

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz



ESCOLA NACIONAL DE SAÚDE PÚBLICA
SERGIO AROUCA
ENSP

Natália Santana Paiva

“Quantas mulheres fizeram aborto provocado nas capitais brasileiras? Resultados de um inquérito nacional com a aplicação de modelos hierárquicos utilizando informações indiretas baseadas na rede de contatos”

Rio de Janeiro

2018

Natália Santana Paiva

“Quantas mulheres fizeram aborto provocado nas capitais brasileiras? Resultados de um inquérito nacional com a aplicação de modelos hierárquicos utilizando informações indiretas baseadas na rede de contatos”

Tese de doutorado apresentada ao Programa de Pós-graduação em Epidemiologia em Saúde Pública, da Escola Nacional de Saúde Pública Sérgio Arouca, na Fundação Oswaldo Cruz, como requisito parcial para a obtenção do título de Doutora em Ciências.

Orientador: Prof. Dr. Francisco Inácio Pinkusfeld Monteiro Bastos

2º Orientador: Prof. Dr. Leonardo Soares Bastos

3º Orientador: Prof. Dr. Daniel Antunes Maciel Villela

Rio de Janeiro

2018

Catálogo na fonte
Fundação Oswaldo Cruz
Instituto de Comunicação e Informação Científica e Tecnológica em Saúde
Biblioteca de Saúde Pública

P149q Paiva, Natália Santana.
Quantas mulheres fizeram aborto provocado nas capitais brasileiras? Resultados de um inquérito nacional com a aplicação de modelos hierárquicos utilizando informações indiretas baseadas na rede de contatos / Natália Santana Paiva. -- 2018.
108 f. : il. color. ; graf. ; mapas ; tab.

Orientadores: Francisco Inácio Pinkusfeld Monteiro Bastos, Leonardo Soares Bastos e Daniel Antunes Maciel Villela.
Tese (doutorado) – Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública Sérgio Arouca, Rio de Janeiro, 2018.

1. Aborto Induzido - estatística & dados numéricos. 2. Métodos de estimação. 3. Modelos hierárquicos. 4. Network Scale-Up Method (NSUM). 5. Transtornos Relacionados ao Uso de Substâncias. 6. Estimativas de População. I. Título.

CDD – 22.ed. – 363.46

Natália Santana Paiva

“Quantas mulheres fizeram aborto provocado nas capitais brasileiras? Resultados de um inquérito nacional com a aplicação de modelos hierárquicos utilizando informações indiretas baseadas na rede de contatos”

Tese de doutorado apresentada ao Programa de Pós-graduação em Epidemiologia em Saúde Pública, da Escola Nacional de Saúde Pública Sérgio Arouca, na Fundação Oswaldo Cruz, como requisito parcial para a obtenção do título de Doutora em Ciências.

Aprovada em: 15/08/2018

Banca Examinadora

Prof. Dra. Sandra Costa Fonseca
UFF

Prof. Dr. Fernando Antônio da Silva Moura
UFRJ

Prof. Dr. Guilherme Loureiro Werneck
UERJ, UFRJ

Prof. Dr. Claudio Jose Struchiner
FIOCRUZ

Prof. Dr. Francisco Inácio Pinkusfeld Monteiro Bastos (Orientador)
FIOCRUZ

Rio de Janeiro

2018

AGRADECIMENTOS

A Deus por me dar força para correr atrás dos meus objetivos e jamais me deixar desistir.

À minha família, pelo carinho e apoio incondicional. Obrigada por acreditarem na educação e me incentivarem a isso.

Aos meus orientadores, Francisco Inácio Bastos e Daniel Villela. Foi uma honra poder aprender com vocês ao longo desses anos.

Ao meu orientador de longa data, Leonardo Bastos, que ajudou MUITO não somente nesta tese, mas me ensina, não só sobre pesquisa, epidemiologia, *ggplot2* ou estatística, mas também sobre a vida. Obrigada por investir na minha formação profissional nesses anos que trabalhamos juntos (e não ter desistido de mim). Obrigada por comemorar cada conquista que tivemos, pois você fez e faz parte desse processo. Você, para mim, é um exemplo de professor e orientador, além de ser um estatístico de inteligência ímpar!

À Secretaria Nacional de Políticas sobre Drogas (SENAD) pelo apoio financeiro à “Pesquisa Nacional sobre o uso do Crack”, à qual esta tese é vinculada.

Aos coordenadores e colaboradores da “Pesquisa Nacional sobre o uso do Crack”.

Ao meu marido, Paulo Victor, que além de discutir modelagem matemática, Epidemiologia II, Aborto e Tuberculose de domingo a domingo, compreendeu todos os dias de estresse e lamúrias e soube me acolher neste momento delicado.

À CAPES pelo apoio financeiro com a bolsa de Doutorado.

Ao corpo social do IESC e aos meus alunos por esses 3 anos de acolhimento e parceria.

E à FIOCRUZ, instituição que me acolheu, desde o ensino técnico profissionalizante, na figura de diversos pesquisadores e professores que pude e posso conviver, tanto na EPSJV, ENSP, PROCC, ICICT e outros institutos, que me fazem compreender sempre a importância de se trabalhar com saúde pública.

Enfim, agradeço a todos os que me ajudaram e incentivaram, de uma forma ou outra, mesmo não tendo sido citados aqui.

Muito obrigada!

[We] cannot know that any statistical technique we develop is useful unless we use it. Major advances in science and in the science of statistics in particular, usually occur, therefore, as the result of the theory-practice iteration.

BOX, 1976

Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad.

BOX, 1976

RESUMO

Introdução: O estudo do aborto provocado é de grande interesse em Saúde Pública, pois sabe-se que o aborto provocado, muitas vezes, inseguro, representa riscos indevidos para a saúde e a vida das mulheres, assim como gera consequências adversas, sociais e financeiras, para o sistema de saúde público e famílias. Dessa forma, é fundamental conhecer sua magnitude com o maior grau de precisão possível. **Objetivo:** Estimar o número de mulheres que fizeram abortos (provocados), nos últimos 12 meses anteriores à avaliação, nas 27 capitais brasileiras, com base em modelos hierárquicos, considerando o desenho amostral e a estrutura demográfica da população, além de informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população em geral. **Materiais e métodos:** O presente projeto tem como fonte de dados a base de dados secundários de acesso restrito oriundos de Inquérito domiciliar nas 27 capitais em 2012, como parte de uma ampla Pesquisa Nacional sobre o uso de crack, já concluída. Os indivíduos amostrados aleatoriamente, via amostragem complexa, na população geral das 27 capitais brasileiras no inquérito domiciliar, que contemplou o emprego da metodologia “Network scale-up”, foram questionados sobre uma gama de perguntas. Dentre elas, a pergunta de interesse do presente projeto, foi: “Quantas mulheres você conhece que residem neste município e tiveram aborto provocado nos últimos 12 meses?”. **Resultados:** Das 13.611.082 mulheres de 15 a 49 anos residentes nas capitais do Brasil, 65.682 mulheres fizeram aborto provocado no período de 2011-2012, resultando numa incidência acumulada de 4,83 para cada 1.000 mulheres de 15 a 49 anos. Em média, as estimativas da incidência acumulada de mulheres que fizeram aborto foram superiores nas capitais das regiões Norte e Nordeste do Brasil, como pode ser visto em Boa Vista (RR) (9,81 IC95% [6,08; 14,60]), Aracaju (SE) (9,26 IC95% [5,66; 14,00]), São Luís (MA) (8,77 IC95% [5,47; 12,95]). No entanto, ao nível de 95% de credibilidade, não houve diferença significativa entre as estimativas nas capitais brasileiras. **Conclusão:** O abortamento provocado é uma prática recorrente entre as mulheres nas capitais brasileiras. A utilização de metodologias de estimação indireta é preferível em contextos de estigma e eventual criminalização, como mulheres que fizeram aborto nas capitais do Brasil, e o método *Network Scale-up* apresenta-se como um método promissor para estimar o tamanho de populações de difícil acesso, em diversas áreas do conhecimento, como a Saúde Pública. **Palavras-chave:** Aborto induzido; Aborto provocado; Métodos de estimação; Modelos hierárquicos; *Network scale-up method (NSUM)*.

ABSTRACT

Introduction: The study about provoked abortion is a major Public Health interest. The provoked abortion, many times unsafe, represents a great risk in health and life for women, as well as generating adverse social and financial consequences for families and public health system. In this way, it is essential to know this magnitude with high accuracy as possible.

Objective: The number of women that performed the provoked abortion in the last 12 months prior to the evaluation considering 27 State capitals in Brazil based on hierarchical models, considering the sample design and demographic structure of the population is estimate. In addition, indirect information based on the social network of each participant randomly choose from the population is considered.

Materials and methods: The data source used is the secondary database of restricted access from household surveys in the 27 capitals in 2012. This database belongs to the National Survey about smoke crack, already accomplished. The individuals sampled randomly, through complex sampling, in the general population of the 27 Brazilian capitals in the household survey, which included the use of the "Network scale-up" methodology, were asked about a range of questions. The main question of interest in this thesis: "How many women do you know who live in this city and have had abortion in the last 12 months?"

Results: From the total of 13,611,082 women aged 15-49 years living in the capitals of Brazil, 65,682 women induced provoked abortion in the period 2011-2012, resulting in a cumulative incidence of 4.83 per 1000 women aged 15-49 years. Cumulative incidence estimates of women that performed provoked abortion are considered high in the regions North and Northeast of Brazil, however, at the 95% level of credibility, there was no significant difference between estimates in Brazilian capitals.

Conclusions: The practice of provoked abortion is recurrent among women in Brazilian. The use of indirect estimation methodologies is preferable in contexts of stigma and eventual criminalization, such women that induce an abortion in Brazilian capitals, and the Network Scale-up method is a promising method to estimate the size of populations that are difficult to access, in several areas of knowledge, such as Public Health.

Keywords: Induced Abortion; Provoked abortion; Estimation methods; Hierarchical modelling; Network scale-up (NSUM).

LISTA DE ILUSTRAÇÕES

Figura 1 - Legislação sobre aborto provocado no mundo, de acordo com a divisão por países	19
Figura 2 - Distribuição (%) da amostra do Inquérito domiciliar (a) e da População (Censo demográfico de 2010) (b) segundo sexo e faixa etária, Brasil	50
Figura 3 - Quantos amigos uma pessoa pode ter?	62
Figura 4 - Autovalidação geral do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 20 subpopulações de tamanhos conhecidos para as 27 capitais brasileiras: estimativas pseudo posteriori da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações em suas capitais.	73
Figura 5 - Autovalidação por subpopulação de tamanho conhecido do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 27 capitais brasileiras: estimativas pseudo posteriori da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações	74
Figura 6 - Autovalidação por capital de tamanho conhecido do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 20 subpopulações de tamanho conhecido: estimativas pseudo posteriori da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações	76
Figura 7 – Incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) nas 27 capitais brasileiras no período de 2011-2012	79
Figura 8 – Média e IC95% a pseudo posteriori da Incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) nas 27 capitais brasileiras, no período de 2011-2012, segundo as Regiões	80
Quadro 1 - Descrição dos estudos cujo objetivo era a estimação do número de abortos induzidos (ou número de mulheres que induziram aborto) no Brasil, regiões e UF, suas técnicas utilizadas e estimativas encontradas	37
Quadro 2 - Modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação (modelo completo) proposto nessa tese e seus casos particulares	57
Quadro 3 - Populações conhecidas que serão utilizadas para estimação do grau da rede de contatos dos indivíduos e fonte de dados de onde foram retiradas	98

LISTA DE TABELAS

- Tabela 1** - Quantitativos da amostra utilizada no inquérito domiciliar da Pesquisa Nacional sobre crack segundo sexo nas 27 capitais brasileiras, 201246
- Tabela 2** - Perfil da amostra geral do inquérito domiciliar da Pesquisa Nacional sobre uso de Crack segundo sexo, faixa etária, raça ou cor, estado civil e escolaridade, 201248
- Tabela 3** - Métricas de comparação para cada uma das 27 capitais considerando os modelos NSUM bayesiano, NSUM bayesiano com pesos amostrais, NSUM bayesiano com pós-estratificação e NSUM bayesiano com pesos amostrais e pós-estratificação e métrica de comparação global (Total do RMSE por modelo) 72
- Tabela 4** - Distribuição da incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres entre 15 e 49 anos) e do número de mulheres que fizeram aborto, e seus intervalos de credibilidade de 95% (IC95%), nas 27 capitais brasileiras no período de 2011-2012..... 77

LISTA DE ABREVIATURAS E SIGLAS

ACASI - AUDIO COMPUTER-ASSISTED SELF-INTERVIEW

DF – DISTRITO FEDERAL

DHS - DEMOGRAPHIC HEALTH SURVEYS

ENSP – ESCOLA NACIONAL DE SAÚDE PÚBLICA SERGIO AROUCA

EQM - ERRO QUADRÁTICO MÉDIO

FIOCRUZ - FUNDAÇÃO OSWALDO CRUZ

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA

IC - INTERVALO DE CREDIBILIDADE

INAMPS - INSTITUTO NACIONAL DE ASSISTÊNCIA MÉDICA DA PREVIDÊNCIA SOCIAL

INFOPEN - INFORMAÇÕES PENITENCIÁRIAS

MCMC - MARKOV CHAIN MONTE CARLO

NHS - NATIONAL HEALTH SYSTEM

NSUM - NETWORK SCALE-UP METHOD

OMS – ORGANIZAÇÃO MUNDIAL DE SAÚDE

ONU - ORGANIZAÇÃO DAS NAÇÕES UNIDAS

PNA – PESQUISA NACIONAL DE ABORTO

RMSE - ROOT-MEAN-SQUARE ERROR

SIH-SUS – SISTEMA DE INFORMAÇÕES HOSPITALARES DO SUS

SUS – SISTEMA ÚNICO DE SAÚDE

TLS - TIME LOCATION SAMPLING

TRA – TEORIA DE RESPOSTA AO AZAR

UF – UNIDADE DA FEDERAÇÃO

LISTA DE SÍMBOLOS

n : tamanho da amostra

N : tamanho da população

N^* : tamanho da população de mulheres de 15-49 anos

N_j : tamanho da população do j -ésimo estrato

n_j : tamanho da amostra do j -ésimo estrato

J : número total de estratos da população

K : número de subpopulações de tamanho conhecido

π_k : prevalência da k -ésima subpopulação de tamanho conhecido (valor previamente conhecido)

N_k : tamanho da subpopulação conhecida k

Y_i : número de mulheres que fizeram aborto (aborto provocado) nos últimos 12 meses anteriores a avaliação que o indivíduo i conhece

X_{ik} : número de pessoas da subpopulação k , cujo tamanho é conhecido, que o indivíduo i conhece

δ_i : grau da rede de contatos do indivíduo i (ou seja, o número total de contatos do indivíduo i)

N_Y : número de mulheres que fizeram aborto nos últimos 12 meses

θ : incidência acumulada de mulheres que fizeram aborto nos últimos 12 meses (por 1.000 mulheres de 15-49 anos)

Y_{ij} : número de mulheres que fizeram aborto (aborto provocado) nos últimos 12 meses que o indivíduo i pertencente ao estrato j conhece

θ_j : incidência acumulada de mulheres que fizeram aborto segundo as informações indiretas baseadas na rede de contatos dos participantes pertencentes ao estrato j

SUMÁRIO

1	INTRODUÇÃO	15
1.1	ESTRUTURA DO TRABALHO	17
2	REVISÃO DE LITERATURA	18
2.1	ABORTO	18
2.1.1	Aborto no mundo.....	19
2.1.2	Aborto no Brasil.....	21
2.2	REVISÃO DOS MÉTODOS DE ESTIMAÇÃO DE ABORTO	25
2.2.1	Métodos baseados em dados secundários.....	26
2.2.2	Pesquisas à beira do leito.....	28
2.2.3	Métodos das questões indiretas.....	29
2.2.4	Métodos diretos.....	29
2.2.5	Métodos indiretos.....	31
2.3	ESTIMATIVAS DE ABORTO PROVOCADO NO BRASIL, REGIÕES E UF	33
3	JUSTIFICATIVA	40
4	OBJETIVOS	42
4.1	OBJETIVO GERAL	42
4.2	OBJETIVOS ESPECÍFICOS	42
5	MATERIAIS E MÉTODOS	43
5.1	PESQUISA NACIONAL SOBRE O USO DO CRACK	43
5.1.1	Análise descritiva do inquérito domiciliar.....	45
5.2	METODOLOGIA NETWORK SCALE-UP	51
5.3	MODELO HIERÁRQUICO	56
5.3.1	Modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação.....	57
5.3.2	Casos particulares.....	62
5.3.3	Inferência e detalhes computacionais.....	65
5.4	AUTOVALIDAÇÃO DOS MODELOS PROPOSTOS	67
5.5	ASPECTOS ÉTICOS	69
6	RESULTADOS	71
6.1	AUTOVALIDAÇÃO	71
6.2	ESTIMATIVAS DO NÚMERO DE MULHERES QUE FIZERAM ABORTO	76
7	DISCUSSÃO	81
8	CONSIDERAÇÕES FINAIS	86
	REFERÊNCIAS	87
	ANEXO A– PARECER DE APROVAÇÃO DA PESQUISA NACIONAL SOBRE O USO DE CRACK E/OU SIMILARES	93
	ANEXO B– TERMO DE CONSENTIMENTO DA PESQUISA NACIONAL SOBRE O USO DE CRACK E/OU SIMILARES	94
	ANEXO C– LISTAGEM DAS SUBPOPULAÇÕES CONHECIDAS QUE FORAM UTILIZADAS PARA ESTIMAÇÃO DO GRAU DA REDE DE CONTATOS	98
	ANEXO D– DECLARAÇÃO DE CESSÃO DO BANCO DE DADOS DA PESQUISA NACIONAL DO CRACK	99
	APÊNDICE – DISTRIBUIÇÃO DA AMOSTRA DO INQUÉRITO DOMICILIAR E DA POPULAÇÃO (CENSO DEMOGRÁFICO 2010) SEGUNDO SEXO E FAIXA ETÁRIA PARA AS 27 CAPITAIS BRASILEIRAS	100

1 INTRODUÇÃO

A elevada incidência de abortos provocados, muitas vezes inseguros, ainda é um problema de saúde pública e um dos principais índices em saúde reprodutiva. A taxa de gravidez indesejada foi de 62 (IC90%[59;72]) por 1.000 mulheres com idades entre 15 e 44 anos em todo o mundo a cada ano no período de 2010 a 2014 e 56% (IC90%[53;60]) de todas as gravidezes indesejadas terminaram em aborto no mesmo período (BEARAK et al., 2018).

Ainda segundo Bearak e cols (2018), no mesmo período, países em desenvolvimento apresentaram a taxa de gravidez indesejada (para cada 1.000 mulheres de 15-44 anos) superior (65 IC90%[62;76]) a de países desenvolvidos (42 IC90%[38;56]). E o percentual das gravidezes indesejadas que terminaram em aborto nos países em desenvolvimento foi 55% (IC90%[52;60]), já nos países desenvolvidos foi 59% (IC90%[54;65]).

Na América do Sul, em 2010-2014, a taxa de gravidez indesejada, por 1.000 mulheres com idades entre 15 e 44 anos, (99 IC90%[38;56]) foi superior a taxa mundial (BEARAK et al., 2018). Do total de gravidezes indesejadas, nos países da América do Sul, em 2010-2014, 47% (IC90%[40;57]) terminaram em aborto.

Segundo Ganatra e cols (2017) aproximadamente 45% (IC90%[40,6;50,1]) dos abortos realizados no mundo, no período de 2010-2014, foram realizados de forma insegura. Além disso, identificaram que há uma disparidade na segurança do aborto entre regiões desenvolvidas e em desenvolvimento no período sob análise. Na América do Sul, 75% (IC90%[46,3;95,4]) dos abortos foram realizados de forma insegura, gerando riscos evitáveis para a saúde e vida da mulher.

Abortos inseguros, ou seja, realizados em condições adversas (como por exemplo, sem assistência médica adequada), continuam sendo uma das principais causas de morbidade e mortalidade materna em todo o mundo (RASTEGARI et al., 2014; SINGH; MADDOW-ZIMET, 2015; WHO, 2015). Na América Latina e Caribe, em 2015, a estimativa da razão de mortalidade materna foi 67 (IC80%[64;77]), já no Brasil esta estimativa foi 44 mortes maternas para cada 1000 nascidos vivos (WHO, UNICEF, UNFPA, WORLD BANK GROUP, 2015).

No entanto, segundo o painel de monitoramento da mortalidade materna do Ministério da Saúde¹, o país continua com valores em torno de 60 por 1000 nascidos vivos, sendo notificadas cerca de 1.600 mortes maternas anualmente.

¹ Disponível em <http://svs.aids.gov.br/dantps/centrais-de-conteudos/paineis-de-monitoramento/mortalidade/materna/>, acessado em 06/08/2018.

Porém, estimar o número de abortos provocados ou de mulheres que fizeram aborto provocado ao menos uma vez ao longo da vida em contextos cuja prática é considerada ilegal, como no Brasil, ainda é um desafio, pois essas mulheres, em geral, evitam relatar tal procedimento, pois além das questões legais, uma vez que o aborto provocado pela gestante ou com seu consentimento é considerado crime contra a vida, previsto no Decreto-Lei nº 2.848, de 7 de dezembro de 1940, do Código Penal (BRASIL, 1940), e prescreve somente em 8 anos, essas mulheres ainda enfrentam julgamentos da sociedade tanto nos aspectos morais e religiosos.

A utilização de técnicas de estimação direta tem limitações, uma vez que o aborto provocado é estigmatizado pela sociedade e considerado uma infração penal (“crime contra a vida”) segundo as leis brasileiras. Uma alternativa é a aplicação de técnicas de estimação indireta que estão em pleno desenvolvimento na literatura científica, não havendo, por ora, um método que seja considerado padrão-ouro (ROSSIER, 2003).

Uma possível metodologia para estimar o tamanho de subpopulações de difícil acesso, que se utiliza de informações indiretas da rede de contatos dos participantes da pesquisa, e vem sendo empregada, em diversos contextos de Saúde Pública, tanto em estudos internacionais (BERNARD et al., 1991, 2010; KADUSHIN et al., 2006; KILLWORTH et al., 1998a; MALTIEL et al., 2015; ZHENG; SALGANIK; GELMAN, 2006) quanto em estudos nacionais (BERTONI, 2014; SALGANIK et al., 2011a), e está em pleno processo de desenvolvimento, é a metodologia NSUM (sigla do inglês para *Network scale-up method*). No entanto, até o presente momento, no Brasil, nenhum estudo empregou esta metodologia para estimar o número de mulheres que fizeram um aborto provocado.

O NSUM permite estimar o tamanho de uma subpopulação de difícil acesso sem que seja necessário selecionar pessoas da subpopulação de interesse, ou seja, não é preciso abordar/questionar mulheres que fizeram um aborto provocado nos últimos 12 meses sobre tal prática. Nessa metodologia, seleciona-se pessoas da população geral, de forma aleatória, muitas vezes via uma amostragem complexa, e questiona-se “quantas pessoas você conhece da subpopulação k ?”. Com isso, é possível (e necessário) perguntar sobre diversas subpopulações, cujos tamanhos são conhecidos (via cadastros nacionais, por exemplo) e desconhecido(s) que se deseja conhecer/estimar.

Diante disso, o escopo deste estudo é estimar o número de mulheres que fizeram abortos provocados, no ano anterior à avaliação, nas 27 capitais brasileiras via modelos hierárquicos

que consideram o desenho amostral e a estrutura demográfica da população, utilizando informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população brasileira.

Esta tese de doutorado é um subprojeto de uma ampla pesquisa nacional, encerrada em 2012, intitulada “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”, aprovada pelo Comitê de Ética da Escola Nacional de Saúde Pública Sergio Arouca (ENSP/FIOCRUZ), sob o número CAAE 0073.0.031.000-11, desenvolvido pela Fundação Oswaldo Cruz (FIOCRUZ), sob a coordenação geral do pesquisador Francisco Inácio Bastos, que além do objetivo principal (o consumo do crack) contemplou uma série de outras questões relacionadas à saúde pública, como a prática do aborto provocado, usuários de drogas ilícitas em geral (exceto a maconha), mulheres profissionais do sexo, entre outras.

Os dados coletados no inquérito domiciliar conduzido na pesquisa nacional “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”, permitem estimar o tamanho de subpopulações de difícil acesso, como, por exemplo, mulheres que fizeram aborto (aborto provocado) no ano anterior à avaliação, mediante a utilização NSUM.

1.1 ESTRUTURA DO TRABALHO

Esta tese de doutorado está dividida em 8 capítulos. O presente capítulo (capítulo 1) apresenta a introdução, o capítulo 2 apresenta uma revisão de literatura sobre os temas abordados na presente pesquisa, o que inclui a questão do aborto, suas definições e dados epidemiológicos, técnicas de estimação indireta do número de abortos provocados, ou de mulheres que provocaram abortos, utilizadas na literatura científica para obtenção dessas estimativas no Brasil e grandes regiões, segundo estudos nacionais e internacionais além das estimativas resultantes. Os capítulos 3 e 4 apresentam a justificativa e os objetivos da presente tese, respectivamente. O capítulo 5 apresenta os materiais e métodos, isto é, a Pesquisa Nacional sobre o uso do Crack (intitulada “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”), a metodologia NSUM e a modelagem proposta através do uso de modelos hierárquicos bayesianos com pesos amostrais e estrutura etária da população. Por fim, o capítulo 6 apresenta os resultados tanto da autovalidação dos modelos propostos como as estimativas do número de mulheres que fizeram aborto nas capitais brasileiras no ano anterior à avaliação. O capítulo 7 apresenta a discussão e o capítulo 8 as considerações da presente tese de doutorado.

2 REVISÃO DE LITERATURA

A revisão de literatura, com a finalidade de analisar criticamente os temas abordados nesta tese, foi subdividida em três partes:

- i. a primeira contempla a temática do aborto e suas definições, o impacto do abortamento provocado em contextos nos quais a prática é ilegal, como no Brasil, sua epidemiologia e taxas de morbidade e mortalidade decorrentes de tal prática;
- ii. a segunda, dedicada aos métodos de estimação indireta do aborto induzido dos quais alguns foram empregados em estudos de abrangência nacional com a finalidade de estimar o número de abortos provocados ou de mulheres que fizeram aborto (“aborto provocado”) pelo menos uma vez na vida no Brasil, grandes regiões e unidades da federação;
- iii. a terceira traça um panorama das estimativas do aborto induzido nas grandes regiões e Brasil, provenientes de estudos nacionais e internacionais desde a década de 80 até os dias atuais.

2.1 ABORTO

Segundo a Organização Mundial de Saúde (OMS), o aborto é definido como a interrupção da gravidez antes que o feto seja viável para sobrevivência extrauterina. E a viabilidade é definida em função da duração gestacional e do peso do feto (WHO, 1970, p. 6).

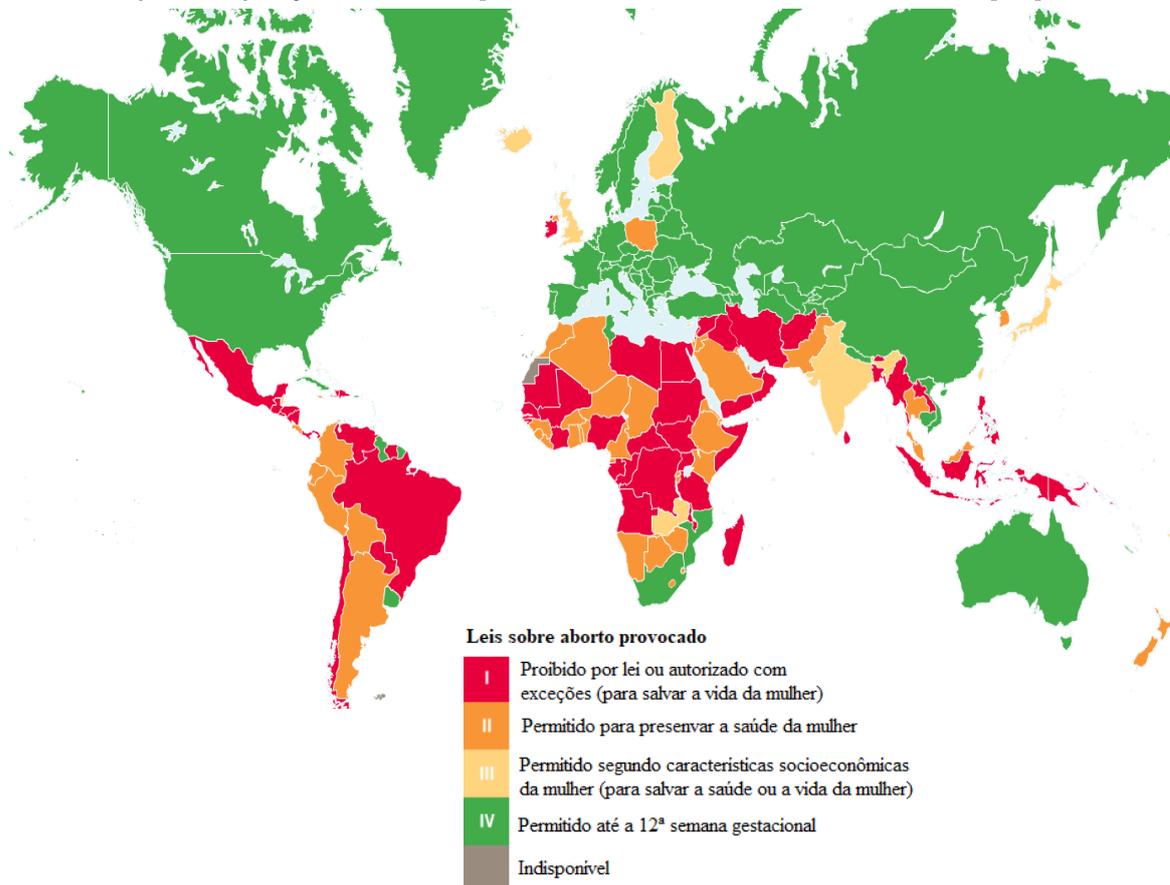
Os abortos podem ser divididos em abortos espontâneos ou induzidos (provocados), e o Relatório da OMS definiu que “abortos induzidos são aqueles iniciados por ação deliberada realizado com a intenção de terminar a gestação. Os demais abortos são considerados como espontâneos, mesmo que alguma causa externa esteja envolvida” (WHO, 1970, p.8).

Muitas vezes os abortos provocados são realizados por pessoas sem as habilidades mínimas necessárias ou em locais que não estão de acordo com os padrões médicos e higiênicos mínimos, ou ambos. Dessa forma, segundo a OMS, os abortos realizados sob tais condições são caracterizados como abortos inseguros e representam riscos evitáveis para a saúde e a vida da mulher. Já abortos realizados de acordo com as normas sancionadas pela OMS e diferentes associações médicas (obviamente, em contextos onde o aborto é legal, e tais normas constituem documentos juridicamente válidos) são considerados seguros e o risco de complicações graves ou a morte é mínima (GANATRA et al., 2014; WHO, 1992, 2011).

2.1.1 Aborto no mundo

No mundo há quatro grupos de países, divididos segundo os marcos legais mais amplos sobre a legalidade ou não do abortamento, como pode ser visto na Figura 1^{2,3}.

Figura 1 Legislação sobre aborto provocado no mundo, de acordo com a divisão por países



O Grupo I (países representados pela cor vermelha na Figura 1) é formado pelos países cujas leis são as mais restritivas, isto é, em que aborto provocado é completamente proibido ou é autorizado com exceções (salvar a vida da gestante). O Grupo II (países representados pela cor laranja na Figura 1) é constituído pelos países que permitem o aborto provocado para preservar a saúde (incluindo a saúde mental, em alguns países do grupo) das gestantes. O Grupo III (países representados pela cor amarela na Figura 1) é composto pelos países que permitem o aborto provocado baseado na situação socioeconômica da gestante, como a idade da mãe, sua

² Disponível em <https://www.reproductiverights.org/sites/crr.civicactions.net/files/documents/AbortionMap2014.PDF>, acessado em 10/03/2018.

³ As respectivas legislações não são necessariamente nacionais, podendo ser de abrangência infranacional em países com marcada estrutura federativa (como é o caso dos EUA, mas NÃO é o caso do Brasil, país com unidades federativas sem qualquer autonomia legislativa nessa e diversas outras matérias). Neste caso, segue-se o critério da legislação predominante na maioria das unidades infranacionais enquanto representativa do conjunto, para fins ilustrativos.

situação econômica, o local onde reside ou seu estado civil, com o objetivo de salvar a vida ou preservar a saúde da gestante. O último grupo, Grupo IV (países representados pela cor verde na Figura 1), é formado pelos países que permitem o aborto provocado, de forma independente da razão subjacente (que não precisa ser explicitada, para fins legais), com relação a conceitos de grávidas com até 12 semanas de gestação.

Em 2014, segundo o Centro de Direitos Reprodutivos (CENTER FOR REPRODUCTIVE RIGHTS, 2014), em 66 países, incluindo o Brasil, o aborto induzido (provocado) não é permitido por lei ou é autorizado com exceções, como, por exemplo, em casos de estupro, má formação fetal e/ou risco de vida da gestante (Grupo I).

Em 59 países, o aborto é permitido em casos de estupro, má formação fetal, risco de vida e/ou preservação da saúde da gestante (Grupo II), sendo que em alguns países desse grupo o aborto também é permitido para preservar a saúde mental das mulheres, como em Gambia, Gana, Nova Zelândia, entre outros.

Em apenas 13 países, o abortamento induzido (provocado) é permitido em casos de estupro, má formação fetal e segundo alguns fatores socioeconômicos (definidores de vulnerabilidades, potencialmente associadas a riscos à saúde), como a idade da mãe, sua situação econômica, o local onde reside ou seu estado civil (Grupo III), com o intuito de preservar a saúde e a vida da mulher.

E a última grande partição/interpretação do marco legal, adotada por 61 países, é a que não impõe restrições legais quanto ao aborto (Grupo IV), ou seja, o ato de abortar (provocar o aborto) é permitido, independente da razão subjacente, até a 12ª semana gestacional.

Dos países e territórios da América Latina e do Caribe, México, Brasil, Chile, República Dominicana, El Salvador, Haiti, Honduras, Nicarágua e Suriname proíbem a prática do aborto em todos os casos (salvo algumas exceções para salvar a vida da mulher), já em Cuba, Porto Rico, Guiana, Guiana Francesa e Uruguai o aborto é descriminalizado até a 12ª semana de gestação. Nos demais países da região, o aborto é considerado uma prática ilegal, com exceções, sendo que estas variam de país a país (CENTER FOR REPRODUCTIVE RIGHTS, 2014).

Segundo a Organização Mundial de Saúde (OMS), em 2008, estimou-se, que cerca de 22 milhões de abortos inseguros foram praticados no mundo, e desses 4,2 milhões foram realizados na América Latina e Caribe (WHO, 2011). Já em 2010-2014, foram estimados 56 milhões de abortos induzidos, realizados a cada ano em todo o mundo, ou seja, uma taxa anual de aborto induzido estimada em 35 abortos por 1.000 mulheres de 15-44 anos, em 2010-2014,

no mundo. Na América Latina e Caribe, estimou-se que, em 2010-2014, 65 milhões de abortos induzidos foram realizados, correspondendo a uma taxa de 44 abortos por 1.000 mulheres de 15-44 anos (GUTTMACHER INSTITUTE, 2016a) .

Segundo os dados do Instituto Alan Guttmacher (2016b), em países onde o aborto é proibido por completo ou permitido apenas para salvar a vida da mulher, estima-se uma taxa de 37 abortos por 1.000 mulheres com idade entre 15 a 44 anos. Já em países onde o aborto é permitido, estima-se uma taxa equivalente a 34 abortos por 1.000 mulheres com idade entre 15 a 44 anos. Dessa forma, conclui-se que, mesmo sendo considerado crime em algumas localidades, o aborto induzido continua sendo praticado, não só no Brasil como no mundo. E, muitas vezes, eles acabam sendo realizados em condições adversas, tornando-se, muitas vezes, abortos inseguros, podendo colocar em risco a saúde da mulher (GUTTMACHER INSTITUTE, 2012).

No período de 2010 a 2014, estimou-se terem ocorrido aproximadamente 56 milhões de abortos, por ano, em todo o mundo, dos quais 25 milhões eram considerados abortos inseguros. No mesmo período, estimou-se a realização de 6,4 milhões de abortos por ano na América Latina e Caribe, dos quais 4,9 milhões foram realizados de forma insegura (2017a).

Estima-se que, a cada ano, mais de um terço das gestações terminam em morte fetal ou aborto, espontâneo ou provocado. Estimativas recentes, usando métodos estatísticos distintos, indicam que o aborto inseguro está associado a 8% a 15% das mortes maternas, durante ou no pós-parto, em todo o mundo (WHO, 2008; GUTTMACHER INSTITUTE, 2016b). Cabe aqui observar que o Brasil, por diferentes razões, muito possivelmente incluindo esta, não atingiu as assim denominadas Metas do Milênio quanto à redução das taxas de morte materna, em consonância com as metas pactuadas pela ONU (Organização das Nações Unidas) (VICTORA et al., 2016).

2.1.2 Aborto no Brasil

No Brasil, o aborto é ilegal, com algumas (poucas) exceções, segundo o Artigo 128 do Código Penal - Decreto Lei 2.848/40 de 07 de Dezembro de 1940, cujo texto segue abaixo.

“Aborto provocado pela gestante ou com seu consentimento

Art. 124. Provocar aborto em si mesma ou consentir que outrem lhe provoque:

Pena - detenção, de um a três anos.

Aborto provocado por terceiro

Art. 125. Provocar aborto, sem consentimento da gestante:

Pena - reclusão, de três a dez anos.

Art. 126. Provocar aborto com o consentimento da gestante:

Pena - reclusão, de um a quatro anos.

Parágrafo único. Aplica-se a pena do artigo anterior, se a gestante não é maior de quatorze anos, ou é alienada ou débil mental, ou se o consentimento é obtido mediante fraude, grave ameaça ou violência.

Forma qualificada

Art. 127. As penas cominadas nos dois artigos anteriores são aumentadas de um terço, se, em consequência do aborto ou dos meios empregados para provocá-lo, a gestante sofre lesão corporal de natureza grave; e são duplicadas, se, por qualquer dessas causas, lhe sobrevém a morte.”

As exceções para as quais o aborto é legalmente autorizado no Brasil são: risco de vida para a gestante, gravidez resultante de estupro e, mais recentemente, segundo a resolução do Conselho Federal de Medicina nº 1.989/2012, os casos de anencefalia.

Informações sobre a incidência de abortos induzidos são cruciais para a formulação e monitoramento de políticas públicas de educação sexual e reprodutiva, e permitem, por exemplo, o planejamento familiar e o incentivo ao uso de métodos anticoncepcionais com o objetivo de reduzir os índices/taxas de gravidez indesejada (RASTEGARI et al., 2014; SINGH; MADDOW-ZIMET, 2015).

No entanto, há uma escassez de informações confiáveis sobre incidência de aborto induzido em determinados locais, pois a tarefa de estimar estas taxas varia, por exemplo, de acordo com o estatuto jurídico da localidade da realização da mesma, isto é, depende se o aborto induzido é permitido por lei ou não, nesta determinada localidade. Em contextos em que os abortos são legais, majoritariamente realizados em unidades públicas de saúde, que contam com registros abrangentes e confiáveis, possivelmente os números registrados se aproximam dos eventos reais, embora persista sempre uma fração (de tamanho bastante variável) de abortos realizados em clínicas privadas, cujos registros não são compartilhados com as autoridades públicas, determinando, portanto, subenumeração dos abortos efetivamente realizados.

Uma síntese dos 20 anos de pesquisa, no Brasil, relacionadas ao aborto pode ser vista no livro “Aborto e Saúde Pública no Brasil: 20 anos” (BRASIL, 2009). Nessa síntese, 2.109 documentos revisados enquanto fontes de informação e referências bibliográficas foram

resgatados e analisados, e pôde-se concluir que dentre artigos de periódicos, livros, monografias e teses, a maioria era de artigos de opinião e peças argumentativas. Das 2.109 fontes, apenas 398 (18,88%) corresponderam a fontes bibliográficas que contemplavam pesquisas empíricas sobre aborto.

Ainda de acordo com a síntese dos 20 anos de pesquisas relacionadas ao aborto e Saúde Pública no Brasil (BRASIL, 2009), estudos descritivos, geralmente, adotaram variáveis sociodemográficas, como idade, classe social e religião potencialmente associados à prática do aborto, e variáveis clínicas e operacionais, como tempo gestacional, tipo de aborto, procedimento abortivo entre outras.

Além de tudo, as principais pesquisas sobre aborto no Brasil comprovaram que a prática ilegal de aborto impacta de forma negativa na saúde da mulher e explicitam as desigualdades sociais subjacentes, ao fazer incidir sobre as mulheres mais pobres as consequências mais graves da assistência inadequada ou simplesmente inexistente (BRASIL, 2009). Corroborando a literatura internacional (BRASIL, 2009), que evidencia que mulheres menos escolarizadas, pobres, adolescentes e/ou mulheres solteiras estão sob um maior risco de recorrer a práticas de aborto inseguro.

Além disto, o estudo “Aborto e Saúde Pública no Brasil: 20 anos” (BRASIL, 2009) ressalta que apenas 20% dos trabalhos analisados eram baseados em evidências de pesquisas empíricas. Isto documenta a dificuldade em estimar a magnitude do aborto induzido no Brasil, pois são raros os estudos que dispõem de evidências científicas, especialmente evidências com um grau adequado de precisão (é frequente, por exemplo, ver trabalhos que se limitam a apresentar estimativas pontuais, sem incluir quaisquer informações acerca dos erros a elas associados), pois os dados confiáveis sobre o abortamento inseguro são escassos, principalmente em locais com leis restritivas quanto ao aborto, como no Brasil.

A estimação do número de abortos inseguros, muitas vezes ilegais/ilícitos, não é trivial, pois métodos de estimação direta não retratam a realidade por se tratar de uma população estigmatizada, além de uma prática criminalizada em diversas sociedades. A estimação direta consiste em perguntar a um indivíduo da população geral se ele faz parte ou não da subpopulação de interesse, como, por exemplo, se fez aborto (aborto provocado) nos últimos 12 meses anteriores à avaliação (BRASIL, 2009).

O aborto envolve aspectos éticos, morais e religiosos. Essas dimensões, complexas e altamente subjetivas, se traduzem em dificuldades adicionais quanto ao relato fidedigno por

parte das mulheres, particularmente em países com leis restritivas quanto ao aborto e onde valores e normas comportamentais de inspiração religiosa tradicional se mostram relevantes, como no Brasil, ou seja, em contextos em que os conceitos de “falha moral” e “infração criminal” se sobrepõem e reforçam. Investigar a prática do aborto nessas condições requer cuidados e a utilização de propostas metodológicas inovadoras, que consigam lidar com as dificuldades mencionadas, e, idealmente, corrigir os erros de estimação delas decorrentes (UNAIDS/WHO, 2010).

Uma possível alternativa para a estimação direta em populações estigmatizadas ou de difícil acesso, como mulheres que fizeram um aborto ou usuárias de drogas ilícitas, é o método de estimação indireta. Tal metodologia reduz o erro secundário à estigmatização (e eventual, criminalização) de comportamento, pois pergunta-se sobre comportamentos de terceiros e não do respondente selecionado, da população geral, para participar do estudo (MENEZES; AQUINO, 2009). Além disso, ao se referir a terceiros que podem estar indisponíveis para contato e fornecimento de informações, tais métodos permitem estimar a fração de pessoas que se encontram em situações extremas e, ao menos temporariamente, inacessíveis aos inquéritos populacionais clássicos (por exemplo, detidas ou presas, hospitalizadas, abrigadas, em situação de rua etc.).

Essas frações, ainda que diminutas frente à população geral, são exatamente aquelas onde tendem a se concentrar diferentes eventos adversos, de uma perspectiva médica, social ou da justiça criminal (esta última no que diz respeito à frequente sobreposição de eventos determinantes da detenção e uma série de eventos infracionais ou ilícitos adicionais, reais ou presumidos; vide MISSE, 2010). Além disso, ao, desafortunadamente, selecionarem segmentos com bases em critérios discriminatórios, sem relação direta com o evento sob análise (no caso em tela, o aborto), podem determinar diferenciais de estimação de difícil ou impossível correção, pois impõe tais medidas restritivas de forma diferenciada a segmentos populacionais com base, por exemplo, no critério “cor de pele”/raça. Tais diferenciais reclamariam fatores de correção estratificados por cor de pele/raça (entre diversos outros, como gênero), cuja base empírica pode ser escassa ou nenhuma (a questão é incidentalmente abordada no contexto brasileiro, quase sempre sob a forma opinativa, a despeito, de, por exemplo, o sobreencarceramento da população negra [se comparada à branca; como documentado nos dados do INFOPEN, disponíveis em: <http://dados.mj.gov.br/dataset/infopen-levantamento-nacional-de-informacoes-penitenciarias/resource/5652dceb-d81a-402f-a5c8-e4d9175241f5>;

infelizmente repletos de lacunas e inconsistências], mas tem merecido análise minuciosas no contexto internacional; vide DRUCKER, 2013).

2.2 REVISÃO DOS MÉTODOS DE ESTIMAÇÃO DE ABORTO

Sabe-se que estimar o número de abortos provocados/induzidos não constitui uma tarefa fácil, pelo fato de ser uma prática estigmatizada pela sociedade em geral, e em muitos lugares, como no Brasil, ser considerado um ato ilícito e passível de apenação (com algumas, poucas exceções).

Devido à ausência de dados confiáveis, os pesquisadores têm, habitualmente, de recorrer a vários métodos de estimação (via de regra, metodologicamente simples e que supõem estruturas lineares de inter-relação entre fenômenos perceptíveis e eventos a serem inferidos e estimados; linearidade esta frequentemente pressuposta, mas não demonstrada) para estimar o número de abortos provocados.

Rossier (2003) procedeu a uma revisão de métodos de estimação da taxa de aborto induzido, tendo a autora encontrado, descrito e discutido oito métodos para estimar a frequência de aborto induzido.

Em 2010, Singh, Remez e Tartaglione (2010) publicaram uma revisão de metodologias para estimação da incidência de abortos e morbidades relacionadas ao aborto provocado. Essa revisão é composta por 14 capítulos, nos quais autores convidados escreveram sobre diferentes técnicas de estimação. Alguns dos métodos descritos nessa revisão já haviam sido citados em 2003 no estudo realizado por Rossier (2003).

A seguir serão descritos os métodos de estimação que se baseiam em dados secundários (tais como Sistemas de Informações em Saúde, pesquisas de base populacional etc), pesquisas à beira do leito, métodos de questões indiretas, métodos diretos e métodos indiretos empregados em estudos nacionais e internacionais para estimar o número de aborto provocado ou de mulheres que provocaram aborto.

Na seção 2.3, serão ilustradas as estimativas resultantes de cada uma das técnicas, descritas nessa seção, para estimar o número de abortos provocados (ou o número de mulheres que fizeram aborto provocado) no Brasil, grandes regiões e UF's.

2.2.1 Métodos baseados em dados secundários

Dentre os métodos de estimação descritos na revisão conduzida por Rossier (2003), tem-se o método estatístico que toma por base as complicações pós-aborto induzido. Esse método estima o número de abortos induzidos a partir do número de pacientes internadas para prestação de cuidados pós-aborto em hospitais. Como mulheres se hospitalizam por complicações de abortos induzidos e espontâneos, é necessário separar os casos de abortos induzidos e espontâneos. Essa separação (correção) é feita mediante o emprego de fatores de correção, conhecidos como multiplicadores. Essa técnica vem sendo amplamente empregada em diversos estudos nacionais e internacionais para estimar o número de abortos induzidos no Brasil, como pode ser visto em (MONTEIRO; ADESSE, 2006; MONTEIRO; ADESSE; DREZETT, 2015; SINGH; WULF, 1991).

Esses fatores de correção podem ser baseados, por exemplo, nas investigações do Instituto Guttmacher (1994), que formularam as hipóteses (e as empregaram como pressupostos a seus procedimentos de estimação) de que no Brasil cerca de 20% das mulheres que induziram aborto tiveram de ser hospitalizadas no pós-aborto, de que há um sub-registro de 12,5% e que a proporção de abortos espontâneos é de 25%.

Tais hipóteses formuladas pelo Instituto Guttmacher (1994) foram adotadas nos estudos de Monteiro & Adesse (2006), cuja fórmula utilizada por para estimar o número de abortos induzidos no Brasil, baseada nos dados internações por complicações decorrentes ao abortamento do SIH-SUS, para os anos de 1992 até 2005 era dada por: Número de abortos induzidos = Número de internações por abortamento $\times 5 \times 1,125 \times 0,75$, e de Monteiro, Adesse e Drezett (2015), em cuja fórmula há uma pequena alteração⁴, já que os resultados incorporam variabilidade apresentados em limites (inferior e superior), e não são apresentados apenas como estimativas pontuais. O limite inferior considerou que 25% das mulheres que induziram o aborto tiveram de ser hospitalizadas por complicações. Já o superior continuou considerando que 20% das mulheres que induziram aborto tiveram de ser hospitalizadas no pós-aborto.

Em 1991, o Instituto Alan Guttmacher conduziu um estudo utilizando-se dessa técnica de estimação baseada em dados de hospitalizações do então Instituto Nacional de Assistência Médica da Previdência Social (INAMPS). Nesse estudo, foram utilizados diversos fatores de

⁴ Limite inferior = Número de internações por abortamento $\times 4 \times 1,125 \times 0,75$ e Limite superior = Número de internações por abortamento $\times 5 \times 1,125 \times 0,75$.

correção tanto para corrigir as subnotificações como visando distinguir os abortos espontâneos dos induzidos.

A limitação óbvia deste método é que, na vigência de complicações, mesmo graves, a demanda por cuidados médicos, e especialmente hospitalares, assim como disponibilidade de recursos assistenciais é claramente diferencial, como discutido pela epidemiologia clássica (WILLIAMS; TORRENS, 2007).

Outra técnica baseada em dados secundários é conhecida como método residual, também descrito em ROSSIER (2003). Esse método é uma adaptação da fórmula de Bongaarts⁵, no qual o nível teórico de fertilidade biológica máxima (TF) é multiplicado por índices que representam os efeitos de redução da fecundidade, como o casamento (Cm), a prática contraceptiva (Cc), o aborto induzido (Ca) e a “infecundidade” pós-parto (Ci), como o período da amamentação e do puerpério, obtendo assim a taxa de fecundidade total (TRF). Esses índices, exceto Ca, são calculados a partir de dados de pesquisas de base populacional, como a série de pesquisas denominadas DHS (sigla do inglês para *Demographic Health Surveys*), realizadas em todo o mundo, com razoável periodicidade (disponíveis em: <https://dhsprogram.com/>).

De acordo com Diniz & Medeiros (2010), as estimativas de aborto para o Brasil, em geral, se baseavam em técnicas que dependiam de informações de internação registradas, como, por exemplo, aquelas disponíveis no Sistema de Informações Hospitalares do SUS (SIH-SUS) e de fatores de correção. Em um contexto de pronunciada flutuação da fração dos procedimentos realizados em unidades privadas (*versus* públicas), dados exclusivamente oriundos do sistema público são habitualmente bastante imprecisos e sujeitos a viés.

A atual dinâmica do “mix” público-privado, com inegável contração dos gastos públicos e desassistência no âmbito do setor público, aliado à migração das famílias de planos privados mais caros (e abrangentes) para modalidades mais restritivas (mas menos onerosas), dentre outros fatores, impõe complicações adicionais à subenumeração estrutural de eventos exclusivamente registrados na esfera pública (SESTELO; SOUZA; BAHIA, 2013).

Em 2017, Ganatra e cols (2017a) utilizaram um modelo hierárquico bayesiano para estimar a proporção de abortos em níveis globais, regionais e sub-regionais segundo o nível de segurança da realização do aborto baseado nas definições da OMS, isto é, abortos foram classificados, por um painel de especialistas, como seguros (sob os cuidados de profissionais

⁵ $TRF = TF \times Cm \times Cc \times Ca \times Ci$

da saúde e utilizando métodos recomendado pela OMS) e inseguros, sendo que os inseguros foram subdivididos em aborto menos seguro ou “less-safe”, ou seja, que se valiam de métodos não recomendados ou utilizavam um método seguro (como por exemplo, misoprostol), mas não acompanhados de informações adequados ou suporte fornecido por um indivíduo qualificado, e o aborto dito de menor segurança/maior risco ou “least-safe”, que se valem de métodos/técnicas invasivas e perigosas, sem o suporte de indivíduos qualificados.

Inicialmente, Ganatra e cols (2017a) sistematizaram dados empíricos de 82 países e regiões listados pelo Departamento de Assuntos Econômicos e Sociais das Nações Unidas, através de pesquisas de saúde reprodutiva, de pesquisas demográficas, de bases de dados oficiais, como do Ministério da Saúde, e buscas manuais em bases de dados bibliográficas. Ao fim da busca ativa, utilizaram dados de 61 países.

No processo de modelagem proposto por Ganatra e cols (2017a) foram utilizadas as covariáveis referentes ao ambiente e à disponibilidade de métodos seguros onde o aborto foi realizado, acesso financeiro a serviços de abortamento seguro, legalidade do aborto no contexto em que fora realizado e desenvolvimento e infraestrutura dos serviços de saúde.

Ganatra e cols (2017a) estimaram a ocorrência de aproximadamente 56 milhões de abortos, por ano, em todo o mundo, no período de 2010 a 2014, dos quais 25 milhões eram considerados abortos inseguros. No mesmo período, estimaram aproximadamente 49 milhões de abortos, por ano, nos países subdesenvolvidos, dos quais 23,4 milhões eram considerados abortos inseguros.

2.2.2 Pesquisas à beira do leito

Pesquisas à beira do leito, com mulheres internadas por complicações do aborto, são aquelas nas quais as histórias de aborto são recuperadas por profissionais de saúde que atuam no serviço médico. Esse tipo de pesquisa, geralmente, investiga o perfil das mulheres que procuraram o serviço de saúde pós-aborto e seus dados clínicos, como descrito em Lucena (2000), que além de traçar o perfil de mulheres atendidas na maternidade do Centro Integrado de Saúde Amaury de Medeiros, Recife - Pernambuco, identificou a proporção de abortos induzidos e gestações não planejadas. Como os dados foram coletados em uma maternidade de Recife, não há como generalizar as estimativas para a população geral.

2.2.3 Métodos das questões indiretas

Outra técnica descrita na revisão executada por Rossier (2003) é o autorrelato de abortos, ou seja, estima-se a taxa de aborto perguntando às mulheres sobre a sua história reprodutiva, incluindo abortos (mas raramente, mencionando apenas a questão específica dos eventuais abortos). A pergunta pode ser feita usando enunciados indiretos, o que não deve ser confundido com métodos indiretos, pois aqui trata-se de alterar o enunciado, mas não a destinatária dos questionamentos, como, por exemplo, perguntar primeiro sobre gravidezes indesejadas da entrevistada ou perda da gravidez, como pode ser visto em Olinto & Moreira Filho (2004).

Alternativamente, técnicas de coleta de informações podem ser refinadas, como na utilização de formulários auto preenchíveis ou entrevistas com base em recursos audiovisuais, disponibilizados em computadores (ACASI e recursos afins).

Dessa forma, sem ser interrogada por entrevistadores, as mulheres tendem a se sentir mais confortáveis para responder sobre perguntas delicadas como a prática de aborto induzido.

2.2.4 Métodos diretos

A Técnica de Resposta ao Acaso (TRA), é um método de estimação direto, que foi empregada com o objetivo de estimar número de abortos ilegais em Vila Madalena, São Paulo, por Silva (1993). Tal técnica, proposta por Warner (1965), consiste em formular a pergunta de interesse com uma probabilidade pré-definida entre 0,5 e 1. Dessa forma, é concedida ao entrevistado a chance de responder a uma de duas questões formuladas sem que o entrevistador tenha o conhecimento de qual pergunta está sendo respondida. A questão a ser formulada, que não a de interesse, deve ser construída de uma forma que sua resposta não cause constrangimento ao respondente, e não esteja relacionada ao assunto de interesse.

Outro método direto de estimação de aborto e/ou mulheres que fizeram é a técnica de urna, também conhecida como *ballot box technique*, que tem como objetivo (tentar) assegurar, no momento da entrevista, o sigilo das respostas, a veracidade do relato de questões sensíveis/estigmatizantes (DINIZ; MEDEIROS, 2010) ou passíveis de criminalização (e, eventual, apenação), uma vez que os indivíduos, muitas vezes, não se sentem confortáveis/seguros em confirmar a prática de atitudes ou orientações julgadas como impróprias, imorais e/ou criminalizáveis, ou seja, o pertencimento a subpopulações como

usuários de drogas ilícitas, homens que fazem sexo com outros homens, mulheres que provocam aborto, entre outras.

No método da urna, as mulheres após responderem a um questionário geral, com questões socioeconômicas, demográficas e sobre saúde sexual e reprodutiva, recebiam um pequeno pedaço de papel (“voto”), contendo questões sobre aborto.

A Técnica da Urna foi utilizada em pesquisas nacionais (DINIZ; MEDEIROS, 2010; DINIZ; MEDEIROS; MADEIRO, 2017) e locais (OLINTO; MOREIRA-FILHO, 2004) para estimar o número de mulheres que fizeram aborto provocado pelo menos uma vez na vida e abortos induzido, respectivamente.

Olinto & Moreira Filho (2004) compararam o método da urna e o método das questões indiretas (isto é, em vez de questionar diretamente a mulher sobre a realização de um aborto induzido, calcula-se a diferença entre o número de gestações relatadas e o número de filhos tidos, exceto aqueles que faleceram após o nascimento) para estimar a frequência de aborto induzido em estudos de base populacional, com uma amostra representativa de 3.002 mulheres de 15 a 49 anos residentes na zona urbana da Cidade de Pelotas, RS, Brasil. No método da urna, além de um questionário geral, as mulheres após responderam questões relacionadas ao abortamento em um pequeno pedaço de papel (“voto”). No método das questões indiretas, as mulheres responderam além do questionário geral, às seguintes questões: você já esteve grávida alguma vez em que não podia ou não queria estar (com as opções de resposta: “sim”, “não”, “não sabe”, “sem informação”); se sim, o que você fez (“nada”, “continuou a gravidez”; “tentou interromper a gravidez, mas não conseguiu”; “tentou interromper a gravidez e conseguiu”; “outra opção”).

Na Pesquisa Nacional de Aborto (PNA), conduzida por Diniz & Medeiros (2010) e por Diniz, Medeiros e Madeiro (2017), que tinha como um dos objetivos estimar o número de mulheres que realizaram um aborto alguma vez na vida, as participantes responderam a dois questionários, um auto-preenchido e depositado na urna e a um segundo, aplicado por uma entrevistadora. Os autores relatam que foi possível parear os dois questionários de uma mesma participante, não sendo, porém, possível identificá-la, evidenciando o caráter anônimo dos respondentes da Pesquisa. O questionário preenchido pelas entrevistadoras contemplava questões socioeconômicas e demográficas, como referentes a escolaridade, ocupação laboral, religião, entre outras. O questionário da urna abordava questões sobre saúde sexual e reprodutiva, como “Você já fez aborto alguma vez?” e, em caso afirmativo, “Ficou internada

em função desse aborto?”. Além da pergunta direta sobre abortamento induzido, três questões adicionais enfatizavam a sondagem do aborto induzido.

A PNA de 2016 conduzida por Diniz, Medeiros e Madeiro (2017) comparou os seus resultados com os da PNA de 2010. Ambas as PNA, executadas em 2010 e 2016, valeram-se do método de um levantamento domiciliar que combinou técnica de urna e entrevistas face-a-face com mulheres alfabetizadas de 18 a 39 anos, exceto as residentes em áreas rurais (ou seja, a amostra corresponde ao que o IBGE habitualmente denomina “Brasil urbano”, *grosso modo*, correspondente a 76% da população total, segundo revisão recente da conceituação proposta e das respectivas estimativas)⁶.

2.2.5 Métodos indiretos

Além dos métodos citados nessa seção, existe o método indireto de estimação, conhecido como método *network scale-up* que vem sendo aplicado em contextos de Saúde Pública com o objetivo de estimar o tamanho de populações de difícil acesso (BERNARD et al., 1991), embora, até o presente momento, não tenha sido aplicado em estudos que tenham como objetivo estimar o número de abortos provocados (ou mulheres que fizeram aborto) no Brasil e grandes regiões. Tal método, pode ser considerado como método de estimação indireta, pois não pergunta ao entrevistado se ele participa de subpopulações de difícil acesso.

No entanto, embora o NSUM já venha sendo empregado na literatura científica internacional e eventualmente nacional, ainda se encontra em pleno desenvolvimento (BERNARD; MCCARTY, 2009). O NSUM tem sido basicamente aplicado em outros países para estimar o tamanho de diversas populações de difícil acesso nos mais diversos contextos.

O primeiro estudo baseado na metodologia *network scale-up* foi proposto por Bernard e cols (1991) após o terremoto de 1985 na Cidade do México e seu objetivo foi estimar o número de pessoas que morreram no terremoto valendo-se de informações indiretas da rede de contatos dos entrevistados. Nos Estados Unidos, o NSUM, vem sendo utilizado com diferentes objetivos, como, por exemplo, estimar a prevalência do HIV na população de rua (KILLWORTH et al., 1998a) ou estimar a prevalência de uso de heroína em 14 comunidades (KADUSHIN et al., 2006).

⁶ Disponível em: https://ww2.ibge.gov.br/home/geociencias/geografia/espacos_rurais_e_urbanos/default.shtm, acessado em 25/06/2018.

No Brasil, Salganik e cols (2011), estimaram o tamanho da subpopulação usuária de drogas ilícitas, que não a maconha, no município de Curitiba, Paraná. Em 2012, Bastos & Bertoni (2014), conduziram o maior inquérito domiciliar empregando a metodologia *network scale-up*, abrangendo todas as capitais brasileiras, com a finalidade de estimar o número de usuários de crack e/ou similares nas capitais brasileiras. E recentemente, o presente estudo, com a finalidade de estimar o número de mulheres que fizeram aborto nas 27 capitais brasileiras no período de 2011-2012.

Apenas um estudo, até o momento, realizado no Irã, aplicou o NSUM para estimar a taxa de aborto, com ou sem indicação médica (RASTEGARI et al., 2014). No Irã, a lei é bastante restritiva quanto à prática do aborto, pois o abortamento só é permitido antes dos 4 meses de gestação caso a vida da gestante estiver em risco ou em casos de anomalias fetais.

No caso do Irã, estimativas diretas se veem às voltas com dificuldades de grande monta, uma vez que a própria estrutura institucional e o marco legal não distinguem de forma minimamente clara o que é uma infração do ponto de vista de um possível código penal laico e a legislação de base religiosa, a sharia. Portanto, não é por acaso que a questão se reveste de uma gravidade e um risco substanciais, que inevitavelmente favorecem o ocultamento e o segredo, além de uma avaliação impossível, para o não especialista, da aplicação prática da sharia em cada contexto, uma vez que o Islamismo contempla inúmeras variantes, interpretações e seitas, com posturas marcadamente distintas em termos não apenas de países e sociedades nacionais, como grupos populacionais e correntes religiosas (SHAPIRO, 2014).

Levando em conta essas questões, Rastegari e cols (2014) discutiram a dificuldade de estimar o número de aborto induzidos no país, principalmente via métodos de estimação direta, e, ao constatar essas dificuldades, sugerem a utilização do NSUM. Essa dificuldade, relatada no estudo do Irã, é comum em todas as localidades cujas leis são restritivas e potencialmente criminalizantes quanto à prática do aborto, como no Brasil.

Segundo Rossier (2003), nenhum desses métodos pode ser considerado padrão-ouro, e até os dias atuais ainda não existe um consenso de qual método seria o mais indicado para estimar o número de abortos induzidos. Todavia, os diferentes métodos são mais ou menos adequados e aplicáveis a diferentes contextos sociopolíticos e servem a diferentes objetivos e questões de pesquisa, que podem enfatizar, por exemplo, a precisão das estimativas abrangentes ou eventos adversos específicos, como eventos adversos graves que resultam em óbitos.

2.3 ESTIMATIVAS DE ABORTO PROVOCADO NO BRASIL, REGIÕES E UF

Alguns autores, ao longo dos anos, valendo-se de diferentes técnicas de estimação de aborto, conseguiram traçar um panorama da situação do aborto provocado ou induzido, muitas vezes inseguro, no Brasil e regiões. Todos os estudos descritos nessa seção estão ilustrados no Quadro 1, onde é possível verificar dados referentes aos autores, ano da publicação, local e período de estudo, técnica de estimação empregada, além das respectivas estimativas, caso a caso.

Embora, vários métodos de estimação indireta de aborto tenham sido empregados na caracterização do aborto provocado no Brasil, as estimativas do número de aborto provocado diferem de acordo com a metodologia e a fonte de dados utilizada (OLINTO; MOREIRA-FILHO, 2004), como pode ser visto nessa seção.

O Instituto Alan Guttmacher, em 1991, estimou os níveis de aborto no Brasil, Colômbia e Peru, utilizando dados de internações hospitalares e oriundos dos levantamentos nacionais. No Brasil, foram utilizados dados do INAMPS (o então, Instituto Nacional de Assistência Médica e Previdência Social, extinto em 1993) e da Pesquisa Nacional Sobre Saúde Materno Infantil e Planejamento Familiar. Nesse estudo, conduzido pelo Instituto Alan Guttmacher, foram geradas estimativas com ajustes “fraco, moderado e forte”, conhecidos na literatura como fatores de correção ou multiplicadores, uma vez que os dados hospitalares não distinguem os abortos espontâneos dos induzidos, principalmente, por se tratar de um ato ilegal, no Brasil. E também foi aplicado um ajuste de “cobertura”, uma vez que assume-se que nem todas as mulheres que fizeram aborto vieram a ser hospitalizadas (SINGH; WULF, 1991).

Singh & Wulf (1991) estimaram os níveis de aborto no Brasil em 1980, baseados nos dados do INAMPS, utilizando os dois tipos de ajustes mencionados anteriormente, aquele relativo à distinção entre abortos espontâneos e induzidos (ajustes “fraco, moderado e forte”) e o de cobertura (assumindo que o INAMPS receberia, à época, 70% de todos os casos de hospitalização por aborto, do país). Baseado nessa fonte de dados e nos ajustes selecionados, estimou-se que o número de abortos induzidos no Brasil, em 1980, variou de 581.280 a 1.356.320.

Singh & Wulf (1991) também estimaram os níveis de aborto induzido em 1985, baseados nos dados da Pesquisa Nacional Sobre Saúde Materno Infantil e Planejamento Familiar, considerando os ajustes “fraco, moderado e forte”, e que 25% das hospitalizações por casos de aborto não são registradas enquanto tal e, portanto, são subnotificadas. Com isso, a

estimativa nacional do número de abortos induzidos, em 1985, variou de 838.878 a 1.957.385. Caso, considerasse que no Brasil a subnotificação era de 50%, a estimativa variaria de 1.411.878 a 3.294.385.

Singh, Henshaw e Haas (1999) apresentaram a estimativa da incidência do aborto induzido, no ano de 1991, segundo dados de internações por complicações decorrentes do abortamento, ao nível mundial. A estimativa para o número de abortos induzidos no Brasil, em 1991, foi de 1.444.000, com uma margem de erro de 1.021.000–2.021.000.

Em 2006, Monteiro & Adesse (2006) publicaram um estudo cujo objetivo foi estimar o número de abortos induzidos no Brasil e nas grandes regiões, por ano, no período de 1992 a 2005. Os autores utilizaram a metodologia de estimação indireta baseada nas internações hospitalares, por abortamento, registradas no SIH-SUS. Esse método se vale de fatores de correção que consideram a proporção de abortos espontâneos (25%), subnotificação (12,5%) e a proporção das mulheres que induziram o aborto e tiveram de ser hospitalizadas por complicações (20%).

Segundo Monteiro & Adesse (2006), a estimativa referente a abortos induzidos no Brasil, em 1992, foi de 1.455.283 ocorrências. Este resultado encontra-se dentro do intervalo das estimativas apresentadas por Singh, Henshaw e Haas (1999) para o ano de 1991.

Ainda em Monteiro & Adesse (2006), é possível verificar que há um declínio no número de abortamentos induzidos até 1998, cujo número estimado de abortos foi de aproximadamente 969.000. De 1998 até 2005, parece haver uma estabilização do número de abortos no Brasil, como sugere a estimativa baseada nos números do SIH-SUS, resultando numa estimativa do número de abortos induzidos no Brasil, em 2005, de aproximadamente 1.054.000.

Diniz & Medeiros (2010), em 2010, publicaram os primeiros resultados da Pesquisa Nacional de Aborto. A Pesquisa Nacional de Aborto foi um inquérito domiciliar nacional referente ao Brasil urbano (ou seja, excluindo a zona rural) que selecionou aleatoriamente mulheres alfabetizadas de 18 a 39 anos, e aplicou questionários preenchidos por entrevistadoras, e, além disso, empregou a técnica de estimação indireta, conhecida como técnica da urna, para estimar o número de abortos induzidos no Brasil.

Segundo a Pesquisa Nacional de Aborto (2010), realizada em 2010, no Brasil urbano, 15% das mulheres alfabetizadas, em idade reprodutiva, entrevistadas relataram ter feito aborto pelo menos uma vez na vida. Como uma mulher pode ter feito mais de um aborto ao longo da

vida, os autores relatam que esse número é menor do que o número de abortos efetivamente realizados. Outro ponto importante, que os autores destacam, é que essa estimativa não considerou a população de mulheres que moram em áreas rurais, mulheres analfabetas, além de mulheres com idade inferior a 18 anos e superior a 39 anos.

Além da estimativa do número de mulheres que fizeram aborto alguma vez na vida em 2010, Diniz & Medeiros (2010), descreveram o perfil dessas mulheres. A prevalência de aborto praticamente não difere segundo a religião das mulheres entrevistadas, no entanto, difere segundo a faixa etária e escolaridade. Mulheres em idades mais avançadas relatam com maior frequência a prática do aborto em relação a mulheres de faixas etárias mais jovens, talvez em função do acúmulo de episódios ao longo da trajetória de vida, embora efeitos associados ao pertencimento a diferentes coortes de nascimento (e suas respectivas escalas de valores e contextos de socialização e de vida) não possam ser descartados. Mulheres com menor escolaridade apresentam uma maior prevalência. Não pode ser descartada aqui a habitual inter-relação entre menor escolaridade, inserção em ocupações mal remuneradas, residência em locais com infraestrutura precária etc., enfim, a conjunção de fatores sociodemográficos e econômicos que aumentam a vulnerabilidade dos indivíduos, especialmente em se tratando de mulheres, em que se somam as consequências adversas das desigualdades de gênero.

Ainda segundo Diniz & Medeiros (2010), dentre as mulheres que admitiram ter feito aborto alguma vez na vida, 55% ficou internada por complicações decorrentes do abortamento. Lançando mão desses achados, pode-se atualizar os fatores de ajuste por internação hospitalar por complicações decorrentes do aborto usados em pesquisas nacionais com métodos de estimação indireta segundo dados de internações hospitalares. No entanto, estudos nacionais adicionais, que se valham de técnicas de estimação direta e indireta, fazem-se necessários para que esse fator de correção retrate a realidade brasileira de forma mais abrangente e fidedigna possível.

No estudo de Martins-Melo e cols (2014) foram analisados os padrões temporais e espaciais do aborto inseguro por unidades da federação, grandes regiões e Brasil, em 1996-2012, baseados nos registros das internações hospitalares de mulheres por abortamento registradas no SIH-SUS e utilização dos fatores de correção propostos pelo Instituto Alan Guttmacher (1994). Foram estimados 16.905.911 abortos inseguros no Brasil, no período de 1996 a 2012, sendo a média anual de abortos inseguros de 994.465.

Ainda segundo Martins-Melo e cols (2014), haveria um declínio no número de abortamentos inseguros ao longo do período sob estudo no país, além de um padrão de concentração espacial, com a presença de conglomerados (*clusters*) nas regiões Norte, Nordeste e Sudeste.

Monteiro, Adesse e Drezett (2015) publicaram a atualização das estimativas da magnitude do aborto induzido, para Brasil e grandes regiões, por ano, para o período de 1995 a 2013. As estimativas dos limites inferior e superior do número de abortos induzidos, entre mulheres de 15 a 49 anos, se basearam nos casos de internação por aborto registradas no SIH-SUS e lançaram mão dos referidos fatores de correção. Os autores assumiram que há um sub-registro de 12,5% e a proporção de abortos espontâneos é de 25%. Dessa forma, tem-se que em 1998 o número de abortos induzidos variou de 752.703 a 946.810, em 2005 variou de 835.849 a 1.050.764, enquanto em 2011 variou de 704.754 a 886.912. Pode-se notar que há um declínio de 1994 a 2013 na taxa nacional de aborto induzido, por 1.000 mulheres, com idades entre 15 e 49 anos.

Diniz, Medeiros e Madeiro (2017) coordenaram novamente a Pesquisa Nacional de Aborto, e compararam os achados mais recentes (2016) com os resultados obtidos em 2010 (DINIZ; MEDEIROS, 2010). Das 2.002 mulheres alfabetizadas, com idades entre 18 e 39 anos, residentes na área urbana, 251 (13%) afirmaram já ter feito ao menos um aborto na vida. Segundo Diniz et al. (2017) não há uma diferença, estatisticamente significativa, entre a proporção estimada pela PNA 2010 (15%) e pela PNA 2016 (13%).

Assim como na PNA 2010 (DINIZ; MEDEIROS, 2010), Diniz, Medeiros e Madeiro (2017) traçam o perfil das mulheres alfabetizadas, com idades entre 18 e 39 anos, residentes na área urbana dos municípios do Brasil, que relataram ter feito pelo menos um aborto ao longo da vida.

Segundo os dados da PNA 2016 (DINIZ; MEDEIROS; MADEIRO, 2017), foi possível verificar que a proporção de mulheres que fizeram aborto pelo menos uma vez na vida era maior entre mulheres nas regiões Norte/Centro-Oeste e Nordeste (19% e 20%) do que nas regiões Sudeste e Sul (12% e 9%). Corroborando os achados da PNA 2010, pois a proporção de mulheres que fizeram aborto pelo menos uma vez na vida era maior entre mulheres nas regiões Norte/Centro-Oeste e Nordeste (15% e 18%) do que nas regiões Sudeste e Sul (11% e 6%).

Quadro 1 - Descrição dos estudos cujo objetivo era a estimação do número de abortos induzidos (ou número de mulheres que induziram aborto) no Brasil, regiões e UF, suas técnicas utilizadas e estimativas encontradas

Título e autor	Técnica de estimação indireta utilizada	Objetivo (principal ou específico)	Local	Período	Estimativas
Estimating Abortion Levels in Brazil, Colombia and Peru, Using Hospital Admissions and Fertility Survey Data (Singh & Wulf, 1991)	Dados do INAMPS e da Pesquisa Nacional Sobre Saúde Materno Infantil e Planejamento Familiar	Estimar os níveis de aborto induzido.	Brasil	1980 e 1985	Brasil (1980) = (581.280; 1.356.320) Brasil (1985 – subnotificação de 25%) = (838.878; 1.957.385) Brasil (1985 – subnotificação de 50%) = (1.411.878; 3.294.385)
O uso da técnica de resposta ao Azar (TRA) na caracterização do aborto ilegal (Silva, 1993)	Técnica de resposta ao Azar e estimativa direta	Comparar o número de abortos provocados (mulheres entre 15 e 49 anos completos) segundo estimativa direta e TRA	Subdistrito de Vila Madalena, São Paulo	1987	Estimativa direta = 8 abortos provocados para cada 1000 mulheres em idade fértil (15 a 49 anos) Estimativa TRA = 41 abortos provocados para cada 1000 mulheres em idade fértil (15 a 49 anos)
La Incidencia del Aborto Inducido a Nivel Mundial (Singh, Henshaw e Haas, 1999)	Dados de internações hospitalares, por abortamento, registradas no SIH-SUS + fatores de correção (Instituto Guttmacher, 1994)	Estimar o número de abortos induzidos.	Brasil	1991	1.444.000 (1.021.000; 2.021.000)
Características de mulheres internadas por aborto em uma maternidade pública em Recife - PE: dimensão do problema e sua relação com a prática contraceptiva (Lucena, 2000)	Pesquisa a beira do leito.	Identificar a proporção de abortos induzidos.	1 maternidade pública de Recife - PE	Julho de 1999 a Janeiro de 2000	Dos 203 casos de aborto, 49 (24,1%) corresponderam a casos de aborto induzido.
Estimativa de aborto induzido: comparação entre duas metodologias (Olinto & Moreira Filho, 2004)	Método da urna e o método das questões indiretas	Estimar o número de mulheres (de 15 a 49 anos residentes na zona urbana da cidade de Pelotas, RS) que já fizeram aborto	Zona urbana da cidade de Pelotas, RS, Brasil.	1995	Método da urna = 7,2% das mulheres entrevistadas relataram ter realizado aborto induzido alguma vez na vida.

		em algum momento da vida.			Método das questões indiretas = 3,8% das mulheres entrevistadas relataram ter realizado aborto induzido alguma vez na vida.
Estimativas de aborto induzido no Brasil e Grandes Regiões (1992-2005) (Monterio & Adesse, 2006)	Dados de internações hospitalares, por abortamento, registradas no SIH-SUS + fatores de correção (Instituto Guttmacher, 1994)	Estimar o número de abortos induzidos e taxas anuais de abortos induzidos por 100 mulheres de 15 a 49 anos	Brasil e grandes regiões	1992-2005 (estimativas por ano)	Estimativas do número de abortos induzidos Brasil (1992) = 1.455.283 Brasil (2005) = 1.054.242 Estimativa das taxas anuais de abortos induzidos por 100 mulheres de 15 a 49 anos Brasil (1992) = 3,69 Brasil (2005) = 2,07 Norte (1992) = 3,55 Norte (2005) = 2,81 Nordeste (1992) = 5,41 Nordeste (2005) = 2,73 Sul (1992) = 1,97 Sul (2005) = 1,28 Sudeste (1992) = 3,39 Sudeste (2005) = 1,82 Centro-Oeste (1992) = 2,78 Centro-Oeste (2005) = 2,01
Aborto no Brasil: uma pesquisa domiciliar com técnica de urna (Diniz & Medeiros, 2010)	Técnica de urna e questionários preenchidos por entrevistadoras	Estimar o número de mulheres (alfabetizadas com idades entre 18 e 39 anos residentes de áreas urbanas) que já fizeram aborto em algum momento da vida.	Brasil “urbano”	2010	15% das mulheres (alfabetizadas com idades entre 18 e 39 anos residentes de áreas urbanas) entrevistadas relataram ter realizado aborto alguma vez na vida.
Tendência temporal e distribuição espacial do aborto inseguro no Brasil, 1996-2012 (Martins-Melo et al., 2014)	Dados de internações hospitalares, por abortamento, registradas no SIH-SUS + fatores de correção (Instituto Guttmacher, 1994)	Estimar o número de abortos inseguros	Brasil, regiões e unidade federativa	1996-2012 (estimativa do período)	Estimativa de abortos inseguros no período de 1996 a 2012: Brasil = 16.905.911 Norte = 1.537.443 Nordeste = 5.923.479 Sul = 1.611.191 Sudeste = 6.630.462

					Centro-Oeste = 1.110.405 Média anual de abortos inseguros: Brasil = 994.465
Atualização das estimativas da magnitude do aborto induzido, taxas por mil mulheres e razões por 100 nascimentos vivos do aborto induzido por faixa etária e grandes regiões. Brasil, 1995 a 2013 (Monteiro, Adesse e Drezett, 2015)	Dados de internações hospitalares, por abortamento, registradas no SIH-SUS + fatores de correção (Instituto Guttmacher, 1994)	Estimar o número de abortos induzidos entre mulheres de 15 a 49 anos	Brasil e grandes regiões	1995-2013 (estimativas por ano)	Estimativas do número de abortos induzidos Brasil (1995) = (864.628; 1.086.708) Brasil (2000) = (824.950; 1.037.125) Brasil (2010) = (735.148; 924.902) Brasil (2011) = (704.754; 886.912) Brasil (2012) = (693.364; 872.678) Brasil (2013) = (687.347; 865.160) Estimativa das taxas anuais de abortos induzidos por 100 mulheres de 15 a 49 anos Norte (2011) = (1,8; 2,2) Nordeste (2011) = (1,6; 2,0) Sul (2011) = (1,0; 1,3) Sudeste (2011) = (1,2; 1,5) Centro-Oeste (2011) = (1,2; 1,4)
Pesquisa Nacional de Aborto 2016 (Diniz, Medeiros e Madeiro, 2017)	Técnica de urna e questionários preenchidos por entrevistadoras	Estimar o número de mulheres (alfabetizadas com idades entre 18 e 39 anos residentes de áreas urbanas) que já fizeram aborto em algum momento da vida.	Brasil "urbano"	2016	13% das mulheres (alfabetizadas com idades entre 18 e 39 anos residentes de áreas urbanas) entrevistadas relataram ter realizado aborto alguma vez na vida.

3 JUSTIFICATIVA

Na década de 90, o Brasil era considerado um dos países com as estatísticas mais deficientes sobre a ocorrência do aborto provocado, em todo o mundo (SILVA, 1993) . Atualmente, a situação não é muito diferente, a despeito de estudos basicamente de abrangência local/regional. Estudos de abrangência nacional ainda são escassos na literatura.

As estatísticas disponíveis provêm, em grande parte, de levantamentos hospitalares, o que impossibilita generalizar as estimativas para a população brasileira e os dados do SIH/SUS, embora sejam de âmbito nacional, exigem o emprego de fatores de correção, pois estão subestimadas. Mesmo levando-se em consideração a hipótese de que no Brasil 20% das mulheres que induziram aborto vieram a ser hospitalizadas no pós-aborto, baseada na investigação do Instituto Alan Guttmacher (1994) , ainda não há um consenso quanto a esses fatores multiplicadores, assim como também não há como extrapolar, com segurança, um fator de correção formulado no início da década de 1990 para eventos que têm lugar, numa sociedade e num sistema de saúde bastante diversos, transcorridos quase 30 anos.

Já em 2010 foram divulgados resultados do inquérito intitulado “Pesquisa Nacional de Aborto” (DINIZ; MEDEIROS, 2010) e concluiu-se que 15% das mulheres brasileiras urbanas alfabetizadas com idade entre 18 e 39 anos já haviam feito aborto pelo menos uma vez na vida, e 55% dessas mulheres que fizeram aborto foram internadas por complicações relacionadas ao aborto. Embora, mulheres analfabetas ou residentes de área rural não tenham sido incluídas na Pesquisa Nacional de Aborto, há uma atualização da estimativa nacional da proporção de internação por complicação pós-aborto. No entanto, acredita-se que a estimação de abortos inseguros via o método de multiplicação baseado em dados de internação hospitalar não é a melhor opção, pois não há um consenso quanto a esses fatores multiplicadores, apresentados pelo Instituto Alan Guttmacher (1994), principalmente numa sociedade heterogênea e com uma estrutura de serviços de saúde, entre outros serviços, extremamente instável (compare-se, a título de exemplo, as oscilações brasileiras, em períodos curtos com as longas séries históricas de sistemas de saúde bastante estáveis, como o NHS [sigla do inglês para *National Health System*], no Reino Unido), transcorridos 24 anos.

A Pesquisa Nacional de Aborto, apresentada em 2010 por Diniz & Medeiros (2010) e a sua edição mais recente, divulgada por Diniz, Medeiros e Madeiro (2017), consistem em levantamentos por amostragem aleatória de domicílios. As estimativas apresentadas pela

Pesquisa Nacional de Aborto de 2010 e 2016 referem-se a abortos feitos em qualquer período da vida das mulheres brasileiras alfabetizadas entre 18-39 anos residentes em áreas urbanas.

Acredita-se que número total de abortos provocados no país é superior ao estimado pelas PNA de 2010 (DINIZ; MEDEIROS, 2010) e de 2016 (DINIZ; MEDEIROS; MADEIRO, 2017), dado que uma mesma mulher pode abortar mais de uma vez. Embora, a pesquisa seja de âmbito nacional (mais precisamente, referente ao Brasil urbano), uma limitação é que mulheres analfabetas, residentes em áreas rurais do Brasil, ou que estivessem internadas ou até mesmo privadas de liberdade, não foram contempladas no inquérito.

Diante desse quadro, o presente projeto de pesquisa, tem como objetivo aprofundar métodos e estimativas acerca do número de mulheres que fizeram um aborto (aborto provocado), com base em um inquérito domiciliar nas 27 capitais brasileiras, realizado entre 2011 e 2012, proveniente de uma ampla Pesquisa Nacional sobre o uso de crack, que, além da questão do consumo de crack em si (seu objeto principal), coletou uma série de informações adicionais, essenciais à aplicação da metodologia *network scale-up* (BASTOS; BERTONI, 2014) ■, lançando mão, em caráter inédito no Brasil, das informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população em geral para estimar o número de mulheres que fizeram aborto (abortos provocados) nas capitais brasileiras no ano anterior à avaliação.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Estimar o número de mulheres que fizeram abortos (abortos provocados), nos últimos 12 meses anteriores à avaliação, nas 27 capitais utilizando informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população em geral.

4.2 OBJETIVOS ESPECÍFICOS

- Descrever e propor modelos hierárquicos que consideram o desenho amostral e a estrutura demográfica da população, com a finalidade de estimar, utilizando-se de informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população em geral, o número de mulheres que fizeram abortos, nos últimos 12 meses anteriores à avaliação, nas 27 capitais;
- Autovalidar os modelos propostos estimando o tamanho de subpopulações conhecidas a partir de perguntas sobre as subpopulações cujos tamanhos são conhecidos para as 27 capitais;
- Comparar o desempenho dos diferentes modelos propostos para estimação usando métricas já estabelecidas na literatura estatística;
- Estimar o número de mulheres que fizeram aborto (aborto provocado), nos últimos 12 meses anteriores à avaliação, nas 27 capitais, usando a estratégia de modelagem proposta aplicada aos dados já coletados do Inquérito domiciliar executado na Pesquisa Nacional sobre o uso de crack, que, além da questão do consumo de crack (seu objeto principal), coletou uma série de informações adicionais, incluindo o aborto provocado.

5 MATERIAIS E MÉTODOS

Este capítulo, com a finalidade de apresentar os materiais e métodos abordados nesta tese, foi subdividido em cinco partes:

- i. a primeira apresenta a Pesquisa Nacional sobre o uso do Crack e o inquérito domiciliar realizado durante a Pesquisa Nacional sobre o uso do Crack, cujo objetivo era estimar o tamanho de diversas subpopulações de interesse (como a de mulheres que provocaram o aborto);
- ii. a segunda contempla o NSUM que vem sendo amplamente empregado em estudos internacionais, em diversos contextos (como, por exemplo, na Saúde Pública) e junto a diversas populações com a finalidade de estimar o tamanho de subpopulações de difícil acesso, por muitas vezes por serem estigmatizadas e/ou criminalizadas, e foi utilizada no inquérito domiciliar da Pesquisa Nacional sobre o uso do Crack;
- iii. a terceira é composta pelo modelo hierárquico pseudo bayesiano, de caráter inédito, que considera o desenho amostral e a estrutura etária da população alvo visando alcançar o objetivo principal deste trabalho, estimar o número de mulheres que fizeram aborto (aborto provocado), nas capitais brasileiras, no entanto pode ser aplicado em qualquer estudo cujo objetivo seja estimar o tamanho de qualquer subpopulação de difícil acesso segundo a metodologia *network scale-up*, além do processo de amostragem e detalhes computacionais;
- iv. a quarta apresenta o processo de autovalidação dos modelos hierárquicos propostos nessa seção através das 20 subpopulações de tamanhos conhecidos presentes no inquérito domiciliar da Pesquisa Nacional sobre o uso do Crack além da métrica de comparação dos modelos;
- v. a quinta é apresenta os aspectos éticos da pesquisa.

5.1 PESQUISA NACIONAL SOBRE O USO DO CRACK

A pesquisa “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” (BASTOS; BERTONI, 2014) , foi uma pesquisa ampla, que além da questão do consumo do crack e/ou similares, seu principal objetivo, coletou uma série de informações adicionais. Além da metodologia *Time Location Sampling* (TLS), que se utiliza da amostragem por conglomerados, baseado no local e no tempo, para acessar os usuários de crack e/ou

similares no Brasil, uma população estigmatizada e de difícil acesso, a pesquisa contemplou um inquérito domiciliar nas 27 capitais, realizado entre Abril e Novembro de 2012, com uma amostra representativa de pessoas com 18 anos ou mais, residentes nas capitais.

A Pesquisa Nacional “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” foi aprovada pelo Comitê de Ética da Escola Nacional de Saúde Pública Sérgio Arouca (ENSP/FIOCRUZ), sob o número CAAE 0073.0.031.000- 11 e suas atividades foram encerradas em 2012 (Anexo A).

A amostragem do inquérito compreendeu três estágios em cada uma das capitais e no DF. São elas: seleção dos setores censitários com probabilidade proporcional ao tamanho, seleção dos domicílios através de uma amostra aleatória sem reposição e seleção do morador do domicílio selecionado, maior de 18 anos, com a data de aniversário mais próxima (posterior) a data da entrevista. Mais detalhes da implementação e logística do Inquérito Domiciliar em Bastos & Bertoni (2014).

Ao morador, em cada domicílio selecionado, era emitido um comunicado sobre a Pesquisa e lido um termo de consentimento (Anexos B) de participação da pesquisa, e dentre algumas das informações contidas no termo, estavam os objetivos e procedimentos da pesquisa, além de informações sobre o sigilo dos dados. Dessa forma, não foi necessário que o entrevistado assinasse nenhum termo, pois o consentimento foi verbal, sugerindo, assim, ao entrevistado que suas respostas não seriam vinculadas ao seu nome (e dados pessoais, como endereço), e com isso, espera-se que as respostas não sejam influenciadas por um possível receio, constrangimento ou medo de ser julgado pelos entrevistadores, uma vez que o inquérito abordou questões delicadas, como aborto provocado, uso de drogas, etc.

Dessa forma, através do inquérito domiciliar nas capitais brasileiras, questões de interesse em Saúde Pública, como aborto provocado, mulheres ou homens que fazem sexo em troca de dinheiro, pessoas que usaram drogas ilícitas, que não a maconha, mais de 25 dias nos últimos 6 meses, entre outras, foram abordadas. No entanto, o presente trabalho, irá se ater, a subpopulação de interesse, mulheres que residem em cada uma das 27 capitais e fizeram um aborto nos últimos 12 meses anteriores à avaliação.

Como o inquérito domiciliar foi baseado na metodologia *network scale-up*, além das questões sobre subpopulações desconhecidas, cujo interesse é a estimação do tamanho de cada uma delas para cada uma das capitais brasileiras, foram feitas outras 20 perguntas do tipo “quantas pessoas residentes neste município você conhece da subpopulação k ?”. Subpopulações

essas cujos tamanhos são enumeráveis e conhecidos através de cadastros nacionais, como, por exemplo, número total de bancários e motoristas de ônibus, para cada uma das 27 capitais. A lista completa das 20 populações conhecidas e suas fontes pode ser vista no Quadro 2 (Anexo C).

O questionário aplicado no Inquérito Domiciliar está disponível nas páginas 201 a 222 do livro digital “Pesquisa Nacional sobre o uso de crack: quem são os usuários de crack e/ou similares do Brasil? quantos são nas capitais brasileiras?”⁷(BASTOS; BERTONI, 2014).

No inquérito domiciliar, baseado na metodologia *network scale-up*, proposto por Bastos & Bertoni (2014) , foi definido que “conhecer alguém” da subpopulação k , significa que são pessoas que moram no mesmo município que o entrevistado, que o entrevistado conhece de vista e de nome, que também o conhecem de vista e de nome, e com as quais o entrevistado entrou em contato, seja pessoalmente, por telefone, correspondência ou e-mail, nos últimos 12 meses, e fazem parte da subpopulação k .

Não houve necessidade de realizar novas entrevistas com os participantes da pesquisa “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” para construção desta tese de doutorado. Os dados para estimar o tamanho da subpopulação de mulheres que fizeram aborto (aborto provocado) no ano anterior à avaliação são secundários restritos oriundos da pesquisa já encerrada “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”.

5.1.1 Análise descritiva do inquérito domiciliar

Nesta seção é uma descrição do perfil dos 24.977 participantes da amostra geral, isto é, de todas as capitais brasileiras, do inquérito domiciliar pesquisa “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”, realizado em 2012, por Bastos & Bertoni (2014). Nesta análise descritiva não foram considerados os pesos amostrais utilizados na metodologia proposta NSUM do Inquérito domiciliar.

A Tabela 1 apresenta os quantitativos da amostra utilizada no inquérito domiciliar da Pesquisa Nacional sobre o uso do Crack, em 2012, para cada uma das capitais. Além do tamanho da amostra por unidade geográfica analisada, é possível verificar os valores absolutos

⁷ Questionário Scale-Up da Pesquisa Nacional sobre o uso de crack: quem são os usuários de crack e/ou similares do Brasil? quantos são nas capitais brasileiras está disponível em <https://www.icict.fiocruz.br/sites/www.icict.fiocruz.br/files/Pesquisa%20Nacional%20sobre%20o%20Uso%20de%20Crack.pdf>, acessado em 10/05/2017.

e relativos (%) da composição dessas amostras segundo sexo. Para cada capital foi informado o número de pessoas cuja variável sexo foi ignorada. Ainda na Tabela 1 é possível verificar a população residente em 2010, segundo o Censo demográfico (IBGE, 2011) , para as 27 capitais.

Em Palmas, pode-se observar que 1,16% dos dados referentes ao sexo dos participantes foram ignorados, sendo o maior percentual dentre as 27 capitais. Algumas capitais como Salvador, Curitiba, Boa Vista entre outras não apresentaram dados ignorados para variável sexo.

Da mesma forma que na amostra geral e na população residente do Brasil, a porcentagem de mulheres, por capital, é superior a porcentagem dos homens. A população residente do Brasil, em 2010, era composta por 97.348.809 (51,03%) mulheres e 93.406.990 (48,97%) homens, totalizando 190.755.799 residentes (IBGE, 2011).

A cidade do Rio de Janeiro, cuja população total em 2010 era de 6.320.446 pessoas, segundo censo demográfico de 2010 (IBGE, 2011), teve uma amostra de 1.547 participantes, selecionados de forma aleatória, segundo o plano amostral adotado no inquérito domiciliar. A amostra da cidade do Rio de Janeiro foi composta por 991 (63,53%) mulheres, 563 (36,09%) homens e 6 (0,38%) de respostas ignoradas para a variável sexo.

Tabela 1 Quantitativos da amostra utilizada no inquérito domiciliar da Pesquisa Nacional sobre crack segundo sexo nas 27 capitais brasileiras, 2012

Região	Estado	Cidade	População em 2010*	Amostra**			
				Total	Mulheres (%)	Homens (%)	Ignorado (%)
Nordeste	AL	Maceió	932.748	804	527 (65,55)	274 (34,08)	3 (0,37)
	BA	Salvador	2.675.656	1.006	605 (60,14)	401 (39,86)	0 (0,00)
	CE	Fortaleza	2.452.185	992	661 (66,63)	329 (33,17)	2 (0,20)
	MA	São Luís	1.014.837	803	456 (56,79)	341 (42,47)	6 (0,74)
	PB	João Pessoa	723.515	804	517 (64,30)	285 (35,44)	2 (0,26)
	PE	Recife	1.537.704	991	631 (63,67)	359 (36,23)	1 (0,10)
	PI	Teresina	814.230	804	500 (62,19)	303 (37,69)	1 (0,12)
	SE	Aracaju	571.149	804	537 (66,79)	261 (32,46)	6 (0,75)

Sul	RN	Natal	803.739	804	525 (65,30)	270 (33,59)	9 (1,11)
	PR	Curitiba	1.751.907	1.001	563 (56,24)	438 (43,76)	0 (0,00)
	SC	Florianópolis	421.240	804	465 (57,84)	335 (41,67)	4 (0,49)
	RS	Porto Alegre	1.409.351	1.005	586 (58,31)	413 (41,09)	6 (0,60)
Centro-Oeste	DF	Brasília	2.570.160	1.000	621 (62,10)	377 (37,70)	2 (0,20)
	GO	Goiânia	1.302.001	1.002	584 (58,28)	415 (41,42)	3 (0,30)
	MT	Cuiabá	551.098	804	505 (62,81)	298 (37,07)	1 (0,12)
	MS	Campo Grande	786.797	800	439 (54,88)	355 (44,38)	6 (0,74)
Sudeste	ES	Vitória	327.801	806	449 (55,71)	357 (44,29)	0 (0,00)
	MG	Belo Horizonte	2 375.151	1.005	623 (61,99)	377 (37,51)	5 (0,50)
	RJ	Rio de Janeiro	6 320.446	1.560	991 (63,53)	563 (36,09)	6 (0,38)
	SP	São Paulo	11.253.503	1.564	1051 (67,2)	511 (32,67)	2 (0,13)
	Norte	AM	Manaus	1.802.014	1.002	553 (55,19)	443 (44,21)
AP		Macapá	398.204	804	457 (56,84)	342 (42,54)	5 (0,62)
PA		Belém	1.393.399	997	594 (59,58)	398 (39,92)	5 (0,50)
RO		Porto Velho	428.527	803	416 (51,81)	387 (48,19)	0 (0,00)
RR		Boa Vista	284.313	804	483 (60,07)	321 (39,93)	0 (0,00)
AC		Rio Branco	336.038	803	524 (65,26)	279 (34,74)	0 (0,00)
TO		Palmas	228.332	601	366 (60,90)	228 (37,94)	8 (1,16)

*Fonte: Censo demográfico de 2010 (IBGE, 2011).

**Fonte: Inquérito Domiciliar (BASTOS; BERTONI, 2014).

A Tabela 2 apresenta a distribuição de frequência absoluta e relativa (%) da amostra geral do inquérito domiciliar segundo sexo, faixa etária, raça ou cor, estado civil e escolaridade, considerando os participantes selecionados nas 27 capitais, em 2012. Para cada uma das variáveis dessa Tabela, é possível visualizar a quantidade de dados ignorados.

Tabela 2 Perfil da amostra geral do inquérito domiciliar da Pesquisa Nacional sobre uso de Crack segundo sexo, faixa etária, raça ou cor, estado civil e escolaridade, 2012

Variáveis	Frequência	%
Sexo		
Feminino	15.229	60,97
Masculino	9.660	38,68
Ignorado	88	0,35
Faixa etária		
18 a 19	1.036	4,15
20 a 29	5.466	21,88
30 a 39	5.655	22,64
40 a 49	4.609	18,45
50 a 59	3.509	14,05
60 a 69	2.566	10,27
70 a 79	1.393	5,58
80 a 89	426	1,7
90 ou mais	43	0,17
Excluído	160	0,64
Ignorado	114	0,46
Raça ou cor		
Branca	11.043	44,21
Preta	3.082	12,34
Amarela	577	2,31
Parda	9.974	39,93
Indígena	162	0,65
Ignorado	139	0,56
Estado Civil		
Casado(a)	10.601	42,44
Desquitado(a) ou Separado(a) judicialmente	675	2,70

Divorciado(a)	1.361	5,45
Viúvo(a)	2.086	8,35
Solteiro(a)	10.171	40,72
Ignorado	83	0,33
Curso de nível mais elevado que frequentou (não necessariamente concluiu)		
Não frequentou nenhum curso	733	2,93
Creche, pré-escolar (maternal e jardim de infância) e classe de alfabetização - CA	87	0,35
Alfabetização de jovens e adultos	154	0,62
Antigo primário, ensino fundamental ou 1º grau (incluindo supletivo)	7.566	30,29
Antigo ginásio, ensino médio, 2º grau... (incluindo supletivo)	10.114	40,49
Superior de graduação ou especialização de nível superior (mínimo de 360 horas) ou mestrado ou doutorado	6.278	25,14
Ignorado	45	0,18

Fonte: Inquérito Domiciliar (BASTOS; BERTONI, 2014).

A amostra geral, formada pelos participantes aleatoriamente selecionados nas 27 capitais, totalizou 24.977 pessoas, sendo 60,97% do sexo feminino, como pode ser visto na Tabela 2. A idade média dos participantes do inquérito domiciliar foi de 42,52 anos com desvio padrão 16,54 anos. Além dos 114 (0,46%) participantes cujas idades não foram relatadas ou computadas (ignorado), foram excluídas 160 (0,64%) pessoas cujas idades, anotadas nos questionários ou digitadas no banco de dados, eram inferiores a 18 anos, já que um dos critérios de inclusão para participar do inquérito era ter 18 anos ou mais. A faixa etária com a maior frequência foi a de 30 a 39 anos com 5.655 (22,64%) participantes, seguida pela faixa etária de 20 a 29 anos com 5.466 participantes (21,88%). A faixa etária de 90 ou mais anos foi composta por 43 (0,17%) participantes, como pode ser visto na Tabela 2.

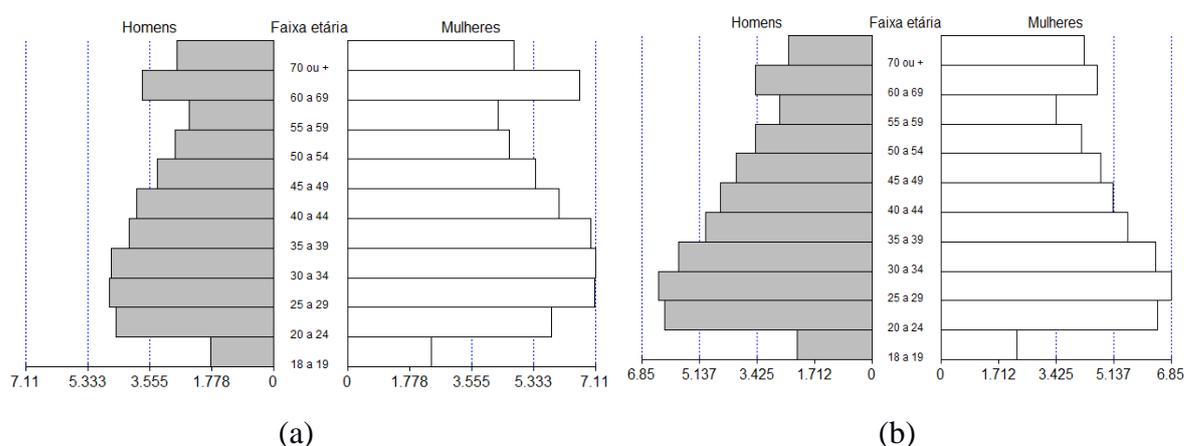
Entre os 24.977 entrevistados no inquérito domiciliar, 11.043 (44,57%) se autodeclararam da raça ou cor branca. A raça ou cor parda obteve a segunda maior frequência, sendo composta por 9.974 participantes (40,25%). Dos 24.977, 139 (0,56%) não responderam a pergunta “a sua raça ou cor é?”, dessa forma, considerou-se como ignorada. Quanto ao estado civil, 10.601 (42,44%) dos respondentes do inquérito domiciliar eram casados(as) e 10.171

solteiros(as) (40,72%). Do total da amostra, 83 (0,33%) respondentes não informaram sua situação conjugal, sendo consideradas como ignorada.

Os participantes foram questionados qual foi o curso de nível mais elevado que frequentou, mas não necessariamente concluiu. 733 pessoas (2,93%) não haviam frequentado nenhum curso. Aproximadamente 30% da amostra geral (7.566) tinham como curso mais elevado, embora não precisassem ter concluído o antigo primário, ensino fundamental ou 1º grau, incluindo supletivo. Dos 24.977 entrevistados, 10.114 (40,49%) tinham como curso mais elevado o antigo ginásio, ensino médio, 2º grau etc. (incluindo supletivo). E aproximadamente ¼ da amostra (6.278 pessoas) tinha o curso superior de graduação, especialização de nível superior (mínimo de 360 horas), mestrado ou doutorado, como curso mais elevado, não necessariamente concluído.

A distribuição da amostra geral do inquérito domiciliar da Pesquisa Nacional sobre uso do Crack por sexo segundo e faixa etária nas 27 capitais, em 2012, pode ser visualizada na Figura 2(a). Embora a distribuição da amostra geral do inquérito domiciliar da Pesquisa Nacional sobre uso do Crack tenha um comportamento semelhante ao da distribuição da população brasileira segundo sexo e faixa etária (acima de 18 anos de idade), como pode ser vista na Figura 2(b), a partir dos dados disponíveis no Censo de 2010 (IBGE, 2011), em termos percentuais, a composição da amostra geral (Figura 2(a)) apresenta um desbalanceamento quando comparada a composição da população brasileira (Figura 2(b)).

Figura 2 - Distribuição (%) da amostra do Inquérito domiciliar (a) e da População (Censo demográfico de 2010) (b) segundo sexo e faixa etária, Brasil



Ao comparar as pirâmides etárias da amostra geral (Figura 2a) e da população brasileira acima de 18 anos de idade (Figura 2b), pode-se notar que a porcentagem dos residentes no

Brasil em 2010, no grupo de 20 a 24 anos, é de 4,25% para o sexo feminino e 6,45 % para o sexo masculino, já na amostra geral é de 5,83% para o sexo feminino e 4,52% para o sexo masculino, nesse mesmo grupo de idade.

Já mulheres da faixa etária 30-34 anos representam aproximadamente 6,40% da população brasileira, em 2010, enquanto na amostra geral do inquérito, as mulheres nessa faixa etária representam aproximadamente 7,10%. Ao analisar o topo da pirâmide etária da Figura 2, pode-se verificar que a porcentagem de homens de 70 anos ou mais na amostra do Inquérito é de 2,77% e na população brasileira 2,47%, segundo Censo de 2010.

Ao analisar a distribuição da população e da amostra, por capital, segundo sexo e faixa etária, é possível verificar que também há um desbalanceamento (disponível no APÊNDICE), e tal fato, deve ser considerado no processo de estimação dos dados.

5.2 METODOLOGIA NETWORK SCALE-UP

Ainda que, as aplicações do NSUM sejam variadas desde seu uso inicial para estimar o número de vítimas de um terremoto no México (BERNARD et al., 1991) , o uso desta metodologia vem sendo amplamente empregado em questões de saúde pública, como nos Estados Unidos, no final da década de 90, para estimar a prevalência de HIV na população de rua em várias localidades, como pode ser vista em Killworth e cols (1998a) . Em 2006, Kadushin e cols (2006) estimaram, via NSUM, o número de usuários de heroína, utilizando informações indiretas da rede de contatos dos 5.892 participantes selecionados de 14 cidades dos Estados Unidos.

No Brasil, em 2011, foi publicado o primeiro estudo baseado na metodologia *network scale-up*, com a finalidade de estimar o tamanho da subpopulação usuária de drogas ilícitas, que não a maconha, no município de Curitiba, Paraná, como pode ser visto em Salganik e cols (2011) . Além, do emprego do NSUM, Salganik e cols (2011) utilizaram outras técnicas de estimação, como a estimação direta para comparar os resultados, e concluiu que as estimativas provenientes do NSUM (e suas generalizações) foram mais elevadas do que as obtidas a partir dos outros métodos utilizados.

Em 2014, Bertoni (2014) estimou o tamanho da população de usuários de crack e/ou similares nas capitais do Brasil, através de duas metodologias de estimação distintas, a direta e o NSUM. Um dos principais resultados do estudo sobre usuários de crack e/ou similares, desenvolvido por Bertoni (2014), é que as estimativas via o método indireto eram mais elevadas

que as estimativas via método direto, evidenciando, assim, que nos contextos de estigmatização e/ou ilegalidade, os resultados obtidos por métodos de estimação direta tendem a subestimar o tamanho de subpopulações de difícil acesso.

O NSUM estima o tamanho de subpopulações de interesse, ou população-alvo, utilizando-se de informações indiretas baseadas na rede social de cada participante do inquérito aleatoriamente selecionado da população em geral (BERNARD et al., 1991; KILLWORTH et al., 1998a, 1998b).

O planejamento de estudos baseados no NSUM requer que as perguntas feitas aos participantes da população geral, aleatoriamente selecionados, sejam do tipo “quantas pessoas você conhece da subpopulação k ?”. Algumas das subpopulações indagadas, devem ter seus tamanhos previamente conhecidos (sendo denotadas como “subpopulações conhecidas”), como, por exemplo, pessoas que residem no município do Rio de Janeiro e são estrangeiras ou pessoas que residem em determinada localidade e se chamam Natália (até recentemente estimado de forma aproximada de registros de óbitos, mas, desde o início de 2016 com precisão bastante maior, a partir da iniciativa do IBGE: <http://censo2010.ibge.gov.br/nomes/#/search>).

Além de questões sobre as subpopulações com os tamanhos previamente conhecidos, pergunta-se sobre subpopulação(ões) desconhecida(s), ou seja, com tamanho(s) desconhecido(s), e que se deseja estimar. Com isso, a vantagem de inquéritos baseados no NSUM, no contexto de uma única pesquisa, pode-se conhecer (estimar) o tamanho de mais de uma subpopulação de interesse.

Além disso, essa metodologia aborda questões da rede de contatos dos participantes selecionados, e não referentes ao próprio respondente, uma vez que se pergunta quantas pessoas o participante conhece de determinada subpopulação, e não se o mesmo integra essa subpopulação ou com ela compartilha determinados hábitos, práticas ou atitudes (como no exemplo das pesquisas acerca de percepção dos riscos decorrentes de ações terroristas). Com isso, espera-se diminuir a proporção de respostas “não verdadeiras”, pois sabe-se que as pessoas tendem a mentir, muitas vezes por vergonha ou medo, quando questionadas se pertencem a subpopulações estigmatizadas ou compartilham alguns dos seus hábitos, práticas ou atitudes.

Assim, além do entrevistado não precisar confirmar ser integrante de subpopulações estigmatizadas, como por exemplo, ser usuário de drogas ilícitas, ter realizado aborto provocado, ser profissional do sexo, entre outros, é possível através do NSUM obter informações sobre pessoas pertencentes a essas subpopulações de interesse que no momento da

pesquisa poderiam não estar disponíveis para seleção (internação hospitalar, privação de liberdade e óbito), uma vez que as informações são baseadas na rede de contatos do entrevistado. Já no método de estimação direta não seria possível acessá-los.

As informações sobre o número de pessoas conhecidas, por indivíduo selecionado aleatoriamente na população geral, para as subpopulações conhecidas, cujos tamanhos podem ser conhecidos por meio de cadastros nacionais, são utilizadas para estimar o “grau” (ou seja, sua estrutura de conexão) na rede de contato desses participantes (KILLWORTH et al., 2003; MCCORMICK; SALGANIK; ZHENG, 2010).

A estrutura de conexão deve ser estimada, uma vez que perguntar ao participante quantas pessoas ele “conhece”, ou seja, perguntar o tamanho de sua rede de contatos interpessoais, em dias atuais, pode-se obter resultados imprecisos pois há uma influência avassaladora das redes virtuais, com base na internet/telefonia, que tendem a se interpenetrar e eventualmente confundir redes reais de contatos com redes virtuais. Mais recentemente, os artigos em psicologia evolutiva têm tentado estabelecer distinções as mais claras possíveis (embora longe de um “ótimo”, em um mundo profundamente interconectado) entre redes reais de contatos interpessoais e das das redes virtuais .

No entanto, a metodologia NSUM ainda apresenta vários tipos de vieses (KILLWORTH et al., 1998a, 1998b) e os principais vieses serão descritos a seguir, com base nos trabalhos Killworth e cols (1998a, 1998b) e McCormick, Salganik e Zheng (2010).

O viés denominado efeito de barreira deve ser considerado uma vez que não se assume as diferentes probabilidades dos participantes conhecerem pessoas de diferentes grupos sociais, pois os participantes têm probabilidades distintas de conhecer pessoas de diferentes grupos sociais, uma vez que tendem a conhecer pessoas com características semelhantes às suas.

O viés de transmissão ocorre quando o participante não contabiliza um de seus contatos (“conhecidos”) como integrante de uma subpopulação abordada no estudo, por exemplo, o entrevistado pode não saber que um de seus contatos pertence a uma dada subpopulação ou terem determinados hábitos, práticas e atitudes, principalmente subpopulações estigmatizadas ou criminalizadas. Uma segunda possibilidade é simplesmente a de jamais ter conversado sobre um dado tema com o membro da sua rede, o que costuma ocorrer com episódios relevantes, não necessariamente estigmatizados, mas habitualmente restritos à esfera familiar, como os episódios de engasgamento de crianças.

O viés de memória refere-se à tendência das pessoas em subestimarem o número de contatos que elas conhecem, que estão inseridas em grupos maiores, pois se esquecem alguns destes “conhecidos”. Por isso, sugere-se utilizar subpopulações conhecidas com baixas prevalências na população geral. Por outro lado, as pessoas também tendem a superestimar o número de pessoas que elas conhecem em grupos pequenos.

A seguir será apresentada a notação da modelagem proposta nessa seção. O objetivo principal deste trabalho é conhecer/estimar o número de mulheres que fizeram aborto provocado nos últimos 12 meses anteriores à avaliação. A notação será para apenas uma capital, e todo o processo de inferência será aplicado para cada uma das capitais separadamente, uma vez, que se supõe independência entre os entrevistados selecionados aleatoriamente de duas capitais distintas. Assim como, será assumido que a rede de contato de um participante da capital A é independente da rede de contato de um participante da capital B ($\forall A \neq B$), para as 27 capitais.

Seja Y_i o número de mulheres que fizeram aborto (aborto provocado) nos últimos 12 meses anteriores a avaliação que o indivíduo i conhece. Assuma que a amostra é composta por n indivíduos, isto é, $i = 1, 2, \dots, n$, selecionados, via amostragem complexa, da população geral cujo tamanho é N .

Seja a variável auxiliar X_{ik} que representa o número de pessoas da subpopulação k , cujo tamanho é conhecido, que o indivíduo i conhece, para $k = 1, 2, \dots, K$. Como essas subpopulações têm seus tamanhos previamente conhecidos, tem-se que a prevalência da k -ésima subpopulação, π_k , também é conhecida, para $k = 1, 2, \dots, K$. Calcula-se $\pi_k = N_k / N$, onde N_k é o tamanho da subpopulação conhecida k e N o tamanho da população geral.

Como a metodologia é baseada na rede social de cada participante, será utilizada a informação do grau da rede de contatos do indivíduo i ($\hat{\delta}_i$), também chamado de tamanho da rede de contatos do indivíduo i , para cada indivíduo selecionado na amostra. O grau da rede social é o número total de contatos do indivíduo i , segundo a definição de “conhecer alguém” adotada na Pesquisa Nacional sobre o uso do Crack.

O estimador do tamanho da subpopulação de interesse, \hat{N}_Y , pelo método *Network Scale-up* é dado por uma função de $\hat{\theta}$, isto é,

$$\hat{N}_Y = \hat{\theta}N = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n \hat{\delta}_i} \times N \quad (1)$$

onde $\hat{\theta}$ é a probabilidade de pertencer a subpopulação de interesse dentre todos os membros da população geral, Y_i é o número de contatos da subpopulação de interesse que o indivíduo i conhece, $\hat{\delta}_i$ é a estimativa do grau da rede de contatos do indivíduo i e N é o tamanho da população geral (KILLWORTH et al., 1998a).

Os pressupostos do NSUM foram considerados na metodologia desta tese. Isto é, algumas suposições acerca dos indivíduos de uma mesma capital: (i) os indivíduos são independentes uns dos outros, isto é, $X_{ik} \perp X_{jk}$ e $Y_i \perp Y_j$, $\forall i \neq j$; (ii) O número de pessoas conhecidas do indivíduo i na subpopulação k é independente do número de pessoas conhecidas do indivíduo i na subpopulação l , isto é, $X_{ik} \perp X_{il}$, $\forall k \neq l$.

Optou-se por utilizar uma estimativa do grau da rede de contatos do indivíduo i , pois tal informação é imprecisa, isto é, geralmente, os indivíduos invés de quantificar todos os seus “conhecidos”, uma vez que esse número pode ser “alto”, gerando uma dificuldade em relatá-lo com exatidão, os indivíduos tendem a “chutar” o tamanho de suas redes, ora subestimando-o, ora superestimando-o. Caso, utilizasse o tamanho da rede de contatos relatadas pelos indivíduos na construção do estimador \hat{Y} , poderia gerar viés na estimativa do tamanho da subpopulação de interesse.

O estimador do grau da rede de contatos para o indivíduo i , pode ser escrito como

$$\hat{\delta}_i = \frac{\sum_k X_{ik}}{\sum_k N_k} \times N \quad (2)$$

onde a variável auxiliar X_{ik} é o número de contatos do indivíduo i da subpopulação conhecida e N_k é o tamanho da subpopulação k , cujo tamanho é conhecido, para $k = 1, 2, \dots, K$. Note que para estimar o grau da rede de contatos de cada indivíduo i selecionado no inquérito domiciliar da Pesquisa Nacional sobre uso de Crack e similares, serão utilizadas as informações indiretas da rede de contatos de cada um dos participantes, ou seja, quantas pessoas cada indivíduo i conhece das K subpopulações conhecidas, cujas prevalências na população geral são conhecidas e pequenas para evitar respostas falsas (conhecidas como “chute”).

Pode ser mostrado que um estimador para a variância do estimador NSUM é dado por

$$\widehat{V(\hat{N}_Y)} = \frac{\hat{\theta}(1 - \hat{\theta})}{\sum_{i=1}^n \hat{\delta}_i} \times N^2 \quad (3)$$

e os intervalos de confiança assintóticos de 95% podem ser obtidos da maneira usual, isto é,

$$\hat{N}_Y \pm 1,96 \sqrt{\widehat{V(\hat{N}_Y)}}.$$

5.3 MODELO HIERÁRQUICO

Nesta seção serão descritos os modelos hierárquicos, sob a ótica bayesiana, para estimar o número de mulheres que fizeram abortos nos últimos 12 meses anteriores à avaliação, baseado nas informações indiretas da rede de contatos dos indivíduos selecionados de forma aleatória, do inquérito domiciliar da Pesquisa Nacional sobre Crack, realizado de abril a novembro de 2012. No entanto, tal modelo pode ser aplicado para estimação do tamanho de quaisquer subpopulações de difícil acesso cuja metodologia da pesquisa seja o NSUM.

Maltiel e cols (2015) propuseram quatro modelos hierárquicos bayesianos para analisar dados com a metodologia *network scale-up*. Nesse trabalho vamos ajustar os modelos de grau aleatório.

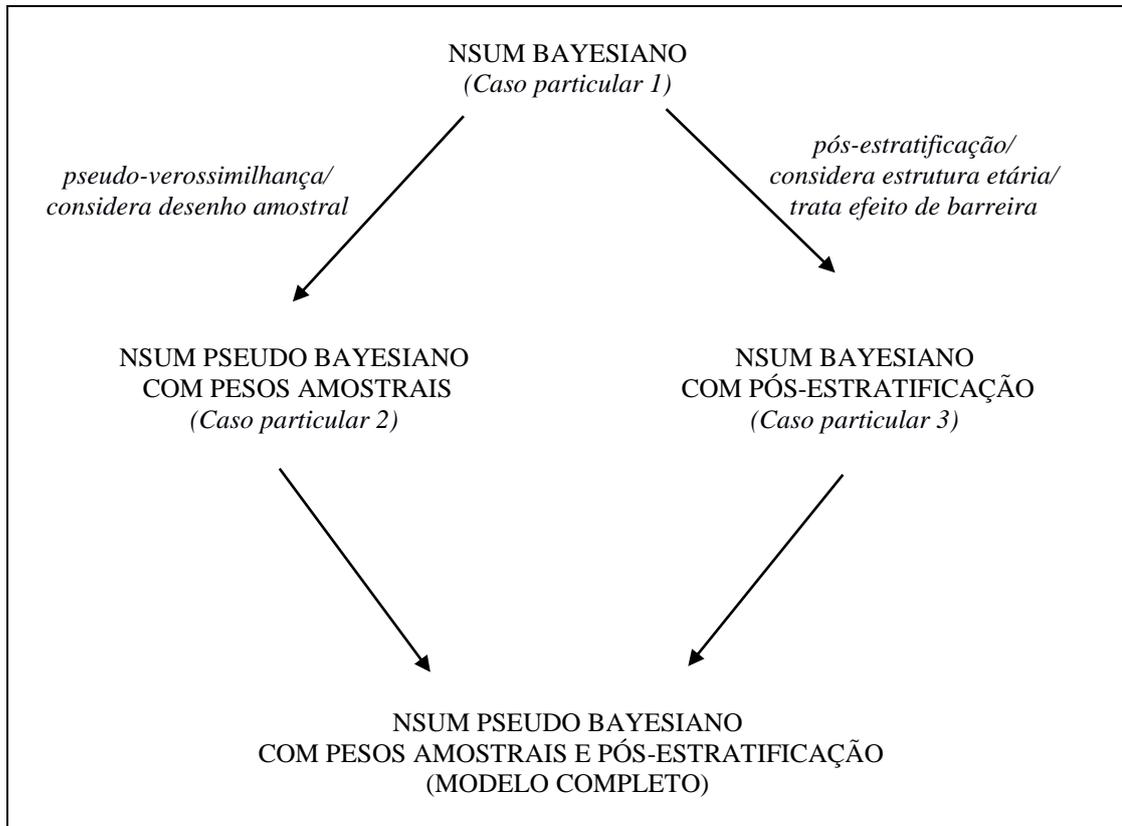
No entanto, os modelos propostos por Maltiel e cols (2015) não incluem o desenho amostral. Salganik e cols (2011) inclui os pesos amostrais em um modelo que também trata do viés de transmissão, mas não considera a pós-estratificação. O modelo hierárquico bayesiano com pesos amostrais e pós-estratificação (modelo completo), que considera o desenho amostral do Inquérito Domiciliar e a estrutura etária da população geral que os participantes foram selecionados, descrito nessa seção, com o objetivo de estimar o número de mulheres que fizeram aborto, é inédito na literatura científica.

Além do modelo completo, outros 3 modelos que são casos particulares também foram descritos nessa seção e podem ser vistos no Quadro 2.

Ainda nesta seção foram apresentadas informações sobre o processo de inferência e detalhes computacionais.

Os 4 modelos propostos nesta seção, foram apresentados e descritos para uma única capital, no entanto, estendem-se para todas as 27 capitais (além de qualquer outra localidade que se deseja estimar o tamanho de quaisquer subpopulações de difícil acesso). Dessa forma, cada modelo foi aplicado para os dados de cada uma das capitais, com objetivo de estimar o número de mulheres que fizeram aborto (aborto provocado) nos últimos 12 meses por capital.

Quadro 2 - Modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação (modelo completo) proposto nessa tese e seus casos particulares



5.3.1 Modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação

Uma alternativa para o estimador NSUM, dado na equação (1), é o uso do modelo hierárquico bayesiano, aqui chamado de NSUM bayesiano (caso particular do modelo hierárquico bayesiano com pesos amostrais e pós-estratificação – Modelo completo), para estimar o tamanho da população de difícil acesso, baseado na metodologia *network scale-up*. O NSUM bayesiano incorpora a incerteza da estimativa do grau dos participantes na estimação do tamanho da subpopulação alvo através do uso de distribuição a priori nas quantidades desconhecidas.

Todavia, houve uma necessidade de considerar o desenho amostral complexo do Inquérito Domiciliar que originou os dados da presente tese, isto é, os pesos amostrais dos participantes da Pesquisa devem ser incorporados no processo de inferência (na modelagem).

A inclusão de pesos amostrais em modelos estatísticos baseados na metodologia *network scale-up* ainda não é tão trivial. No entanto, alguns estudos recentes já vêm

incorporando o desenho amostral no NSUM, tais como em Kuniyama e cols (2016), Savitsky & Toth (2016) e Si, Pillai e Gelman (2015).

Como visto na Figura 2 (e nas figuras do ANEXO B) houve um desbalanceamento entre a distribuição da amostra e população geral das capitais, segundo sexo e faixas de idade. Então, para minimizar esse desbalanceamento e para incorporar a incerteza proveniente do efeito de barreira associado à metodologia NSUM, no modelo completo utilizou a pós-estratificação (segundo sexo e idade), pois além de corrigir o desbalanceamento, acredita-se que a chance de uma mulher jovem conhecer uma mulher que fez aborto (aborto provocado) nos últimos 12 meses anteriores à avaliação não é a mesma que a de um homem idoso, por exemplo.

Dessa forma, nessa seção será apresentado o modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação (modelo completo), e posteriormente seus casos particulares. A notação será dada para uma única capital, mas pode ser estendida para qualquer outra localidade sem perda de generalidade.

Notação

Para o processo de pós-estratificação, a amostra de tamanho n foi subdividida em estratos de faixa etária e sexo. Dessa forma, o modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação apresentado o nesta tese foi aplicado em cada um dos estratos.

A estimativa de interesse, isto é, o número de mulheres que fizeram aborto no local e período especificados pelo estudo (N_Y) foi obtida através de uma média ponderada entre as estimativas da incidência acumulada de mulheres que fizeram aborto segundo as informações indiretas baseadas na rede de contatos dos participantes pertencentes ao estrato j (θ_j) e o tamanho da população de cada estrato (N_j). Com isso, pode-se calcular a incidência acumulada de mulheres que fizeram um aborto (θ) através da proporção entre N_Y e N^* (tamanho da população de mulheres de 15 a 49 anos).

Cada estrato de faixa etária e sexo, da pirâmide etária da população geral, foi denotado por $j=1,2,\dots,J$, onde $J = 36$. Chame $j = 1$ o estrato composto por mulheres com 18 a 19 anos (pois menores de 18 anos não foram considerados no Inquérito Domiciliar), $j = 2$ o estrato composto por mulheres com 20 a 24 anos e assim por diante, até $j = 18$ o estrato composto por mulheres com 100 ou mais anos de idade. O estrato $j = 19$ é composto por homens com 18 a 19 anos, o estrato $j = 20$ é composto por homens com 20 a 24 anos e assim por diante, até $j = 36$ que é o estrato composto por homens com 100 ou mais anos de idade.

Seja Y_{ij} o número de mulheres que fizeram aborto (aborto provocado) num determinado período de tempo que o indivíduo i pertencente ao estrato j conhece. Assuma que a amostra do estrato j é composta por n_j indivíduos $\forall j$, isto é, $i = 1, 2, \dots, n_j$, selecionados, via amostragem complexa, da população geral cujo tamanho é N . Temos que a amostra geral (n) é dada pela soma da amostra dos estratos. Isto é, $n = n_1 + n_2 + \dots + n_J$, para $J=36$.

Nesta tese, foram consideradas $K = 20$ subpopulações de tamanhos previamente conhecidos, como motoristas de ônibus, bancários, estrangeiros etc e $J = 36$ estratos da população divididos segundo sexo e faixa etária.

Dessa maneira, pode-se escrever o estimador do número de mulheres que fizeram um aborto nos últimos 12 meses (\widehat{N}_Y) como

$$\widehat{N}_Y = \sum_{j=1}^J \theta_j \times N_j \quad (4)$$

onde N_j é o tamanho da população do estrato j , disponível no censo demográfico de 2010 (IBGE, 2011), e θ_j é a proporção da subpopulação de interesse estimada usando informação indireta dos contatos dos indivíduos selecionados no inquérito que pertencem ao estrato $j = 1, 2, \dots, J$.

Com isso, a partir de N_Y , pode-se escrever o estimador da incidência acumulada de mulheres que fizeram um aborto nos últimos 12 meses ($\hat{\theta}$) por 1.000 mulheres de 15 a 49 anos

$$\hat{\theta} = \frac{N_Y}{N^*} \times 1.000 \quad (5)$$

onde N^* é o tamanho da população de mulheres de 15 a 49 anos disponíveis no censo demográfico de 2010 (IBGE, 2011).

Seja a variável auxiliar X_{ikj} que representa o número de pessoas da subpopulação k que o indivíduo i pertencente ao estrato j conhece, para $k = 1, 2, \dots, K$. A prevalência da k -ésima subpopulação, π_k , é conhecida e pode ser escrita como uma fração do tamanho da subpopulação k (N_k) pelo tamanho da população geral (N). Isto é, $\pi_k = N_k / N$.

No NSUM as informações indiretas da rede de contato dos participantes da amostra são utilizadas para estimar do grau da rede de contatos do indivíduo i (δ_i), também chamado de tamanho da rede de contatos do indivíduo i . O grau da rede social é o número total de contatos do indivíduo, segundo a definição de “conhecer alguém” adotada na Pesquisa.

A inclusão do desenho amostral complexo foi dada através do peso amostral (w) que pode ser calculado como o inverso da probabilidade de inclusão.

Especificação do modelo completo

No processo de modelagem dos dados assumiu-se que a população geral não era homogênea, no entanto, nos estratos havia homogeneidade, e a amostra era probabilística.

O modelo completo considerou, além dos pesos amostrais, a composição da faixa etária segundo sexo da população de cada capital, numa tentativa de corrigir possíveis desbalanceamentos entre a distribuição segundo faixa etária e sexo da amostra de cada capital e da população para cada capital. Além disso, acredita-se que estratos da população tendem a ter chances distintas de conhecer indivíduos participantes da subpopulação de interesse.

O número de mulheres que fizeram um aborto nos últimos 12 meses anteriores a pesquisa que o indivíduo i pertencente ao estrato j da população conhece, segue uma distribuição binomial com parâmetros δ_i e θ_j . Ou seja,

$$Y_{ij} | \delta_i, \theta_j, w_i \sim \text{Binomial}(\delta_i, \theta_j), \quad (6)$$

para $i=1,2,\dots,n$ e $j=1,2,\dots,J$.

A variável auxiliar X_{ikj} , número de contatos que o indivíduo i pertencente ao estrato j conhece da subpopulação de tamanho conhecido k , segue uma distribuição Binomial com parâmetros δ_i e π_k , para $i=1,2,\dots,n$; $k=1,2,\dots,K$ e $j=1,2,\dots,J$. Ou seja,

$$X_{ik} | \delta_i, \pi_k, w_i \sim \text{Binomial}(\delta_i, \pi_k). \quad (7)$$

Sabe-se que os participantes do Inquérito Domiciliar foram amostrados usando um desenho amostral complexo. Ou seja, cada indivíduo teve uma probabilidade, s_i , de ser incluído na amostra. O peso amostral é então dado pelo inverso da probabilidade de inclusão, $w_i = \frac{1}{s_i}$.

Nesta tese, o peso amostral foi incorporado como expoente na verossimilhança do modelo completo, resultando, assim, na pseudo verossimilhança. A utilização da pseudo verossimilhança tem como objetivo ponderar a contribuição cada indivíduo amostrado (SAVITSKY; TOTH, 2016). Dessa forma, a pseudo verossimilhança ($LP(\theta, \delta_1, \dots, \delta_n)$) pode ser escrita como

$$LP(\theta_j, \delta_1, \delta_2, \dots, \delta_{n_j}) = \prod_{i=1}^{n_j} \left[p(y_{ij} | \delta_i, \theta_j) \prod_k p(x_{ik} | \delta_i, \pi_k) \right]^{w_i} \quad (8)$$

onde $w_i = \frac{w_i}{\sum_i w_i} \times n_j$, garantindo que $\sum_i w_i = n_j$, $\pi_k = N_k / N$ é conhecido, $p(y_{ij} | \delta_i, \theta_j)$ é a função de probabilidade da variável aleatória $Y_{ij} | \delta_i, \theta_j, w_i$ dada na equação (6) e $p(x_{ik} | \delta_i, \pi_k)$ é a função de probabilidade da variável aleatória $X_{ik} | \delta_i, \pi_k, w_i$ dada na equação (7).

A pseudo verossimilhança não é a única forma de incluir os pesos amostrais em um modelo bayesiano. Por exemplo, Si, Pillai e Gelman (2015) utilizaram processos gaussianos que podem incorporar os pesos amostrais na modelagem e Kuniyama e cols (2016) utilizaram processos Dirichlet com o mesmo intuito.

Distribuições a priori

Dado o modelo completo, proposto nesta seção, temos que as quantidades $\{\theta_j, \delta_1, \delta_2, \dots, \delta_{n_j}\}$ são desconhecidas. Assumiu-se que a distribuição a priori de todas as quantidades desconhecidas eram independentes, isto é,

$$p(\theta_j, \delta_1, \delta_2, \dots, \delta_{n_j}) = p(\theta_j) \prod_{i=1}^{n_j} p(\delta_i) \quad (9)$$

Optou-se por atribuir distribuições a priori Beta não informativa para a proporção da subpopulação de interesse e Gama informativa para o grau da rede de contatos do indivíduo i . Com isso, pode-se escrever

$$\theta_j \sim \text{Beta}(a_\theta, b_\theta) \quad (10)$$

$$\delta_i \sim \text{Beta}(a_{\delta_i}, b_{\delta_i}), \text{ para } i = 1, 2, \dots, n_j, \quad (11)$$

onde as quantidades $a_\theta = b_\theta = 1$ foram escolhidas de forma que a distribuição fosse não informativa e $a_{\delta_i}, b_{\delta_i}$ foram baseados no número de Dunbar (HILL; DUNBAR, 2003).

O tamanho da rede de contatos de cada indivíduo, δ_i , é o número total de contatos ou amigos do indivíduo i . Acredita-se que há um número limitado de contatos/amigos que se pode estabelecer em um relacionamento social estável (HILL; DUNBAR, 2003), chamado número de Dunbar.

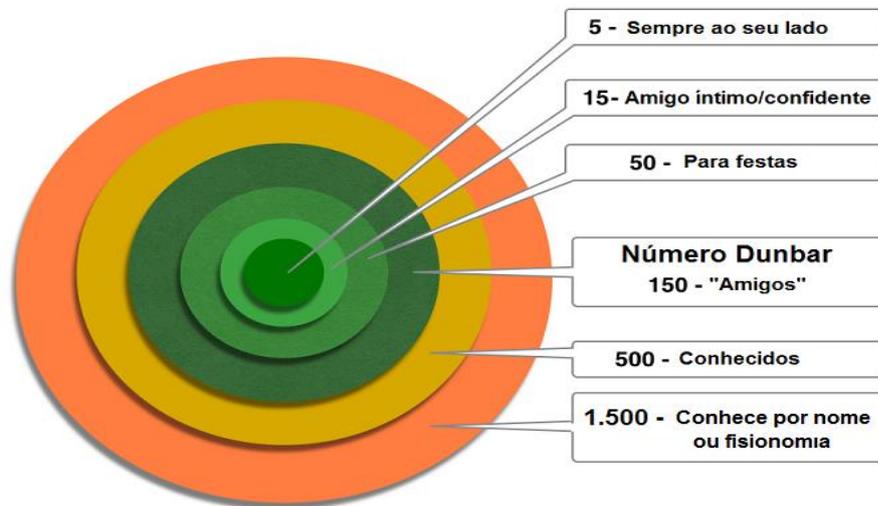
Segundo o antropologista britânico Robin Dunbar, 150 é o número médio de “amigos” uma pessoa pode ter. Além disto, sugere com quantos “amigos” uma pessoa pode ter uma relação de amizade, desde amigos íntimos, como amigos para festas e conhecidos, como pode

ser visto na Figura 3, que foi adaptada de “*How social networks distort reality: you will always appear below average in your group*”⁸.

Então, baseado no número de Dunbar, as quantidades a_{δ_i} e b_{δ_i} foram definidas tais que, para $i=1,2,\dots,n_j$, $\text{moda}(\delta_i) = 100$ e $P(\delta_i < 300) = 0,95$.

Dessa forma, para o grau da rede de contatos atribuiu-se a priori uma probabilidade baixa para valores muito grandes/altos para o número total de amigos de um indivíduo e concentra o grau da rede de contatos em torno de 100, para cada indivíduo.

Figura 3 - Quantos amigos uma pessoa pode ter?



5.3.2 Casos particulares

Os casos particulares do modelo completo, ou seja, o modelo hierárquico pseudo bayesiano com pesos amostrais e pós-estratificação, são os modelos (i) modelo hierárquico bayesiano (NSUM bayesiano), (ii) modelo hierárquico pseudo bayesiano com pesos amostrais (NSUM bayesiano com pesos amostrais) e (iii) modelo hierárquico bayesiano com pós-estratificação (NSUM bayesiano com pós-estratificação).

A descrição de cada um dos modelos particulares desta seção foi feita para uma única capital sem perda de generalidade.

⁸ Disponível em: https://iomedica.org/?page_id=2685&id=2753, acessado em 10/05/2016.

Caso particular 1 – Modelo hierárquico bayesiano (NSUM bayesiano)

O modelo hierárquico bayesiano é o caso particular mais simples. Pode ser considerado como uma versão bayesiana do estimador NSUM (clássico) dado na equação (1). Assumiu-se que a população é homogênea e foi realizada uma amostra aleatória simples (AAS) para a seleção dos participantes do inquérito.

A diferença entre o estimador clássico dado na equação (1) (chamado de NSUM) e o estimador bayesiano (aqui chamado de NSUM bayesiano) é que no contexto bayesiano o processo de inferência incorpora a variabilidade das quantidades desconhecidas através da distribuição a priori, além de considerar a variabilidade na estimação do grau da rede de contatos. Esse estimador bayesiano (chamado de “*random degree model*”) foi proposto em Killworth e cols (1998a, 1998b).

Isto é, nesse modelo não são considerados o desenho amostral (não inclui pesos amostrais, isto é, $w_i = 1, \forall i$) nem a composição da população, isto é, não há o processo de pós-estratificação da amostra). Isto é, a verossimilhança pode ser escrita como

$$L(\theta, \delta_1, \dots, \delta_n) = \prod_{i=1}^n \left[p(y_i | \delta_i, \theta) \prod_k p(x_{ik} | \delta_i, \pi_k) \right] \quad (12)$$

onde $p(y_i | \delta_i, \theta)$ é a função de probabilidade da variável aleatória $Y_i | \delta_i, \theta \sim \text{Binomial}(\delta_i, \theta)$ e $p(x_{ik} | \delta_i, \pi_k)$ é a função de probabilidade da variável aleatória $X_{ik} | \delta_i, \pi_k \sim \text{Binomial}(\delta_i, \pi_k)$.

A distribuição a priori das quantidades desconhecidas foram

$$\theta \sim \text{Beta}(a_\theta, b_\theta) \quad (13)$$

$$\delta_i \sim \text{Gama}(a_\delta, b_\delta), \text{ para } i = 1, 2, \dots, n, \quad (14)$$

onde as quantidades $a_\theta = b_\theta = 1$ foram escolhidas de forma que a distribuição fosse não informativa e a_δ, b_δ foram baseados no número de Dunbar (HILL; DUNBAR, 2003).

Note que a incidência acumulada de mulheres que fizeram aborto nos últimos 12 meses (θ) e o tamanho da rede de contato dos indivíduos (δ) são desconhecidos. E após estimar θ , pode-se obter o número de mulheres que fizeram aborto provocado nos últimos 12 meses (N_Y), pois

$$\theta = \frac{N_Y}{N^*},$$

onde N^* é o número de mulheres em idade reprodutiva (e neste estudo, considerou-se mulheres em idade reprodutiva àquelas mulheres entre 15-49 anos).

Caso particular 2 – Modelo hierárquico pseudo bayesiano com pesos amostrais (NSUM pseudo bayesiano com pesos amostrais)

O modelo hierárquico pseudo bayesiano com pesos amostrais é o caso particular do modelo completo, pois considera apenas do desenho amostral da pesquisa. Ou seja, não considera a composição da população (não há o processo de pós-estratificação).

Este modelo se enquadra em Inquéritos baseados na metodologia *network scale-up* (cujo objetivo é estimar populações de difícil acesso) tal que o desenho amostral é complexo.

No caso particular 2 considerou-se o desenho amostral, da mesma forma que no modelo completo, mas não o processo de pós-estratificação da amostra).

Os pesos amostrais foram incluídos na pseudo verossimilhança com o intuito de ponderar a contribuição cada indivíduo amostrado (SAVITSKY; TOTH, 2016). Dessa forma, a pseudo verossimilhança ($LP(\theta, \delta_1, \dots, \delta_n)$) pode ser escrita como

$$LP(\theta, \delta_1, \dots, \delta_n) = \prod_{i=1}^n \left[p(y_i | \delta_i, \theta) \prod_k p(x_{ik} | \delta_i, \pi_k) \right]^{w_i} \quad (15)$$

onde w_i é o peso amostral, $p(y_i | \delta_i, \theta)$ é a função de probabilidade da variável aleatória $Y_i | \delta_i, \theta \sim \text{Binomial}(\delta_i, \theta)$ e $p(x_{ik} | \delta_i, \pi_k)$ é a função de probabilidade da variável aleatória $X_{ik} | \delta_i, \pi_k \sim \text{Binomial}(\delta_i, \pi_k)$.

A distribuição a priori das quantidades desconhecidas foram

$$\theta \sim \text{Beta}(a_\theta, b_\theta) \quad (16)$$

$$\delta_i \sim \text{Gama}(a_\delta, b_\delta), \text{ para } i = 1, 2, \dots, n, \quad (17)$$

onde as quantidades $a_\theta = b_\theta = 1$ foram escolhidas de forma que a distribuição fosse não informativa e a_δ, b_δ foram baseados no número de Dunbar (HILL; DUNBAR, 2003).

A incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) pode ser escrita como

$$\theta = \frac{N_Y}{N^*},$$

onde N^* é o número de mulheres em idade reprodutiva (e neste estudo, considerou-se mulheres em idade reprodutiva àquelas mulheres entre 15-49 anos).

Caso particular 3 - Modelo hierárquico bayesiano com pós-estratificação (NSUM bayesiano com pós-estratificação)

O modelo hierárquico bayesiano com pós-estratificação é o caso particular do modelo completo pois não considera o desenho amostral (não inclui pesos amostrais, isto é, $w_i = 1, \forall i$), no entanto, considera a composição da população no processo de inferência.

Ou seja, é como se para cada estrato (de sexo e idade) da amostra aplicasse um NSUM bayesiano (caso particular 1).

A função de verossimilhança pode ser escrita como:

$$L(\theta_j, \delta_1, \delta_2, \dots, \delta_{n_j}) = \prod_{i=1}^{n_j} \left[p(y_{ij} | \delta_i, \theta_j) \prod_k p(x_{ik} | \delta_i, \pi_k) \right] \quad (18)$$

onde $p(y_{ij} | \delta_i, \theta_j)$ é a função de probabilidade da variável aleatória $Y_{ij} | \delta_i, \theta_j \sim \text{Binomial}(\delta_i, \theta_j)$ e $p(x_{ik} | \delta_i, \pi_k)$ é a função de probabilidade da variável aleatória $X_{ik} | \delta_i, \pi_k \sim \text{Binomial}(\delta_i, \pi_k)$.

A distribuição a priori das quantidades desconhecidas foram

$$\theta_j \sim \text{Beta}(a_\theta, b_\theta) \quad (19)$$

$$\delta_i \sim \text{Gama}(a_\delta, b_\delta), \text{ para } i = 1, 2, \dots, n_j, \quad (20)$$

onde as quantidades $a_\theta = b_\theta = 1$ foram escolhidas de forma que a distribuição fosse não informativa e a_δ, b_δ foram baseados no número de Dunbar (HILL; DUNBAR, 2003).

O número de mulheres que fizeram aborto e a incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) podem ser escritos através das equações (4) e (5), respectivamente.

5.3.3 Inferência e detalhes computacionais

A inferência para os parâmetros desconhecidos do modelo completo proposto nesta tese foi realizada sob o paradigma bayesiano. Toda a inferência foi baseada na distribuição pseudo posteriori das quantidades desconhecidas, $\theta_j, \delta_1, \dots, \delta_n$.

A inferência para os parâmetros desconhecidos dos modelos casos particulares do modelo completo, NSUM bayesiano, NSUM bayesiano com pesos amostrais e NSUM bayesiano com pós-estratificação, foi realizada sob o paradigma bayesiano. Toda a inferência

foi baseada na distribuição posteriori (NSUM bayesiano e NSUM bayesiano com pós-estratificação) e pseudo posteriori (NSUM bayesiano com pesos amostrais) das quantidades desconhecidas.

Amostras da distribuição pseudo posteriori dos modelos foram obtidas através de métodos computacionais intensivos, como MCMC (sigla do inglês para *Markov chain Monte Carlo methods*) (GAMERMAN; LOPES, 2006). As rotinas foram implementadas no software livre R (R DEVELOPMENT CORE TEAM, 2018).

Pseudo posteriori e condicional completa

O processo de inferência do modelo completo foi baseado na distribuição conjunta a pseudo posteriori de todos os parâmetros, definida por Savitsky e Toth (2016), que pode ser obtida via teorema de Bayes, ou seja,

$$p\left(\theta_j, \delta_1, \dots, \delta_{n_j} \mid \underline{y}, \underline{x}, \underline{w}\right) \propto p(\theta_j) \prod_{i=1}^{n_j} p(\delta_i) \prod_{i=1}^{n_j} \left\{LP\left(\theta_j, \delta_1, \dots, \delta_{n_j}\right)\right\} \quad (21)$$

onde $j=1,2,\dots,J$, $\underline{y} = (y_1, y_2, \dots, y_{n_j})$, $\underline{x} = (x_{11}, x_{21}, \dots, x_{n_1}, \dots, x_{1K}, x_{2K}, \dots, x_{nK})$, $\underline{w} = (w_1, w_2, \dots, w_{n_j})$ e $p(\theta_j)$ e $p(\delta_i)$ são as distribuições a priori das quantidades desconhecidas, dadas nas equações (10) e (11), respectivamente, e $LP\left(\theta_j, \delta_1, \dots, \delta_{n_j}\right)$ é a função de pseudo verossimilhança dada na equação (8) que sob independência condicional foi dada pelo produto da função de probabilidade conjunta.

Dessa forma, a pseudo posteriori pode ser escrita por

$$p\left(\theta_j, \delta_1, \dots, \delta_{n_j} \mid \underline{y}, \underline{x}, \underline{w}\right) \propto p(\theta) \prod_{i=1}^{n_j} \left\{p(\delta_i) \left[p(y_{ij} \mid \delta_i, \theta_j) \prod_k p(x_{ik} \mid \delta_i, \pi_k) \right]^{w_i}\right\} \quad (22)$$

A distribuição condicional completa a posteriori de θ_j foi dada por

$$\theta_j \mid \delta_1, \dots, \delta_n, \underline{y}, \underline{w} \sim \text{Beta}(a_\theta^*, b_\theta^*) \quad (23)$$

onde $a_\theta^* = a_\theta + \sum_{i=1}^{n_j} w_i \times y_{ij}$ e $b_\theta^* = b_\theta + \sum_{i=1}^{n_j} w_i \times \delta_i - \sum_{i=1}^{n_j} w_i \times y_{ij}$.

Por simplificação, se o tamanho da rede for grande e a proporção da subpopulação de interesse for pequena, pode-se aproximar a distribuição de $Y_{ij} \mid \theta_j, \delta_i$ dada na equação (6) e $X_{ik} \mid \pi_k, \delta_i$ dada na equação (7) por uma distribuição de Poisson, isto é,

$$Y_{ij} \mid \theta_j, \delta_i \sim \text{Poisson}(\theta_j \delta_i), i = 1, 2, \dots, n_j \quad (24)$$

$$X_{ik} \mid \delta_i \sim \text{Poisson}(\pi_k \delta_i), i = 1, 2, \dots, n_j; k = 1, 2, \dots, K \quad (25)$$

Com isso, a distribuição condicional completa a posteriori de $\delta_i \mid \underline{y}, \underline{x}, \underline{w}$ é uma Gama Truncada com parâmetros $a_{\delta_i}^*, b_{\delta_i}^*$. Ou seja,

$$\delta_i \mid \underline{y}, \underline{x}, \underline{w} \sim \text{GammaTruncada}_{(c_i, \infty)}(a_{\delta_i}^*, b_{\delta_i}^*), \quad (26)$$

para $i = 1, 2, \dots, n_j$, onde $a_{\delta_i}^* = a_{\delta_i} + w_i \times (y_{ij} + \sum_k x_{ik})$, $b_{\delta_i}^* = b_{\delta_i} + w_i \times (\theta_j + \sum_k \pi_k)$ e $c_i = \max(y_{ij}, x_{i1}, x_{i2}, \dots, x_{iK})$.

A distribuição condicional completa a posteriori de $\delta_i \mid \underline{y}, \underline{x}, \underline{w}$ é uma Gama truncada para garantir que a estimação do tamanho da rede de contatos do indivíduo i seja superior ou igual ao valor máximo de “conhecidos” pertencentes entre as K subpopulações cujos tamanhos são conhecidos e na subpopulação alvo, uma vez que o grau da rede de contatos é o número total de conhecidos do indivíduo i .

5.4 AUTOVALIDAÇÃO DOS MODELOS PROPOSTOS

Esta seção descreve o processo de autovalidação dos modelos hierárquicos propostos neste capítulo. A autovalidação de um modelo estatístico verifica a capacidade deste em estimar o valor verdadeiro do parâmetro sem a presença de vieses.

Além disso, a autovalidação neste projeto permite verificar se as rotinas computacionais, de autoria própria, implementadas estão corretas e a se a distribuição a priori foi elicitada de forma que não interfira no resultado a posteriori. No entanto, só foi possível autovalidar os modelos pois os valores verdadeiros dos parâmetros de interesse eram conhecidos.

Então, no processo de autovalidação dos quatro modelos hierárquicos propostos nesse capítulo foram utilizadas as subpopulações cujos tamanhos são previamente conhecidos (a listagem com as subpopulações de tamanhos conhecidos está disponível no ANEXO C). Nessa seção foi descrito o processo de inferência para estimar o tamanho de apenas uma subpopulação “conhecida” (K), mas esse processo foi feito para todas as K subpopulações de tamanhos conhecidos para cada uma das 27 capitais.

Suponha X_{ik} , para $i = 1, 2, \dots, n$ e $k = 1, 2, \dots, K$, onde X_{ik} é o número de pessoas conhecidas do indivíduo i que pertencem a subpopulação de tamanho conhecido k . Nosso interesse, agora,

é estimar π_K , prevalência da subpopulação K na população geral e consequentemente N_K , tamanho da subpopulação K.

No processo de inferência, as quantidades π_K e N_K são consideradas desconhecidas e as informações indiretas baseada na rede de contatos dos participantes que foram utilizadas para estimar o grau são provenientes das informações dos participantes referentes as perguntas do tipo “quantas pessoas residentes neste município você conhece da subpopulação k?” para as K-1 subpopulações de tamanho conhecido, $X_{-K} = (X_1, X_2, \dots, X_{(K-1)})$.

A inferência para os parâmetros desconhecidos dos modelos hierárquicos propostos nesse capítulo foi feita sob o paradigma bayesiano. Toda a inferência foi baseada na distribuição a posteriori, ou pseudo posteriori dependendo do modelo, das quantidades desconhecidas, $(\pi_K, \delta_1, \dots, \delta_n)$, que via Teorema de Bayes pode ser escrito como

$$p(\pi_K, \delta_1, \dots, \delta_n | x_K, \underline{x}_{-K}, \underline{w}) \propto p(\pi_K, \delta_1, \dots, \delta_n) \times LP(\pi_K, \delta_1, \dots, \delta_n) \quad (27)$$

onde $p(\pi_K, \delta_1, \dots, \delta_n)$ é a distribuição a priori das quantidades desconhecidas, $x_K = (x_{1K}, x_{2K}, \dots, x_{nK})$, $\underline{x}_{-K} = (x_{1k}, x_{2k}, \dots, x_{nk}, \dots, x_{1(K-1)}, \dots, x_{2(K-1)}, \dots, x_{n(K-1)})$ e $LP(\pi_K, \delta_1, \dots, \delta_n)$ é a função de pseudo verossimilhança (e se $w_i = 1 \forall i$ temos a função de verossimilhança usual dos casos particulares, como os modelos NSUM bayesiano e NSUM bayesiano com pós-estratificação) que sob independência condicional é dada pelo produto da função de probabilidade conjunta.

Após o processo de inferência, o objetivo é comparar os valores verdadeiros dos parâmetros, isto é, a prevalência (“verdadeira”) da subpopulação K de tamanho conhecido (π_K) com a prevalência estimada da subpopulação K de tamanho conhecido, como por exemplo, a média a posteriori, ou pseudo posteriori, $(E[\pi_K | \underline{\delta}, x_K, \underline{x}_{-K}, \underline{w}])$, assim como o intervalo de credibilidade (IC), para todas subpopulações, nas 27 capitais para cada um dos quatro modelos propostos.

A seguir serão descritas as duas métricas de comparação de cada um dos modelos propostos no processo de autovalidação. Isto é, tais medidas foram calculadas para cada um dos modelos separadamente.

(i) *RMSE* (sigla do inglês para *Root Mean Square Error*) por capital (*s*)

$$RMSEs = \sqrt{EQMs} \quad (28)$$

onde $s = 1, \dots, S$ é referente a capital s e o EQMs (sigla para Erro Quadrático Médio da capital s) pode ser escrito como

$$EQM_s = \sum_{k=1}^K \frac{(E[\pi_{ks} | \underline{\delta}, x_{ks}, \underline{x}_{-ks}, W] - \pi_{ks})^2}{K}$$

onde π_{ks} é a prevalência “verdadeira” da subpopulação conhecida k na capital s .

(ii) *RMSE* (sigla do inglês para *Root Mean Square Error*) global do modelo

$$RMSE = \sqrt{EQM} \quad (29)$$

e o EQM (sigla para Erro Quadrático Médio) global pode ser escrito como

$$EQM = \sum_{s=1}^S \sum_{k=1}^K \frac{(E[\pi_{ks} | \underline{\delta}, x_{ks}, \underline{x}_{-ks}, W] - \pi_{ks})^2}{S \times K}$$

onde π_{ks} é a prevalência “verdadeira” da subpopulação conhecida k na capital s .

Para considerar um dos 4 modelos como o melhor modelo para estimar o tamanho de subpopulações de difícil acesso é desejável que o modelo apresente o menor erro quadrático médio, e conseqüentemente o menor RMSE, ou seja, deseja-se um modelo que, de forma geral, gere estimativas mais próximas do valor verdadeiro.

5.5 ASPECTOS ÉTICOS

A pesquisa proposta nesta tese foi desenvolvida de acordo com os princípios éticos de respeito pela pessoa, beneficência e justiça, seguindo as diretrizes e normas regulamentares da resolução nº 466/2012 do Ministério da Saúde. No presente estudo não considerou-se necessário o uso de Termo de Consentimento Livre e Esclarecido (TCLE), pois utilizou apenas com dados secundários restritos, sem a informação nominal dos participantes, da Pesquisa Nacional intitulada “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”, aprovada pelo Comitê de Ética da ENSP/FIOCRUZ, sob o número CAAE 0073.0.031.000-11, desenvolvido pela FIOCRUZ, sob a coordenação geral do pesquisador Francisco Inácio Bastos, tendo suas atividades encerradas em 2012. Portanto, ressalta-se que em nenhum momento houve contato do pesquisador responsável deste projeto de pesquisa com qualquer sujeito de pesquisa.

Diante disso, o pesquisador responsável da Pesquisa Nacional intitulada “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” (sob o número CAAE 0073.0.031.000-11) emitiu uma declaração de cessão do banco de dados do Inquérito Nacional (ANEXO D) para análises deste projeto de pesquisa, cujo objetivo principal é estimar o tamanho da subpopulação de mulheres que fizeram aborto (aborto provocado) nas capitais brasileiras segundo as informações indiretas da rede de contatos dos participantes do Inquérito.

A pesquisadora responsável pela tese assinou um Termo de Responsabilidade de Utilização de Dados, de tal forma que garante o compromisso com a privacidade e a confidencialidade dos dados utilizados, preservando integralmente o anonimato dos participantes, principalmente por não ter tido acesso a informação nominal dos participantes.

Finalmente, mediante essas garantias o projeto foi encaminhado e aprovado pelo Comitê de Ética em Pesquisa da Escola Nacional de Saúde Pública sob parecer CAAE: 84961418.4.0000.5240.

6 RESULTADOS

Este capítulo descreve os resultados da autovalidação, assim como as estimativas da média a posteriori para a incidência acumulada de mulheres que fizeram aborto no ano anterior ao Inquérito Domiciliar por capitais brasileiras bem como os intervalos de credibilidade de 95% (IC95%).

6.1 AUTOVALIDAÇÃO

A autovalidação desta seção teve como objetivo verificar se os modelos hierárquicos apresentados no capítulo anterior estavam gerando estimativas não viesadas, além de verificar se havia algum erro na rotina implementada no software estatístico R utilizada nas análises. Desta forma, realizou-se uma análise de validação interna.

O processo de validação interna foi realizado para cada uma das 27 capitais brasileiras considerando cada um dos 4 modelos descritos anteriormente (NSUM bayesiano, NSUM pseudo bayesiano com pesos amostrais, NSUM bayesiano com pós-estratificação e NSUM pseudo bayesiano com pesos amostrais e pós-estratificação).

Isto é, para um determinado modelo hierárquico descrito nesta tese e para uma determinada capital considerou-se que cada uma das 20 subpopulações de tamanhos conhecidos utilizadas neste estudo (como por exemplo, Meninas menores de 5 anos) tinha seu tamanho “desconhecido” e utilizou-se as informações das 19 subpopulações de tamanhos conhecidos restantes para estimar o seu tamanho. Por fim, a estimativa gerada foi comparada com o tamanho real de tal subpopulação.

Nas linhas da Tabela 3 é possível observar as métricas de comparação, dadas nas equações (28) e (29), de cada capital para os 4 modelos. A métrica de comparação global avaliou o desempenho global dos modelos considerando as estimativas de todas as subpopulações nas 27 capitais. Ou seja, a métrica de comparação global está apresentada na linha referente ao Total da Tabela 3.

Foi possível verificar que todos os modelos apresentaram um RMSE pequeno para todas as capitais, além da métrica global também pequena, isto é, o RMSE Global para o modelo “NSUM bayesiano” = 0,0146; modelo “NSUM pseudo bayesiano com pesos amostrais” = 0,0148; modelo “NSUM bayesiano com pós-estratificação” = 0,0148 e modelo “NSUM pseudo bayesiano com pesos amostrais e pós-estratificação” = 0,0149 (Tabela 3). Com isso, conclui-se

que os modelos são razoáveis para estimar o tamanho de subpopulações desconhecidas nas capitais brasileiras, pois os modelos não estão enviesados e as rotinas implementadas no software R estão corretas.

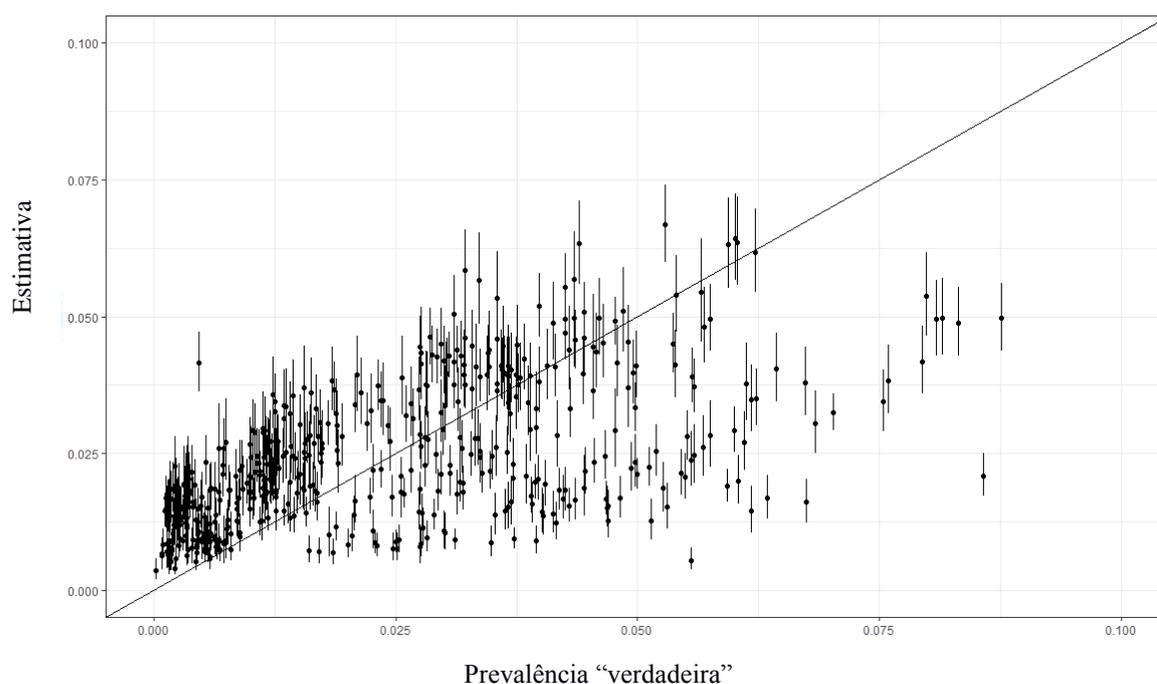
Tabela 3 - Métricas de comparação para cada uma das 27 capitais considerando os modelos NSUM bayesiano, NSUM bayesiano com pesos amostrais, NSUM bayesiano com pós-estratificação e NSUM bayesiano com pesos amostrais e pós-estratificação e métrica de comparação global (Total do RMSE por modelo)

Região	Estado	Cidade	NSUM bayesiano	NSUM bayesiano com pesos amostrais	NSUM bayesiano com pós-estratificação	NSUM bayesiano com pesos amostrais e pós-estratificação
Nordeste	AL	Maceió	0,021	0,022	0,022	0,022
	BA	Salvador	0,016	0,016	0,016	0,016
	CE	Fortaleza	0,016	0,016	0,017	0,017
	MA	São Luís	0,014	0,013	0,014	0,014
	PB	João Pessoa	0,017	0,018	0,018	0,018
	PE	Recife	0,017	0,018	0,017	0,018
	PI	Teresina	0,018	0,018	0,018	0,021
	RN	Natal	0,017	0,017	0,017	0,017
	SE	Aracaju	0,015	0,016	0,016	0,016
Norte	AC	Rio Branco	0,017	0,017	0,017	0,017
	AM	Manaus	0,012	0,013	0,012	0,013
	AP	Macapá	0,015	0,015	0,014	0,014
	PA	Belém	0,012	0,011	0,012	0,012
	RO	Porto Velho	0,017	0,017	0,017	0,017
	RR	Boa Vista	0,017	0,017	0,017	0,016
	TO	Palmas	0,012	0,012	0,012	0,013
Centro-Oeste	DF	Brasília	0,010	0,010	0,010	0,010
	GO	Goiânia	0,012	0,012	0,012	0,012
	MS	Campo Grande	0,014	0,014	0,015	0,015
	MT	Cuiabá	0,013	0,013	0,013	0,013
Sudeste	ES	Vitória	0,015	0,014	0,015	0,015
	MG	Belo Horizonte	0,014	0,013	0,014	0,014
	RJ	Rio de Janeiro	0,012	0,013	0,013	0,013
	SP	São Paulo	0,012	0,012	0,012	0,012
Sul	PR	Curitiba	0,014	0,013	0,013	0,013
	RS	Porto Alegre	0,016	0,016	0,016	0,016
	SC	Florianópolis	0,011	0,011	0,011	0,011
TOTAL			0,0146	0,0148	0,0148	0,0149

Embora os modelos descritos nessa tese não tenham apresentado diferenças expressivas nas métricas de comparação (Tabela 3) o modelo hierárquico completo com pesos amostrais e pós-estratificação foi considerado o melhor modelo pois incorporou o desenho amostral complexo do inquérito além de minimizar um dos vieses do NSUM ao pós-estratificar a população segundo sexo e faixa etária e o desbalanceamento da amostra em relação a população geral.

A seguir, as Figuras 4-6 ilustram os resultados da autovalidação do modelo hierárquico “pseudo bayesiano” com pesos amostrais e pós-estratificação, chamado de modelo hierárquico completo. A Figura 4 apresenta um resultado geral da autovalidação, a Figura 5 apresenta a autovalidação por subpopulação e a Figura 6 por capital.

Figura 4 - Autovalidação geral do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 20 subpopulações de tamanhos conhecidos para as 27 capitais brasileiras: estimativas pseudo posteriori da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações em suas capitais.

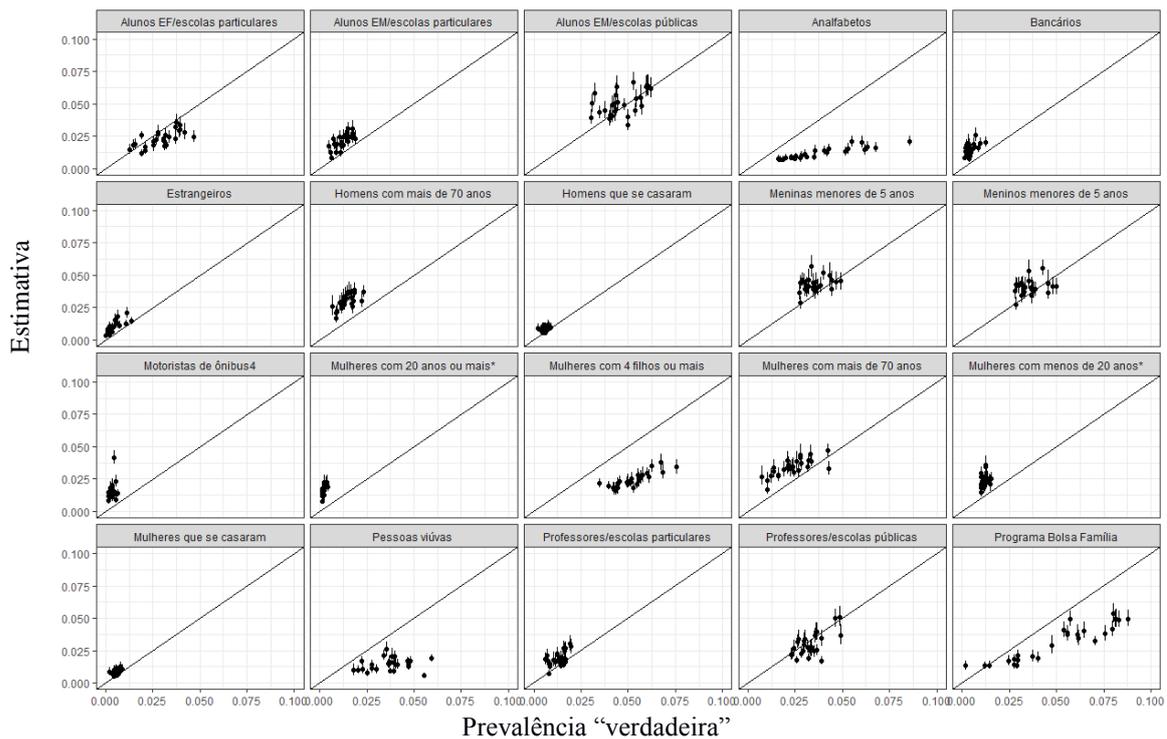


A partir da Figura 4, foi possível verificar que o modelo hierárquico “pseudo bayesiano” baseado nas informações indiretas da rede de contatos dos participantes selecionados de forma aleatória da população geral das 27 capitais (que além de considerar o desenho amostral complexo do inquérito domiciliar também considerou a composição da população) apresentou resultados satisfatórios de uma forma geral. Espera-se que os intervalos de credibilidade das estimativas pseudo posteriori das prevalências das subpopulações de tamanhos conhecidos, na

autovalidação, estejam próximos (e cruzem) a linha diagonal da figura 4. Caso isso ocorra, podemos considerar que o modelo está estimando as prevalências de forma razoável com 95% de credibilidade.

Na Figura 5, pode-se observar a autovalidação das estimativas pseudo a posteriori das prevalências *versus* as prevalências verdadeiras por cada uma das 20 subpopulações de tamanho conhecido. Para cada uma das subpopulações pode-se verificar que os resultados (estimativas pseudo posteriori) nas 27 capitais são homogêneos, pois não há, por subpopulação, resultados muito distintos entre si evidenciando um mal comportamento do modelo em determinada capital. O modelo completo, é um bom modelo para estimar o tamanho/prevalência de subpopulações de difícil acesso (com baixa prevalência na população de origem) em todas as capitais.

Figura 5 - Autovalidação por subpopulação de tamanho conhecido do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 27 capitais brasileiras: estimativas pseudo posteriori da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações



Nota: * Mulheres que tiveram bebê nos últimos 12 meses; EF: Ensino fundamental (5ª-8ª séries); EM: Ensino médio

Nesse estudo de autovalidação, embora não haja nenhuma subpopulação de tamanho conhecido que seja criminalizada, como a subpopulação de interesse, mulheres que fizeram aborto, temos subpopulações de tamanhos conhecidos que podem causar possíveis vieses na

estimação NSUM como pessoas viúvas, (isto é, homens ou mulheres cujo último cônjuge faleceu e não se casaram no civil novamente), mulheres com 4 filhos ou mais (apenas filhos biológicos vivos), mulheres com mais de 70 anos entre outros, pois as pessoas selecionadas na amostra do estudo podem conhecer pessoas integrantes dessas subpopulações citadas mas não terem tido conhecimento prévio desse pertencimento. O mesmo ocorre com abortamento provocado.

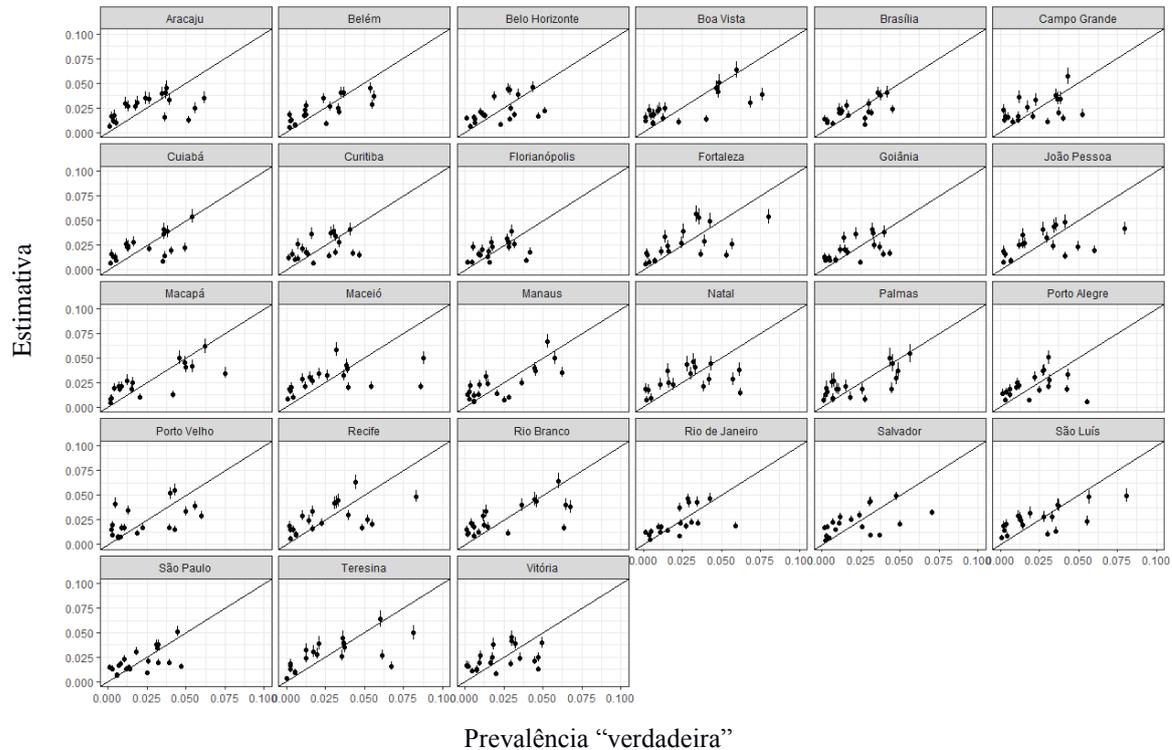
No entanto, o processo de pós-estratificação na modelagem teve como objetivo, além de ponderar o desbalanceamento da amostra das capitais, de minimizar esses vieses que podem ser gerados pela estrutura etária da população, uma vez que se espera que certos estratos da população tenham mais chance de conhecer pessoas de determinadas subpopulações.

Ainda na Figura 5 foi possível observar que em todas as capitais, a estimativa pseudo a posteriori da subpopulação “analfabetos” (isto é, pessoas com 15 anos ou mais e que não sabem ler ou escrever) subestima a prevalência verdadeira. Vale ressaltar que essa subpopulação de tamanho conhecido “analfabetos” é uma subpopulação estigmatizada na nossa sociedade, assim como mulheres que fizeram aborto, embora o estigma com o abortamento provocado perpassa julgamentos sociais e envolva aspectos morais, éticos e religiosos e moral.

Na Figura 6 pode-se observar as estimativas pseudo posteriori das prevalências para cada uma das 27 capitais. Para cada uma das capitais pode-se verificar que os resultados (estimativas pseudo posteriori) nas 20 subpopulações de tamanhos conhecidos são homogêneos.

Todavia, esperava-se que o modelo hierárquico completo tivesse um bom desempenho na estimativa pseudo posteriori das prevalências de subpopulações (com uma prevalência baixa na população geral). Dado esse estudo de autovalidação do modelo hierárquico completo, considerou-se que as estimativas pseudo posteriori da prevalência dessas subpopulações de difícil acesso foram satisfatórias. Não houve por capital nem por subpopulação resultados enviesados ou considerados discrepantes, vide a métrica de comparação (RMSE) para todas as capitais e a métrica global (Tabela 3) além do desempenho de estimação do modelo hierárquico completo nas Figuras 4-6.

Figura 6 - Autovalidação por capital de tamanho conhecido do modelo NSUM bayesiano com pesos amostrais e pós-estratificação utilizando as 20 subpopulações de tamanho conhecido: estimativas pseudo posteriores da prevalência e seus respectivos IC95% versus a prevalência “verdadeira” das subpopulações



6.2 ESTIMATIVAS DO NÚMERO DE MULHERES QUE FIZERAM ABORTO

Valendo-se do contexto bayesiano, nesta seção, serão descritas as estimativas da média pseudo posteriori para a incidência acumulada nos últimos 12 meses anteriores ao Inquérito Domiciliar de mulheres que fizeram aborto bem como os intervalos de credibilidade de 95% (IC95%) para o modelo hierárquico completo, isto é, NSUM pseudo bayesiano com pesos amostrais e pós-estratificação (proposto no capítulo 5) para cada uma das 27 capitais brasileiras.

Na Tabela 4 estão descritas as estimativas da média a pseudo posteriori da incidência acumulada de mulheres que fizeram aborto provocado no ano anterior ao Inquérito Domiciliar por capital, para cada 1.000 mulheres de 15 a 49 anos, segundo o modelo hierárquico completo, construído a partir da metodologia *Network scale-up*, considerando o desenho amostral e a estrutura etária da população segundo sexo, para as 27 capitais brasileiras no período de 2011-2012.

Tabela 4 - Distribuição da incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres entre 15 e 49 anos) e do número de mulheres que fizeram aborto, e seus intervalos de credibilidade de 95% (IC95%), nas 27 capitais brasileiras no período de 2011-2012

Região	Cidade	Incidência acumulada de mulheres que fizeram aborto ^{a,b} (IC95%)		Número de mulheres que fizeram aborto ^b (IC95%)	
Nordeste	Aracaju (SE)	9,26	(5,66; 14,00)	1.684	(1.030; 2.547)
	Fortaleza (CE)	5,79	(3,06; 9,70)	4.424	(2.342; 7.417)
	João Pessoa (PB)	4,29	(2,02; 7,60)	957	(450; 1.696)
	Maceió (AL)	7,72	(4,84; 11,35)	2.220	(1.393; 3.262)
	Natal (RN)	4,91	(2,21; 8,92)	1.211	(544; 2.200)
	Recife (PE)	6,13	(3,63; 9,49)	2.839	(1.681; 4.392)
	Salvador (BA)	5,29	(3,67; 7,26)	4.523	(3.138; 6.207)
	São Luís (MA)	8,77	(5,47; 12,95)	2.900	(1.807; 4.283)
	Teresina (PI)	5,34	(2,87; 8,77)	1.382	(743; 2.270)
Norte	Rio Branco (AC)	4,64	(2,40; 7,88)	465	(240; 789,44)
	Belém (PA)	7,23	(4,79; 10,36)	3.098	(2.049; 4.438)
	Boa Vista (RR)	9,81	(6,08; 14,60)	830	(514; 1.236)
	Macapá (AP)	7,64	(4,67; 11,58)	899	(550; 1.364)
	Manaus (AM)	7,86	(5,08; 11,54)	4.264	(2.756; 6.256)
	Palmas (TO)	4,11	(1,56; 8,40)	304	(115; 621)
	Porto Velho (RO)	4,52	(2,70; 6,86)	572	(341; 869)
Centro-Oeste	Brasília (DF)	3,40	(1,76; 5,75)	2.749	(1.425; 4.653)
	Campo Grande (MS)	5,81	(3,29; 9,37)	1.352	(765; 2.181)
	Cuiabá (MT)	3,94	(1,95; 6,81)	662	(328; 1.146)
	Goiânia (GO)	4,46	(2,44; 7,34)	1.805	(987; 2.973)
Sudeste	Belo Horizonte (MG)	2,99	(1,53; 5,03)	2.124	(1.087; 3.577)
	Rio de Janeiro (RJ)	4,70	(2,93; 7,02)	8.260	(5.150; 12.348)
	São Paulo (SP)	3,29	(1,80; 5,31)	10.864	(5.943; 17.516)
	Vitória (ES)	3,90	(1,92; 6,68)	382	(188; 653)
Sul	Curitiba (PR)	3,04	(1,37; 5,66)	1.582	(714; 2.947)
	Florianópolis (SC)	5,76	(3,44; 8,88)	730	(437; 1.125)
	Porto Alegre (RS)	6,58	(3,86; 10,21)	2.601	(1.524; 4.032)
Total ^c				65.682	

^a por 1.000 mulheres de 15 a 49 anos; ^b Estimativas do número médio; ^c Total representa o número médio de mulheres que fizeram aborto nas capitais brasileiras.

Verificou-se que, em média, nas Regiões Norte e Nordeste a incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres entre 15-49 anos) foi maior do que nas demais Regiões do Brasil, no período de 2011-2012.

Nas capitais do Brasil, das 136.110.82 mulheres de 15 a 49 anos residentes, 65.682 mulheres fizeram aborto, em média, no período de 2011-2012, resultando numa incidência

acumulada média de 4,83 para cada 1.000 mulheres de 15 a 49 anos. Vale ressaltar, que no Brasil, isto é, incluindo todas as cidades (não somente as capitais), o número de mulheres que abortaram, no mesmo período de análise, foi ainda maior.

Não é possível aproximar o número de abortamentos provocados nas capitais brasileiras pelo número de mulheres que fizeram aborto, pois uma mulher pode ter provocado mais de um aborto em um período de 12 meses. Com isto, nas capitais brasileiras, em 2011-2012, o número de abortos provocados pode ser superior a 65.682 abortos provocados.

A ilustração dos resultados, por capital, pode ser vista na Figura 7. As categorias de incidência acumulada de mulheres que fizeram aborto nas capitais brasileiras na Figura 7 foram construídas baseada nos quartis. A menor incidência acumulada média foi 2,99 IC95% [1,53; 5,03], referente a Belo Horizonte (MG) e a maior foi 9,81 IC95% [6,08; 14,60] referente a Boa Vista (RR).

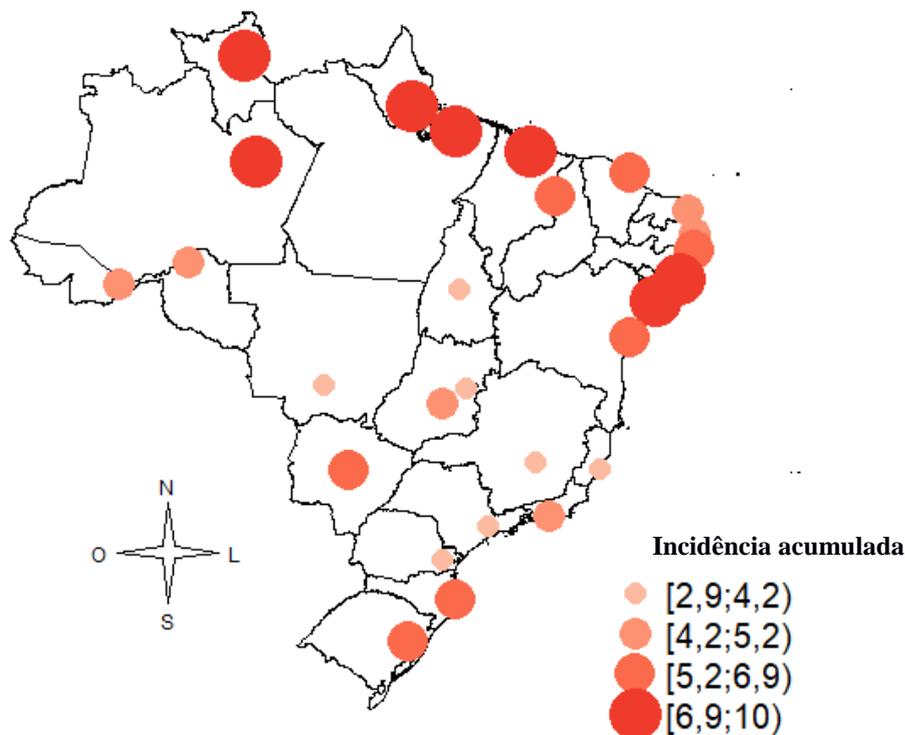
Majoritariamente as capitais da Região Sudeste apresentaram, em média, incidências acumuladas inferiores a 4,2 para cada 1.000 mulheres de 15-49 anos (quartil 1), exceto a cidade do Rio de Janeiro (RJ) (4,70 IC95% [2,93; 7,02]).

Em geral, as capitais da Região Norte apresentaram, em média, incidências acumuladas iguais ou superiores a 6,9 para cada 1.000 mulheres de 15-49 anos (quartil 3), exceto as cidades Rio Branco (AC) (4,64 IC95% [2,40; 7,88]), Palmas (TO) (4,11 IC95% [1,56; 8,40]) e Porto Velho (RO) (4,52 IC95% [2,70; 6,86]).

Das 9 capitais da região Nordeste, 3 tiveram uma incidência acumulada média de mulheres que provocaram abortos nos últimos 12 meses igual ou superior 6,9 em cada 1.000 mulheres de 15 a 49 anos (quartil 4). São elas: Aracaju (SE) (9,26 IC95% [5,66; 14,00]), Maceió (AL) (7,72 IC95% [4,84; 11,35]) e São Luís (MA) (8,77 IC95% [5,47; 12,95]).

As capitais que tiveram a maior incidência acumulada a posteriori de mulheres que fizeram aborto provocado (para cada 1.000 mulheres de 15-49 anos) foram das regiões Norte e Nordeste. São elas: Boa Vista (RR) (9,81 IC95% [6,08; 14,60]), Aracaju (SE) (9,26 IC95% [5,66; 14,00]), São Luís (MA) (8,77 IC95% [5,47; 12,95]), Manaus (AM) (7,86 IC95% [5,08; 11,54]), Maceió (AL) (7,72 IC95% [4,84; 11,35]).

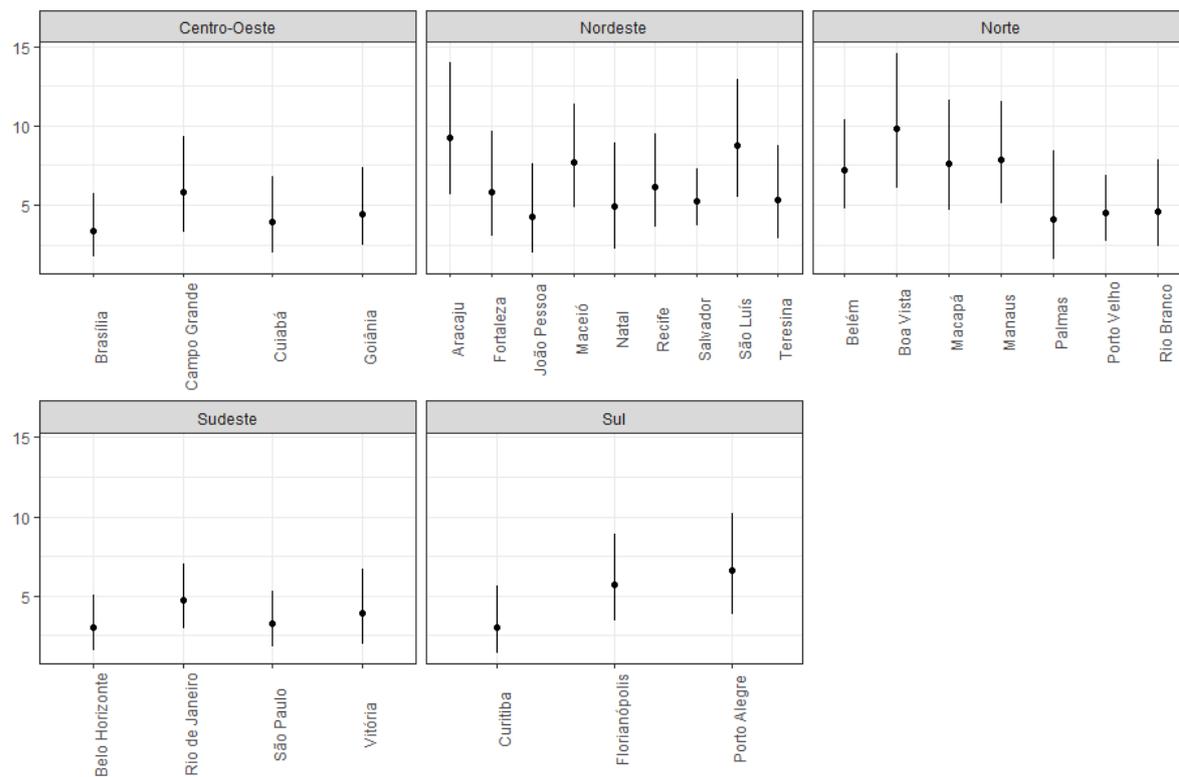
Figura 7 – Incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) nas 27 capitais brasileiras no período de 2011-2012



Na Figura 8, as estimativas da média e do IC95% da incidência acumulada de mulheres que fizeram aborto provocado, para cada 1.000 mulheres de 15 a 49 anos, no período de 2011-2012 nas 27 capitais brasileiras são apresentadas por Região.

A distribuição da média de incidência acumulada de mulheres que fizeram aborto na Região Centro-Oeste é homogênea entre as capitais, assim como na Região Sudeste. Na região Sul, Curitiba apresentou a menor incidência acumulada de mulheres que fizeram aborto. As Regiões Norte e Nordeste foram as regiões que apresentaram uma maior heterogeneidade nas médias de incidência acumulada de mulheres que fizeram aborto, no entanto, não houve diferença significativa entre as incidências acumuladas estimadas, como pode-se observar através dos intervalos de credibilidade de 95% (IC95%) apresentados na Figura 8.

Figura 8 – Média e IC95% a pseudo posteriori da Incidência acumulada de mulheres que fizeram aborto (por 1.000 mulheres de 15-49 anos) nas 27 capitais brasileiras, no período de 2011-2012, segundo as Regiões



7 DISCUSSÃO

O ato de provocar o aborto, segundo legislação brasileira, configura crime previsto no Decreto-Lei nº 2.848 do Código Penal (BRASIL, 1940), está vigente desde 1940, e só prescreve após 8 anos. No entanto, sabe-se que mulheres (independentemente de cor ou raça, religião, estado civil, situação socioeconômica etc.) vêm praticando tal ato (DINIZ; MEDEIROS, 2010; DINIZ; MEDEIROS; MADEIRO, 2017), muitas vezes de maneira insegura em todo o território brasileiro.

Com isso, há uma subnotificação e sub-registros das informações dos procedimentos para a interrupção da gravidez, muitas vezes realizados em “clínicas” clandestinas, de forma insegura, ou com a autoadministração de misoprostol, por exemplo.

O número real de mulheres que fizeram aborto (aborto provocado) no Brasil, seja em um período específico ou ao longo da vida, não é conhecido pois as mulheres tendem a ocultar tal informação tanto pelo estigma (de cunho moral, social, religioso ou ético) como pela ilegalidade, uma vez que abortamento provocado, salvo poucas exceções, no Brasil, configura crime previsto em lei.

Ganatra e cols (2017) evidenciaram que essa é a realidade em todos os países que têm leis restritivas quanto a interrupção da gestação através do aborto induzido/provocado, isto é, a criminalização do aborto e a restrição ao acesso aos serviços de saúde adequados para a realização do abortamento de forma segura, não garante que estas mulheres não irão recorrer a meios escusos para realização do aborto. Com isso, nesse contexto de criminalização e restrição ao acesso aos serviços de saúde, só aumenta os riscos evitáveis para a saúde e a vida destas mulheres (WHO, 1992, 2011).

Portanto, o norteamento da discussão sobre abortamento inseguro, muitas vezes provocado, e suas consequências, assim como a (re)formulação de políticas públicas e de medidas voltadas para a saúde da mulher (como serviços de saúde adequados e com profissionais de saúde que atuem com postura ética, garantindo o respeito aos direitos humanos das mulheres atuem independentemente dos seus preceitos morais e religiosos⁹) devem ser pautados em um conjunto amplo de evidências empíricas. No entanto, é essencial o

⁹ Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. PNDS 2006: Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher. Brasília (DF); 2008

conhecimento da magnitude de abortamentos inseguros, do número de mulheres que provocaram abortos e de suas características.

Nesse contexto, o NSUM é de grande importância, pois é um método promissor, que está em pleno desenvolvimento e vem sendo utilizado na literatura científica para estimar o tamanho de subpopulações de difícil acesso, muitas vezes estigmatizadas e/ou criminalizadas, em especial na área de Saúde Pública. Todavia, ainda é pouco aplicado em estudos com objetivo de estimar o tamanho de subpopulações brasileiras.

Salganik e cols (2011a) utilizaram o método NSUM para estimar a prevalência de HIV/AIDS em Curitiba (PR), Brasil. Em 2014, Bastos & Bertoni (2014) aplicaram a mesma metodologia *network scale-up*, em caráter inédito, para estimar o número de usuários de crack e/ou similares nas capitais brasileiras.

Em 2015, Maltiel e cols (2015) aplicaram o método NSUM nos mesmos dados utilizados por Salganik et al. (2011a) com o objetivo estimar a subpopulação de HIV/AIDS, em Curitiba (PR), Brasil, no entanto, estenderam o método clássico desenvolvido por Killworth e cols (1998a) incluindo efeitos aleatórios para tratar o grau da rede de contatos e minimizar vieses do NSUM.

A presente tese é pioneira na aplicação do NSUM para estimar a incidência acumulada de mulheres (e o número de mulheres) que fizeram aborto nas capitais brasileiras. Além de lançar mão, em caráter inédito, da aplicação de modelos hierárquicos bayesianos utilizando informações indiretas baseadas na rede de contatos considerando o desenho amostral complexo e a estrutura etária da população para estimar o tamanho de subpopulações de difícil acesso.

A aplicação do NSUM permitiu a autovalidação das estimativas a posteriori (ou pseudo posteriori) no processo de inferência, comparando e contrastando estimativas e cadastros disponíveis relativos a cada uma das subpopulações cujos tamanhos são conhecidos que estão incluídas na pesquisa, o que fornece parâmetros de avaliação referente à subpopulação (às subpopulações) considerada(s) “desconhecida(s)” e estimada(s) por meio deste método. Assim, consegue-se saber o quão próximo (ou distante) a estimativa está do valor real. Killworth e cols (1998a; 1998b) sugerem que sejam usadas cerca de 20 subpopulações de tamanho conhecido, para estimar o grau da rede de contatos nesta metodologia, apesar das críticas de Snidero e cols (2004), no sentido de que o estimador estaria sujeito a vícios, seja qual for o número desses grupos.

Por ser necessário contar com esses 20 cadastros de subpopulações ditas “conhecidas”, é indispensável que, em cada localidade pesquisada, esses cadastros sejam similares e tenham elevada qualidade (completude e informações consistentes).

Dessa forma, o estudo de autovalidação do modelo hierárquico pseudo bayesiano, resultante dessa tese, contribuiu para a verificar o desempenho do modelo no processo de inferência, pois havia a necessidade de observar se em diferentes cenários, ou seja, nas 20 subpopulações, disponíveis, cujos tamanhos eram conhecidos, as estimativas a pseudo posteriori não estavam enviesadas e/ou discrepantes, além de testar a rotina de autoria própria que foi implementada no software livre R (R DEVELOPMENT CORE TEAM, 2018) , uma vez que o modelo hierárquico pseudo bayesiano que considera o desenho amostral e a estrutura etária da população, baseado nas informações indiretas baseado na rede de contatos dos participantes, ainda não havia sido empregado em estudos similares a esta tese de doutoramento.

Então, no estudo de autovalidação, observou-se que mesmo dada a imprecisão intrínseca a estes processos de estimação através da metodologia *network scale-up*, considerando a variabilidade associada a diferentes subpopulações, contextos e localidades, já era esperado que a correlação entre estimativas e valores verdadeiros não fosse perfeita, podendo, assim, considerar que as estimativas pseudo posteriori do modelo hierárquico completo com pesos amostrais e pós-estratificação se mostraram razoáveis quando comparadas com os valores reais da prevalência destas subpopulações.

No entanto, outras limitações do NSUM ainda precisam ser exploradas com mais afinco, como por exemplo, o viés de memória e o de transmissão. Maltiel e cols (2015) trataram o viés de transmissão na modelagem hierárquica bayesiana para estimar a prevalência de HIV/AIDS em Curitiba (PR), Brasil, baseando-se na estimativa do viés de transmissão de Salganik e cols (2011b) que usaram o método do jogo de contatos, pois segundo Maltiel e cols (2015) é importante usar informações externas sobre viés de transmissão, sempre que possível. E no presente estudo, não houve a possibilidade da utilização de informações externas, como resultados de jogos de contatos, por exemplo, onde os pesquisadores entrevistariam mulheres que provocaram aborto para estimar a proporção de suas redes que estão cientes do seu abortamento provocado nas capitais brasileiras.

O presente estudo tentou corrigir alguns vieses do pressuposto básico do método *network scale-up*. Por exemplo, ao considerar a variabilidade no grau da rede de contatos

através de distribuição de probabilidade a priori (informativa baseada em estudos científicos através da utilização do número Dunbar) minimiza o pressuposto de que o grau de contatos do indivíduo é uma fração baseada em dados de uma amostra aleatória. Outro exemplo, é a utilização da técnica de pós-estratificação que ajudou a minimizar o viés de barreira, pois ponderou as estimativas segundo os estratos da população, uma vez que participantes tendem a conhecer pessoas com características semelhantes às suas e o aborto está ligado diretamente a certos estratos de sexo e faixa etária.

Sabe-se que essas dificuldades intrínsecas à metodologia podem comprometer a consistência dos resultados e gerar erros nas estimativas. No entanto, estimações indiretas, mesmo com suas limitações, são preferíveis quando comparadas as estimações diretas no contexto de estigmatização e/ou criminalização dessas subpopulações de interesse, como a subpopulação de mulheres que fizeram aborto nas capitais brasileiras, pois não expõe os participantes ao constrangimento quando questionados sobre sua participação nas subpopulações ditas de difícil acesso.

Embora algumas metodologias de estimação do número de abortos, como o emprego de multiplicadores em dados de Sistema de Informações, venham sendo empregadas na literatura científica e recomendada pela OMS para estimar o número de abortos provocados em UF's, grandes regiões e Brasil nas últimas décadas, acredita-se que é necessária uma atualização desses multiplicadores, propostos no início da década de 90 pelo Instituto Alan Guttmacher (SINGH; WULF, 1991).

Além de transcorridos quase 30 anos desde sua formulação, utilizar um único fator de correção para estimar o número de abortos induzidos em um país de proporções continentais, como o Brasil, que tem uma sociedade e um sistema de saúde bastante heterogêneos, deve ser reavaliado.

O presente estudo não tem comparabilidade direta com estimativas de estudos empíricos da incidência acumulada de mulheres que provocaram aborto nas capitais brasileiras, por conta do ineditismo do método de estimação, do local e período de análise.

Nas estimativas resultantes desta tese, observou-se que há uma disparidade regional da incidência acumulada de mulheres que provocaram aborto no período de estudo. Pode-se observar que, em média, a incidência média acumulada de mulheres que fizeram aborto (aborto provocado), por 1.000 mulheres de 15 a 49 anos, é superior nas regiões Norte e Nordeste, no

entanto, ao considerar a variabilidade das estimativas, a partir dos intervalos de credibilidade de 95%, não houve diferença significativa nas incidências acumuladas nas capitais brasileiras.

Segundo Diniz, Medeiros e Madeiro (2017), as taxas de aborto (na vida) em mulheres alfabetizadas de 18 a 39 anos são maiores entre mulheres residentes das regiões Norte/Centro-Oeste e Nordeste quando comparadas as das mulheres residentes nas regiões Sudeste e Sul. Além disso, Diniz, Medeiros e Madeiro (2017) apontam uma possível diferença na taxa de aborto (na vida) entre mulheres que residem em capitais e não-capitais. No entanto, o evento (aborto provocado pelo menos uma vez na vida), o local (municípios não-rurais) e o recorte de mulheres alfabetizadas de 18 a 39 anos, não permitem uma comparação direta com os resultados desta tese.

Martins-Melo e cols (2014) também identificaram diferenças entre regiões além de altos valores de abortos inseguros, especialmente em estados das regiões Norte, Nordeste e Sudeste, no entanto, consideraram todos os municípios brasileiros e espera-se, de fato, que haja uma diferença na taxa de abortamento entre capitais e não-capitais.

Limitações do estudo

Além das limitações já discutidas nesse capítulo, vale ressaltar outra importante limitação presente em nosso estudo, que foi o fato de ter sido desenhado apenas para as capitais brasileiras. O Inquérito domiciliar da pesquisa “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” originária dos dados da presente tese como do estudo realizado por Bertoni (2014) e por Bastos & Bertoni (2014) foi executado apenas nas capitais brasileiras e não compreendeu todos os municípios, uma vez que para execução do NSUM são necessários os cadastros confiáveis e completos, de todas as subpopulações conhecidas e, no momento da implementação do Inquérito, os pesquisadores responsáveis não tinham acesso a esses cadastros confiáveis e completos por municípios. Mesmo assim, cabe destacar, que até o presente momento, esse foi o maior estudo com a aplicação da metodologia *network scale-up* já realizado no mundo.

O estudo “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil” foi uma demanda do governo federal e da sociedade de identificar e quantificar o número de usuários de crack, mas a metodologia *network scale-up* permitiu que no mesmo inquérito fossem questionadas diversas subpopulações de difícil acesso, tais como, mulheres que provocaram aborto.

8 CONSIDERAÇÕES FINAIS

Em pleno século XXI, ainda há uma necessidade de ampliação de esforços para o desenvolvimento de pesquisas, criação de novas metodologias e aprimoramento das técnicas já existentes para estimar o número de mulheres que provocaram aborto (ou realizaram abortos inseguros), principalmente em localidades cujo abortamento provocado é considerado crime contra a vida, além de identificar características dessas mulheres e o real impacto do aborto provocado sobre as vidas dessas mulheres, de suas famílias e do setor saúde.

Deve-se ressaltar que as metodologias já existentes de estimação de aborto, como a estimação baseada em dados de Sistemas de Informações sobre internações com a utilização de fatores de correção, devem ser utilizadas, principalmente para construção de uma série histórica do número de abortos inseguros (ou induzidos), no entanto, há uma necessidade de garantir a qualidade desses dados secundários e uma atualização desses fatores de correção para o Brasil, regiões e UF's.

A aplicação do NSUM, para estimar o número de mulheres que fizeram aborto, seja em um período de tempo específico, como do presente estudo, ou pelo menos uma vez na vida, abrange de forma ampla todas as mulheres que fizeram aborto e contaram para membros da sua rede de contatos, independentemente da idade, nível socioeconômico, escolaridade e situação (como por exemplo, privada de liberdade, moradora de rua, já falecida), pois o NSUM é realizado através da implementação de um Inquérito domiciliar com a população geral e as perguntas do tipo “quantas pessoas você conhece na subpopulação de interesse” podem ser inseridas em Inquéritos regulares já existentes.

Com isso, a presente tese de doutoramento pode ser vista como uma linha de base (*baseline*) para estudos futuros com a utilização de mesma metodologia, com o propósito de gerar séries históricas consistentes e confiáveis para dimensionar o, ainda atual, problema de Saúde Pública aborto provocado (ou mulheres que fizeram aborto) nas capitais do país

REFERÊNCIAS

BASTOS, F I; BERTONI, N. (org.) **Uso da metodologia Network Scale-up Method (NSUM) para estimar o número de usuários de crack e/ou similares nas capitais do Brasil – implementação e logística do Inquérito Domiciliar.** In: *Pesquisa Nacional sobre o uso de crack: quem são os usuários de crack e/ou similares do Brasil? Quantos são nas capitais brasileiras?* ICICT/FIOCRUZ, ed., Rio de Janeiro. (pp. 117 – 130)., 2014.

BEARAK, J., POPINCHALK, A., ALKEMA, L., & SEDGH, G.. **Global, regional, and subregional trends in unintended pregnancy and its outcomes from 1990 to 2014: estimates from a Bayesian hierarchical model.** *The Lancet Global Health*, 6(4), e380-e389, 2018.

BERNARD, H. R. et al. **Estimating the size of an average personal network and of an event subpopulation: some empirical results.***Social Science Research*, 20, pp.109–121, 1991.

BERNARD, H. R. et al. **Counting hard-to-count populations: the network scale-up method for public health.** *Sexually Transmitted Infections*, v. 86, n. Suppl 2, p. ii11-ii15, 2010.

BERNARD, H. R.; MCCARTY, C. **The Network Scale-Up Method: Background and Theory.** *Evaluation Review*, 2009.

BERTONI, N. **Quantos usuários de crack e/ou similares existem nas capitais brasileiras? Resultados de um inquérito nacional com a utilização da metodologia Network Scale-Up.** Escola Nacional de Saúde Pública Sergio Arouca, 2014.

BRASIL. Código Penal. Decreto-Lei 2.848, Art 124-128, de 07 de dezembro de 1940. . 1940.

BRASIL. **Aborto e saúde pública no Brasil: 20 anos.** Brasília: Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia, 2009.

BOX, G. E. P. **Science and Statistics** , *Journal of the American Statistical Association*, 71: 791–799, 1976.

CENTER FOR REPRODUCTIVE RIGHTS. **The World’s Abortion Laws.** Disponível em: <<http://worldabortionlaws.com/map/>>. Acesso em: 24 maio. 2016.

DINIZ, D.; MEDEIROS, M. Aborto no Brasil: uma pesquisa domiciliar com técnica de urna. **Ciência & Saúde Coletiva**, v. 15, p. 959–966, 2010.

DINIZ, D.; MEDEIROS, M.; MADEIRO, A. Pesquisa Nacional de Aborto 2016. **Ciência & Saúde Coletiva**, vol.22, n.2, pp.653-660, 2017.

DRUCKER, E. M. **A plague of prisons : the epidemiology of mass incarceration in America.** The New Press; Reprint edition (May 28, 2013), 2013.

GAMERMAN, D.; LOPES, H. F. *Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference.* (2ª edition), Chapman & Hall/CRC, Boca Raton, FL, 2006.

GANATRA, B. et al. **From concept to measurement: Operationalizing WHO’s definition**

of unsafe abortion. Bulletin of the World Health Organization, 2014.

GANATRA, B. et al. **Global, regional, and subregional classification of abortions by safety, 2010–14: estimates from a Bayesian hierarchical model.** *The Lancet*, v. 390, n. 10110, p. 2372–2381, 2017.

GUTTMACHER INSTITUTE. **Clandestine Abortion: A Latin American Reality.** New York, 1994 .

GUTTMACHER INSTITUTE. **Facts on Induced Abortion Worldwide.** Disponível em: <http://www.who.int/reproductivehealth/publications/unsafe_abortion/induced_abortion_2012.pdf>. Acesso em: 19 maio. 2016.

GUTTMACHER INSTITUTE. **Restrictive laws do not stop women from having abortions. They mainly make the procedure clandestine and often unsafe.** Disponível em: <<https://www.guttmacher.org/infographic/2016/restrictive-laws-do-not-stop-women-having-abortions>>. Acesso em: 19 maio. 2016a.

GUTTMACHER INSTITUTE. **Induced Abortion Worldwide.** New York, . Disponível em: <<papers3://publication/uuid/E73D169E-5FAD-4D15-904A-EEAFE8D81683>>.

HILL, R. A; DUNBAR, R. I. M. Social network size in humans. **Human Nature-an Interdisciplinary Biosocial Perspective**, v. 14, n. 1, p. 53–72, 2003.

IBGE. **CENSO DEMOGRÁFICO 2010. Características da população e dos domicílios: resultados do universo.** Rio de Janeiro, 2011.

KADUSHIN, C. et al. Scale-up methods as applied to estimates of heroin use. **Journal of Drug Issues**, v. 36, n. 2, p. 417–440, 2006.

KILLWORTH, P. D. et al. Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. **Evaluation review**, v. 22, n. 2, p. 289–308, 1998a.

KILLWORTH, P. D. et al. A social network approach to estimating seroprevalence in the United States. **Social Networks**, 1998b.

KILLWORTH, P. D. et al. Two interpretations of reports of knowledge of subpopulation sizes. **Social Networks**, v. 25, n. 2, p. 141–160, 2003.

KUNIHAMA, T. et al. Nonparametric Bayes modeling with sample survey weights. **Statistics and Probability Letters**, 2016.

LAURENTI, R.; JORGE, M. H. P. DE M.; GOTLIEB, S. L. D. A mortalidade materna nas capitais brasileiras: algumas características e estimativa de um fator de ajuste. **Revista Brasileira de Epidemiologia**, 2004.

LUCENA, R. C. B. **Características de mulheres internadas por aborto em uma maternidade pública em Recife - PE: dimensão do problema e sua relação com a prática contraceptiva.** [s.l.] Centro de Pesquisas Aggeu Magalhães, 2000.

MALTIEL, R. et al. Estimating Population Size Using the Network Scale Up Method. **The Annals of Applied Statistics Institute of Mathematical Statistics in The Annals of Applied Statistics**, 2015.

MARTINS-MELO, F. R. et al. Temporal trends and spatial distribution of unsafe abortion in Brazil, 1996-2012. **Revista de Saude Publica**, v. 48, n. 3, p. 508–520, 2014.

MCCORMICK, T. H.; SALGANIK, M. J.; ZHENG, T. How Many People Do You Know?: Efficiently Estimating Personal Network Size. **Journal of the American Statistical Association**, v. 105, n. 489, p. 59–70, 2010.

MENEZES, G.; AQUINO, E. M. L. Pesquisa sobre o aborto no Brasil : avanços e desafios para o campo da saúde coletiva Research on abortion in Brazil : gaps and challenges for the public health field. **Cad. Saúde Pública**, v. 25, n. 2, p. 193–204, 2009.

MISSE, M. Crime, Sujeito e Sujeição Criminal: aspectos de uma contribuição analítica sobre a categoria “bandido”. **Lua Nova: Revista de Cultura e Política**, 2010.

MONTEIRO, M.; ADESSE, L. Estimativas de aborto induzido no Brasil e Grandes Regiões (1992-2005). **XV Encontro Nacional de Estudos Populacionais, ABEP**, 2006.

MONTEIRO, M. F. G.; ADESSE, L.; DREZETT, J. Atualização das estimativas da magnitude do aborto induzido, taxas por mil mulheres e razões por 100 nascimentos vivos do aborto induzido por faixa etária e grandes regiões. Brasil, 1995 a 2013. **Reprodução & Climatério**, v. 30, n. 1, p. 11–18, 2015.

OLINTO, M. T. A.; MOREIRA-FILHO, D. DE C. Estimativa de aborto induzido: comparação entre duas metodologias. **Revista Panamericana de Salud Pública**, v. 15, n. 5, p. 331–336, maio 2004.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2018.

RASTEGARI, A. et al. Estimating the Annual Incidence of Abortions in Iran Applying a Network Scale-up Approach. **Iranian Red Crescent Medical Journal**, v. 16, n. 10, p. 1–7, 2014.

ROSSIER, C. Induced Abortion Rates : Estimating A Review. v. 34, n. 2, p. 87–102, 2003.

RUSSELL BERNARD, H. et al. Estimating the size of an average personal network and of an event subpopulation: Some empirical results. **Social Science Research**, 1991.

SALGANIK, M. J. et al. Assessing network scale-up estimates for groups most at risk of HIV/AIDS: Evidence from a multiple-method study of heavy drug users in Curitiba, Brazil. **American Journal of Epidemiology**, v. 174, n. 10, p. 1190–1196, 2011a.

SALGANIK, M. J. et al. The game of contacts: Estimating the social visibility of groups. **Social Networks**, 2011b.

SAVITSKY, T. D.; TOTH, D. Bayesian estimation under informative sampling. **Electronic Journal of Statistics**, 2016.

SESTELO, J. A. DE F.; SOUZA, L. E. P. F. DE; BAHIA, L. Saúde suplementar no Brasil: abordagens sobre a articulação público/privada na assistência à saúde. **Cadernos de Saúde Pública**, v. 29, n. 5, p. 851–866, maio 2013.

SHAPIRO, G. K. **Abortion law in Muslim-majority countries: An overview of the Islamic discourse with policy implications** *Health Policy and Planning*, 2014.

SI, Y.; PILLAI, N. S.; GELMAN, A. Bayesian nonparametric weighted sampling inference. **Bayesian Analysis**, 2015.

SILVA, R. DE S. O uso da Técnica de Resposta ao Azar (TRA) na caracterização do aborto ilegal. **Revista Brasileira de Estudos Populacionais**, v. 10, n. 1/2, p. 41–56, 1993.

SINGH, S.; HENSHAW, S. K.; HAAS, T. La Incidencia del Aborto Inducido a Nivel Mundial. **Perspectivas Internacionales en Planificación Familiar**, p. 16–24, 1999.

SINGH, S.; MADDOW-ZIMET, I. Facility-based treatment for medical complications resulting from unsafe pregnancy termination in the developing world, 2012: a review of evidence from 26 countries. **BJOG: An International Journal of Obstetrics & Gynaecology**, p. n/a-n/a, 2015.

SINGH, S.; REMEZ, L.; TARTAGLIONE, A. **Methodologies for Estimating Abortion Incidence and Abortion-Related Morbidity: A Review** *Guttmacher Institute*. [s.l.: s.n.]. Disponível em: <<http://www.guttmacher.org/pubs/compilations/IUSSP/abortion-methodologies.pdf>>.

SINGH, S.; WULF, D. Estimating Abortion Levels in Brazil, Colombia and Peru, Using Hospital Admissions and Fertility Survey Data. **International Family Planning Perspectives**, v. 17, n. 1, p. 8–13+24, 1991.

SNIDERO, S., MORRA, B., CORRADETTI, R., GREGORI, D. **Use of the scale-up methods in injury prevention research: An empirical assessment to the case of choking in children**. *Social Networks* 29, 527–538, 2007.

UNAIDS/WHO. **Guidelines on Estimating the Size of Populations Most at Risk to HIV**. Genebra, 2010.

VICTORA, C. G. et al. Countdown to 2015: A decade of tracking progress for maternal, newborn, and child survival. **The Lancet**, 2016.

WARNER, S. L. Randomized response: A survey technique for eliminating evasive answer bias. **Journal of the American Statistical Association**, v. 60, n. 309, p. 63–69, 1965.

WHO, UNICEF, UNFPA, WORLD BANK GROUP, THE U. N. P. D. **Trends in Maternal Mortality : 1990 to 2015**. Organization. 2015.

WHO. **Aborto espontáneo y provocado: informe de un Grupo Científico de la OMS**.

Genebra, 1970.

WHO. The prevention and management of unsafe abortion. Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. Geneva: World Health Organization. Genebra, 2015.

WHO. Unsafe abortion: global and regional estimates of the incidence of unsafe abortion and associated mortality in 2008. Genebra, 2008.

WHO. Health worker roles in providing safe abortion care and post-abortion contraception. Genebra, 2015. Disponível em: <http://apps.who.int/iris/bitstream/10665/181041/1/9789241549264_eng.pdf?ua=1&ua=1>.

WILLIAMS, STEPHEN J.; TORRENS, P. R. Introduction to Health Services. 7th. ed. Cengage Learning, 2007.

WORLD HEALTH ORGANIZATION. World Health Statistics 2018- Monitoring Health for the SDG's (Sustainable development goals), Gebenra, 2018.

ZHENG, T.; SALGANIK, M. J.; GELMAN, A. How Many People Do You Know in Prison? **Journal of the American Statistical Association**, v. 101, n. February 2015, p. 409–423, 2006.

ANEXOS

ANEXO A – PARECER DE APROVAÇÃO DA PESQUISA NACIONAL SOBRE O USO DE CRACK E/OU SIMILARES



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz
Escola Nacional de Saúde Pública Sergio Arouca
Comitê de Ética em Pesquisa



Rio de Janeiro, 03 de maio de 2011.

O Comitê de Ética em Pesquisa da Escola Nacional de Saúde Pública Sergio Arouca – CEP/ENSP, constituído nos Termos da Resolução CNS nº 196/96 e, devidamente registrado na Comissão Nacional de Ética em Pesquisa - CONEP, recebeu, analisou e emitiu parecer sobre a documentação referente ao Protocolo de Pesquisa, conforme abaixo, discriminado:

PROTOCOLO DE PESQUISA CEP/ENSP - Nº 62/11 CAAE: 0073.0.031.000-11

Título do Projeto: “Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil”

Classificação no Fluxograma: Grupo III

Será encaminhado à Conep (áreas temáticas especiais) e, portanto, deve aguardar a apreciação final desta para início da execução? Não

Pesquisadora Responsável: Monica Siqueira Malta

Instituição onde se realizará: Escola Nacional de Saúde Pública Sergio Arouca – ENSP/Fiocruz

Data de recebimento no CEP-ENSP: 22 / 03 / 2011

Data de apreciação: 06 / 04 / 2011

Parecer do CEP/ENSP: Aprovado.

Ressaltamos que a pesquisadora responsável por este Protocolo de Pesquisa deverá apresentar a este Comitê de Ética um relatório das atividades desenvolvidas no período de 12 meses a contar da data de sua aprovação (*item VII.13.d, da resolução CNS/MS Nº 196/96*) de acordo com o modelo disponível na página do CEP/ENSP na internet.

Esclarecemos, que o CEP/ENSP deverá ser informado de quaisquer fatos relevantes (incluindo mudanças de método) que alterem o curso normal do estudo, devendo a pesquisadora justificar caso o mesmo venha a ser interrompido.

Ângela Esher
Coordenadora do
Comitê de Ética em Pesquisa
CEP/ENSP
Metr. SIAPE1 1355884

ANEXO B – TERMO DE CONSENTIMENTO DA PESQUISA NACIONAL SOBRE O USO DE CRACK E/OU SIMILARES¹⁰

COMUNICADO SETORES



Rio de Janeiro, xx de xxxxxxxx de 2012.

Prezado(a) Sr(a),

A Fundação Oswaldo Cruz (FIOCRUZ) do Ministério da Saúde e a Secretaria Nacional de Políticas sobre Drogas (SENAD) estão realizando uma pesquisa de opinião para coletar dados sobre a rede de contatos das pessoas, com o objetivo de calcular o tamanho de alguns grupos de pessoas com determinadas características que outras pesquisas não conseguiram calcular. Esta pesquisa auxiliará também na formulação de políticas de saúde, como políticas de enfrentamento ao crack e outras drogas no país. A pesquisa é de natureza acadêmica e utiliza a metodologia "Scale up".

Neste estudo, estamos realizando entrevistas com a população em geral (indivíduos com 18 anos ou mais) em alguns domicílios sorteados aleatoriamente, a partir dos registros de endereços disponibilizados pelo Censo 2010 do Instituto Brasileiro de Geografia e Estatística (IBGE). Dessa forma, nem todos os domicílios de sua vizinhança serão pesquisados. Esta etapa está ocorrendo simultaneamente em todas as capitais do país e no Distrito Federal, nos meses de XXX a XXX de 2012.

O(A) senhor(a) está recebendo esta carta pois o seu endereço foi selecionado em nossa amostra para participar da pesquisa. Dessa forma e pedimos a gentileza de receber nosso entrevistador em seu domicílio, a fim de que possamos realizar com sucesso esta pesquisa. A identificação (nome, documento de identificação e foto) do entrevistador responsável por realizar as entrevistas em sua região está anexa a esta carta.

O XXXXXXX foi contratado para realizar esta coleta de dados. Para tanto, seus entrevistadores estarão identificados por crachás e demais materiais com o nome e logotipo da empresa. Em anexo, uma cópia do crachá do entrevistador responsável por realizar a pesquisa em seu domicílio, juntamente com **os telefones de contato, para fazer o agendamento da entrevista**, que deverá ser presencial e irá seguir alguns critérios de seleção de um dos moradores de seu domicílio, que o entrevistador lhe explicitará em tempo oportuno.

Caso necessite de maiores informações e esclarecimentos sobre a pesquisa, o(a) senhor(a) poderá entrar em contato com os pesquisadores da FIOCRUZ, responsáveis pela realização e coordenação da pesquisa, nos endereços/telefones abaixo listados. E também com o XXXXXXX responsável pela coleta de dados, através do telefone (xx)xxxx-xxxx e do e-mail xxxxxxxxxxxxxxxxxxxxxxxx .

Certos de sua colaboração, subscrevemo-nos.

Atenciosamente,

Francisco Inácio Bastos
Pesquisador Senior
Coordenador Nacional da Pesquisa
FIOCRUZ
Av. Brasil 4365
Biblioteca de Manguinhos sala 229
Rio de Janeiro - RJ – CEP 21041-210
Tel: 55-21-3865-3231
Email: bastos@icict.fiocruz.br

Neilane Bertoni dos Reis
Pesquisadora Colaboradora ICICT
Coordenação Nacional da Pesquisa
FIOCRUZ
Av. Brasil 4365
Biblioteca de Manguinhos sala 229
Rio de Janeiro - RJ – CEP 21041-210
Tel: 55-21-3865-3231
Email: nbertoni@icict.fiocruz.br

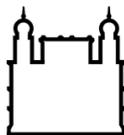
Pavilhão Hailly Moussatché - Avenida Brasil 4365 - RJ CEP 21045 900 Tel: 21 3865 3231
Fax: 21 2270-2668 E-mail: icict@icict.fiocruz.br www.icict.fiocruz.br

10

Disponível em

<https://www.icict.fiocruz.br/sites/www.icict.fiocruz.br/files/Pesquisa%20Nacional%20sobre%20o%20Uso%20e%20Crack.pdf>, acessado em 10/05/2017.

COMUNICADO SINDICO



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Comunicado

Rio de Janeiro, XX de XXXX de 2012.

Assunto: Realização da Pesquisa Domiciliar sobre Redes de Contatos

Prezado(a) Sr(a),

A Fundação Oswaldo Cruz (FIOCRUZ) do Ministério da Saúde está realizando uma pesquisa sobre "Redes de Contatos", que é de natureza acadêmica e que visa conhecer como as pessoas reconhecem, contam e classificam os indivíduos que fazem parte de seus contatos pessoais no cotidiano. As informações coletadas nesta pesquisa serão utilizadas para pensar estratégias sociais e de saúde em nosso país.

Neste estudo, estamos realizando entrevistas com a população em geral em alguns domicílios selecionados. Esta etapa está ocorrendo simultaneamente em todas as capitais do país e no Distrito Federal, nos meses de Março a Maio de 2012.

O XXXXXXXXXXXXXXX, foi o contratado para realizar esta coleta de dados. Para tanto, seus entrevistadores estarão identificados por crachás e demais materiais com o nome e logotipo da empresa. Na sua região, o(a) entrevistador(a) responsável é: XXXXXXXXXXXXX.

O(A) senhor(a) está recebendo este comunicado pois este endereço foi selecionado em nossa amostra: XXXXXXXXXXXXXXX.

Pedimos a gentileza de permitir acesso de nossos entrevistadores aos moradores destes domicílios, a fim de que possamos realizar com sucesso esta pesquisa.

Caso necessite de maiores informações e esclarecimentos, o(a) senhor(a) poderá entrar em contato com os pesquisadores da Fiocruz nos endereços/telefones abaixo listados.

Certos de sua compreensão, subscrevemo-nos.

Atenciosamente,

Francisco Inácio Bastos

Pesquisador Senior
Coordenador Nacional da Pesquisa
FIOCRUZ
Av. Brasil 4365
Biblioteca de Manguinhos sala 229
Rio de Janeiro - RJ – CEP 21041-210
Tel: 55-21-3865-3231
Email: bastos@icict.fiocruz.br

Neilane Bertoni dos Reis

Pesquisadora Colaboradora ICICT
Coordenação Nacional da Pesquisa
FIOCRUZ
Av. Brasil 4365
Biblioteca de Manguinhos sala 229
Rio de Janeiro - RJ – CEP 21041-210
Tel: 55-21-3865-3231
Email: nbertoni@icict.fiocruz.br

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO PARA PARTICIPAÇÃO NA ENTREVISTA

PESQUISA: Rede de Contatos

Entrevistador: Este termo de consentimento deve ser apresentado e explicado, de maneira clara, para todo participante no início de cada entrevista. As pessoas que concordarem em participar devem apenas consentir oralmente ao entrevistador.

OBJETIVOS DA PESQUISA

Esta é uma pesquisa de natureza acadêmica que visa conhecer como as pessoas reconhecem, contam e classificam os indivíduos que fazem parte de seus contatos pessoais no cotidiano.

PROCEDIMENTOS UTILIZADOS

Esta entrevista se constitui na aplicação de um questionário de aproximadamente 60 perguntas, onde participam somente o(a) entrevistador(a) e o(a) entrevistado(a). Não existem respostas certas ou erradas, pois o que importa é a percepção pessoal de cada entrevistado(a) sobre sua rede de contatos pessoais no cotidiano.

A participação é absolutamente voluntária.

O tempo de duração da entrevista é em média de 30 minutos, e o(a) entrevistado(a) tem o direito de responder ou não as questões colocadas pelo(a) entrevistador(a), bem como finalizar a entrevista no momento em que desejar.

CONFIDENCIALIDADE

Será garantido anonimato e sigilo absoluto por parte dos(as) pesquisadores(as). A cada entrevistado(a) será atribuído um número de código quando os questionários forem analisados, assim as respostas permanecerão confidenciais e nomes não serão associados a elas.

DESCONFORTOS E BENEFÍCIOS

Ao participar deste estudo, um desconforto possível é do entrevistado se sentir ansioso ou angustiado por estar respondendo a questões pessoais e sensíveis. Além disso, a duração da entrevista pode ser longa.

O benefício principal desta pesquisa é utilizar as informações coletadas de todos os entrevistados para que sejam pensadas estratégias sociais e de saúde no país.

DÚVIDAS – PESQUISADORES RESPONSÁVEIS

Qualquer dúvida poderá ser tirada no momento da entrevista. Uma cópia do cartão de contato dos pesquisadores da instituição responsável por este estudo deverá ser oferecida para o(a) participante, para caso haja dúvida em que a pessoa demande confirmação sobre a seriedade do estudo e de suas intenções para com o(a) entrevistador(a).

Francisco Inácio P. M. Bastos & Neilane Bertoni

Coordenadores Nacionais do projeto

ICICT - Instituto de Comunicação e Informação Científica e Tecnológica em Saúde

FIOCRUZ - Fundação Oswaldo Cruz

Av. Brasil, 4365. Biblioteca de Manguinhos #229

Rio de Janeiro 21040- 900 - Brasil

Telefone: 55 21 3865 3231 / 55 21 3865 3292

Fax: 55 21 2290 1696

CONSENTIMENTO

Você declara que está de acordo com os objetivos, riscos e procedimentos deste estudo e concorda em participar de forma voluntária?

Entrevistador: *Se o voluntário concordou em participar da pesquisa, siga para o Questionário Individual, anotando o seu consentimento.*

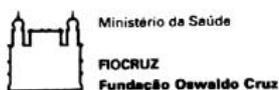
ANEXO C – LISTAGEM DAS SUBPOPULAÇÕES CONHECIDAS QUE FORAM UTILIZADAS PARA ESTIMAÇÃO DO GRAU DA REDE DE CONTATOS

Quadro 3 - Populações conhecidas que serão utilizadas para estimação do grau da rede de contatos dos indivíduos e fonte de dados de onde foram retiradas

População
Meninas menores de 5 anos ¹
Meninos menores de 5 anos ¹
Mulheres com mais de 70 anos ¹
Homens com mais de 70 anos ¹
Pessoas com 15 anos ou mais e que não sabem ler ou escrever ¹
Mulheres menores de 20 anos que tiveram bebês nos últimos 12 meses ¹
Mulheres com 20 anos ou mais que tiveram bebês nos últimos 12 meses ¹
Mulheres com 4 filhos ou mais (apenas filhos biológicos vivos) ¹
Homens que se casaram no civil nos últimos 12 meses ¹
Mulheres que se casaram no civil nos últimos 12 meses ¹
Pessoas viúvas (isto é, homens ou mulheres cujo último cônjuge faleceu e não se casaram no civil novamente) ¹
Estrangeiros ¹
Mães que recebem auxílio do Programa Bolsa Família ²
Estudantes de 5 ^a a 8 ^a . Séries (6 ^o ao 9 ^o ano) de escolas particulares ³
Estudantes de ensino médio de escolas públicas ³
Estudantes de ensino médio de escolas particulares ³
Professores de escolas públicas no ensino médio ou fundamental ³
Professores de escolas particulares no ensino médio ou fundamental ³
Motoristas de ônibus ⁴
Bancários ⁴

Fontes de dados: ¹ Censo 2010 – IBGE, ² Ministério do Desenvolvimento Social e Combate à Fome - 2010, ³ Censo Escolar 2011 – INEP, ⁴ Ministério do Trabalho e Emprego – 2011

ANEXO D – DECLARAÇÃO DE CESSÃO DO BANCO DE DADOS DA PESQUISA NACIONAL DO CRACK



Declaração de cessão de banco de dados

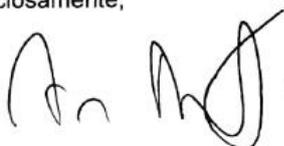
Na qualidade de Investigador Principal e Coordenador do Projeto intitulado "Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil", coordenado pela Fiocruz/MS e financiado pela SENAD/MJ, autorizo a aluna Natália Santana Paiva, CPF 116.144.387-81, a ter acesso ao banco de dados do referido projeto para subsidiar a realização das análises estatísticas referentes ao seu estudo de doutorado.

O projeto de pesquisa da aluna tem como título "Quantas mulheres fizeram aborto no ano anterior à avaliação nas capitais brasileiras? Resultados de um inquérito nacional com a aplicação de modelos hierárquicos bayesianos utilizando informações indiretas baseadas na rede de contatos", e objetiva estimar o número de mulheres que fizeram abortos (abortos provocados), nos últimos 12 meses, nas 26 capitais e DF via modelos hierárquicos, com ou sem pesos amostrais e estrutura demográfica, utilizando informações indiretas baseadas na rede social de cada participante aleatoriamente selecionado da população em geral.

O projeto "Perfil dos usuários de crack nas 26 capitais, DF, 9 regiões metropolitanas e Brasil", foi submetido e aprovado pelo Comitê de Ética em Pesquisa da Escola Nacional de Saúde Pública da Fundação Oswaldo Cruz (ENSP/FIOCRUZ), sob o CAAE 0073.0.031.000-11, tendo suas atividades encerradas em 2012.

Cabe ressaltar que, conforme previsto em acordo com financiador, em todo material acadêmico produzido a partir deste projeto deverá constar a informação de financiamento e agradecimento a SENAD/MJ em sessão específica.

Atenciosamente,



Francisco Inácio P. M. Bastos
Mat: 655062 - LIS/ICICT/FIOCRUZ
Pesquisador Titular

Francisco Inácio Pinkusfeld Monteiro Bastos
Pesquisador Sênior
Coordenador Nacional da Pesquisa
FIOCRUZ
Av. Brasil, 4365
Biblioteca de Manguinhos sala 229
Rio de Janeiro - RJ – CEP 21041-900
Tel: 55-21-3865-3231
Email: francisco.inacio.bastos@hotmail.com

APÊNDICE – DISTRIBUIÇÃO DA AMOSTRA DO INQUÉRITO DOMICILIAR E DA POPULAÇÃO (CENSO DEMOGRÁFICO 2010) SEGUNDO SEXO E FAIXA ETÁRIA PARA AS 27 CAPITAIS BRASILEIRAS

REGIÃO NORTE

Figura I: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Porto Velho/RO

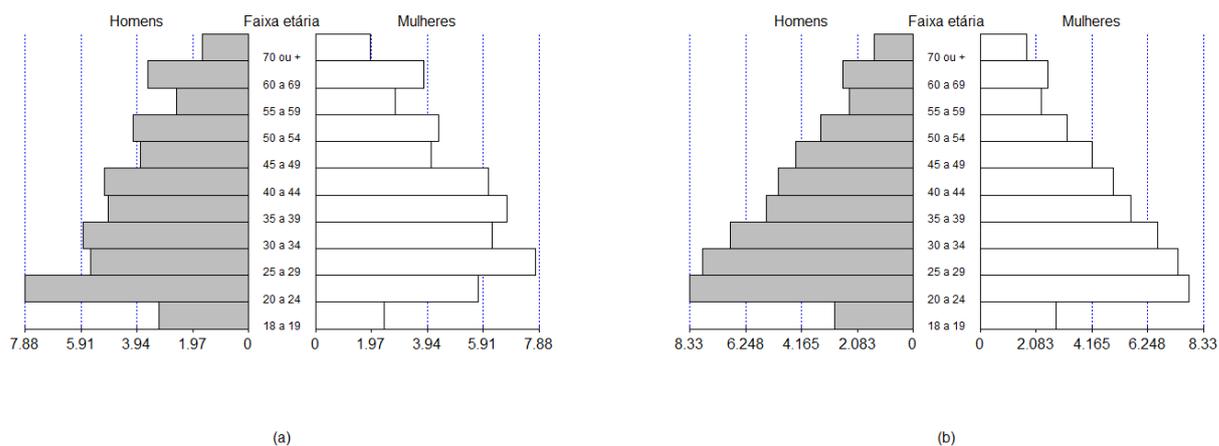


Figura II: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Rio Branco/AC

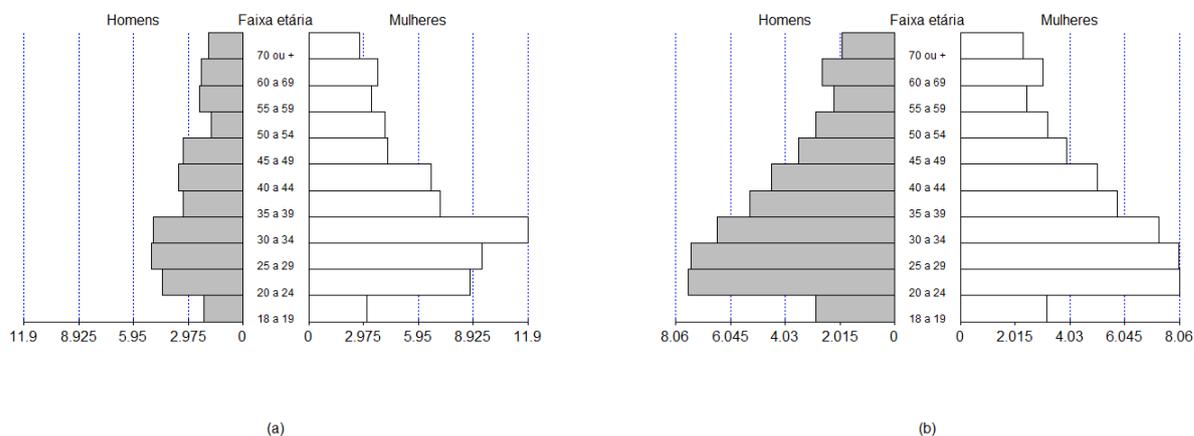


Figura III: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Manaus/AM

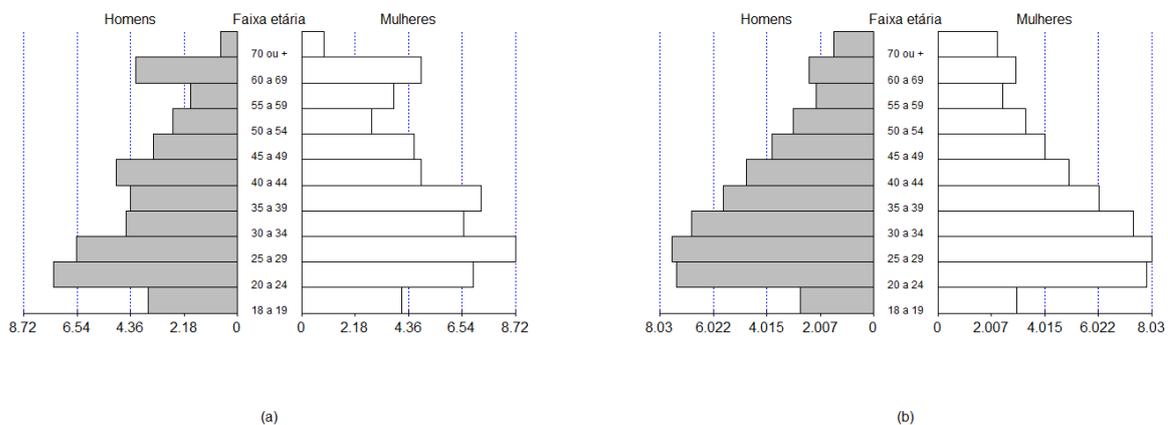


Figura IV: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Boa Vista/RR

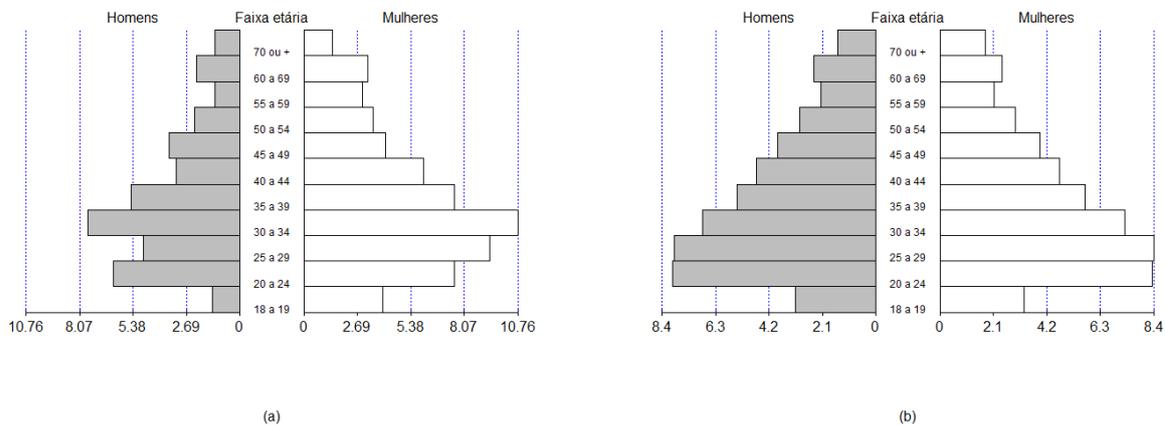


Figura V: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Belém/PA

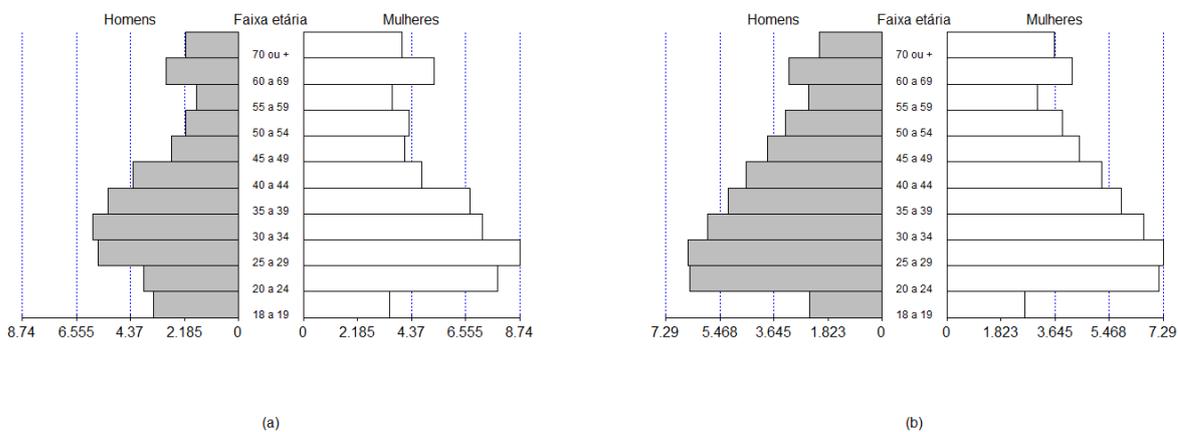


Figura VI: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Macapá/AP

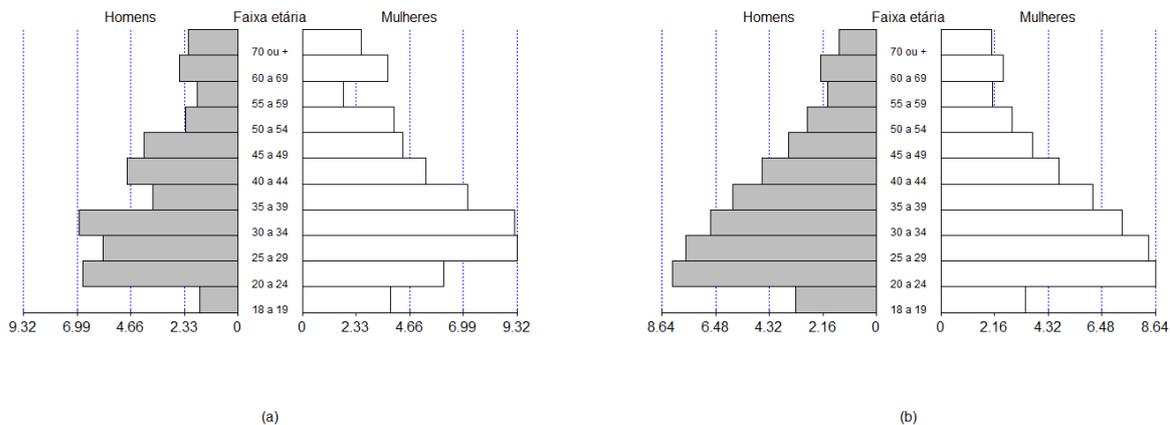
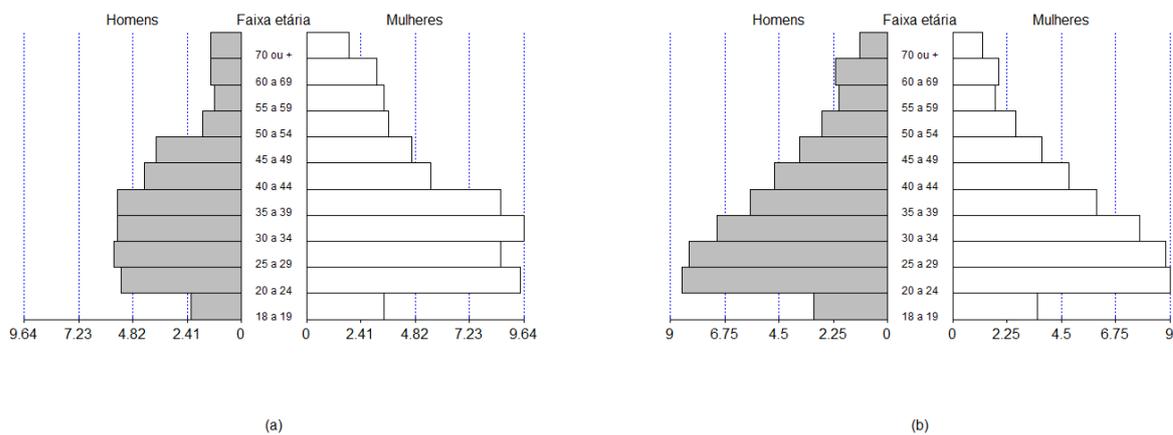


Figura VII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Palmas/TO



REGIÃO NORDESTE

Figura VIII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em São Luís/MA

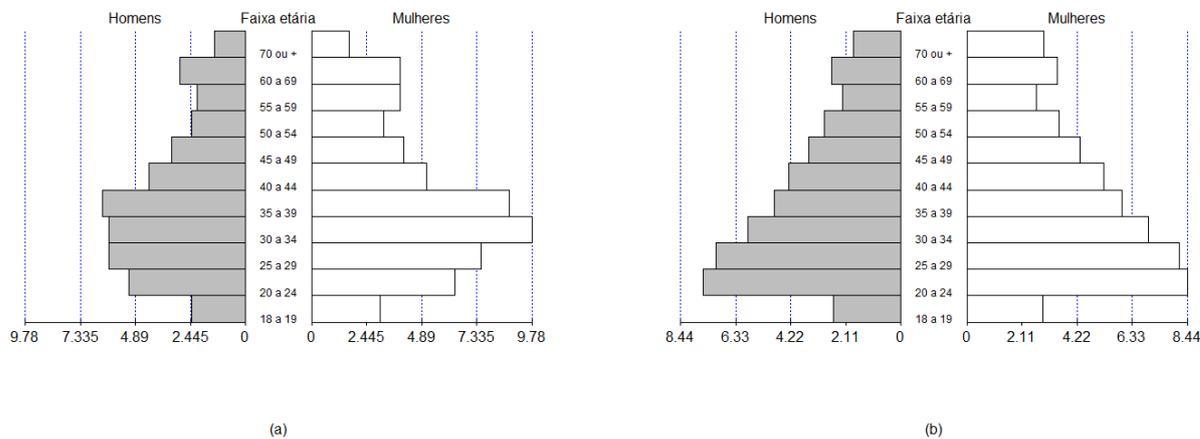


Figura IX: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Teresina/PI

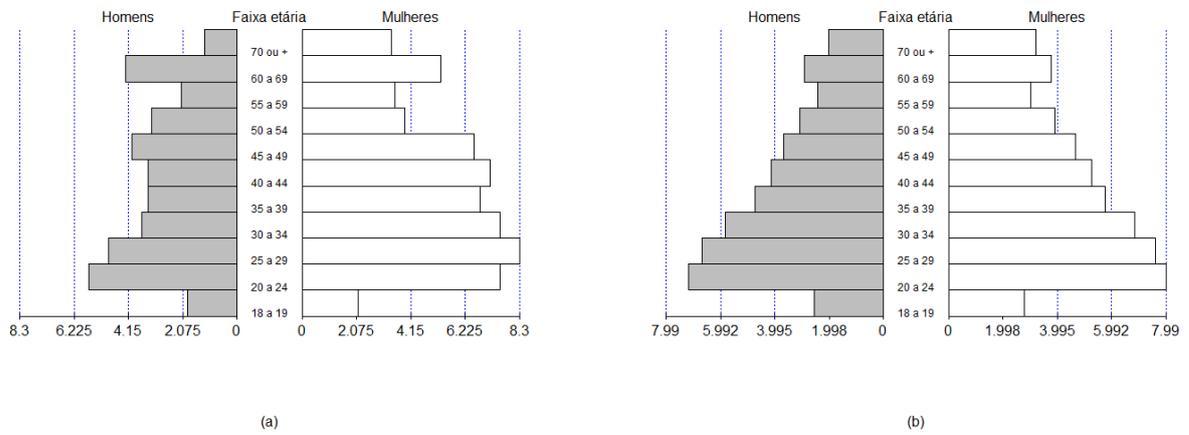


Figura X: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Fortaleza/CE

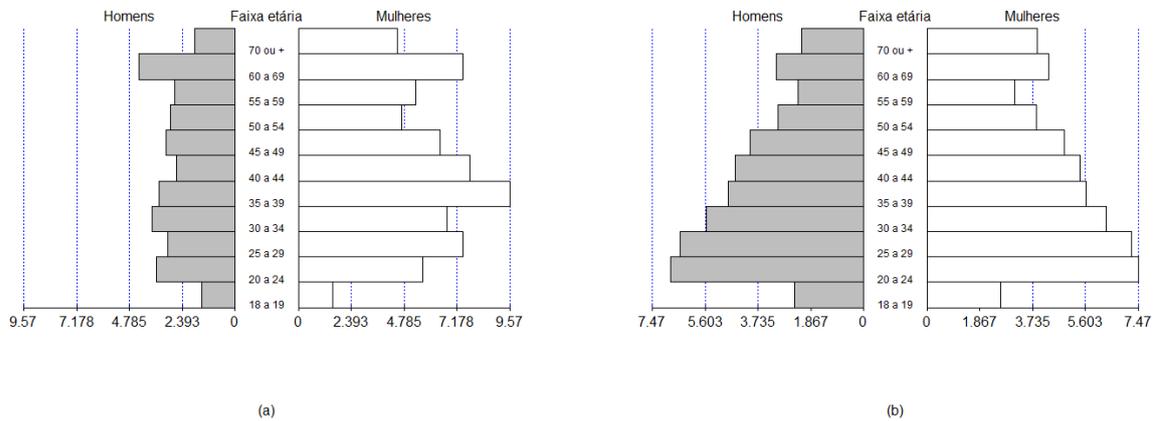


Figura XI Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Natal/RN

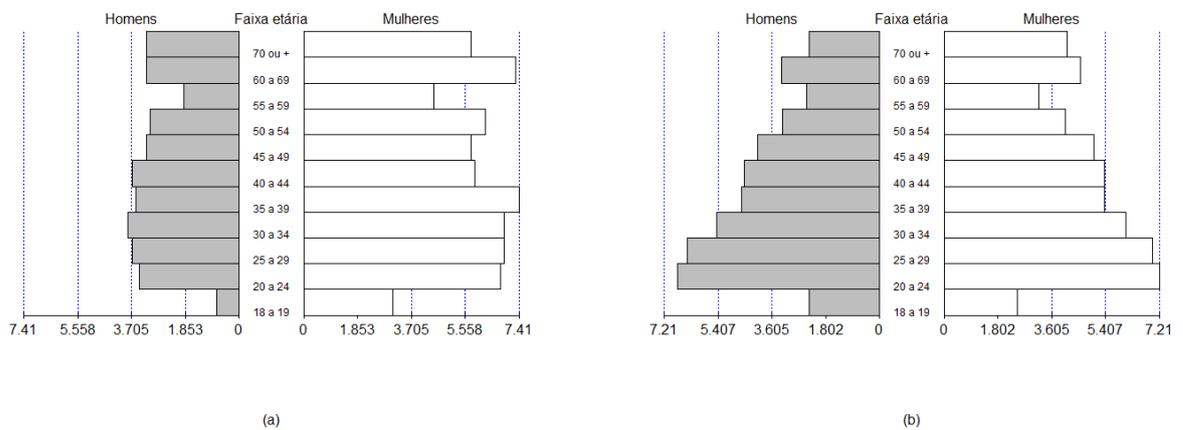


Figura XII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em João Pessoa/PB

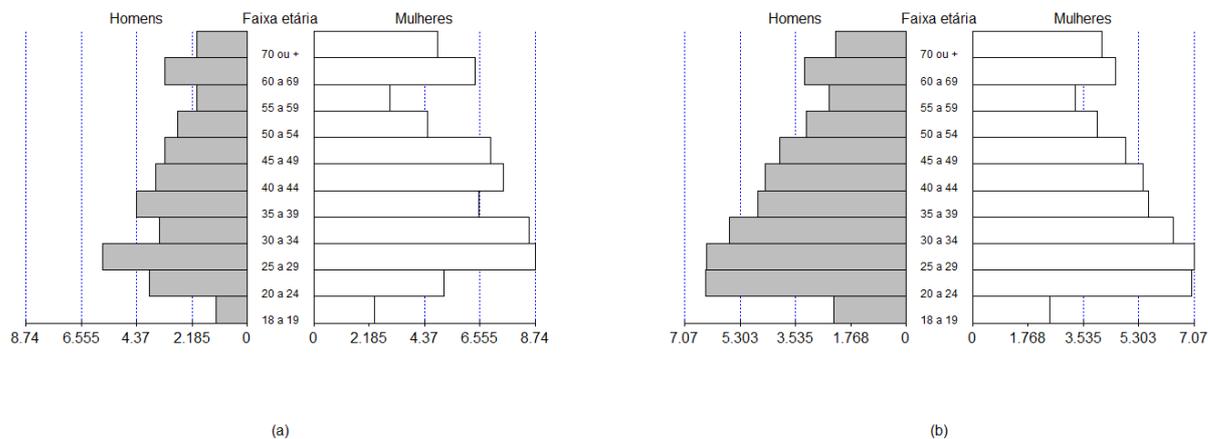


Figura XIII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Recife/PE

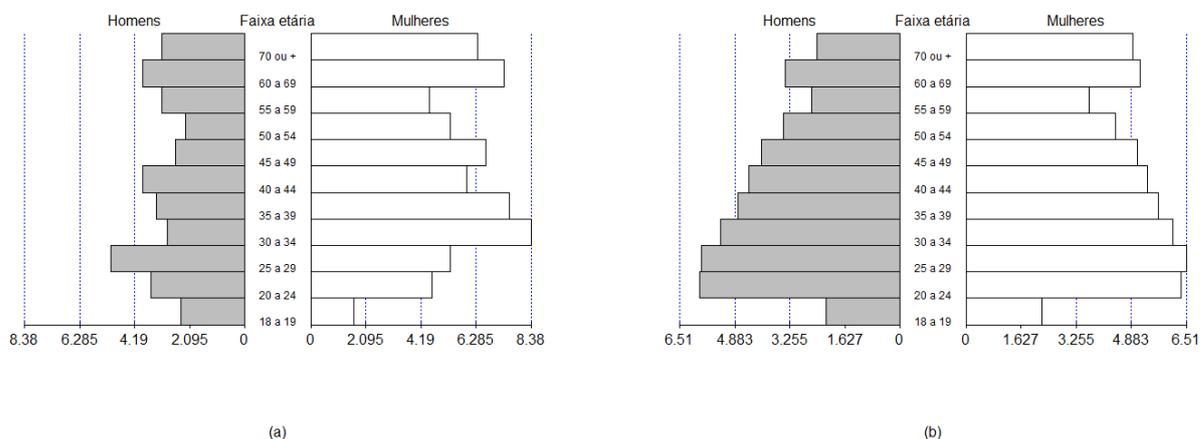


Figura XIV: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Maceió/AL

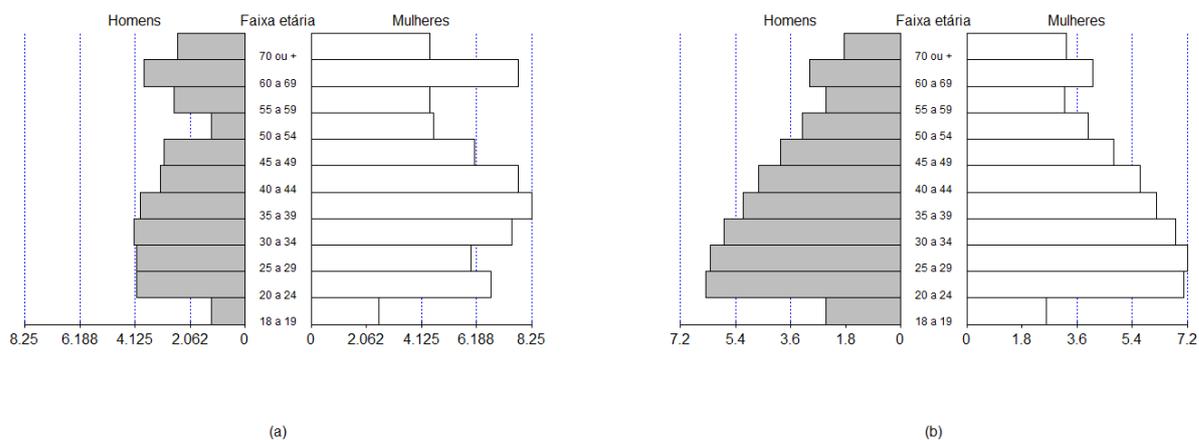


Figura XV: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Aracaju/SE

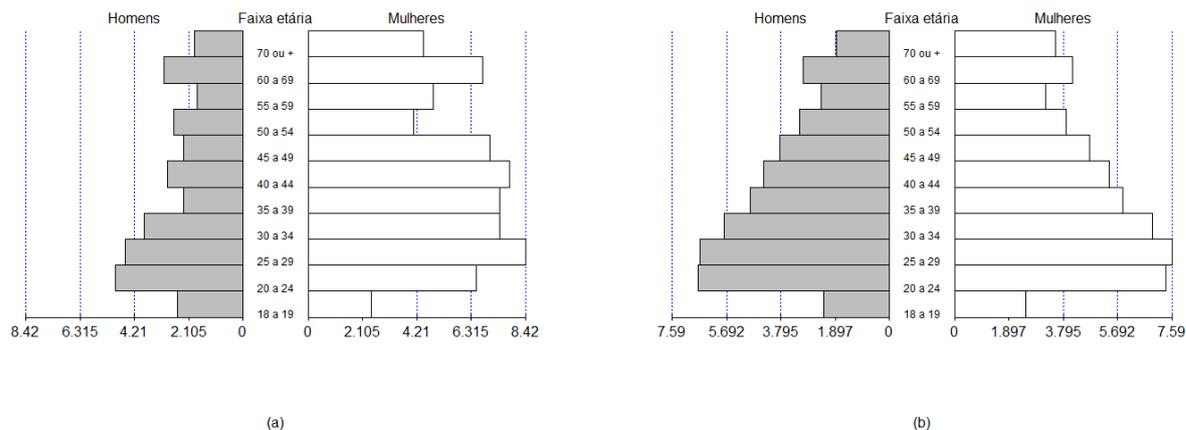
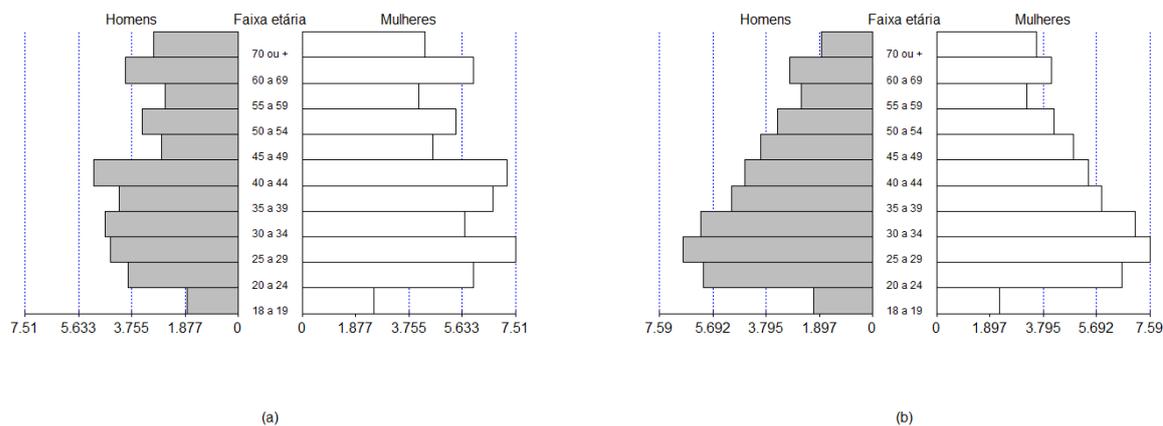


Figura XVI: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Salvador/BA



REGIÃO SUDESTE

Figura XVII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Belo Horizonte/MG

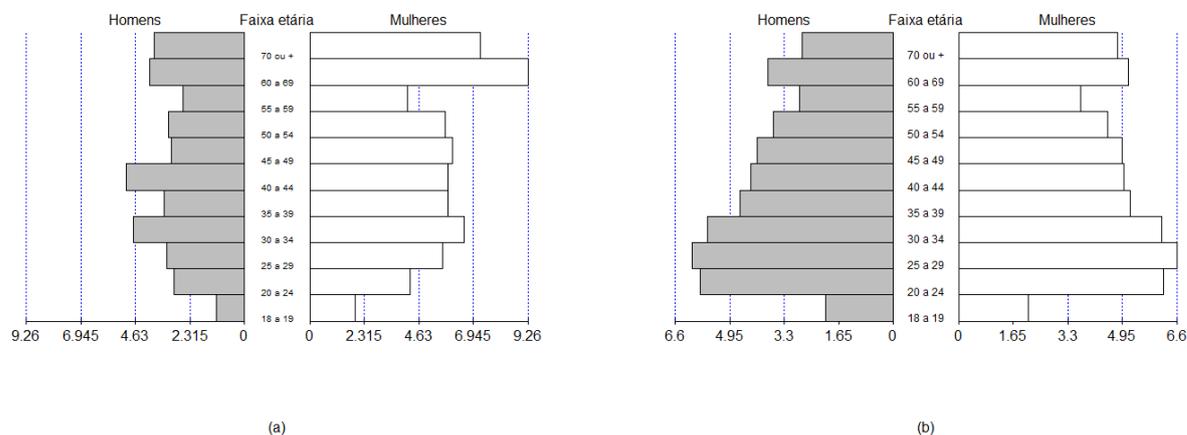


Figura XVIII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Vitória/ES

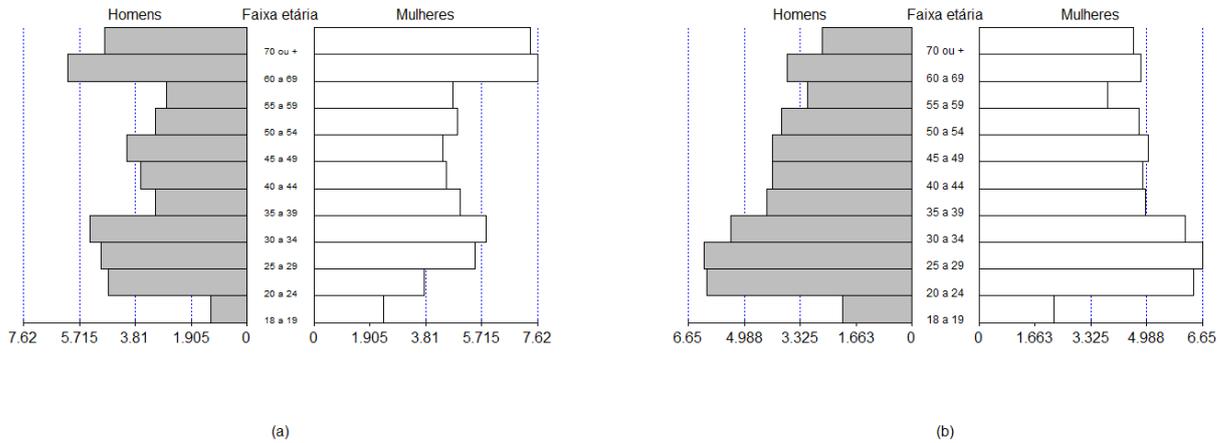


Figura XIX: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Rio de Janeiro/RJ

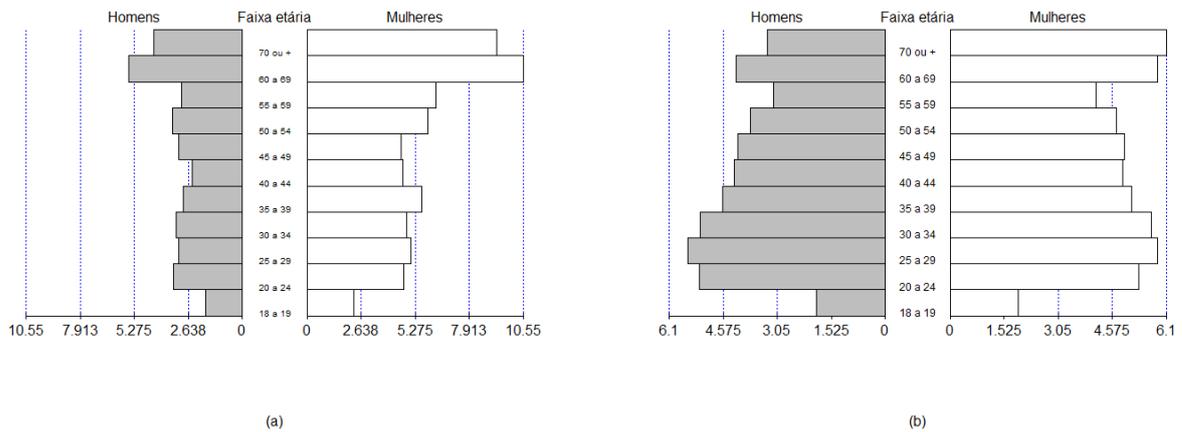


Figura XX: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em São Paulo/SP

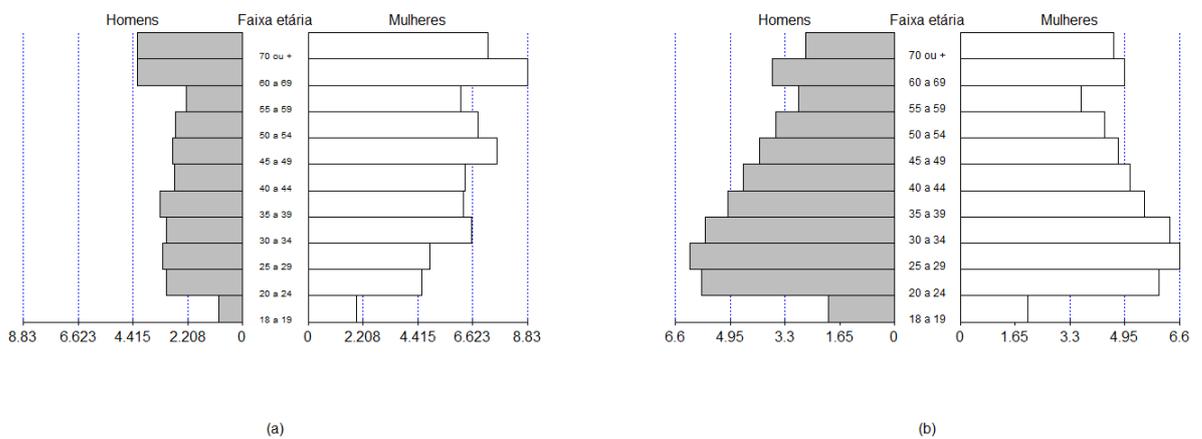


Figura XXI: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Curitiba/PR

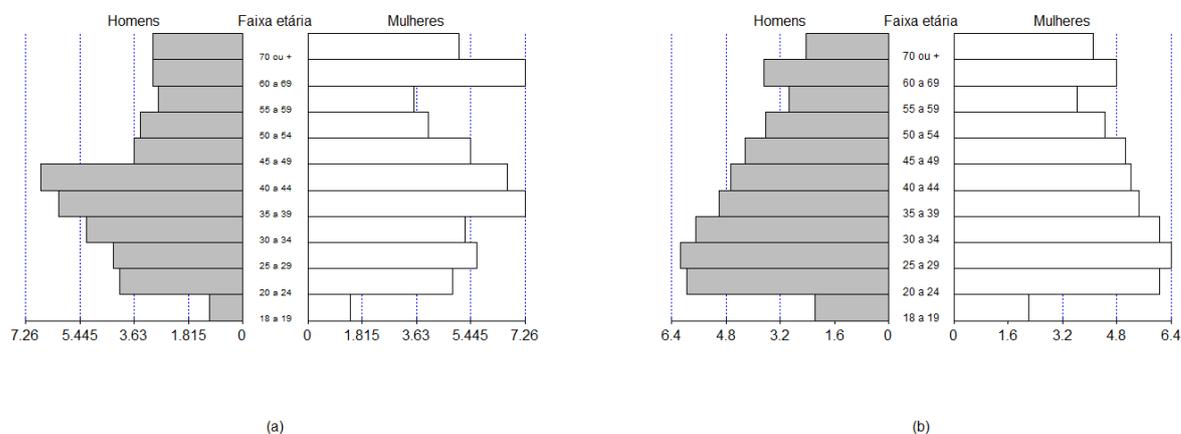


Figura XXII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Florianópolis/SC

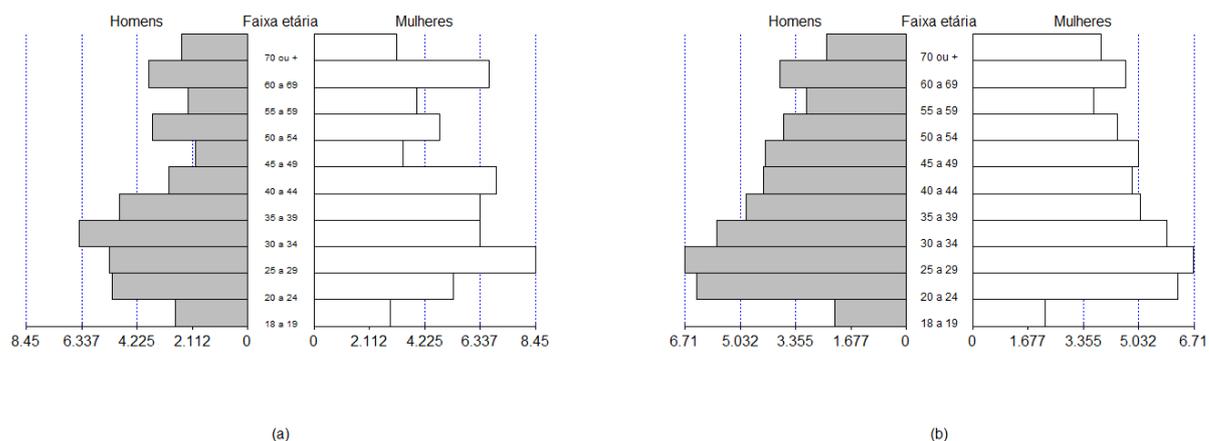
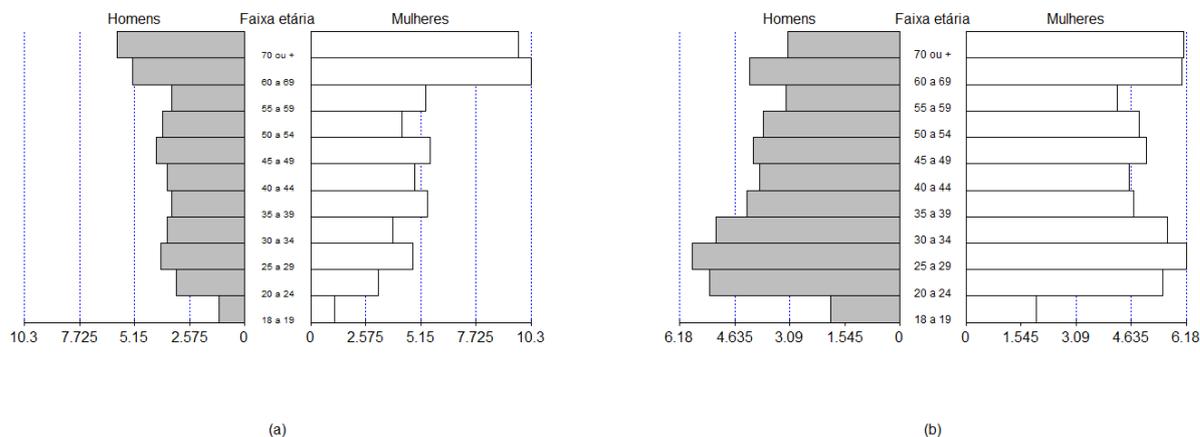


Figura XXIII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Porto Alegre/RS



REGIÃO CENTRO OESTE

Figura XXIV: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Campo Grande/MS

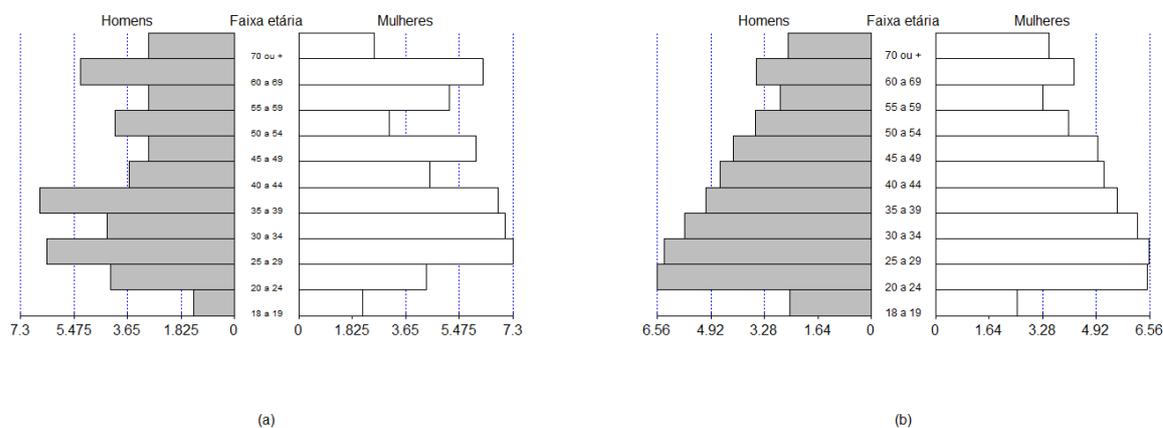


Figura XXV: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Cuiabá/MT

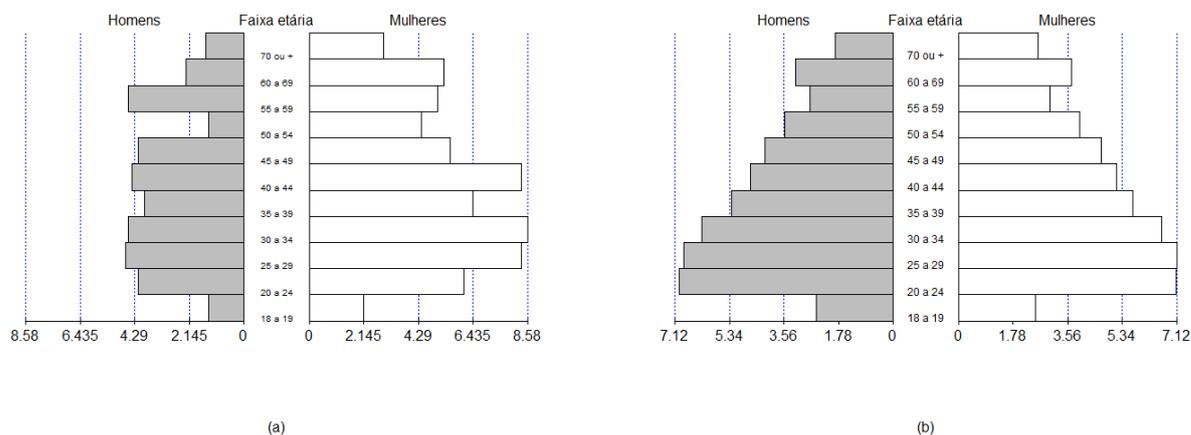


Figura XXVI: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Goiânia/GO

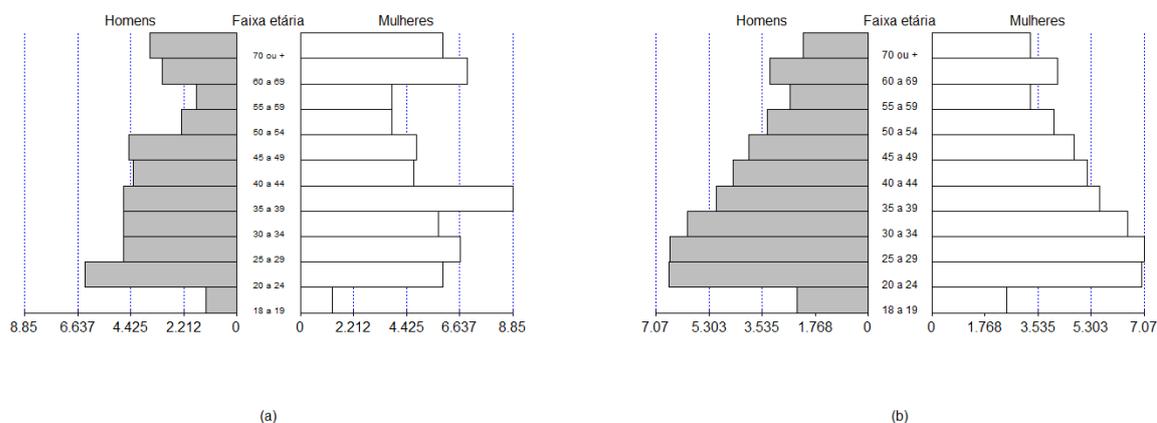
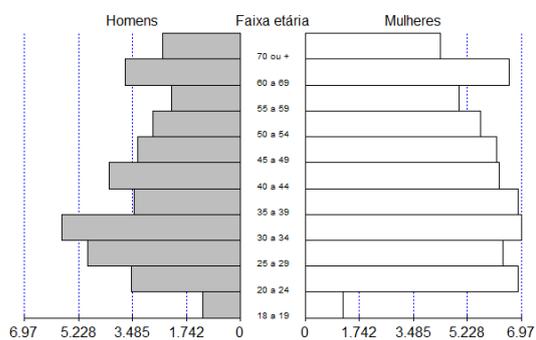
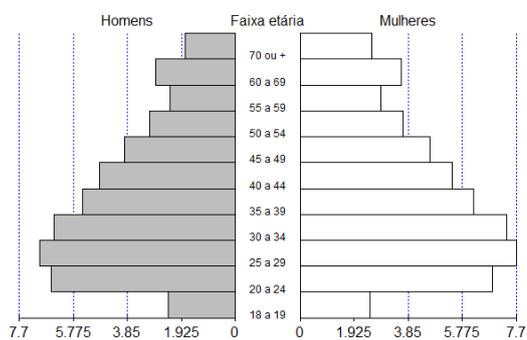


Figura XXVII: Distribuição (%) da amostra do Inquérito domiciliar (a) e da População – Censo demográfico de 2010 - (b) segundo sexo e faixa etária em Brasília/DF



(a)



(b)