo
2. Artigo
3. Conclusões e Considerações Finais

Documentos referentes ao Ensaio Clínico Randomizado, Aprovação do Comitê de Ética, Termo de Consentimento Livre e Esclarecido e Recordatório Alimentar de 24 horas, estão apresentados nos anexos.

## 2 Introdução

As mudanças nos padrões alimentares globais em termos de fontes de alimentos, modos de processamento e distribuição nas últimas décadas levaram ao elevado consumo de alimentos e bebidas altamente processados. Essas mudanças, que décadas atrás começaram a surgir em países de alta renda, hoje atingem bilhões de pessoas em todo o mundo (Popkin, 2011). No Brasil, segundo a última Pesquisa de Orçamento Familiar (POF), a disponibilidade de alimentos ultraprocessados passou de 12,6% para 18,4% entre os anos de 2002 e 2018, para todas as classes de renda (IBGE, 2020). Na população infantil, o consumo de alimentos ultraprocessados também é alto. Estudos em diferentes faixas etárias apontam consumo de ultraprocessados que variam de 38% até 47% do valor energético total diário (Sparrenberger et al., 2015; Fonseca et al., 2018; Rauber et al., 2015).

Evidências científicas apontam que o aumento nas taxas de excesso de peso e doenças crônicas não transmissíveis é decorrente, dentre outros fatores, dessa mudança nos padrões alimentares (Enes e Slater, 2010). Estudos mostram a relação entre o consumo de alimentos ultraprocessados e diversos desfechos relacionados à saúde tais como ganho de peso e obesidade (Canella et al., 2014; Louzada et al., 2015a; Canhada et al., 2020), hipertensão (Mendonça et al., 2017) e até aumento no risco de câncer (Fiolet et al., 2018).

Considerando a relação entre o processamento de alimentos, mais especificamente seu tipo e intensidade, e o aumento da obesidade e de doenças crônicas não transmissíveis, o Guia Alimentar para a População Brasileira (Brasil, 2014) recomenda a classificação NOVA (Monteiro et al., 2010, 2018), um novo método de avaliação do consumo por graus de processamento dos alimentos. Esta classificação divide os alimentos em quatro grupos de acordo com seu grau de processamento e a categoria de maior grau é a dos chamados produtos ultraprocessados (PUP) (Monteiro et al., 2018).

Para a mensuração de consumo alimentar, é importante frisar que não existe um método padrão-ouro, porém existem diversas ferramentas que podem ser adequadas a diferentes contextos, tais como os Recordatórios Alimentares de 24 horas (R24h), Questionários de Frequência Alimentar (QFA) e Registros Alimentares (RA). O principal objetivo a se alcançar com tais ferramentas é conseguir estimar o consumo médio a longo prazo, também conhecido como consumo habitual, o que geralmente requer modelagem estatística para levar em conta as variabilidades intra e interindividuais (Laureano et al., 2016). Apesar disso, no contexto de estimação do consumo de PUP, ainda é muito utilizada a média de repetidas medições como correção para variabilidade intraindividual (Fonseca et al., 2018; Berti et al., 2019).

Dentre os principais métodos estatísticos para estimação da distribuição do consumo

habitual podem-se destacar: *Statistical Program to Assess Dietary Exposure* (SPADE) (Dekkers et al., 2014), *Multiple Source Method* (MSM) (Haubrock et al., 2011), *National Cancer Institute Method* (NCI) (Tooze et al., 2006) e *Iowa State University Method* (ISU) (Nusser et al., 1996). Embora a metodologia geral seja similar, os métodos diferem em alguns detalhes, tais como os modelos estatísticos utilizados, suposições e transformações utilizadas (Souverein et al., 2011).

No presente estudo é utilizado o método SPADE, que lida com consumos provenientes de diversas fontes alimentares, como no caso dos PUP, obtidos a partir de repetidas medições de instrumentos à curto prazo (R24h) e é implementado no *software* livre R. Além disso, o método conta com diferentes modelos onde são utilizadas duas abordagens aplicadas na estimação dos consumos, conhecidas como *first-add-then-shrink* e *first-shrink-then-add* (Slob et al., 2010; Verkaik-Kloosterman et al., 2011), que não foram ainda comparadas na literatura no contexto de alimentos ultraprocessados. Tais abordagens utilizam metodologias diferentes para a estimação do consumo habitual, possivelmente apresentando resultados distintos para um mesmo consumo.

Considerando essas possíveis diferenças nas estimativas advindas da utilização de ambas as abordagens, o objetivo deste trabalho é comparar empiricamente através de um estudo realizado em crianças o efeito das abordagens *first-add-then-shrink* e *first-shrink-then-add* aplicadas na estimação do consumo de PUP e, além disso, compará-las com a média de repetidas medições como correção para variabilidade intraindividual.

## 3 Revisão de Literatura

Tendo em vista o conceito de consumo habitual, definido como o consumo médio diário, a longo prazo, de certo nutriente ou alimento (Freedman et al., 2004), neste capítulo são apresentados conceitos em relação aos tipos de erros que ocorrem em dados de consumo e seus respectivos impactos, assim como estratégias para lidar com tais impactos, além de metodologias e abordagens utilizadas na estimação do consumo habitual. Os dados de consumo podem ser medidos por diversos instrumentos e dentre os mais comuns estão o QFA, R24h e o RA. O QFA tem como propósito obter informações da frequência e, em alguns casos, das porções de alimentos e bebidas consumidos ao longo de um período específico de tempo, tipicamente um mês ou ano. O RA é um autorrelato utilizado para obter informações detalhadas sobre todos os alimentos e bebidas consumidos ao longo de um período de um ou mais dias. Por fim, o R24h, instrumento aplicado para coletar os dados utilizados neste trabalho, consiste em um questionário estruturado por refeição com o objetivo de obter informações detalhadas sobre todos os alimentos e bebidas consumidos em um dia específico (National Cancer Institute, 2015).

Visto que, neste trabalho, o interesse principal é na estimação do consumo de ultraprocessados, também é introduzido o conceito de alimentos ultraprocessados e sua contribuição cada vez maior às dietas nos dias atuais.

A revisão de literatura é estruturada em cinco seções. Primeiramente, versa sobre os erros de medida, com suas definições sob a visão de diferentes referências, até seus impactos. Logo após, é apresentado o método SPADE de estimação do consumo habitual juntamente dos dois tipos de abordagens para estimação que podem ser utilizadas no contexto de produtos ultraprocessados, e por fim, uma apresentação desse grupo.

### 3.1 Erros de medida

Erro de medida se refere, de maneira geral, à diferença entre o valor obtido em uma medição e o verdadeiro valor do que está sendo medido. Para dados de consumo, o erro de medida se refere à diferença entre a ingestão nutricional reportada - durante certo período de tempo - e a verdadeira (National Cancer Institute, 2015).

Diversas fontes de erro podem interferir na medição do consumo habitual, sendo as principais o entrevistador, o método de coleta e o próprio entrevistado (Fisberg et al., 2005). Segundo Willett (1998), os erros são, em geral, divididos em dois tipos: aleatórios e sistemáticos.

O erro aleatório leva este nome por se dever ao acaso, ou seja, se origina de variações

naturais inerentes ao que está sendo medido. A ocorrência deste reduz a precisão das estimativas, mas seus efeitos podem ser reduzidos aumentando o número de observações por indivíduo ou o número de indivíduos (Rutishauser, 2005). Um exemplo é a flutuação diária que pode ocorrer no relato de consumo de certo alimento para um mesmo indivíduo. Na Figura 1, pode-se ver a diferença entre a verdadeira distribuição do consumo e uma distribuição estimada com erro aleatório.

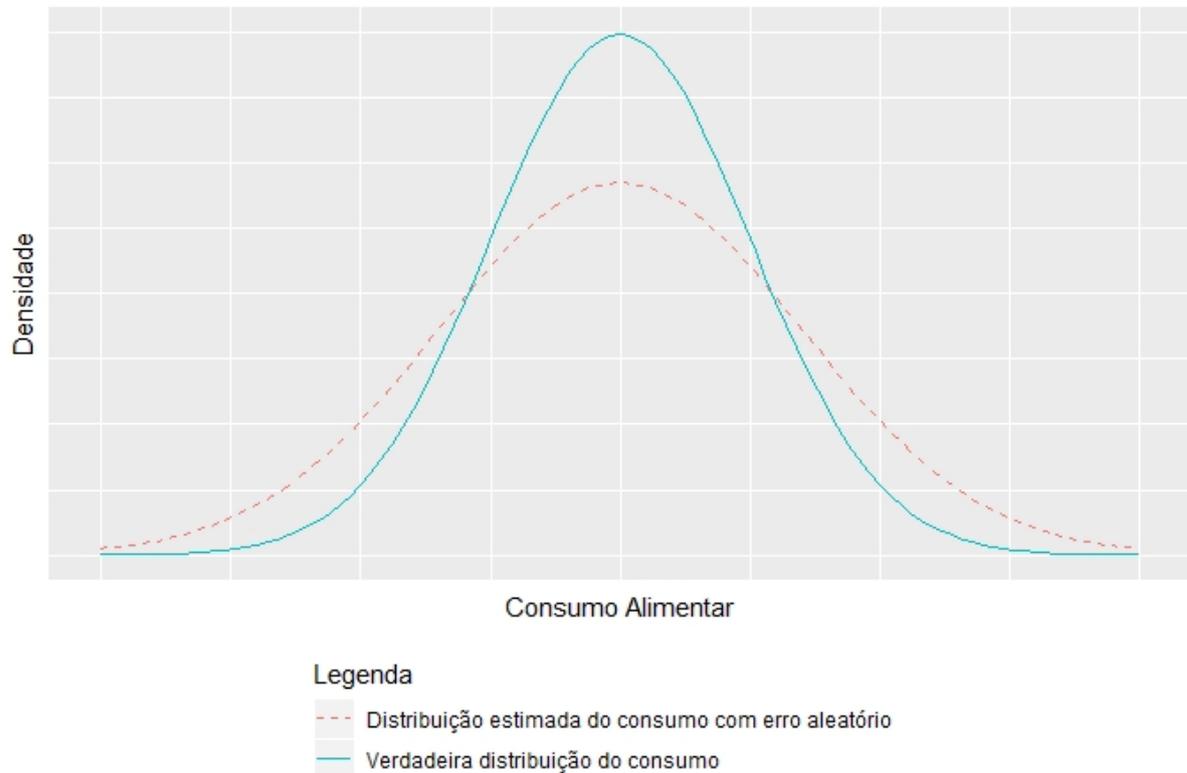


Figura 1 – Distribuição de consumo alimentar com erro aleatório.

Fonte: Elaborado pela autora

Já o erro sistemático se refere a um erro consistente, geralmente causado por instrumentos mal calibrados ou medições incorretas (National Cancer Institute, 2015), não podendo, assim, ter seus efeitos diminuídos com a repetição da medida, visto que ele voltaria a ocorrer (Rutishauser, 2005). Um exemplo é o de uma balança que ao invés de marcar zero quando não há nada sobre ela, marca cem gramas. Desse modo, todas as medições feitas nesta balança terão o acréscimo de cem gramas, caracterizando um erro sistemático. Pode-se notar, pela Figura 2, que as medições de consumo com este tipo de erro se afastam do centro da distribuição verdadeira na mesma direção.

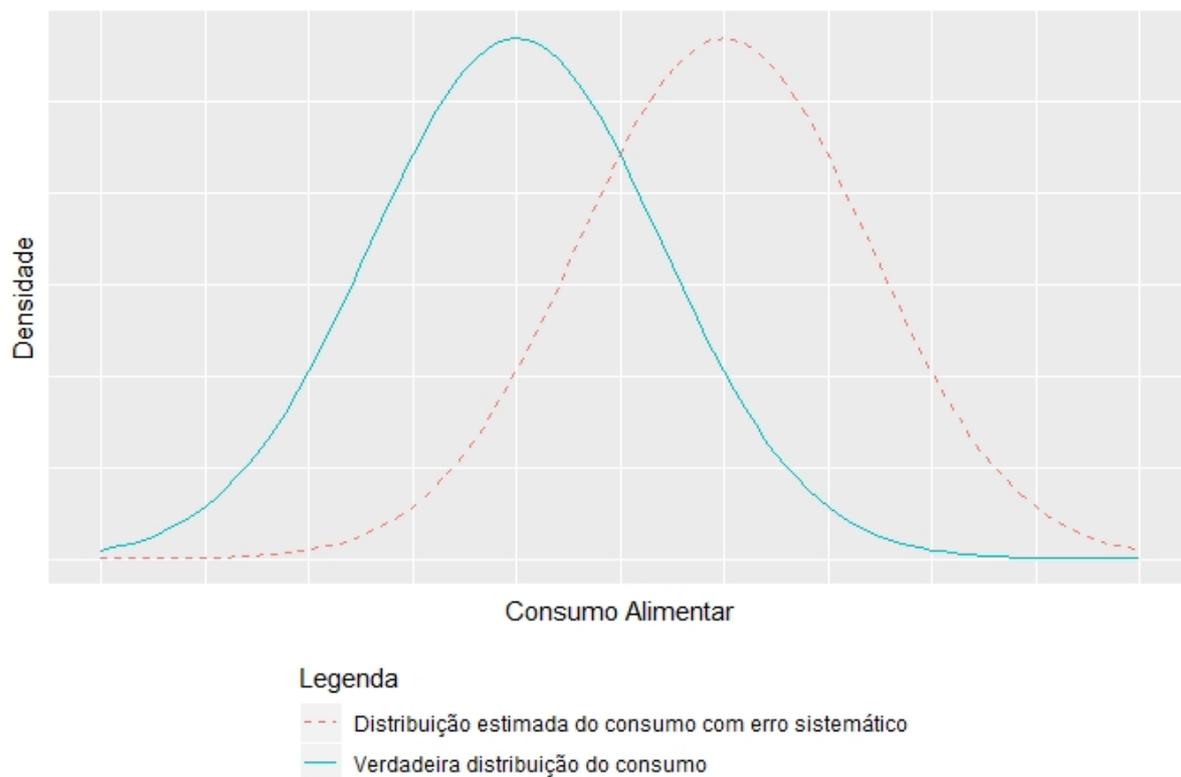


Figura 2 – Distribuição de consumo alimentar com erro sistemático.

Fonte: Elaborado pela autora

Segundo Willett (1998), os erros ocorrem em dois diferentes níveis: interindivíduos (entre os indivíduos) e intraindivíduos (dentro dos indivíduos). Portanto, podem existir ao menos quatro tipos de erros, isto é, erro aleatório intra e interindividual e erro sistemático intra e interindividual. As definições de cada erro, segundo Bennett et al. (2017), são as seguintes:

- O erro aleatório intraindividual é definido como uma variação que é observada no consumo alimentar usando um instrumento específico medido repetidamente no mesmo indivíduo. Por exemplo, a variação diária no consumo alimentar relatado usando vários Recordatórios Alimentares de 24 horas (R24h) para um indivíduo. A variação diária pode ser aleatória e, portanto, resulta em uma estimativa do consumo habitual que é não viesada, o que significa que a verdadeira distribuição do consumo habitual do indivíduo é estimada sem viés, em média, com várias medições repetidas, embora com alguma imprecisão.
- O erro aleatório interindividual, definido em função da diferença entre o consumo do indivíduo e da população, resulta em uma estimativa não viesada do consumo habitual médio para a população de interesse. Mesmo com erro aleatório intraindi-

vidual presente é possível calcular uma estimativa não viesada para a população, equilibrando a superestimação de alguns indivíduos com subestimação para outros.

- O erro sistemático intraindividual ocorre de maneira específica ao indivíduo e se revela como uma diferença positiva ou negativa no consumo relatado. Por exemplo, ocasionalmente alguns indivíduos podem usar suplementos alimentares que podem levar a um erro sistemático aditivo indicando que uma quantidade constante é adicionada a cada consumo diário reportado. Isso pode levar a uma super ou subestimação para esses indivíduos na mesma quantidade. Esse erro geralmente é constante dentro de um indivíduo e permaneceria independentemente de quantas medições repetidas fossem realizadas, e a média das medidas repetidas não se aproximaria à média verdadeira. Então, o erro sistemático intraindividual pode estar relacionado a características individuais, como a conveniência social/cultural, que afeta como um indivíduo reporta seu consumo alimentar e também com instrumentos mal calibrados/utilizados.
- O erro sistemático interindividual pode ser aditivo ou multiplicativo e difere do sistemático intraindividual pois não depende do indivíduo, mas sim do instrumento ou mesmo do entrevistados. O aditivo pode ocorrer quando um instrumento causa medidas muito maiores ou muito menores do que o consumo verdadeiro por uma quantidade constante. Por exemplo, se este for negativo, cada indivíduo relataria um consumo menor do que o consumo verdadeiro usando o instrumento alimentar de interesse. O multiplicativo pode ocorrer quando em vez de relatar seu verdadeiro consumo, todos os participantes relatam um múltiplo fixo deste. Isso pode ser pensado como um viés relacionado ao consumo, onde há um desvio sistemático da verdade devido a uma correlação entre erros no instrumento alimentar de interesse e o verdadeiro consumo. Tais erros, aditivo e multiplicativo, afetam todos os indivíduos na mesma direção, mas não no mesmo grau, pois o consumo dos alimentos difere entre os sujeitos. Causas comuns podem ser a omissão de alimentos frequente consumidos no instrumento utilizado e também, utilização de valores nutritivos incorretos para alguns alimentos.

De acordo com o National Cancer Institute (2015), os dados de consumo são principalmente afetados pelos dois tipos de erros sistemáticos presentes na definição de Willett (1998) e Bennett et al. (2017) e por apenas um tipo de erro aleatório, o intraindividual (do inglês, *within-person random error*). Os erros sistemáticos são definidos com uma nomenclatura diferente: o erro relacionado ao consumo (do inglês, *intake-related bias*) e o erro específico do indivíduo (do inglês, *person-specific bias*), se referindo ao interindividual e ao intraindividual, respectivamente.

O erro relacionado ao consumo consiste na parte do erro sistemático que é uma função do verdadeiro consumo habitual. Um exemplo comum é o fenômeno da “inclinação achatada” (do inglês, *flattened-slope*), no qual pessoas com consumo maior tendem a reportar menos e vice-versa. O erro específico do indivíduo é então, a diferença entre o erro sistemático total e o erro relacionado ao consumo, podendo estar relacionado a características que afetam a maneira com a qual o indivíduo reporta seu consumo (National Cancer Institute, 2015).

No próximo subcapítulo será apresentada a equação do consumo habitual, definida em Kipnis et al. (1999) e no webinar intitulado *Measurement Error Webinar Series* (EGRP/NCI, 2011), com o intuito de apresentar a estrutura dos erros e como eles se relacionam ao consumo. É interessante notar que é apresentada, ainda, outra nomenclatura para os erros.

## 3.2 Definição da equação do consumo habitual

A equação do consumo habitual, de acordo com Kipnis et al. (1999) e EGRP/NCI (2011), é formada por diversos termos e será apresentada de forma crescente para melhor entendimento. Por meio dela, pode-se ver a maneira pela qual os erros de medida se relacionam ao consumo habitual.

### 3.2.1 Erro aditivo constante

Existe um erro a nível de grupo, ou seja, que afeta a todos os indivíduos da mesma maneira, que é o aditivo constante. Este configura um tipo de erro sistemático interindividual, conforme a definição de Willett (1998).

A equação 3.1 define a relação entre consumo reportado e consumo habitual na presença do erro aditivo constante, que ocorre quando um instrumento faz com que todas as medições de consumo sejam aumentadas ou diminuídas de acordo com uma quantidade constante:

$$R_{ij} = \beta_0 + T_i, \quad (3.1)$$

onde  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ ,  $R_{ij}$  é o consumo reportado do indivíduo  $i$  no dia  $j$ ,  $\beta_0$  é o erro aditivo constante e  $T_i$  é o verdadeiro valor do consumo do indivíduo  $i$ .

### 3.2.2 Erro multiplicativo

A seguir, tem-se a equação 3.2, que define a relação entre consumo reportado e consumo habitual na presença do erro aditivo e do erro multiplicativo, que ocorre

quando ao invés de relatar sua verdadeira ingestão habitual, todos os participantes relatam um múltiplo fixo desse consumo. Este erro configura outro tipo de erro sistemático interindividual (Willett, 1998; Bennett et al., 2017), análogo ao relacionado ao consumo (National Cancer Institute, 2015).

$$R_{ij} = \beta_0 + \beta_1 T_i, \quad (3.2)$$

onde  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ ,  $R_{ij}$  é o consumo reportado do indivíduo  $i$  no dia  $j$ ,  $\beta_0$  é o erro aditivo constante,  $\beta_1$  é o erro multiplicativo e  $T_i$  é o verdadeiro valor do consumo do indivíduo  $i$ .

### 3.2.3 Erro específico do indivíduo

Na equação 3.3, tem-se a relação entre consumo reportado e consumo habitual na presença dos erros aditivo e multiplicativo e do erro específico do indivíduo, que é um erro a nível individual, determinada a seguir:

$$R_{ij} = \beta_0 + \beta_1 T_i + u_i, \quad (3.3)$$

onde  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ ,  $R_{ij}$  é o consumo reportado do indivíduo  $i$  no dia  $j$ ,  $\beta_0$  é o erro aditivo constante,  $\beta_1$  é o erro multiplicativo,  $T_i$  é o verdadeiro valor do consumo do indivíduo  $i$  e  $u_i$  é o erro específico do indivíduo  $i$ .

Apesar de ser apresentado como sistemático em Kipnis et al. (1999) e EGRP/NCI (2011) por representar uma quantidade constante adicionada as várias medidas do mesmo indivíduo, pode ser considerado uma quantidade aleatória, pois varia de indivíduo para indivíduo, se relacionando assim com o erro aleatório interindividual (Willett, 1998).

### 3.2.4 Erro aleatório intraindividual

Por fim, na equação 3.4 é definida a relação entre consumo reportado e consumo habitual na presença dos erros anteriores e do erro aleatório intraindividual, completando a equação geral de estimação do consumo habitual:

$$R_{ij} = \beta_0 + \beta_1 T_i + u_i + \epsilon_{ij}, \quad (3.4)$$

onde  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ ,  $R_{ij}$  é o consumo reportado do indivíduo  $i$  no dia  $j$ ,  $\beta_0$  é o erro aditivo constante,  $\beta_1$  é o erro multiplicativo,  $T_i$  é o verdadeiro valor do consumo do indivíduo  $i$ ,  $u_i$  é o erro específico do indivíduo  $i$  e  $\epsilon_{ij}$  é o erro aleatório intraindividual. O índice  $j$  indica o número da medição feita no indivíduo. Dessa forma, tem-se que um termo

é referente ao erro aleatório intraindividual ( $\epsilon_{ij}$ ) e três referentes aos erros sistemáticos ( $\beta_0$ ,  $\beta_1$  e  $u_i$ ).

### 3.3 Efeitos dos erros de medida

É fundamental entender que não se pode sempre controlar e evitar a ocorrência dos erros, sendo importante conhecer seus efeitos, com o objetivo de limitar e mensurar seu impacto, evitando assim conclusões e recomendações equivocadas (Willett, 1998). Um estudo pode sofrer diferentes efeitos a partir de erros de medida, dependendo do seu objetivo e do tipo de erro que o acomete.

Estudos epidemiológicos nutricionais, em geral, podem ter como objetivo: estimar a distribuição de consumo de um componente alimentar, isto é, alimentos e/ou nutrientes e analisar relações de dieta-saúde em modelos de regressão, onde o componente alimentar pode ser tanto o desfecho quanto o preditor. Dito isso, nas próximas subseções serão abordados esses objetivos com maiores detalhes e como a presença dos erros de medida podem impactar em seus resultados. Serão apresentados também possíveis estratégias para lidar com tais efeitos e, ao final, um quadro resumindo tais informações.

#### 3.3.1 Estimação da distribuição

Estudos deste tipo têm como objetivo a estimação de médias, percentis e proporções de consumo acima ou abaixo de certos pontos, bem como estimar a distribuição de tal consumo em certa população. Tais medidas são importantes também para estimação de prevalência do consumo inadequado de certo nutriente (Slater et al., 2004).

Na teoria, utilizar dados provenientes de um instrumento afetado apenas por erro aleatório intraindividual leva a uma distribuição estimada do consumo habitual que é mais larga comparada à verdadeira, ou seja, com maior variabilidade (National Cancer Institute, 2015). Podemos ver na Figura 3 (que assume não haver erro sistemático), que aumentar o número de dias medidos diminui aos poucos a variabilidade da distribuição do consumo, diminuindo assim o efeito do erro aleatório intraindividual, porém é importante frisar que não é o suficiente para removê-lo.

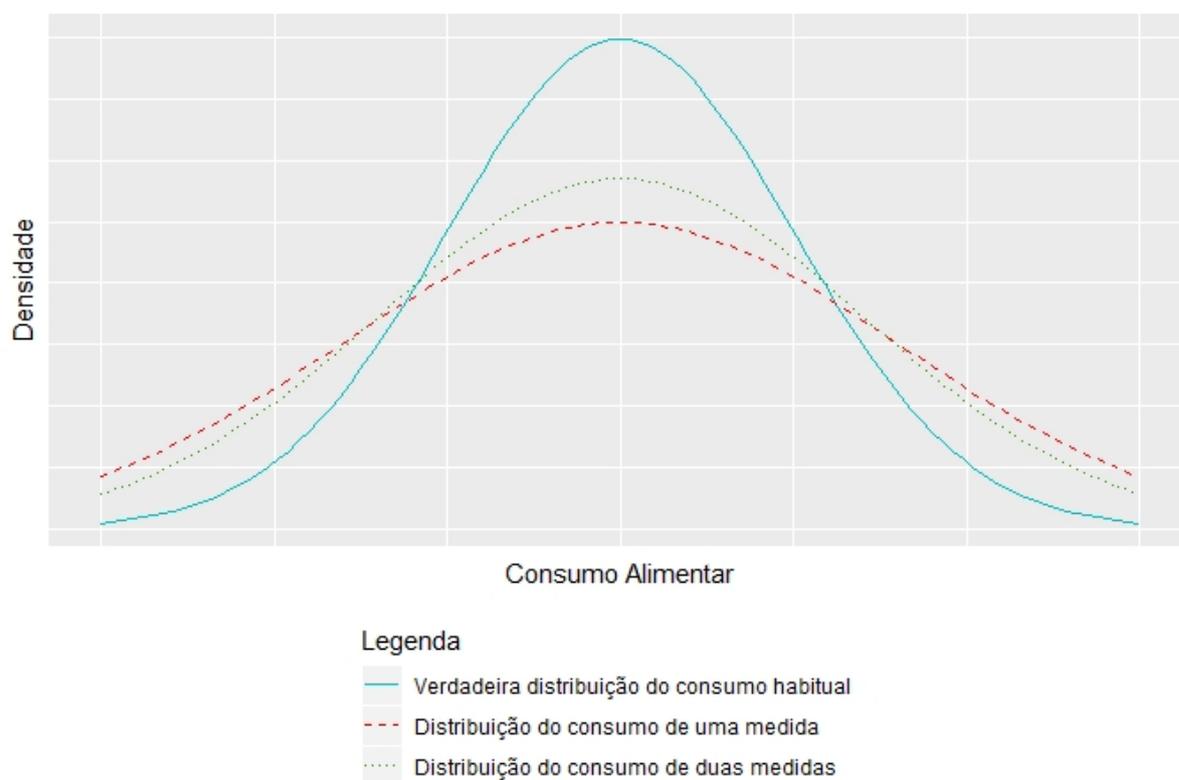


Figura 3 – Distribuições de consumo alimentar sem a presença de erro sistemático.

Fonte: Elaborado pela autora

Uma das principais implicações da ocorrência do erro aleatório na estimação da distribuição é a superestimação das probabilidades caudais, afetando assim as estimativas de frações da população com consumo de certo alimento ou nutriente acima ou abaixo de algum limite. Porém, se o objetivo for apenas estimar a média do consumo habitual, não é necessário o ajuste do erro aleatório, pois a média de uma única medida é igual a média populacional se os dados estiverem razoavelmente distribuídos entre estações e dias da semana (Thompson et al., 1986; Subar et al., 1994).

Por outro lado, o uso de dados provenientes de um instrumento que apresenta erro sistemático leva a uma distribuição do consumo habitual com média e formato diferente da verdadeira. Como visto na Figura 4, igualmente como na ocorrência do erro aleatório, o erro sistemático também afeta as estimativas de frações da população acima ou abaixo de algum limite, porém, ele também acaba afetando a média da população, levando a uma estimativa viesada.

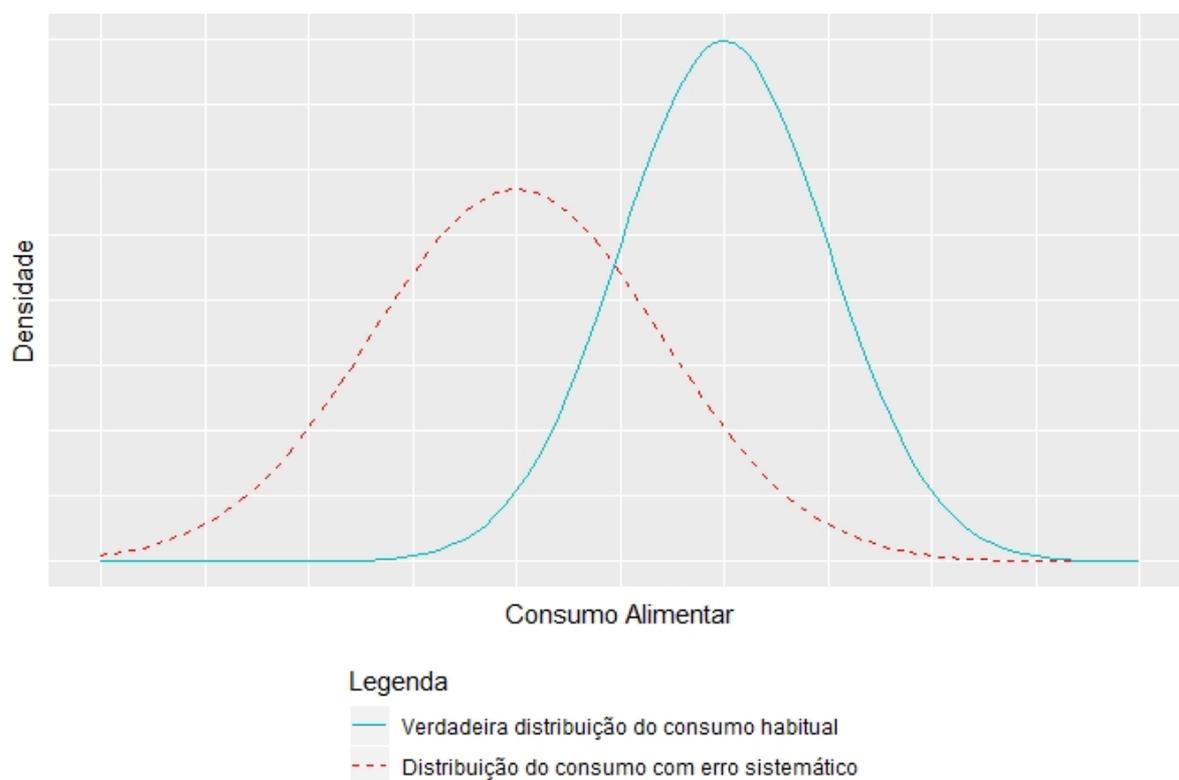


Figura 4 – Distribuição de consumo alimentar com a presença de erro sistemático.

Fonte: Elaborado pela autora

Na prática, dados de instrumentos como o R24h medem o consumo com menos erro sistemático em relação a outros instrumentos e, apesar de ainda ocorrer algum erro, eles representam o melhor cenário possível na ausência de instrumentos de referência (National Cancer Institute, 2015).

### 3.3.2 Em modelos de regressão

Neste caso, o interesse principal é estudar as relações entre exposições relacionadas a dieta, tais como consumo de diferentes nutrientes e desfechos em saúde, bem como a relação entre possíveis intervenções e desfechos relacionados ao consumo. No caso em que a dieta é a exposição relacionada a algum desfecho em saúde, é importante levar em conta se o modelo inclui uma ou mais exposições para determinar os efeitos dos erros.

Quando uma única exposição é medida apenas com erro aleatório intraindividual, a associação entre a exposição e desfecho é sempre atenuada, ou seja, viesada em direção à nulidade. O erro aleatório causa também perda de poder, então mesmo que uma relação exista na população, seria necessário um tamanho de amostra maior para detectá-la. Por fim, a probabilidade do erro tipo I não é afetada, permanecendo controlada ao nível

especificado. Quando ocorre a presença de erro aleatório e sistemático simultaneamente, a associação pode estar viesada em qualquer direção, isto é, atenuada ou inflada. Isso ocorre porque o erro aleatório intraindividual e o erro específico do indivíduo atenuam a associação, mas o erro relacionado ao consumo exagera a associação (National Cancer Institute, 2015).

A Figura 5 mostra o efeito, em geral, dos erros de medida quando o erro aleatório intraindividual e o erro específico do indivíduo predominam sobre o erro relacionado ao consumo ao modelar relações entre dieta e desfechos em saúde, levando à atenuação da associação, ou seja, sua subestimação.

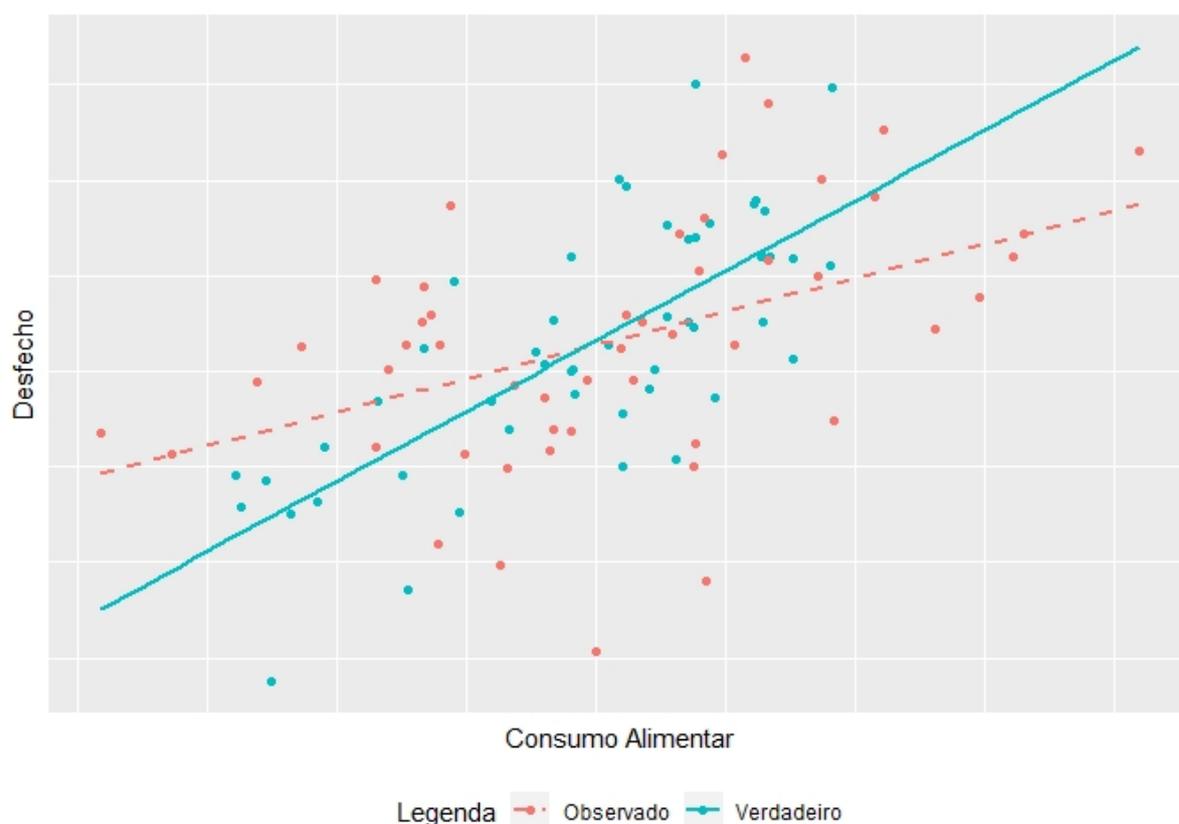


Figura 5 – Efeito do erro de medida em modelagem de relações entre dieta e desfechos em saúde.

Fonte: Elaborado pela autora

Quando o modelo contém outras exposições os efeitos do erro de medida se tornam mais imprevisíveis, devido a possível existência de confundimento. Neste caso, ainda, a probabilidade do erro tipo I pode ser afetada, em contraste com os casos anteriores (National Cancer Institute, 2015).

Por fim, a avaliação do efeito de intervenções em desfechos de consumo na presença de erro de medida pode resultar em estimativas viesadas e falta de poder para detecção de associações. Além do mais, a própria intervenção pode mudar a maneira de reportar o consumo, tornando as interpretações difíceis. Tal diferença pode levar a resultados espúrios,

como, por exemplo, um grupo de intervenção reportar menor consumo de gorduras a um nível maior que um grupo controle, levando a estimativas espúrias do efeito da intervenção no consumo de gordura (National Cancer Institute, 2015).

O Quadro 1 resume os efeitos dos erros para os diferentes casos acima descritos e cita algumas maneiras possíveis de lidar com tais efeitos, de acordo com o National Cancer Institute (2015):

Quadro 1 – Resumo dos erros, seus respectivos efeitos e maneiras de lidar

Tipos/Subtipos	Efeitos	Maneiras de lidar
<b>Erro sistemático</b>	Falta de acurácia: estimativas viesadas para média e distribuição	Utilização de um instrumento de referência
Relacionado ao consumo	Relações exageradas	
Específico do indivíduo	Relações atenuadas; Falta de poder	
<b>Erro aleatório (intraindividual)</b>	Falta de precisão: estimativas dispersas ao redor do verdadeiro valor; Distribuição muito larga, superestimação de probabilidades caudais; Relações atenuadas; Perda de poder	Utilizar várias medidas ou modelar estatisticamente caso o número de repetições seja limitado

Nesta seção foram apresentados os efeitos dos erros de medida e como eles podem afetar os diferentes tipos de objetivos em estudos nutricionais. No presente trabalho são utilizados dados provenientes de R24h, e diversos avanços em modelagem estatística foram desenvolvidos para o melhor uso desse instrumento nos diversos tipos de estudos nutricionais (National Cancer Institute, 2015). Na seguinte seção é apresentado detalhadamente o método para estimação do consumo habitual utilizado neste trabalho, bem como a justificativa para sua utilização.

### 3.4 Método SPADE

Dodd et al. (2006) realizaram uma revisão a respeito da estimação do consumo habitual. Segundo os autores, ele não é bem retratado com a aferição de um único dia de consumo e, por isso, começou-se a utilizar mais de uma medição de R24h por sujeito, ao passo que a estimação da distribuição de consumo habitual passou a ser feita pela distribuição empírica das médias individuais de consumo. Entretanto, tais medidas são insuficientes para estimar os percentis da distribuição de consumo habitual, sendo necessário a utilização de métodos estatísticos que lidem com a variabilidade intraindividual.

Além disso, é importante frisar que o consumo habitual, em geral, possui distribuição assimétrica. Neste caso, não é possível utilizar métodos estatísticos baseados na suposição de normalidade. Portanto, a estratégia mais utilizada nos métodos de estimação do consumo

consiste na transformação dos dados, cuja distribuição, assim, se aproxima da Normal, tornando possível estimar os parâmetros de consumo da população (Dodd et al., 2006).

Diversos métodos foram propostos para a estimação da distribuição do consumo habitual. Esses métodos se diferenciam em relação à forma de estimação dos parâmetros (média e variâncias) e da transformação dos dados. Alguns trabalhos já fizeram a comparação de tais métodos (Souverein et al., 2011; Dodd et al., 2009; Nusser et al., 1996; Laureano et al., 2016).

O *Statistical Program to Assess Dietary Exposure* apresentado em Dekkers et al. (2014) é um método implementado no *software* livre R (R Core Team, 2017) para a estimação do consumo habitual de componentes alimentares a partir de múltiplas fontes, isto é, alimentos e nutrientes ou mesmo diversos alimentos como acontece no caso dos ultraprocessados. Por ser implementado no R e pelas características acima descritas, optou-se pelo método SPADE para este trabalho. O SPADE está disponível para *download* gratuito por meio do pacote SPADE.RIVM (RIVM, 2014).

Os componentes alimentares podem ser divididos em dois tipos: os que são consumidos regularmente por quase todos os indivíduos, definidos como de consumo diário (Nusser et al., 1996; Slob, 1993), e componentes consumidos com menos frequência, definidos como de consumo episódico (Tooze et al., 2006; Souverein et al., 2011). O SPADE conta com opções para modelar ambos os tipos. Além disso, conta com opções para modelar consumos de diferentes fontes, como por exemplo de alimentos e suplementos ou de duas ou mais fontes de alimentos. Por isso, esta seção está dividida em dois subcapítulos, onde o primeiro aborda a estimação a partir de um componente alimentar e o segundo aborda a partir de múltiplos componentes alimentares.

### 3.4.1 Um componente alimentar

Nessa subseção é apresentado os dois tipos de estimação do consumo habitual de um componente alimentar, quanto à frequência de consumo: o consumo diário, o qual está descrito no modelo de uma parte e o consumo episódico, o qual está descrito no modelo de duas partes.

#### 3.4.1.1 Modelo de uma parte

O modelo de uma parte estima o consumo habitual de componentes alimentares de consumo diário. Neste caso, praticamente todos (ou quase todos) os indivíduos devem ter um consumo maior que zero em dois dias observados. Conforme descrito em Dekkers et al. (2017), este modelo consiste nos seguintes passos:

1. Tem-se as distribuições dos consumos observados.

2. Os consumos observados são transformados via transformação de Box-Cox.
3. O consumo transformado é modelado como função da idade e são estimados os parâmetros do modelo, incluindo as variabilidades inter e intraindividuais, posteriormente usadas para obter a distribuição “encolhida” (do inglês, *shrinkage*).
4. Os percentis da distribuição são transformados de volta à escala original e a variabilidade intraindividual é eliminada.

Na figura 6, apresentada em Dekkers et al. (2014), é possível ter uma visão geral do funcionamento acima descrito do SPADE.

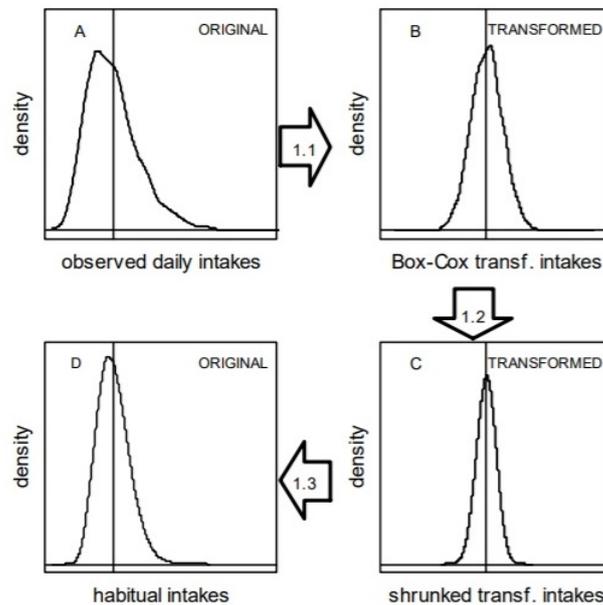


Figura 6 – Esquema de funcionamento do SPADE.

Fonte: Dekkers et al. (2014)

O método de Box-Cox estima o parâmetro de transformação ( $\lambda$ ) pelo método da máxima verossimilhança, resultando em dados simétricos e normalmente distribuídos após a transformação, descrita na equação 3.5.

$$f(x) = \begin{cases} \frac{(x^\lambda - 1)}{\lambda} & \text{para } \lambda \neq 0 \\ \ln(x) & \text{para } \lambda = 0 \end{cases}, \quad (3.5)$$

onde  $f(x)$  é a função de transformação, para  $x > 0$ .

Após a transformação dos consumos diários individuais, o consumo médio individual é modelado por meio de uma regressão polinomial fracionária, dada pelas equações 3.6 e 3.7.

$$Y_{ij} = a + bZ_i^p + cZ_i^q + \varepsilon_{ij}, \quad (i = 1, 2, \dots, n, p \neq q) \quad \text{ou} \quad (3.6)$$

$$Y_{ij} = a + bZ_i^p + cZ_i^q \ln(Z_i) + \varepsilon_{ij}, \quad (i = 1, 2, \dots, n, p = q), \quad (3.7)$$

onde  $n$  é o número de observações,  $p$  e  $q$  as potências do polinômio fracionário  $Y(Z_i)$ ,  $Z_i$  é a idade do indivíduo  $i$  e  $Y_{ij}$  o consumo transformado. As potências  $p$  e  $q$  podem assumir os valores  $\{-2, -1, -0.5, 0, 0.5, 1, 2\}$ , desta maneira os consumos estimados são descritos como uma função da idade por uma família de curvas de no máximo três parâmetros e as frações ideais de  $p$  e  $q$  são estimadas, bem como  $a$ ,  $b$  e  $c$ .

Devido ao fato de termos repetições de medidas de consumo para o mesmo indivíduo, as equações 3.6 e 3.7 precisam ser reajustadas com um Modelo de Efeitos Mistos. Então, a equação para quando  $p \neq q$  é a seguinte:

$$Y_{ij} = a + \alpha_i + bZ_i^p + cZ_i^q + \varepsilon_{ij}, \quad (3.8)$$

onde  $Y_{ij}$  é o consumo transformado do indivíduo  $i$  no dia  $j$ . Assim, cada indivíduo é visto como um “grupo” de observações, permitindo estimar as variâncias interindividual ( $\sigma_b^2$ ) e intraindividual ( $\sigma_w^2$ ), com  $\alpha_i \sim N(0, \sigma_b^2)$  e  $\varepsilon_{ij} \sim N(0, \sigma_w^2)$ . Os resíduos  $\varepsilon_{ij}$  são assumidos como normalmente distribuídos com variância constante ao longo da idade.

Após tais ajustes, é necessário a investigação de possíveis *outliers*, os quais podem influenciar seriamente nas estimativas. No caso de haver necessidade de eliminar *outliers*, os passos de estimação devem ser refeitos até que não exista mais a presença dos mesmos.

Por fim, ocorre a transformação inversa do consumo médio individual pela Quadratura Gaussiana, com o objetivo de obter estimativas na escala original eliminando a variabilidade intraindividual. Após isso, são calculados a média e os percentis que formam a distribuição do consumo alimentar habitual. Para maiores detalhes, sugere-se Waijers et al. (2006).

#### 3.4.1.2 Modelo de duas partes

O modelo de duas partes estima o consumo habitual de componentes alimentares de consumo episódico, que são caracterizados por diversos zeros, ou seja, uma baixa frequência de consumo. Esse número elevado de consumos zero tornam o modelo de uma parte não apropriado, pois não existe transformação de Box-Cox que consiga lidar com esse problema (Dekkers et al., 2017). Neste modelo, por padrão, todos os indivíduos são considerados potenciais consumidores.

A primeira parte do modelo, que visa calcular as probabilidades de consumo, é a seguinte:

1. São estimadas as probabilidades de consumo como uma função da idade via modelo Beta-Binomial (de Boer et al., 2009). Neste caso, todos os indivíduos são utilizados para a estimação da distribuição de probabilidade do consumo.

Considere  $M$  o número de dias entrevistado e  $Y_i^+$  o número de dias com um consumo positivo do indivíduo  $i$ . Assume-se que  $Y_i^+$  segue distribuição Binomial com probabilidade individual  $p_i$ . Para explicar a variabilidade, assume-se um  $p_i$  aleatório e que segue uma distribuição Beta. A distribuição resultante de  $Y_i^+$  é uma distribuição Beta-Binomial com média  $\mu$  e desvio padrão  $\theta$ :

$$P(Y_i^+ = k | \mu, \theta) = \frac{\Gamma(M+1)}{\Gamma(k+1)\Gamma(M-k+1)} \frac{\Gamma(\frac{1}{\theta})\Gamma(k + \frac{\mu}{\theta})\Gamma(M + \frac{1-\mu}{\theta} - k)}{\Gamma(M + \frac{1}{\theta})\Gamma(\frac{\mu}{\theta})\Gamma(\frac{1-\mu}{\theta})}, \quad (3.9)$$

onde  $P(Y_i^+ = k | \mu, \theta)$  representa a probabilidade da variável  $Y_i^+$  assumir um valor  $k$  dados seus parâmetros  $\mu$  e  $\theta$  com  $k = 0, 1, \dots, M$ ,  $0 < \mu < 1$  e  $\theta > 0$ , e  $\Gamma$  representa a função matemática Gama.

Tem-se que  $E(Y_i^+) = M\mu$  e  $Var(Y_i^+) = M\mu(1-\mu)(1 + \frac{\theta}{1+\theta}(M-1))$ . Como assumido anteriormente, a probabilidade de um indivíduo  $i$  ter consumo positivo é  $p_i$  e este segue uma distribuição Beta( $\alpha, \beta$ ), com média e variância, respectivamente:

$$\mu = \frac{\alpha}{\alpha + \beta} \quad (3.10)$$

$$\theta^2 = \frac{\alpha\beta(\alpha + \beta + M)}{(\alpha + \beta)^2(\alpha + \beta + M)} \quad (3.11)$$

Para vincular os parâmetros da distribuição Beta aos da Beta-Binomial considera-se  $\mu = \frac{\alpha}{\alpha + \beta}$  e  $\theta = 1 + \frac{1}{(\alpha + \beta)}$ . Também, que  $\hat{\mu}$  e  $\hat{\theta}$  são as estimativas de máxima verossimilhança dos parâmetros da distribuição Beta-Binomial. Com isso, tem-se as estimativas dos parâmetros da distribuição Beta padronizada:

$$\hat{\alpha} = \frac{\hat{\mu}}{\hat{\theta}} \quad (3.12)$$

$$\hat{\beta} = \frac{1 - \hat{\mu}}{\hat{\theta}} \quad (3.13)$$

Então, tomando  $\pi_i$  como  $P(Y_i^+ = k)$ , o modelo resultante para a frequência de consumo do indivíduo  $i$  é:

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + p_i \quad (3.14)$$

Ao considerar  $p_i$  como uma função da idade e o uso de splines cúbicos da idade  $Z_i$  para haver mais flexibilidade, tem-se que a equação 3.14 se transforma em:

$$\text{logit}(\pi_i) = \beta_0 + cs(Z_i; 2), \quad (3.15)$$

o que significa que o parâmetro  $\mu$  do modelo Beta-Binomial é um spline cúbico da idade e possui dois graus de liberdade.

A segunda parte do modelo trata das quantidades em dias com consumo e neste caso, apenas indivíduos com no mínimo um dia de consumo positivo são envolvidos, pois dois consumos zero não fornecem nenhuma informação sobre o consumo habitual. O modelo consiste nos seguintes passos:

1. Selecionam-se os consumos positivos de todos os indivíduos com apenas um consumo positivo. Dos indivíduos com dois consumos positivos, seleciona-se um aleatoriamente. Indivíduos com dois consumos zero são omitidos. Tais consumos são transformados via Box-Cox (descrito na Equação 3.5).
2. Os consumos transformados são modelados como função da idade (Equações 3.6 e 3.7) e todos os parâmetros são estimados, incluindo a variância residual total deste modelo ( $\sigma^2$ ).
3. As variâncias intra e interindividuais ( $\sigma_{Tw}^2$  e  $\sigma_{Tb}^2$ , respectivamente) são estimadas baseadas nos dados de indivíduos com dois consumos positivos e nos resíduos do modelo ajustado no passo anterior, calculados na Equação 3.16.

$$r_{ij} = y_{ij} - \{\hat{a} + \hat{b}z_i^{\hat{p}} + \hat{c}z_i^{\hat{q}}\} \quad (3.16)$$

4. Ocorre a aplicação da razão de ambas as variâncias obtidas no passo 3 à variância residual total do passo 2 como uma estimativa para as variabilidades intra e interindividuais, como mostram as Equações 3.17 e 3.18.

$$\sigma_w^2 = \frac{\sigma_{Tw}^2 \sigma^2}{(\sigma_{Tw}^2 + \sigma_{Tb}^2)} \quad (3.17)$$

e

$$\sigma_b^2 = \frac{\sigma_{Tb}^2 \sigma^2}{(\sigma_{Tw}^2 + \sigma_{Tb}^2)} \quad (3.18)$$

Com tais variâncias, o modelo ajustado no passo 2 contém toda informação necessária para ocorrer a transformação de volta para a escala original. Os consumos habituais

são então derivados como a probabilidade de consumo multiplicada pela quantidade de consumo, obtidas via simulação de Monte Carlo (Dekkers et al., 2014).

Para os modelos acima, o SPADE considera todos os indivíduos como “potenciais consumidores”, que são aqueles que não são “não-consumidores”, podendo vir a ter consumo zero em todos os dias avaliados, porém têm probabilidade positiva de ter consumo positivo. Os “não-consumidores” são indivíduos que indicaram que nunca consomem certo alimento. Quando tal informação está disponível o modelo a leva em consideração. Para maiores detalhes sobre este tipo de modelagem, recomenda-se Dekkers et al. (2017).

### 3.4.2 Múltiplos componentes alimentares

Nessa subseção são apresentados mais dois tipos de estimação do consumo, agora a partir de múltiplos componentes alimentares: o modelo de três partes e o modelo multiparte para duas ou mais fontes. O modelo de três partes pode ser útil para estimação do consumo de vitaminas, por exemplo, pois além dos alimentos sua ingestão pode vir de suplementos vitamínicos. O modelo multiparte, por sua vez, é adequado para casos em que se pretende estimar um consumo que é formado por diversos componentes alimentares, que é o caso dos ultraprocessados.

#### 3.4.2.1 Modelo de três partes

Com o objetivo de estimar o consumo total proveniente de alimentos e suplementos alimentares foi criado o modelo de três partes, baseado em Verkaik-Kloosterman et al. (2011). O modelo faz uso do modelo de duas partes (apresentado anteriormente) e do modelo de consumo de suplementos alimentares, descrito em detalhes em Dekkers et al. (2014). Além disso, utiliza-se da amostragem de “pseudo-indivíduos” apresentada em Tooze et al. (2006). A ideia básica deste modelo é dada nos três seguintes passos:

1. Estima-se, com o modelo de duas partes, o consumo habitual de alimentos como função da idade somente para não-consumidores de suplementos.
2. Estima-se, também com o modelo de duas partes, o consumo habitual de alimentos como função da idade somente para consumidores de suplementos.
3. Estima-se, com o modelo de duas partes de consumo de suplementos alimentares, a quantidade de consumo de suplementos.

A amostragem de pseudo-indivíduos, ou seja, a simulação de indivíduos gerados baseados nos parâmetros dos indivíduos presentes na amostra, é utilizada para:

- a. Consumo habitual de alimentos de indivíduos que não consomem suplementos.

- b. Consumo habitual de alimentos e de suplementos de indivíduos que consomem suplementos e a seguir, somam-se ambos os consumos para cada indivíduo para obter o consumo habitual total.

Combinam-se a e b para obter a amostra de pseudo-indivíduos completa. Os passos 1, 2 e 3 são a parte de correção (do inglês, *shrinkage*) e a e b são a parte de adição (do inglês *add*). Este método, apresentado em Slob et al. (2010) e Dodd et al. (2009), é chamado de abordagem *first-shrink-then-add* em Verkaik-Kloosterman et al. (2011), a qual consiste em aplicar uma correção aos consumos e após combiná-los em um consumo total.

Segundo Dekkers et al. (2014), o modelo de três partes consiste em estimar primeiro o consumo habitual de alimentos em geral, seja em usuários e não-usuários de suplementos e posteriormente estimar apenas o consumo de suplementos separadamente. A estimação em cada parte se dá de maneira análoga aos modelos anteriores, porém as estimativas das três partes são posteriormente combinadas constituindo a estimativa final do consumo habitual, caracterizando assim a abordagem *first-shrink-then-add*.

Verkaik-Kloosterman et al. (2011) mostram que este é o método correto para evitar a multimodalidade que pode ser causada pelas diferentes fontes consideradas, que vem a ser um problema pois, para distribuições multimodais, as transformações geralmente utilizadas para tornar a distribuição simétrica falham com frequência (Slob et al., 2010). Por outro lado, o consumo habitual proveniente de alimentos e suplementos pode ser estimado de maneira mais simples, somando-se as quantidades antes de aplicar uma correção para a variabilidade intraindividual, caracterizando assim a abordagem chamada de *first-add-then-shrink*. Tal abordagem não leva em conta as potenciais diferenças nos parâmetros ou variâncias do modelo que se devem à utilização de consumos de diferentes fontes e não evita a presença de multimodalidade (Verkaik-Kloosterman et al., 2011).

#### 3.4.2.2 Modelo multiparte de duas ou mais fontes

O modelo multiparte também utiliza a abordagem *first-shrink-then-add* e consiste da combinação dos diferentes tipos de modelagens apresentados, sendo todos opcionais, porém ao menos um dos seguintes deve ser utilizado:

1. Uma ou mais fontes (para consumos diários ou episódicos).
2. Uma fonte com não-consumidores (consumo episódico).
3. Suplementos.

Todas as fontes são modeladas separadamente (correção) e ao final todos os consumos para pseudo-indivíduos são adicionados (adição) o que demonstra a utilização da abordagem *first-shrink-then-add*.

Ambas as abordagens descritas nas subseções anteriores podem ser utilizadas de acordo com o consumo que se deseja estimar. Segundo Verkaik-Kloosterman et al. (2011) e Slob et al. (2010), a abordagem *first-shrink-then-add* é mais adequada para consumos que são provenientes de múltiplas fontes, enquanto que a abordagem *first-add-then-shrink* não leva em conta as diferenças advindas das fontes, porém é mais simples de se aplicar. Com o intuito de atender o objetivo do presente estudo, as duas abordagens serão utilizadas na estimação do consumo dos alimentos ultraprocessados, a serem apresentados em maiores detalhes na seguinte seção.

### 3.5 Produtos Ultraprocessados

Grande parte das avaliações nutricionais disponíveis, assim como as recomendações dietéticas em nível mundial, fazem uso de classificações de alimentos que por vezes ignoram ou minimizam o impacto do processamento industrial dos alimentos (Monteiro et al., 2010; Kahn et al., 1997). Em Monteiro (2011), tal impacto é discutido e é ressaltada a importância de fazer distinções nos diferentes tipos de processamento, assim surgindo o termo “produtos ultraprocessados” (PUP).

Os alimentos ultraprocessados são definidos segundo a classificação NOVA (Monteiro et al., 2010, 2018), a qual agrupa alimentos de acordo com a extensão e propósito do seu processamento industrial (Monteiro et al., 2019). A classificação divide os alimentos em quatro grandes grupos, em ordem crescente de grau de processamento: alimentos *in natura* ou minimamente processados; ingredientes culinários processados; alimentos processados; e, por fim, os alimentos ultraprocessados. No Guia Alimentar para a População Brasileira (Brasil, 2014) é recomendada a utilização desta classificação, e os possíveis impactos do consumo excessivo de ultraprocessados são discutidos, desde impactos sociais e culturais até doenças crônicas não transmissíveis, como obesidade e diabetes.

O grupo 1 é composto por alimentos *in natura* e minimamente processados, que são alimentos com alterações mínimas tais como remoção de partes não-comestíveis, desidratação, fracionamento, pasteurização, refrigeração e fermentação não-alcóolica. Nenhum desses processos adiciona nenhum tipo de substância ao alimento original e tem como principal objetivo estender a vida de grãos, vegetais, leites e carnes permitindo seu armazenamento por mais tempo e, por vezes, facilitar a sua preparação.

O grupo 2 consiste em ingredientes culinários processados, que são substâncias obtidas diretamente do grupo 1 ou da natureza, como óleos e gorduras, açúcares e sal. São criados via processos industriais tais como pressão, refinamento e extração e seu uso principal é na preparação, tempero e cozimento dos alimentos do grupo 1.

O grupo 3 compreende os alimentos processados. Estes são produzidos com adição de sal, açúcar ou outras substâncias encontradas nos grupos anteriores e com a utilização

de métodos de preservação e conserva, tais como enlatados e engarrafados e, no caso de pães e queijos, há a utilização de fermentação não-alcóolica. Os processamentos deste grupo tem o objetivo de aumentar a durabilidade dos alimentos do grupo 1 e torná-los mais agradáveis modificando ou acentuando suas qualidades sensoriais.

O grupo 4 compreende os alimentos ultraprocessados. São formulações de ingredientes de uso exclusivo da indústria que resultam de uma série de diferentes processamentos. Alguns processos empregados na produção de ultraprocessados incluem fracionamento de alimentos em substâncias, que posteriormente são submetidos a hidrólise e outras modificações químicas. Após isso, tais substâncias são recombinadas via processamentos tais como extrusão e pré-fritura. Ao final são adicionados conservantes e aditivos cosméticos (corantes, aromas, saborizantes, realçadores de sabor) e tais produtos geralmente são finalizados com embalagens sofisticadas feitas com materiais sintéticos (Monteiro et al., 2019).

Os processos e ingredientes usados na produção de ultraprocessados têm como foco a criação de produtos convenientes (prontos para consumir), de elevada palatabilidade e altamente lucrativos, visto que possuem prazos de validade mais extensos e geralmente resultam em produtos finais de baixo custo, além de serem produtos de fácil acesso, não sujeitos a sazonalidade.

Uma maneira fácil de identificar um ultraprocessado é conferir se na sua lista de ingredientes está presente pelo menos um item característico do grupo de ultraprocessados da NOVA que são substâncias alimentícias raramente (ou nunca) utilizadas na cozinha (tais como óleos hidrogenados e proteínas hidrolisadas) ou classes de aditivos destinados a fazer o produto final mais palatável (saborizadores, corantes e emulsificantes).

Por fim, é importante também notar que, se comparados com outros tipos de alimentos, eles possuem menos fibras e proteínas, além de serem de alta densidade calórica, quando sólidos (Monteiro, 2011; Moubarac et al., 2013). Também, costumam ser ricos em açúcares adicionados, gorduras saturadas e sal (Monteiro et al., 2019).

Trabalhos sobre estimação do consumo de PUP têm sido mais frequentes nos últimos anos devido à mudanças nos padrões alimentares e também a sua caracterização como possíveis fatores de risco para desfechos relacionados à saúde (Enes e Slater, 2010; Monteiro et al., 2010). Nesses trabalhos, é frequente a utilização da média de diversas medições como método de correção para variabilidade intraindividual (Sparrenberger et al., 2015; Fonseca et al., 2018; Adjibade et al., 2019; Schnabel et al., 2019) e a abordagem *first-add-then-shrink* também acaba aparecendo como opção mais usual em relação à *first-shrink-then-add* (Louzada et al., 2015b). Tendo em vista os conceitos de PUP e as abordagens acima citadas, na seguinte seção são apresentados os objetivos deste trabalho.

## 4 Objetivos

Comparar empiricamente através de um estudo realizado em crianças com idades entre 4 e 7 anos o efeito das abordagens *first-add-then-shrink* e *first-shrink-then-add* e, também, compará-las com a média de repetidas medições como correção para variabilidade intraindividual. Além disso, divulgar e expandir a utilização das abordagens citadas para uma maior gama de aplicações, não se restringindo apenas ao já existente na literatura, buscando esclarecer para a comunidade acadêmica as diferenças entre ambas abordagens.

# REFERÊNCIAS

- Adjibade, M., Julia, C., Allès, B., Touvier, M., Lemogne, C., Srour, B., Hercberg, S., Galan, P., Assmann, K., e Kesse-Guyot, E. (2019). Prospective association between ultra-processed food consumption and incident depressive symptoms in the french nutrinet-santé cohort. *BMC Medicine*, 17:78. Citado na página 32.
- Bennett, D., Landry, D., Little, J., e Minelli, C. (2017). Systematic review of statistical approaches to quantify, or correct for, measurement error in a continuous exposure in nutritional epidemiology. *BMC Medical Research Methodology*. Citado 3 vezes nas páginas 15, 16 e 18.
- Berti, T., Rocha, T., Curioni, C., Verly, E., Bezerra, F., Canella, D., e Faerstein, E. (2019). Consumo alimentar segundo o grau de processamento e características sociodemográficas: Estudo pró-saúde. *Revista Brasileira de Epidemiologia*, 22. Citado na página 11.
- Brasil, M. d. S. (2014). Guia alimentar para a população brasileira. Citado 2 vezes nas páginas 11 e 31.
- Canella, D., Levy, R., Martins, A., Claro, R., Moubarac, J.-C., Baraldi, L., Cannon, G., e Monteiro, C. (2014). Ultra-processed food products and obesity in brazilian households (2008–2009). *PLoS ONE*, 9(3). Citado na página 11.
- Canhada, S., Luft, V., Giatti, L., Duncan, B., Chor, D., Fonseca, M., Matos, S., Molina, M., Barreto, S., Levy, R., e Schmidt, M. (2020). Ultra-processed foods, incident overweight and obesity, and longitudinal changes in weight and waist circumference: the brazilian longitudinal study of adult health (elsa-brasil). *Public Health Nutrition*, 23 (6):1076–1086. Citado na página 11.
- de Boer, W., van der Voet, H., Bokkers, B., Bakker, M., e Boon, P. (2009). Comparison of two models for the estimation of usual intake addressing zero consumption and non-normality. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess*, 26:1433–49. Citado na página 27.
- Dekkers, A., Verkaik-Kloosterman, J., e Ocké, M. (2017). *SPADE: Statistical Program to Assess habitual Dietary Exposure. User's manual*. RIVM. Citado 4 vezes nas páginas 24, 26, 29 e 57.
- Dekkers, A., Verkaik-Kloosterman, J., van Rossum, C., e Ocké, M. (2014). Spade, a new statistical program to estimate habitual dietary intake from multiple food sources and dietary supplements. *The Journal of Nutrition*, 144:2083–91. Citado 5 vezes nas páginas 12, 24, 25, 29 e 30.

- Dodd, K., Bailey, R., Wilger, J., Sempos, C., Dwyer, J., Radimer, K., McDowell, M., Johnson, C., e Picciano, M. (2009). Estimating distributions of usual total nutrient intake: a comparison of available methods. In *7th International Conference on Diet and Activity Methods (ICDAM 7)*. Citado 2 vezes nas páginas 24 e 30.
- Dodd, K. W., Guenther, P., L.S., F., Subar, A., V., K., Midthune, D., Tooze, J., e Krebs-Smith, S. (2006). Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. *Journal of The American Dietetic Association*, 106(10):1640–50. Citado 2 vezes nas páginas 23 e 24.
- EGRP/NCI, E. G. R. P. (2011). Measurement error webinar series. Citado 2 vezes nas páginas 17 e 18.
- Enes, C. e Slater, B. (2010). Obesidade na adolescência e seus principais fatores determinantes. *Revista Brasileira de Epidemiologia*, 13:163–71. Citado 2 vezes nas páginas 11 e 32.
- Fiolet, T., Srour, B., Sellem, L., Kesse-Gyout, E., Alles, B., Mejean, C., Deschasaux, M., Fassier, P., Latino-Martel, P., Beslay, M., Hercberg, S., Lavalette, C., Monteiro, C., Julia, C., e Touvier, M. (2018). Consumption of ultra-processed foods and cancer risk: results from nutrinet-santé prospective cohort. *British Medical Journal*, 360. Citado na página 11.
- Fisberg, R., Slater, B., Marchioni, D., e Martini, L. (2005). *Inquéritos Alimentares Métodos e bases científicas*. Manole, 1 edition. Citado na página 13.
- Fonseca, P., Ribeiro, S., Andreoli, C., Carvalho, C., Pessoa, M., Novaes, J., Priore, S., e Franceschini, S. (2018). Association of exclusive breastfeeding duration with consumption of ultra-processed foods, fruit and vegetables in brazilian children. *European Journal of Nutrition*, 7:2887–2894. Citado 2 vezes nas páginas 11 e 32.
- Freedman, L., Midthune, D., Carroll, R., Krebs-Smith, S., Subar, A., Troiano, R., Dodd, K., Schatzkin, A., Ferrari, P., e Kipnis, V. (2004). Adjustments to improve the estimation of usual dietary intake distributions in the population. *Journal of Nutrition*, 134:1836–43. Citado na página 13.
- Haubrock, J., Nöthlings, U., Volatier, J., Dekkers, A., Ocké, M., e Harttig, U. (2011). Estimating usual food intake distributions by using the multiple source method in the epic-potsdam calibration study. *Journal of Nutrition*, 141:914–20. Citado na página 12.
- IBGE, C. d. T. e. R. (2020). Pesquisa de orçamentos familiares 2017-2018: Avaliação nutricional da disponibilidade domiciliar de alimentos no brasil. Citado na página 11.

- Kahn, H., Tatham, L., Rodriguez, C., Calle, E., Thun, M., e Heath, C. (1997). Stable behaviors associated with adults' 10-year change in body mass index and likelihood of gain at the waist. *American Journal of Public Health*, 87:747–54. Citado na página 31.
- Kipnis, V., Carroll, R., Freedman, L., e Li, L. (1999). Implications of a new dietary measurement error model for estimation of relative risk: Application to four calibration studies. *American Journal of Epidemiology*, 150:642–651. Citado 2 vezes nas páginas 17 e 18.
- Laureano, G., Torman, V., Crispim, S., Dekkers, A., e Camey, S. (2016). Comparison of the isu, nci, msm and spade methods for estimating usual intake: a simulation study of nutrients consumed daily. *Nutrients*, 8:166. Citado 2 vezes nas páginas 11 e 24.
- Louzada, M., Baraldi, L., Steele, E., Martins, A., Canella, D., Moubarac, J.-C., Levy, R., Cannon, G., Afshin, A., Imamura, F., Mozaffarian, D., e Monteiro, C. (2015a). Consumption of ultra-processed foods and obesity in brazilian adolescents and adults. *Preventive Medicine*, 81:9–15. Citado na página 11.
- Louzada, M., Martins, A., Canella, D., Baraldi, L., Levy, R., Claro, R., Moubarac, J., Cannon, G., e Monteiro, C. (2015b). Alimentos ultraprocessados e perfil nutricional da dieta no brasil. *Revista de Saúde Pública*, 49:38. Citado na página 32.
- Mendonça, R., Lopes, A., Pimenta, A., Gea, A., Martinez-Gonzalez, M., e Bes-Rastrollo, M. (2017). Ultra-processed food consumption and the incidence of hypertension in a mediterranean cohort: The seguimiento universidad de navarra project. *American Journal Hypertension*, 30(4):358–366. Citado na página 11.
- Monteiro, C. (2011). Processing. the good, the bad, and the toxic. *WPHNA*, 2. Citado 2 vezes nas páginas 31 e 32.
- Monteiro, C., Cannon, G., Levy, R., Moubarac, J.-C., Louzada, M., Rauber, F., Khandpur, N., Cediel, G., Neri, D., Martinez-Steele, E., Baraldi, L., e Jaime, P. (2019). Ultra-processed foods: What they are and how to identify them. *Public Health Nutr.*, 22:936–41. Citado 2 vezes nas páginas 31 e 32.
- Monteiro, C., Cannon, G., Moubarac, J.-C., Levy, R., Louzada, M., e Jaime, P. (2018). The un decade of nutrition, the nova food classification and the trouble with ultra-processing. *Public Health Nutrition*, 21 (1):5–17. Citado 2 vezes nas páginas 11 e 31.
- Monteiro, C., Levy, R., Claro, R., e Castro, I. (2010). A new classification of foods based on the extent and purpose of their processing. *Caderno de Saúde Pública*, 26:2039–49. Citado 3 vezes nas páginas 11, 31 e 32.

- Moubarac, J.-C., Martins, A., Claro, R., Levy, R., Cannon, G., e Monteiro, C. (2013). Consumption of ultra-processed foods and likely impact on human health. evidence from canada. *Public Health Nutr.*, 16:2240–48. Citado na página 32.
- National Cancer Institute, N. (2015). Dietary assessment primer. Acessado em maio de 2019. Citado 9 vezes nas páginas 13, 14, 16, 17, 18, 19, 21, 22 e 23.
- Nusser, S., Carriquiry, A., Dodd, K., e Fuller, W. (1996). A semiparametric transformation approach to estimating usual daily intake distributions. *Journal of the American Statistical Association*, 91:1440–49. Citado 2 vezes nas páginas 12 e 24.
- Popkin, B. (2011). Contemporary nutritional transition: determinants of diet and its impact on body composition. *National Institute of Health*, 70:82–91. Citado na página 11.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL=<https://www.R-project.org/>. Citado na página 24.
- Rauber, F., Campagnolo, P. D. B., Hoffman, D. J., e Vitolo, M. R. (2015). Consumption of ultra-processed food products and its effects on children’s lipid profiles: A longitudinal study. *Nutrition, Metabolism and Cardiovascular Diseases*, 25:116–122. Citado na página 11.
- RIVM (2014). National institute for public health and the environment, spade package. Acessado em abril de 2019. Citado na página 24.
- Rutishauser, I. (2005). Dietary intake measurements. *Public Health Nutrition*, pages 1100–1107. Citado na página 14.
- Schnabel, L., Kesse-Guyot, E., Allès, B., Touvier, M., Srour, B., Hercberg, S., Buscail, C., e Julia, C. (2019). Association between ultraprocessed food consumption and risk of mortality among middle-aged adults in france. *JAMA Internal Medicine*, 179(4):490–498. Citado na página 32.
- Slater, B., Marchioni, D., e Fisber, R. (2004). Estimating prevalence of inadequate nutrient intake. *Revista Saúde Pública*. Citado na página 19.
- Slob, W. (1993). Modeling long-term exposure of the whole population to chemicals in food. *Risk Analysis*, 13:525–30. Citado na página 24.
- Slob, W., de Boer, W., e van der Voet, H. (2010). Can current dietary exposure models handle aggregated intake from different foods? a simulation study for the case of two foods. *Food Chem Toxicol.*, 48:178–86. Citado 3 vezes nas páginas 12, 30 e 31.

- Souverein, O., Dekkers, A., Geelen, A., Haubrock, J., de Vries, J., Ocké, M., Harttig, U., Boeing, H., e van 't Veer, P. (2011). Comparing four methods to estimate usual intake distributions. *European Journal of Clinical Nutrition*, 65:92–101. Citado 2 vezes nas páginas 12 e 24.
- Sparrenberger, K., Friedrich, R., Schiffner, M., Schuch, I., e Wagner, M. (2015). Consumo de alimentos ultraprocessados entre crianças de uma unidade básica de saúde. *Jornal de Pediatria*, 91:535–542. Citado 2 vezes nas páginas 11 e 32.
- Subar, A., Frey, C., Harlan, L., e Kahle, L. (1994). Differences in reported food frequency by season of questionnaire administration: the 1987 national health interview survey. *Epidemiology*, 5:226–233. Citado na página 20.
- Thompson, F., Larkin, F., e Brown, M. (1986). Weekend-weekday differences in reported dietary intake: The nationwide food consumption survey, 1977-78. *Nutrition Research*, 6:647–662. Citado na página 20.
- Tooze, J., Midthune, D., Dodd, K., Guenther, P., Freedman, L., Subar, A., Kipnis, V., e Krebs-Smith, S. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of American Diet Association*, 106:1575–87. Citado 3 vezes nas páginas 12, 24 e 29.
- Verkaik-Kloosterman, J., Dodd, K., Dekkers, A., Veer, P., e Ocké, M. (2011). A three-part, mixed-effects model to estimate the habitual total vitamin d intake distribution from food and dietary supplements in dutch young children. *Journal of Nutrition*, 141:2055–63. Citado 4 vezes nas páginas 12, 29, 30 e 31.
- Waijers, P., Dekkers, A., J.M.A., B., Boshuizen, H., e van Rossum, C. (2006). The potential of age mode, an age-dependent model, to estimate usual intakes and prevalences of inadequate intakes in a population. *The Journal of Nutrition*, 136:2916–2920. Citado na página 26.
- Willett, W. (1998). *Nutritional epidemiology*, volume 30. Oxford University Press. Citado 6 vezes nas páginas 13, 15, 16, 17, 18 e 19.

## 5 Artigo

Formas de correção de consumo de produtos ultraprocessados  
mensurados a partir de Recordatórios Alimentares de 24  
horas

Luana Giongo Pedrotti<sup>1</sup>

Bruna Lazzeri<sup>1,2</sup>

Betina Soldateli Paim<sup>2</sup>

Michele Drehmer<sup>1,2</sup>

Vanessa Bielefeldt Leotti<sup>1</sup>

Artigo a ser submetido à Revista Brasileira de Epidemiologia

---

<sup>1</sup>Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

<sup>2</sup>Programa de Pós-Graduação em Alimentação, Nutrição e Saúde, Departamento de Nutrição, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

## Resumo

**OBJETIVO:** Comparar empiricamente, utilizando dados de um estudo realizado em crianças, o efeito das abordagens *first-add-then-shrink* e *first-shrink-then-add* aplicadas na estimação do consumo de produtos ultraprocessados (PUP).

**MÉTODOS:** Ensaio clínico randomizado com dados coletados em 2012 e 2013. A amostra final envolveu 155 crianças de 4 a 7 anos de idade. O consumo alimentar foi avaliado por meio de recordatórios alimentares de 24h e os alimentos consumidos foram agrupados segundo a classificação NOVA. O consumo habitual foi estimado através do método SPADE (*Statistical Program to Assess Dietary Exposure*), em calorias e percentual do valor energético total (VET), utilizando duas abordagens distintas, *first-add-then-shrink* e *first-shrink-then-add*. Também, foi utilizado o método de correção pela média para comparação com as abordagens.

**RESULTADOS:** Observou-se um aumento do consumo médio de PUP com a idade partindo de 31,57% e 32,63% aos 4 anos para 46,68% e 45,42% aos 7 anos para as abordagens *first-add-then-shrink* e *first-shrink-then-add* respectivamente. Para percentual do VET e para calorias respectivamente, os intervalos interquartis da abordagem *first-shrink-then-add* são mais amplos se comparado à *first-add-then-shrink*: [30,90; 45,28] e [528,16; 826,20] contra [34,18; 44,63] e [587,01; 820,72]. Porém, são menos amplos quando comparados às estimativas corrigida pela média, implicando maior variabilidade na estimação com este método.

**CONCLUSÕES:** A abordagem *first-shrink-then-add* foi menos precisa, apresentando maior amplitude nas estimativas quando comparada à *first-add-then-shrink*, em todos os casos. Ambas as abordagens apresentaram distribuições estimadas com menor variabilidade quando comparadas à utilização da média. Reforça-se a necessidade de utilizar métodos que diminuam o impacto da variabilidade intraindividual para obter estimativas mais precisas e recomendações adequadas de consumo de PUP tendo em vista seus possíveis efeitos na saúde.

**Palavras-chave:** *First-add-then-shrink*. *First-shrink-then-add*. Produtos Ultraprocessados. Classificação NOVA. Epidemiologia Nutricional.

## Abstract

**OBJECTIVE:** Empirically compare the effect of the first-add-then-shrink and first-shrink-then-add approaches on the ultra-processed products (UPP) intake estimation using data from a study carried out in children.

**METHODS:** Randomized clinical trial with data collected in 2012 and 2013. The final sample enrolled 155 children aged 4 to 7 years. The intake was assessed by 24-hour dietary recalls (R24h) and the consumed foods were grouped using the NOVA classification. Usual intake was estimated through the SPADE (Statistical Program to Assess Dietary Exposure) method, in calories and percentage contribution to the total energy value (TEV), using two different approaches, first-add-then-shrink and first-shrink-then-add. Also, the mean-corrected estimate was compared with the approaches.

**RESULTS:** There is an increase in UPP average intake by age, starting at 31,57% and 32,63% at 4 years and getting to 46,68% and 45,42% at 7 years, for the first-add-then-shrink and first-shrink-then-add approaches respectively. For both percentage of TEV and calories, the interquartile ranges for the first-shrink-then-add approach are wider than for the first-add-then-shrink: [30,90; 45,28] and [528,16; 826,20] against [34,18; 44,63] and [587,01; 820,72]. However, they are narrower when compared to the mean-corrected estimate, implying that this method estimates intake with larger variability.

**CONCLUSIONS:** The first-shrink-then-add approach was less precise, with wider range estimates when compared to the first-add-then-shrink. Both approaches resulted in estimated distributions with less variability when compared to the mean-corrected method. We reinforce the need of using statistical methods to reduce the within-person variability impact to obtain more accurate estimates and adequate UPP intake recommendations given their possible effects on health outcomes.

**Keywords:** First-add-then-shrink. First-shrink-then-add. Ultra-processed products. NOVA classification. Nutritional Epidemiology.

## 1 INTRODUÇÃO

As mudanças nos padrões alimentares globais em termos de fontes de alimentos, modos de processamento e distribuição nas últimas décadas levaram ao elevado consumo de alimentos e bebidas altamente processados. Essas mudanças, que décadas atrás começaram a surgir em países de alta renda, hoje atingem bilhões de pessoas em todo o mundo (POPKIN, 2011). No Brasil, segundo a última Pesquisa de Orçamento Familiar (POF), a disponibilidade de alimentos ultraprocessados passou de 12,6% para 18,4% entre os anos de 2002 e 2018, para todas as classes de renda (IBGE, 2020). Na população infantil, o consumo de alimentos ultraprocessados também é alto. Estudos em diferentes faixas etárias apontam consumo de ultraprocessados que variam de 38% até 47% do valor energético total diário (VET) (SPARRENBERGER et al., 2015; FONSECA et al., 2018; RAUBER et al., 2015).

Considerando a relação entre o processamento de alimentos, mais especificamente seu tipo e intensidade, e o aumento da obesidade e de doenças crônicas não transmissíveis, o Guia Alimentar para a População Brasileira (BRASIL, 2014) recomenda a classificação NOVA (MONTEIRO et al., 2010; MONTEIRO et al., 2018), um novo método de avaliação do consumo por graus de processamento dos alimentos. Esta classificação divide os alimentos em quatro grupos de acordo com seu grau de processamento e a categoria de maior grau é a dos chamados produtos ultraprocessados (PUP) (MONTEIRO et al., 2018).

Para a mensuração de consumo alimentar, é importante frisar que não existe um método padrão-ouro, porém existem diversas ferramentas que podem ser adequadas a diferentes contextos, tais como os Recordatórios Alimentares de 24 horas (R24h), Questionários de Frequência Alimentar e Registros Alimentares. O principal objetivo a se alcançar com tais ferramentas é conseguir estimar o consumo médio a longo prazo, também conhecido como consumo habitual, o que geralmente requer modelagem estatística para levar em conta as variabilidades intra e interindividuais (LAUREANO et al., 2016). Apesar disso, no contexto de estimação do consumo de PUP, ainda é muito utilizada a média de repetidas medições como correção para variabilidade intraindividual (FONSECA et al., 2018; BERTI et al., 2019).

Dentre os principais métodos estatísticos para estimação da distribuição do consumo habitual podem-se destacar: *Statistical Program to Assess Dietary Exposure* (SPADE) (DEKKERS et al., 2014), *Multiple Source Method* (MSM) (HAUBROCK et al., 2011), *National Cancer Institute Method* (NCI) (TOOZE et al., 2006) e *Iowa State University Method* (ISU) (NUSSER et al., 1996). Embora a metodologia geral seja similar, os métodos diferem em alguns detalhes, tais como os modelos estatísticos utilizados, suposições e transformações utilizadas (SOVEREIN et al., 2011).

No presente estudo é utilizado o método SPADE, que considera consumos provenientes de diversas fontes alimentares, que é o caso dos PUP, obtidos a partir de repetidas medições de instrumentos à curto prazo (R24h) e é implementado no *software* livre R. Além disso, o método conta com diferentes modelos onde são utilizadas duas abordagens aplicadas na estimação dos consumos, conhecidas como *first-add-then-shrink* e *first-shrink-then-add* (SLOB et al., 2010; VERKAIK-KLOOSTERMAN et al., 2011), que não foram ainda comparadas na literatura no contexto de alimentos ultraprocessados. Tais abordagens utilizam metodologias diferentes para a estimação do consumo habitual, possivelmente apresentando resultados distintos para um mesmo consumo.

Considerando essas possíveis diferenças nas estimativas advindas da utilização de

ambas as abordagens, o objetivo deste artigo é comparar empiricamente através de um estudo realizado em crianças o efeito das abordagens *first-add-then-shrink* e *first-shrink-then-add* aplicadas na estimação do consumo de PUP e, além disso, compará-las com a média de repetidas medições como correção para variabilidade intraindividual.

## 2 MÉTODOS

### 2.1 Delineamento e amostra

Neste trabalho são utilizados dados de um ensaio clínico randomizado (ECR) que avaliou uma intervenção educacional de alimentação saudável realizada em mães adolescentes nos primeiros meses de vida dos seus bebês, com seguimento relativo à medida do consumo de alimentos ultraprocessados aos 4 a 7 anos de vida das crianças, mensurada em 2012 e 2013. O consumo alimentar foi avaliado por meio de até três recordatórios alimentares de 24h, em dias não consecutivos. Nas coletas, o entrevistador solicitou ao responsável pela criança para listar todos os alimentos e bebidas ingeridas no dia anterior à entrevista, incluindo suas quantidades. Os dados obtidos foram digitados com dupla verificação no *software* ADSNutri (ADSNUTRI, 2006). Para detalhes referentes as fases do ECR, sugere-se Schwartz et al. (2015), Soldateli et al. (2016) e Nunes et al. (2017).

Ao final do seguimento, 194 crianças compuseram a amostra. Devido ao fato da implementação computacional das abordagens utilizadas no modelo multiparte do SPADE ter a restrição de trabalhar com apenas duas observações por indivíduo (DEKKERS et al., 2017), foram excluídas 39 crianças que possuíam apenas um R24h, e para as 139 que possuíam três R24h foi realizado um sorteio aleatório para obter dois. Como haviam 16 crianças com dois R24h coletados, o tamanho da amostra final foi de 155 crianças.

### 2.2 SPADE

O *Statistical Program to Assess Dietary Exposure* apresentado em Dekkers et al. (2014) é um método implementado no *software* livre R (R Core Team, 2017) para a estimação do consumo habitual de componentes alimentares a partir de múltiplas fontes, isto é, alimentos e nutrientes ou mesmo diversos alimentos como acontece no caso dos ultraprocessados. Por ser implementado no R e pelas características acima descritas, optou-se pelo método SPADE para este trabalho. O SPADE está disponível para *download* gratuito por meio do pacote SPADE.RIVM (RIVM, 2014).

Os componentes alimentares podem ser divididos em dois tipos, os que são consumidos regularmente por quase todos os indivíduos, definidos como de consumo diário (NUSSER et al., 1996; SLOB, 1993) e componentes consumidos com menos frequência, definidos como de consumo episódico (TOOZE et al., 2006; SOUVEREIN et al., 2011). O SPADE conta com opções para modelar ambos os tipos de consumo e, além disso, as diversas fontes possíveis, tais como alimentos e suplementos:

- Modelo de uma parte: modela componentes alimentares consumidos diariamente por todos (ou quase todos) os indivíduos.
- Modelo de duas partes: modela componentes alimentares consumidos episodicamente.
- Modelo de três partes: modela especificamente o consumo habitual total de componentes alimentares, que pode ser formado pelo consumo de alimentos e suplementos alimentares, como por exemplo, o consumo de vitaminas.

- Modelo multiparte para duas ou mais fontes: modela o consumo habitual proveniente de diferentes fontes ou grupos alimentares, como por exemplo, o valor energético total.

Neste trabalho é utilizado o modelo multiparte pois o consumo de PUP é constituído de diferentes fontes, como pode ser visto em Monteiro et al. (2010) e Monteiro et al. (2018), tornando assim este modelo o mais adequado. Maiores detalhes sobre a metodologia dos modelos do SPADE estão descritos em Pedrotti (2020).

### 2.3 Abordagens

Sabendo que os PUP são constituídos de diferentes fontes, deve-se considerar a maneira de lidar com o consumo proveniente de cada fonte e como estes se combinam para compor o consumo habitual que está sendo estimado. A estimação de um consumo habitual proveniente de múltiplas fontes pode apresentar problemas como a multimodalidade da distribuição do consumo, que faz com que as transformações para tornar a distribuição simétrica, geralmente utilizadas nos modelos estatísticos, falhem com certa frequência (SLOB et al., 2010).

Neste contexto, existem duas abordagens apresentadas na literatura por Verkaik-Kloosterman et al. (2011), Slob et al. (2010) e Dodd et al. (2009), que procedem de maneiras diferentes na estimação do consumo. A primeira abordagem, mais direta, consiste em somar a ingestão dos alimentos consumidos para cada participante e posteriormente ajustar para a variabilidade inerente ao indivíduo. Tal abordagem é conhecida como *first-add-then-shrink*. Segundo Verkaik-Kloosterman et al. (2011), sua utilização não evita problemas de multimodalidade, e além do mais, ela não leva em conta as potenciais diferenças nos parâmetros ou variâncias do modelo que se devem à utilização das diferentes fontes de consumo.

O método adequado para evitar tais problemas é, então, a segunda abordagem, conhecida como *first-shrink-then-add*, que consiste em modelar separadamente cada alimento que compõe o consumo habitual que está sendo estimado para posteriormente combiná-los numa estimativa geral (SLOB et al., 2010; DODD et al., 2009). O modelo de três partes e o modelo multiparte para duas ou mais fontes do SPADE contam com a utilização desta abordagem nos seus procedimentos de estimação, além da *first-add-then-shrink*, presente em todos os modelos.

### 2.4 Análise estatística

As análises foram conduzidas no *software* R versão 3.4.3 (R Core Team, 2017) e o SPADE foi utilizado para estimar a distribuição do consumo habitual de PUP em calorias e também em percentual do VET. É importante frisar aqui que neste trabalho não é explorado o efeito da intervenção no ECR em função de não fazer parte dos objetivos. A categoria de ultraprocessados, foi agrupada em subcategorias de maneira subjetiva inicialmente definidas por duas pesquisadoras do estudo individualmente, e após o confronto das opiniões acerca da classificação dos alimentos e/ou preparações foi realizado um painel com grupo de especialistas da área para definição final. As subcategorias encontradas na amostra estão descritas na tabela S1. Todos os alimentos que a compõe foram considerados de consumo episódico, pois todos apresentaram mais que 10% de consumos zero na amostra em estudo (NCI, 2015; GUENTHER et al., 2014).

Foi, então, utilizado o modelo multiparte para duas ou mais fontes (onde as subcategorias de ultraprocessados caracterizam as diferentes fontes) considerando a idade da criança, contido na função *f.spade4*, que apresenta separadamente os resultados para as abordagens *first-add-then-shrink* e *first-shrink-then-add*. Como a idade das crianças variou de 4 a 7 anos, cinco diferentes distribuições do consumo habitual foram estimadas, tanto para calorias quanto para percentual do VET. Assim, a média, os quartis e os percentis 5% e 95% do consumo de PUP foram apresentados para a faixa geral de idade e para cada idade separadamente. Além disso, para comparação entre as abordagens foram utilizados os gráficos de suas funções de distribuição acumulada empírica (ECDF), gerados pelo próprio método SPADE. Também, para via de comparação, foi calculada a média entre os dois R24h de cada criança visando verificar a distribuição do consumo estimada por este método.

### 3 RESULTADOS

Nas figuras S1 e S2 é possível ver a distribuição do consumo em calorias e em percentual do VET para cada subcategoria dos PUP, ficando evidente a predominância do baixo consumo em cada uma individualmente, caracterizando-as assim como de consumo episódico.

A tabela 1 apresenta a descrição da amostra e medidas do consumo corrigidas para a variabilidade intraindividual apenas pela média. A média de idade na amostra foi de 6,1 ( $\pm 0,5$ ) anos e 53,5% das crianças eram do sexo masculino. As três maiores médias de calorias de consumo de PUP foram, respectivamente, nas categorias de bolos e sobremesas industrializados, bebidas gaseificadas - refrigerantes e biscoitos, com 95,3 ( $\pm 143,7$ ), 89,5 ( $\pm 79,9$ ) e 76,2 ( $\pm 106,6$ ) calorias e a média das calorias totais ingeridas foi de 1741,5 ( $\pm 467,9$ ) calorias.

Na tabela 2, consta a distribuição do consumo habitual de PUP em percentual do VET geral e para cada idade separadamente. De maneira geral, pode-se observar o aumento do consumo com a idade, partindo de 31,57% e 32,63% aos 4 anos para 46,68% e 45,42% aos 7 anos para as abordagens *first-add-then-shrink* e *first-shrink-then-add* respectivamente. Na distribuição geral o intervalo interquartil da abordagem *first-shrink-then-add* [30,90; 45,28] é mais amplo quando comparado à abordagem *first-add-then-shrink* [34,18; 44,63], porém menos amplo quando comparada à estimativa corrigida pela média [28,5; 49,0]. Isso se aplica também para cada idade separadamente.

Na tabela 3, estão apresentadas as distribuições consumo habitual de PUP em calorias. Assim como para o consumo medido em percentual do VET, a abordagem *first-add-then-shrink* apresentou médias e medianas maiores em relação à abordagem *first-shrink-then-add*, no geral e para as idades de 5 e 7 anos. Porém, novamente o intervalo interquartil da abordagem *first-shrink-then-add* é mais amplo em comparação com *first-add-then-shrink*, [528,16; 826,20] contra [587,01; 820,72], respectivamente. Isso se aplica a todas as idades, exceto para 4 e 5 anos. A estimativa corrigida pela média apresentou intervalo interquartil mais amplo que o de ambas as abordagens, [463,8; 875,8].

Nas figuras 1 e 2 estão apresentadas as distribuições acumuladas empíricas para ambas as abordagens na faixa geral de idade, para o consumo habitual em percentual do VET e em calorias, respectivamente. O mesmo comportamento é visto em ambas as figuras. A abordagem *first-shrink-then-add* apresenta uma distribuição mais inclinada e um pouco deslocada à esquerda. Até aproximadamente o percentil 70 a abordagem

*first-shrink-then-add* estimou percentuais e calorias menores que a *first-add-then-shrink* e após esse ponto há uma inversão. Notam-se as maiores diferenças entre as abordagens nos percentis mais baixos e mais altos.

## 4 DISCUSSÃO

Neste trabalho visou-se comparar empiricamente as abordagens *first-add-then-shrink* e *first-shrink-then-add* e com base nos resultados foi possível notar que elas produziram estimativas e distribuições diferentes. Além disso, ambas apresentaram distribuições estimadas com menos variabilidade quando comparadas às resultantes da correção pela média. Em termos de calorias, as abordagens *first-add-then-shrink* e *first-shrink-then-add* estimaram, para a distribuição geral, uma média de 711,09 e 696,05 calorias diárias, respectivamente. Já para o percentual do VET, novamente para a distribuição geral, obteve-se uma média de 39,54% e 38,75%, respectivamente.

De acordo com Martins et al. (2013), em 2009, os PUP representavam em média 25,4% da dieta dos brasileiros (sem restrição de idade), já em Sparrenberger et al. (2015), as crianças brasileiras com idade de 2 a 10 anos consumiam 47% de PUP na sua dieta, e em Fonseca et al. (2018) os PUP representavam 38% da dieta em crianças com idade de 4 a 7 anos. Portanto, os resultados de consumo de PUP deste estudo, estimados pelo método SPADE para ambas as abordagens, vão ao encontro do que se tem nas pesquisas acima citadas.

O método de correção da variabilidade intraindividual pela média é geralmente utilizado em estudos (BERTI et al., 2019; FONSECA et al., 2018; SPARRENBERGER et al., 2015) e apenas o seu uso pode não ser capaz de lidar com a variabilidade intraindividual, fato reforçado pelos resultados obtidos no presente estudo, visto que os intervalos interquartis das estimativas corrigidas pela média foram sempre mais amplos que os resultantes do método SPADE. Por isso, os modelos estatísticos são de suma importância, bem como a definição da abordagem adequada a ser utilizada para a estimação do consumo habitual em questão, visto que ambas as abordagens apresentaram estimativas mais precisas quando comparadas apenas à utilização da média.

No que diz respeito à aplicação em outros métodos, é necessário levar em conta que a abordagem *first-add-then-shrink* pode ser realizada em qualquer método, pois havendo uma soma de consumos para uma posterior aplicação em uma modelagem de uma ou duas partes, caracteriza-se a sua utilização. Por outro lado, a abordagem *first-shrink-then-add* foi descrita nos métodos SPADE e NCI, sendo que em Verkaik-Kloosterman et al. (2011) foram comparadas as abordagens na estimação de consumo habitual de vitamina D proveniente de alimentos e suplementos em crianças.

Sobre as limitações deste estudo, não é possível ter conclusões mais definitivas sobre as duas abordagens em relação a sua acurácia e precisão, visto que não é possível saber a verdadeira distribuição do consumo habitual da população em estudo. Para trabalhos futuros, tem-se como sugestão a simulação de dados de consumo para proceder a comparação das abordagens com uma referência de consumo habitual verdadeiro. Além disso, há um certo grau de adição na composição das subcategorias da NOVA neste trabalho, ou seja, de certa maneira foi utilizada a abordagem *first-add-then-shrink* nessa etapa. Essa adição se fez necessária para evitar uma quantidade muito alta de consumos zero em diversos alimentos. Com um tamanho de amostra maior, haveria a possibilidade de testar a influência de colocar cada alimento separadamente no modelo multiparte, ficando como outra sugestão

para trabalhos futuros.

Em face das evidentes diferenças entre as estimativas apresentadas no presente trabalho, que mostram uma possível ocorrência de problemas de sub ou superestimação, ressalta-se a importância da consideração da utilização das abordagens. Visando políticas e recomendações alimentares adequadas é de suma importância obter boas e fidedignas estimativas de consumo de PUP, pois há uma crescente de evidências que associam este consumo com diversos desfechos relacionados à saúde tais como ganho de peso e obesidade (CANELLA et al., 2014; LOUZADA et al., 2015; CANHADA et al., 2020), hipertensão (MENDONÇA et al., 2017) e até aumento no risco de câncer (FIOLET et al., 2018).

Assim, reforça-se a necessidade de utilizar métodos que diminuam o impacto da variabilidade intraindividual na estimação da distribuição do consumo habitual, bem como escolher a abordagem mais adequada ao tipo de consumo que se quer estimar, para obter resultados mais fidedignos e assim possibilitar uma avaliação mais clara do estado nutricional atual da população.

## Referências

- ADSNUTRI. *ADS Nutri: Ferramenta para Análise de Dietas Online*. 2006. Disponível em: <<http://heufpel.com.br/sistemas/login/index.php?caminho=nutricao&programa=nutricao>>.
- BERTI, T. et al. Consumo alimentar segundo o grau de processamento e características sociodemográficas: Estudo pró-saúde. *Revista Brasileira de Epidemiologia*, v. 22, 2019.
- BRASIL, M. d. S. Guia alimentar para a população brasileira. Brasília; Ministério da Saúde, 2014.
- CANELLA, D. et al. Ultra-processed food products and obesity in brazilian households (2008–2009). *PLoS ONE*, v. 9(3), 2014.
- CANHADA, S. et al. Ultra-processed foods, incident overweight and obesity, and longitudinal changes in weight and waist circumference: the brazilian longitudinal study of adult health (elsa-brasil). *Public Health Nutrition*, v. 23 (6), p. 1076–1086, 2020.
- DEKKERS, A.; VERKAIK-KLOOSTERMAN, J.; OCKÉ, M. *SPADE: Statistical Program to Assess habitual Dietary Exposure. User's Manual*. [S.l.], 2017.
- DEKKERS, A. et al. Spade, a new statistical program to estimate habitual dietary intake from multiple food sources and dietary supplements. *The Journal of Nutrition*, v. 144, p. 2083–91, 2014.
- DODD, K. et al. Estimating distributions of usual total nutrient intake: a comparison of available methods. In: *7th International Conference on Diet and Activity Methods (ICDAM 7)*. [S.l.: s.n.], 2009.
- FIOLET, T. et al. Consumption of ultra-processed foods and cancer risk: results from nutrinet-santé prospective cohort. *British Medical Journal*, v. 360, 2018.

FONSECA, P. et al. Association of exclusive breastfeeding duration with consumption of ultra-processed foods, fruit and vegetables in brazilian children. *European Journal of Nutrition*, v. 7, p. 2887–2894, 2018.

GUENTHER, P. M. et al. The healthy eating index-2010 is a valid and reliable measure of diet quality according to the 2010 dietary guidelines for americans. *The Journal of Nutrition*, v. 144, p. 399–407, 2014.

HAUBROCK, J. et al. Estimating usual food intake distributions by using the multiple source method in the epic-potsdam calibration study. *Journal of Nutrition*, v. 141, p. 914–20, 2011.

IBGE, C. d. T. e. R. Pesquisa de orçamentos familiares 2017-2018: Avaliação nutricional da disponibilidade domiciliar de alimentos no brasil. Rio de Janeiro; Instituto Brasileiro de Geografia e Estatística, 2020.

LAUREANO, G. et al. Comparison of the isu, nci, msm and spade methods for estimating usual intake: a simulation study of nutrients consumed daily. *Nutrients*, v. 8, p. 166, 2016.

LOUZADA, M. et al. Consumption of ultra-processed foods and obesity in brazilian adolescents and adults. *Preventive Medicine*, v. 81, p. 9–15, 2015.

MARTINS, A. P. B. et al. Participação crescente de produtos ultraprocessados na dieta brasileira (1987-2009). *Revista de Saúde Pública*, v. 47, n. 4, p. 656–65, 2013.

MENDONÇA, R. et al. Ultra-processed food consumption and the incidence of hypertension in a mediterranean cohort: The seguimiento universidad de navarra project. *American Journal Hypertension*, v. 30(4), p. 358–366, 2017.

MONTEIRO, C. et al. The un decade of nutrition, the nova food classification and the trouble with ultra-processing. *Public Health Nutrition*, v. 21(1), p. 5–17, 2018.

MONTEIRO, C. et al. A new classification of foods based on the extent and purpose of their processing. *Caderno de Saúde Pública*, v. 26, p. 2039–49, 2010.

NCI, N. C. I. *Dietary Assessment Primer*. 2015. Acessado em maio de 2019. Disponível em: <<https://dietassessmentprimer.cancer.gov/concepts/>>.

NUNES, L. et al. Efeito de intervenção no cumprimento das recomendações alimentares no primeiro ano de vida: ensaio clínico randomizado com mães adolescentes e avós maternas. *Cadernos de Saúde Pública*, v. 33(6), 2017.

NUSSER, S. et al. A semiparametric transformation approach to estimating usual daily intake distributions. *Journal of the American Statistical Association*, v. 91, p. 1440–49, 1996.

PEDROTTI, L. *Formas de correção de consumo de produtos ultraprocessados mensurados a partir de Recordatórios Alimentares de 24 horas*. Dissertação (Mestrado em Epidemiologia) — Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

POPKIN, B. Contemporary nutritional transition: determinants of diet and its impact on body composition. *National Institute of Health*, v. 70, p. 82–91, 2011.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. URL=<https://www.R-project.org/>.

RAUBER, F. et al. Consumption of ultra-processed food products and its effects on children's lipid profiles: A longitudinal study. *Nutrition, Metabolism and Cardiovascular Diseases*, v. 25, p. 116–122, 2015.

RIVM. *National Institute for Public Health and the Environment, SPADE package*. 2014. Acessado em abril de 2019. Disponível em: <<http://rivm.nl/en/spade>>.

SCHWARTZ, R. et al. The effect of a pro-breastfeeding and healthy complementary feeding intervention targeting adolescent mothers and grandmothers on growth and prevalence of overweight of preschool children. *PLoS ONE*, v. 10(7), 2015.

SLOB, W. Modeling long-term exposure of the whole population to chemicals in food. *Risk Analysis*, v. 13, p. 525–30, 1993.

SLOB, W. et al. Can current dietary exposure models handle aggregated intake from different foods? a simulation study for the case of two foods. *Food Chem Toxicol.*, v. 48, p. 178–86, 2010.

SOLDATELI, B.; VIGO, A.; GIUGLIANI, E. Adesão a recomendações alimentares em pré-escolares: ensaio clínico com mães adolescentes. *Revista de Saúde Pública*, v. 50:83, 2016.

SOUVEREIN, O. et al. Comparing four methods to estimate usual intake distributions. *European Journal of Clinical Nutrition*, v. 65, p. 92–101, 2011.

SPARRENBERGER, K. et al. Consumo de alimentos ultraprocessados entre crianças de uma unidade básica de saúde. *Jornal de Pediatria*, v. 91, p. 535–542, 2015.

TOOZE, J. et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of American Diet Association*, v. 106, p. 1575–87, 2006.

VERKAIK-KLOOSTERMAN, J. et al. A three-part, mixed-effects model to estimate the habitual total vitamin d intake distribution from food and dietary supplements in dutch young children. *Journal of Nutrition*, v. 141, p. 2055–63, 2011.

## TABELAS E FIGURAS

Tabela 1 – Descrição da amostra e consumo de PUP e suas categorias (n=155)

Características	Média ± Desvio padrão
Idade (anos)	6,1 ± 0,5
Sexo <sup>1</sup>	
Masculino	83 (53,5)
Feminino	72 (46,5)
Calorias diárias (kcal/dia)	1741,5 ± 467,9
<b>Subcategorias dos produtos ultraprocessados (calorias diárias)</b>	
Bolos e sobremesas industrializados	95,3 ± 143,7
Bebidas gaseificadas – refrigerantes	89,5 ± 79,9
Biscoitos	76,2 ± 106,6
Achocolatado em pó e similares	61,8 ± 94,5
Alimentos prontos para consumo	53,6 ± 145,4
Produtos à base de carne reconstituída	48,0 ± 84,3
Pães industrializados	45,4 ± 80,7
Margarina/geleias/queijos ultraprocessados	42,4 ± 50,8
Alimentos prontos para aquecer – refeições congeladas	40,3 ± 128,7
Bebidas à base de frutas – incluindo suco em pó	40,2 ± 38,3
Salgadinhos/batatas fritas embaladas	37,9 ± 76,2
Produtos de confeitaria – chocolates e doces em geral	31,1 ± 112,1
Bebidas à base de leite/soja	28,4 ± 57,1
Cereais matinais	16,9 ± 61,7
Molhos e condimentos prontos	1,5 ± 3,9
Outros produtos ultraprocessados	1,8 ± 10,6

<sup>1</sup>n (%)

Tabela 2 – Distribuição do consumo habitual do percentual de PUP para crianças da faixa de 4 a 7 anos e para cada idade (n=155)

Idade	Método	Média	Percentis				
			5%	25%	50%	75%	95%
Geral	First-add-then-shrink	39,54	27,27	34,18	39,30	44,63	52,66
	First-shrink-then-add	38,75	23,60	30,90	37,39	45,28	58,04
	Correção pela média	39,25	17,97	28,53	38,56	49,01	66,07
4 anos (n=5)	First-add-then-shrink	31,57	21,33	27,01	31,29	35,86	42,78
	First-shrink-then-add	32,63	21,46	26,79	31,55	36,97	47,82
	Correção pela média	31,93	19,85	27,19	29,56	42,2	42,58
5 anos (n=50)	First-add-then-shrink	36,34	25,43	31,53	36,08	40,87	48,17
	First-shrink-then-add	34,41	22,28	28,33	33,29	39,20	49,68
	Correção pela média	36,31	17,11	27,47	34,73	43,70	66,23
6 anos (n=97)	First-add-then-shrink	41,38	29,84	36,32	41,13	46,17	53,80
	First-shrink-then-add	42,54	28,49	35,57	41,59	48,33	59,73
	Correção pela média	41,21	18,77	29,81	42	50,84	66,20
7 anos (n=3)	First-add-then-shrink	46,68	34,54	41,39	46,44	51,71	59,66
	First-shrink-then-add	45,42	30,71	37,97	44,21	51,34	64,43
	Correção pela média	51,43	44,26	45,76	47,64	55,21	61,26

Tabela 3 – Distribuição do consumo habitual de calorias de PUP para crianças da faixa de 4 a 7 anos e para cada idade (n=155)

Idade	Método	Média	Percentis				
			5%	25%	50%	75%	95%
Geral	First-add-then-shrink	711,09	447,77	587,01	698,00	820,72	1019,06
	First-shrink-then-add	696,05	396,71	528,16	661,58	826,20	1100,33
	Correção pela média	710,30	265,72	463,77	630,00	875,78	1485,42
4 anos (n=5)	First-add-then-shrink	511,60	326,86	423,08	500,91	588,46	732,77
	First-shrink-then-add	534,99	360,30	448,54	516,91	599,39	767,95
	Correção pela média	579,80	256,27	440,96	445,60	776,04	976,15
5 anos (n=50)	First-add-then-shrink	626,54	411,69	524,54	614,94	715,89	880,95
	First-shrink-then-add	585,64	377,78	483,63	571,89	666,91	848,44
	Correção pela média	625,70	267,87	404,70	512,36	775,83	1405,92
6 anos (n=97)	First-add-then-shrink	758,82	511,25	642,26	746,29	861,72	1049,09
	First-shrink-then-add	776,76	514,40	651,05	755,35	881,31	1097,49
	Correção pela média	748,40	273,33	491,10	717,10	886,27	1482,49
7 anos (n=3)	First-add-then-shrink	909,81	626,90	777,60	896,35	1027,34	1238,61
	First-shrink-then-add	887,38	584,20	730,61	856,73	1008,89	1299,20
	Correção pela média	1109,1	592,46	853,79	1180,46	1400,03	1575,69

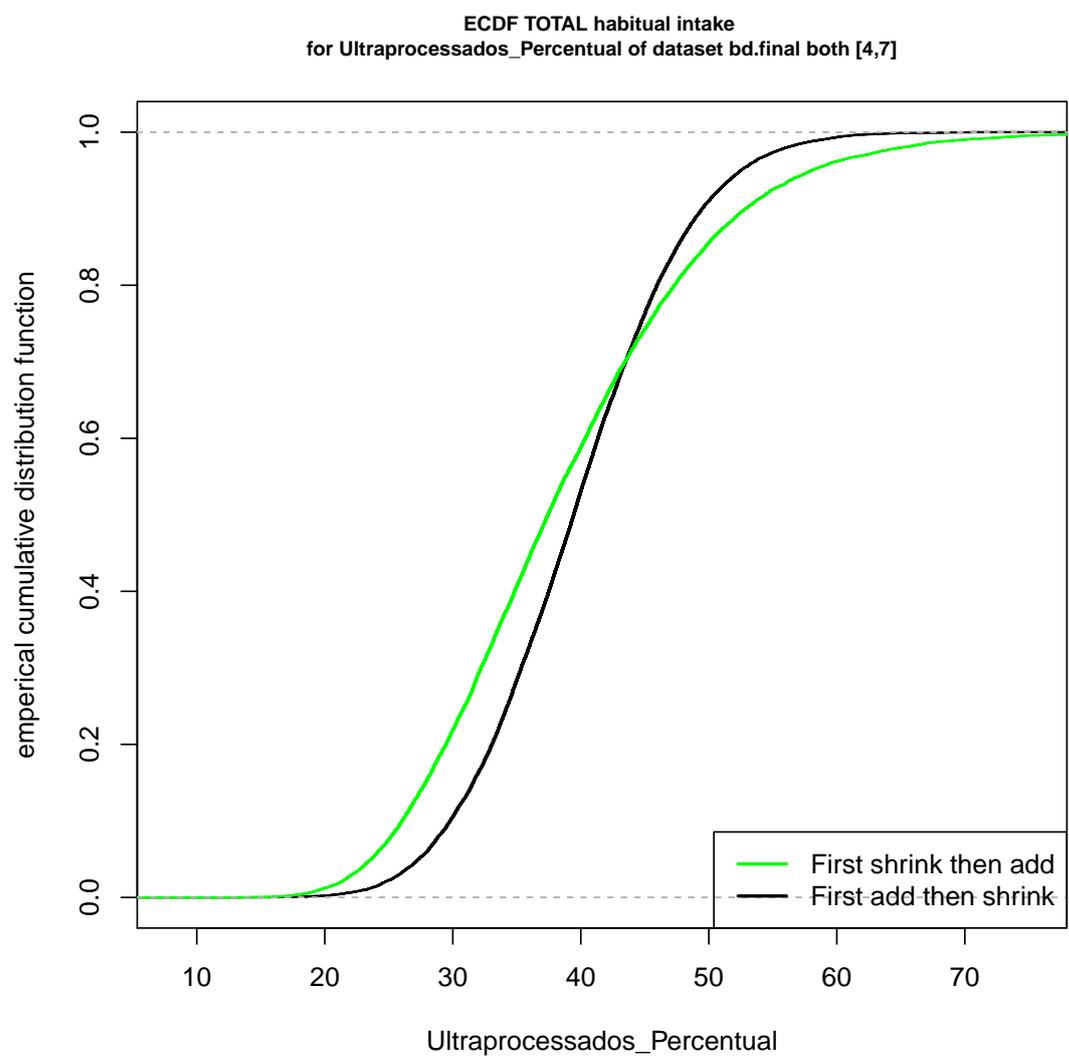


Figura 1 – Função de distribuição acumulada empírica (ECDF) do consumo habitual em percentual do VET para ambas as abordagens.

Fonte: *Output* SPADE

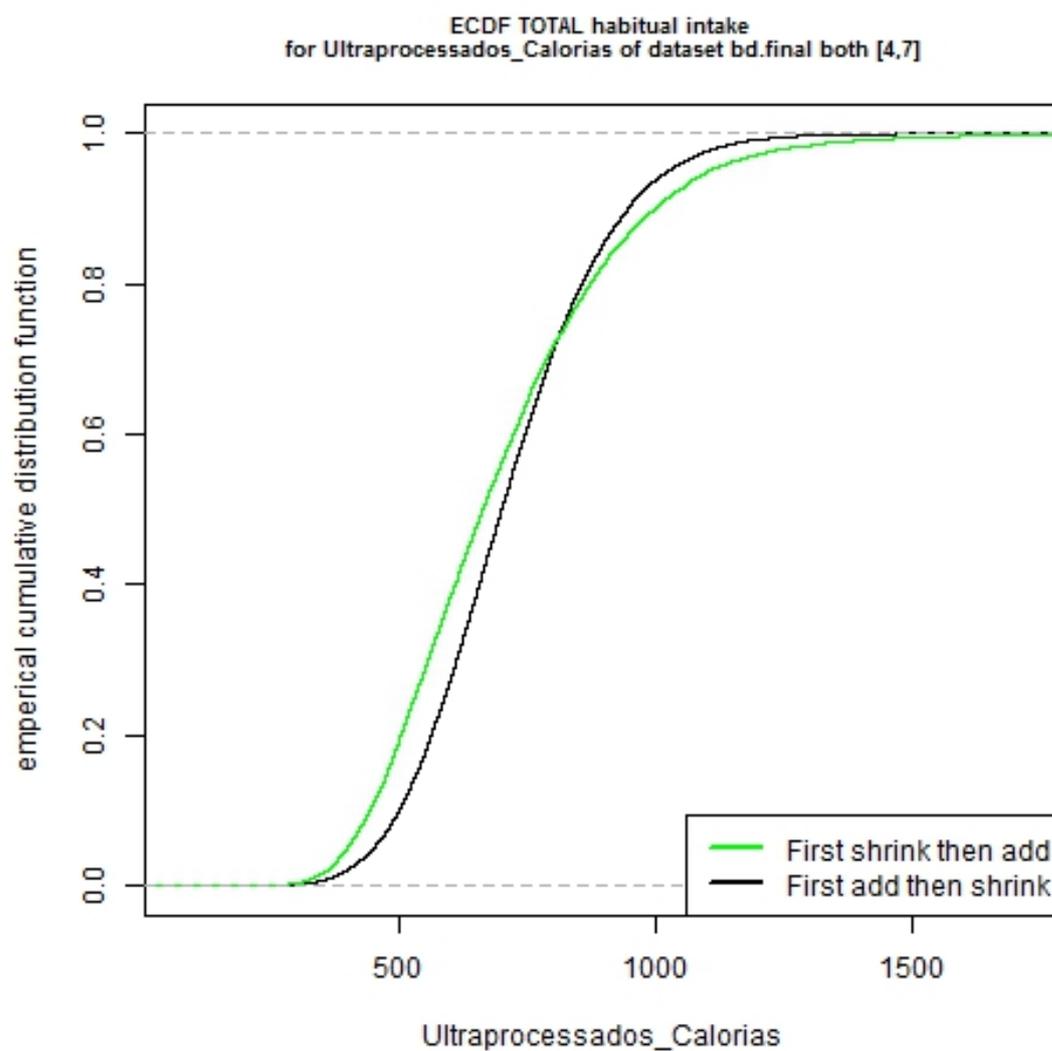


Figura 2 – Função de distribuição acumulada empírica (ECDF) do consumo habitual em calorias para ambas as abordagens.

Fonte: *Output SPADE*

---

## MATERIAL SUPLEMENTAR

Tabela S1 – Subcategorias que compõe a categoria de PUP

---

### Subcategorias de PUP

---

- 401 - Bebidas gaseificadas – refrigerantes
  - 402 - Bebidas à base de frutas (não naturais) – incluindo suco em pó
  - 404 - Bebidas à base de leite/soja
  - 406 - Pães industrializados
  - 407 - Cereais matinais
  - 408 - Produtos de confeitaria – chocolates e doces em geral
  - 409 - Biscoitos/bolachas embaladas
  - 410 - Salgadinhos/batatas fritas embaladas
  - 411 - Bolos e sobremesas industrializados
  - 412 - Produtos à base de carne reconstituída
  - 413 - Alimentos prontos para consumo
  - 414 - Alimentos prontos para aquecer – refeições congeladas
  - 415 e 416 - Margarina/geleias/queijos ultraprocessados
  - 417 - Molhos e condimentos prontos
  - 419 e 421 - Achocolatado em pó e similares
  - 422 - Outros produtos ultraprocessados
-

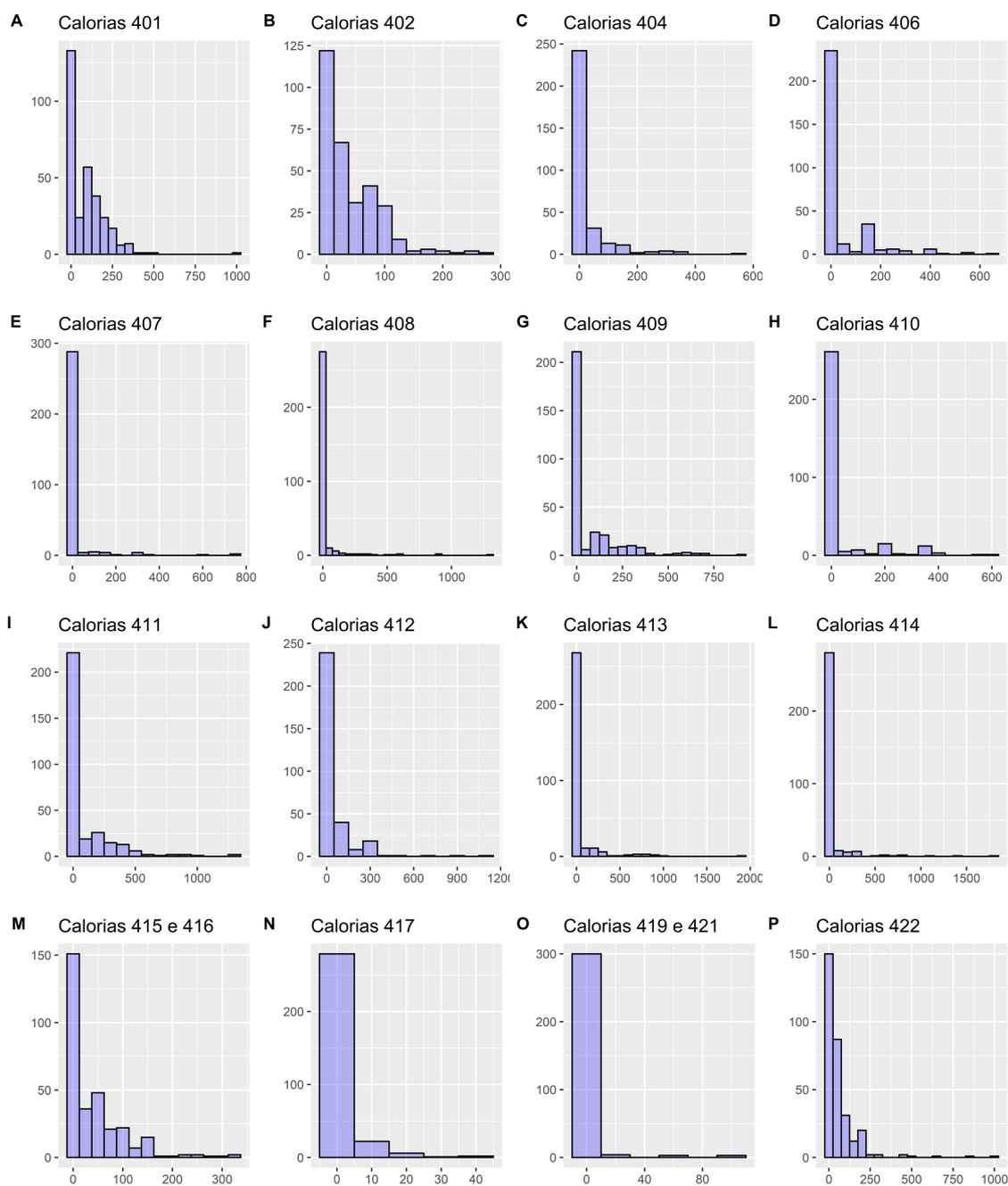


Figura S1 – Distribuição de consumo de calorias para cada subcategorias de alimentos ultraprocessados conforme Tabela S1.

Fonte: Elaborado pela autora

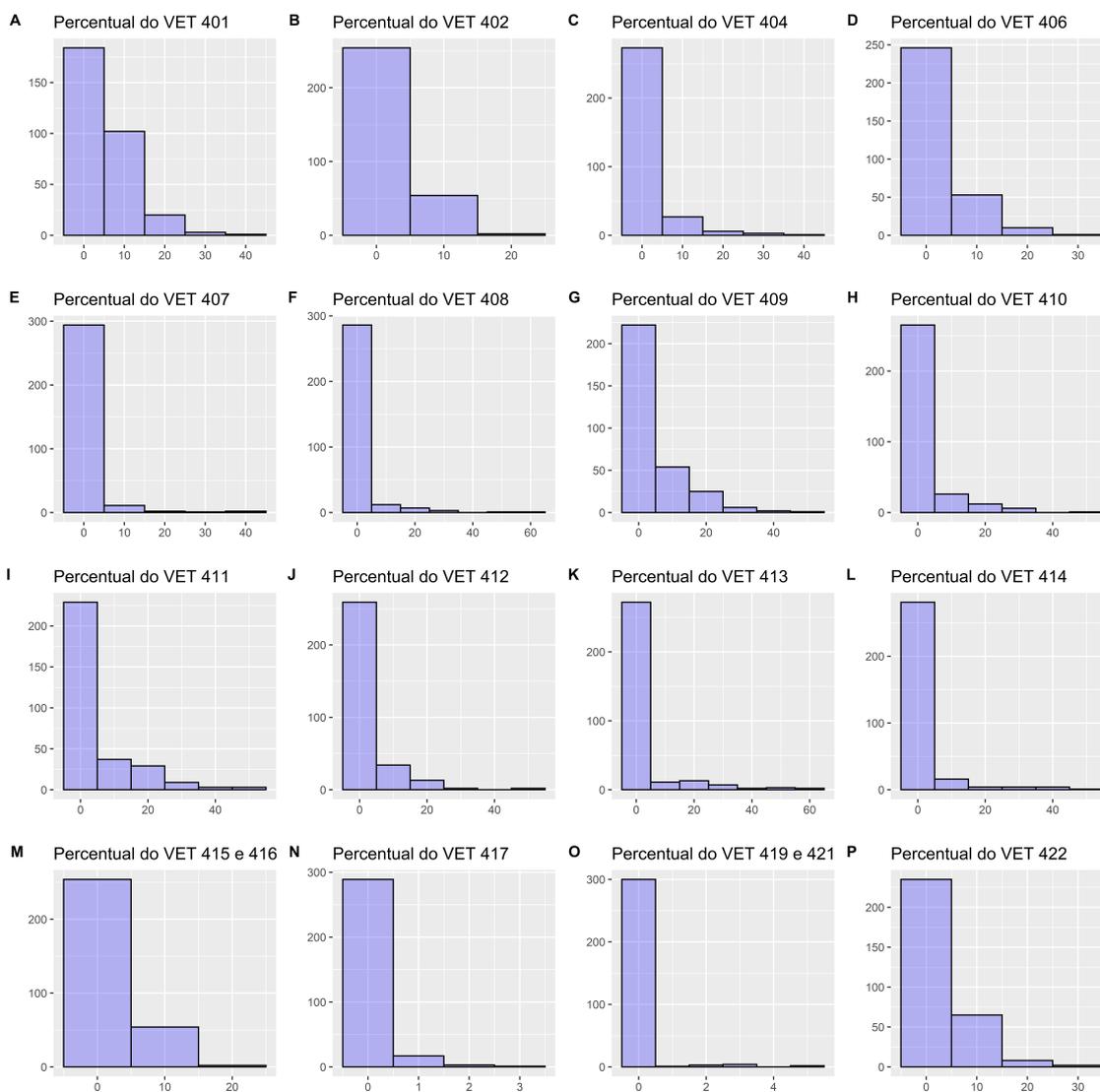


Figura S2 – Distribuição de consumo em percentual do VET para cada subcategorias de alimentos ultraprocessados conforme Tabela S1.

Fonte: Elaborado pela autora

## 6 Conclusões e Considerações Finais

Neste trabalho foram estimados consumos habituais de ultraprocessados por meio do método SPADE utilizando duas diferentes abordagens, *first-add-then-shrink* e *first-shrink-then-add*. O objetivo deste trabalho foi comparar empiricamente as duas abordagens aplicadas no consumo de ultraprocessados em crianças de 4-7 anos. Também, visou-se contrapor o seu uso aos métodos mais comumente utilizados, como a simples correção pela média.

As abordagens apresentaram melhores resultados se comparadas às estimativas corrigidas pela média, levando em conta que a última apresentou intervalos interquartis mais amplos que as outras. Por sua vez, a abordagem *first-shrink-then-add* apresentou intervalos interquartis mais amplos que a *first-add-then-shrink*. Tal resultado pode ser devido ao fato de que a abordagem *first-shrink-then-add* estima o consumo habitual levando em conta as diferentes fontes que compõe os produtos ultraprocessados separadamente, e para cada uma delas ocorre uma transformação. Em contraste, para a *first-add-then-shrink* ocorre apenas uma transformação, o que pode explicar a maior amplitude da abordagem anterior (Dekkers et al., 2017).

Levando em conta os resultados obtidos, recomenda-se a utilização dos métodos estatísticos de estimação em detrimento ao uso de apenas a correção pela média e, ainda, a escolha da abordagem mais adequada ao consumo de interesse. Porém, ainda se faz necessária a utilização de um estudo de simulação para poder comparar as abordagens em relação a acurácia e precisão, tendo em mãos um consumo habitual verdadeiro como parâmetro para comparação.

# Anexos

## Anexo A - Aprovação do Comitê de Ética



HCPA - HOSPITAL DE CLÍNICAS DE PORTO ALEGRE  
GRUPO DE PESQUISA E PÓS-GRADUAÇÃO

COMISSÃO CIENTÍFICA

A Comissão Científica do Hospital de Clínicas de Porto Alegre analisou o projeto:

**Projeto:** 120249

**Data da Versão do Projeto:**

**Pesquisadores:**

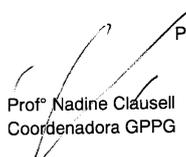
ELSA REGINA JUSTO GIUGLIANI  
LEANDRO MEIRELLES NUNES  
LUCIANA DIAS DE OLIVEIRA  
RENATA SCHWARTZ

**Título:** IMPACTO NO MÉDIO PRAZO DO ACONSELHAMENTO EM ALEITAMENTO MATERNO E ALIMENTAÇÃO COMPLEMENTAR SAUDÁVEL, REALIZADO NOS PRIMEIROS QUATRO MESES DE VIDA DA CRIANÇA, NOS HÁBITOS ALIMENTARES E ESTADO NUTRICIONAL: UM ENSAIO CLÍNICO RANDOMIZADO ENVOLVENDO MÃES ADOLESCENTES E AVÓS

Este projeto foi APROVADO em seus aspectos éticos, metodológicos, logísticos e financeiros para ser realizado no Hospital de Clínicas de Porto Alegre.  
Esta aprovação está baseada nos pareceres dos respectivos Comitês de Ética e do Serviço de Gestão em Pesquisa.

- Os pesquisadores vinculados ao projeto não participaram de qualquer etapa do processo de avaliação de seus projetos.
- O pesquisador deverá apresentar relatórios semestrais de acompanhamento e relatório final ao Grupo de Pesquisa e Pós-Graduação (GPPG)

Porto Alegre, 02 de agosto de 2012.

  
Prof. Nadine Clausell  
Coordenadora GPPG

# Anexo B - Termo de Consentimento Livre e Esclarecido

## Termo de Consentimento Livre e Esclarecido

Pesquisadora responsável: Elsa Regina Justo Giugliani

Endereço: Centro de Pesquisa Clínica – HCPA. Rua Ramiro Barcelos, 2350 - 3º andar, sala 21307.

Porto Alegre, RS. Fone: (51) 33596326

Comitê de Ética em Pesquisa – HCPA. Endereço: Ramiro Barcelos 2350, Porto Alegre, RS. Fone: (51) 33598304

Instituição: Universidade Federal do Rio Grande do Sul/ Hospital de Clínicas de Porto Alegre.

Prezada:

A senhora está sendo convidada a participar da pesquisa “**IMPACTO NO MÉDIO PRAZO DO ACONSELHAMENTO EM ALEITAMENTO MATERNO E ALIMENTAÇÃO COMPLEMENTAR NOS PRIMEIROS QUATRO MESES DE VIDA NOS HÁBITOS ALIMENTARES E ESTADO NUTRICIONAL DE PRÉ-ESCOLARES: UM ENSAIO CLÍNICO RANDOMIZADO ENVOLVENDO MÃES ADOLESCENTES E AVÓS**” que será realizado com as mães e crianças que participaram de pesquisa anterior cujo objetivo era verificar o efeito de orientações dadas na maternidade e em casa a mães adolescentes e avós maternas sobre a alimentação dos bebês no primeiro ano de vida. A pesquisa que estamos desenvolvendo agora tem como objetivo avaliar o impacto dessas orientações sobre os hábitos alimentares e estado nutricional dessas crianças em fase pré-escolar e de seus irmãos menores. Para tanto, precisamos entrevistar e conversar com as mães e seus filhos que participaram do estudo anterior e seus irmãos menores nascidos após a pesquisa.

A pesquisa é vinculada ao Curso de Pós-Graduação em Saúde da Criança e do Adolescente da Universidade Federal do Rio Grande do Sul (UFRGS) e compreende as seguintes etapas: (1) pesagem, verificação da altura das crianças e aplicação de questionário com informações sobre a alimentação da criança que será realizada no Centro de Pesquisa Clínica do Hospital de Clínicas de Porto Alegre (HCPA) ou no domicílio da criança caso não haja possibilidade de deslocamento por parte da família (2) realização de dois contatos telefônicos para aplicação de questionário sobre a alimentação da criança (se não for possível o contato por telefone, a mãe receberá a visita de um pesquisador em sua casa ou no local de sua preferência, para a aplicação dos questionários).

É importante que a Sra. saiba que:

- Não há riscos conhecidos na participação da pesquisa;
- Benefícios previstos: os resultados desta pesquisa poderão contribuir para a melhoria da saúde das crianças do nosso País no que se refere à alimentação;
- Dados pessoais coletados não serão divulgados evitando constrangimentos ou prejuízos aos participantes;
- A Sra. poderá desistir de participar do estudo a qualquer momento e terá ampla oportunidade de fazer perguntas, esclarecendo plenamente suas dúvidas, antes e durante o estudo.
- Não haverá prejuízos a quem se recusar a participar da pesquisa assim como não haverá prejuízos de acompanhamento ambulatorial regular se for o caso da criança ser atendida por este serviço;
- Não haverá despesa decorrente da participação na pesquisa, bem como, não haverá remuneração pela participação na mesma;

Salientamos que os dados obtidos serão utilizados única e exclusivamente para o fim desta pesquisa. Todas as informações obtidas neste estudo poderão ser publicadas com finalidade científica, de forma anônima, ou seja, manteremos o caráter confidencial das informações.

Caso queira mais informações, você poderá entrar em contato com os pesquisadores no endereço e fones citados acima ou ainda através dos fones abaixo:

**Renata Schwartz fone: 81863634**

**Cristiano Silva fone: 82106868**

**Betina Soldateli fone: 98181223**

**Luciana Oliveira fone: 99335828**

Agradecemos seu apoio e atenção!

Declaro que li e recebi cópia do presente Termo de Consentimento Livre e Esclarecido e concordo em participar da pesquisa em questão.

_____	_____	Porto Alegre ____/____/____
Nome da participante	Assinatura da participante	Data
_____	_____	Porto Alegre ____/____/____
Nome do pesquisador	Assinatura do pesquisador	Data



**Período da tarde:**

Alimento/ bebida	Quantidade (em medidas caseiras)

**Jantar:**

Alimento/ bebida	Quantidade (em medidas caseiras)

**Ceia:**

Alimento/ bebida	Quantidade (em medidas caseiras)

Data da entrevista: DATA1

Entrevistador ENTRV1

Entrevistadores: 1) Betina 2) Cristiano 3) Renata