



UNIVERSIDADE FEDERAL DE PELOTAS
FACULDADE DE MEDICINA



PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

**PESQUISA EPIDEMIOLÓGICA BASEADA NA WEB:
ESTRATÉGIAS E CARACTERÍSTICAS PESSOAIS ASSOCIADAS
À ADESÃO E PARTICIPAÇÃO DE ADULTOS JOVENS**

TESE DE DOUTORADO

CAUANE BLUMENBERG

Pelotas, RS

2018

CAUANE BLUMENBERG

**PESQUISA EPIDEMIOLÓGICA BASEADA NA WEB:
ESTRATÉGIAS E CARACTERÍSTICAS PESSOAIS ASSOCIADAS
À ADESÃO E PARTICIPAÇÃO DE ADULTOS JOVENS**

Tese apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal de Pelotas, como requisito para a conclusão do Doutorado em Epidemiologia.

Prof. Aluísio J D Barros

Orientador

Pelotas, RS

2018

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

B658p Blumenberg, Cauane

Pesquisa epidemiológica baseada na web : estratégias e características pessoais associadas à adesão e participação de adultos jovens / Cauane Blumenberg ; Aluísio Jardim Dornellas de Barros, orientador. — Pelotas, 2018.

179 f. : il.

Tese (Doutorado) — Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal de Pelotas, 2018.

1. Epidemiologia. 2. Metodologia de pesquisa. 3. Recrutamento de participantes. 4. Internet. 5. Taxa de resposta. I. Barros, Aluísio Jardim Dornellas de, orient. II. Título.

CDD : 614.4

CAUANE BLUMENBERG

**PESQUISA EPIDEMIOLÓGICA BASEADA NA WEB:
ESTRATÉGIAS E CARACTERÍSTICAS PESSOAIS ASSOCIADAS
À ADESÃO E PARTICIPAÇÃO DE ADULTOS JOVENS**

Banca examinadora:

Professor Aluísio J D Barros (orientador)

Universidade Federal de Pelotas

Professor Adenauer Correa Yamin (examinador)

Universidade Federal de Pelotas

Professor Fernando César Wehrmeister (examinador)

Universidade Federal de Pelotas

Professor Rafael Moreira Claro (examinador)

Universidade Federal de Minas Gerais

“Primeiro, lembrem-se de olhar para as estrelas lá no alto e não para seus pés lá embaixo. Dois, nunca desistam do seu trabalho. O trabalho lhe dá sentido e propósito, e a vida é vazia sem isso. Três, se você for afortunado a ponto de encontrar amor, lembre-se de que ele está ali e nunca o jogue fora.”

Stephen Hawking

Agradecimentos

Há sete anos atrás o ainda professor, e agora reitor, Pedro Hallal, me convidava para integrar a equipe do Centro de Pesquisas Epidemiológicas. Na época eu ainda era um estudante de Ciência da Computação, que não tinha a mínima ideia do que fazer quando terminasse o curso de graduação. Aquela oferta de emprego feita durante os tradicionais churrascos de quinta-feira parecia um sonho se tornando realidade, e na segunda-feira seguinte eu já estava lá, trabalhando no meu primeiro emprego. Sou muito grato ao Pedrinho pela oportunidade que me deu, pois fez com que eu conhecesse e fizesse parte de um dos mais renomados centros de pesquisas em saúde do Brasil e do mundo.

Já com algum tempo de “casa”, passei a conhecer os pesquisadores que tornavam o Programa de Pós-graduação em Epidemiologia referência mundial. Pesquisadores estes que são referência em diferentes áreas do conhecimento, mas que tem uma coisa em comum: a amizade. Obrigado pela recepção, parceria e simplicidade de cada um de vocês! Obrigado pela confiança, pelas risadas e pelos trabalhos super complicados, que a gente tinha gosto de fazer acontecer. Vocês são o exemplo perfeito de como podemos atingir o ápice em nossas carreiras, sem deixar de lado o nosso lado humano. Muito obrigado por todos os ensinamentos e pela amizade de cada um de vocês!

Falando em pesquisador amigo, não poderia deixar de dedicar um agradecimento especial ao Aluísio. Lá em 2014, quando eu ainda pensava em sair do Brasil para fazer doutorado em Computação, ele surgiu com uma ideia: “Cauane, por que você não faz doutorado aqui com a gente?”. Minha primeira reação foi de espanto. Sair da área de Ciência da Computação e migrar para a Epidemiologia? Pensando bem, por que não? O Aluísio confiou em mim, me deu suporte para que esse projeto maluco de integrar computação e epidemiologia desse certo. E parece que deu... Aluísio, tenho uma profunda admiração por ti. Admiro a tua coragem de inovar e de sempre se reinventar. Admiro o apoio que tu dá para os teus orientandos. E agradeço, de coração, por teres instigado um *nerd* da Computação a fazer doutorado em Epidemiologia. Afinal, “chega daqueles estudos de fatores associados de x com y”, não é?

Não poderia deixar de agradecer a todos os colegas do Centro de Pesquisas Epidemiológicas, que fazem com que a máquina funcione a todo vapor! Tenho orgulho de dizer que a minha relação com cada um de vocês é muito boa e saudável. Muito obrigado pelos cafés, festas e conversas que foram muito importantes para a minha formação. Um agradecimento muito especial aos meus irmãos Tiago e Mateus. Meus companheiros de sala, de churrasco, de risadas, de jogos e de tudo que bons amigos podem desfrutar juntos. Esses dois sempre acreditaram em mim e estiveram do meu lado. Espero ter dado para vocês pelo menos um terço da alegria que vocês me dão. Valeu, TI CPE!

Gostaria de agradecer também à equipe que fez parte do projeto *coortesnaweb*. Muito obrigado à Cíntia, que sempre conseguiu solucionar os meus problemas de combinação de cores. Obrigado por ter me emprestado esse teu talento enorme para o design gráfico. Como te disse pessoalmente, a plataforma ficou linda! Também queria agradecer à Patrícia, que foi minha fiel escudeira durante todo o processo de pesquisa da *coortesnaweb*. Mais que a secretária do projeto, ela foi fundamental para que tudo desse certo. Pati, o “chefinho” (como ela carinhosamente me chama) te agradece de coração!

Em 2017 tive a oportunidade de passar seis meses em Turim (Itália), como parte de um doutorado sanduíche. Foi um período de muito aprendizado e trabalho. Porém, o que fica na memória foram as amizades que construí por lá. Queria agradecer, e muito, ao Lorenzo, Costanza, Maja, Dani, Lele e Mila, que me fizeram sentir em casa! Grazie a tutti!

Falando em amigos, um obrigado especial para a galera do UFE, aqueles amigos de fé que cresceram comigo e continuam presentes até hoje. Obrigado pelos churrascos de quarta-feira, pelas risadas e pela força que vocês me dão. A vida é muito mais leve quando a gente tem amigos que a gente gosta do nosso lado!

Obrigado à minha família, a pedra fundamental de tudo isso. Gostaria de agradecer aos meus avós, tios, primos, sogros, cunhados. Vocês foram muito importantes em todo esse processo! Um obrigado especial aos meus pais, Luis Fernando e Viviane, e à minha irmã, Catuama. Eles sempre estiveram presentes com palavras de apoio e carinho. São aquelas pessoas que conseguem ver o lado bom e positivo de todo problema que a gente encontra. Obrigado por apoiarem minhas decisões, por mais malucas que elas fossem. Obrigado por me fazer seguir em frente. Amo todos vocês!

Por fim, um obrigado com muito amor à minha companheira de vida, Rafaela. É muito bom saber que posso contar contigo. É muito bom te ter do meu lado, dividindo o nosso lar, os estudos e todas as aventuras malucas que a gente inventa. Obrigado por ser exatamente como és, uma pessoa de uma bondade extrema, e que está sempre disposta a progredir e melhorar. Assim como no mestrado, no doutorado tivemos que lidar com um período de distância, mas nosso amor é maior que qualquer obstáculo. E mais, nosso tempo desfrutando a Itália, juntos, foi fora da realidade! Minha caminhada é muito mais fácil contigo do meu lado! Que nossa vida siga sendo leve e divertida, como só nós conseguimos fazer. Te amo, muito!

Resumo

Blumenberg, Cauane. **Pesquisa epidemiológica baseada na web: estratégias e características pessoais associadas à adesão e participação de adultos jovens**. Tese de doutorado. Programa de Pós-graduação em Epidemiologia. Universidade Federal de Pelotas; 2015.

Nos últimos anos, as pesquisas epidemiológicas têm sofrido com a constante redução das taxas de resposta. Este problema pode comprometer a validade dos achados em virtude da maior probabilidade de ocorrência do viés de não resposta. Diversos fatores podem comprometer a taxa de resposta de estudos epidemiológicos, incluindo o método utilizado para coletar os dados (artigo 1). Aproveitando o avanço da cobertura e do acesso à internet, muitos estudos em países de alta renda estão utilizando este meio de comunicação para o levantamento de dados epidemiológicos, porém a condução de pesquisas epidemiológicas em países de baixa ou média renda ainda é pouco explorada. Com isso, o primeiro objetivo desta tese foi desenvolver uma plataforma, chamada *coortesnaweb*, para a condução de pesquisas epidemiológicas através da internet entre os membros da Coorte de Nascimentos de Pelotas de 1993. A *coortesnaweb* compreende todas as etapas do processo de levantamento de dados epidemiológicos, como recrutamento, coleta de dados e retorno de resultados, além de funções extra como estratégias de gamificação. Através desta plataforma, foram analisadas a eficiência e as desigualdades de três métodos de recrutamento *online* e gratuitos para convidar os membros da Coorte 1993 para se registrar na *coortesnaweb* (artigo 2). Os membros elegíveis foram alocados aleatoriamente para serem recrutados por e-mail, mensagens de Facebook ou Whatsapp. A taxa de recrutamento foi calculada para cada um dos métodos e de acordo com as características sócio demográficas dos membros elegíveis. Também foram analisadas as desigualdades absoluta e relativa por método de recrutamento e de acordo com escolaridade e renda. A taxa de recrutamento geral foi 26,8%, sendo o Facebook o método de recrutamento mais eficiente (taxa de recrutamento = 30,6%). As maiores desigualdades no recrutamento, com relação à renda e escolaridade, foram observadas utilizando e-mail. O Facebook foi o método que apresentou as menores desigualdades. Também através da plataforma *coortesnaweb*, foram analisadas a influência do tamanho do questionário e da frequência de envio dos lembretes sobre a taxa de resposta a cinco questionários aplicados através da *web*, utilizando um desenho fatorial 2x2 (artigo 3). No momento do registro na plataforma *coortesnaweb*, os participantes foram aleatoriamente alocados para responder a questionários longos ou curtos, e para receber lembretes com alta ou baixa frequência. Além da taxa de resposta por questionário, também foi analisado o número total de questionários respondidos de acordo com as características sócio demográficas dos participantes. Considerando os cinco questionários, a taxa de resposta média foi 54,3%. O envio mais frequente de lembretes foi associado a uma maior taxa de resposta para os primeiros dois questionários aplicados, enquanto o tamanho dos questionários não teve influência na taxa de

resposta. Mulheres e participantes que estudaram 12 anos ou mais tiveram, respectivamente, 13,0% (IC95% =1,03; 1,23) e 28,0% (IC95% =1,09; 1,50) maior risco de responder um questionário a mais, comparando com as categorias de referência. Em geral, nossos resultados mostram que é viável conduzir pesquisas epidemiológicas através da internet no contexto brasileiro, mesmo que a cobertura de internet não seja universal. Ainda assim, a maior participação de alguns grupos sócio demográficos deve ser levada em consideração na interpretação dos achados produzidos por pesquisas *web*, principalmente quanto às estimativas de prevalência.

Palavras-chave: Epidemiologia; Metodologia de pesquisa; Recrutamento de participantes; Internet; Taxa de resposta

Abstract

Blumenberg, Cauane. **Pesquisa epidemiológica baseada na web: estratégias e características pessoais associadas à adesão e participação de adultos jovens**. Thesis (PhD). Postgraduate Program in Epidemiology. Universidade Federal de Pelotas; 2015.

In the recent years, epidemiologic surveys are being compromised by the reduction on the response rates. This problem can jeopardize the validity of the findings due to the higher probability of occurrence of the nonresponse bias. Several factors contribute to the declining response rates of epidemiologic studies, including the data collection method used (paper 1). With the growing internet coverage and access around the globe, some studies conducted in high-income countries are using the internet to collect epidemiologic data. However, using the internet to conduct surveys in low- and middle-income countries is still in its infancy. Thus, the first objective of this thesis was to develop the *coortesnaweb*, a platform used to conduct internet-based epidemiologic surveys with the members of the 1993 Pelotas birth cohort. The *coortesnaweb* cover all the steps through recruitment, data collection and dissemination of results, including some extra features such gamification strategies. Using the platform, we analyzed the efficiency and the inequalities of three online and free-to-use recruitment methods, used to invite the members of the 1993 birth cohort to register into the platform (paper 2). The eligible cohort members were randomly allocated to be recruited using e-mail, Facebook or Whatsapp messages. The recruitment rate was calculated for each method and according to socio demographic characteristics of the sample. The absolute and relative inequalities on recruitment were also analyzed according to schooling and socioeconomic position. The overall recruitment rate was 26.8%, being Facebook the most efficient recruitment method (recruitment rate = 30.6%). The highest inequalities on recruitment according to schooling and socioeconomic position were seen using e-mail as recruitment method. The Facebook was the methods that presented the lowest inequalities. Also using the *coortesnaweb* platform, we analyzed the influence of questionnaire length and frequency of reminders on the response rates of five questionnaires using a 2x2 factorial design (paper 3). At the moment of registration into the *coortesnaweb* platform, the participants were randomly allocated to respond to short or long questionnaires, and to receive reminders with low or high frequency. Besides calculating the response rate for each questionnaire, we also analyzed the total number of questionnaires responded according to the socio demographic characteristics of the sample. Considering the five questionnaires together, the mean response rate was 54.3%. Sending reminder more frequently was positively associated with the response rates of the first two questionnaires, but the questionnaire length had no influence on the response rates. Women and participants that studied for 12 years or more had, respectively, 13.0% (95%CI = 1.03, 1.23) e 28.0% (95%CI = 1.09, 1.50) increased risk of responding an additional questionnaire compared to the reference

groups. Our findings show that it is feasible to conduct internet-based epidemiologic surveys in the Brazilian context, even though the internet coverage is not universal. Additionally, the higher participation of some specific groups should be considered in order to interpret the findings from web-based survey, especially regarding prevalence estimates.

Keywords: Epidemiology; Survey methodology; Recruitment of survey participants; Internet; Web surveys; Response rate

Sumário

1 – Projeto de Pesquisa.....	13
2 – Ajustes ao projeto durante o curso do trabalho	65
3 – Trabalho de campo	68
4 – Artigo 1	81
5 – Artigo 2	96
6 – Artigo 3	116
7 – Nota de imprensa.....	132
8 – Anexos.....	134

Apresentação

A presente tese de doutorado, exigência para obtenção do título de Doutor em Epidemiologia pelo Programa de Pós-Graduação Em Epidemiologia da Universidade Federal de Pelotas, é organizada da seguinte forma.

1. Projeto de pesquisa apresentado no dia 13 de setembro de 2016 como requisito para qualificação de doutorado. Este projeto conta com as modificações sugeridas pela banca avaliadora, composta pelos professores Adenauer Correa Yamin e Fernando César Wehrmeister.
2. Relatório com ajustes referentes à logística e análise da proposta original do projeto de pesquisa.
3. Descrição do trabalho de campo ao longo do período de doutoramento, incluindo a participação em dois acompanhamentos das Coortes de Nascimentos de Pelotas, o período de doutoramento sanduíche junto ao grupo de pesquisa da Universidade de Turim (Itália) e também a condução do acompanhamento dos membros da Coorte 1993 através da plataforma coortesnaweb.
4. Artigo de revisão intitulado “*Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature*”, publicado no periódico *International Journal of Public Health*.
5. Artigo original 1, intitulado “*How different online recruitment methods impact on recruitment rates for the web-based coortesnaweb project: a randomized trial*”, a ser submetido para o periódico *International Journal of Epidemiology*.
6. Artigo original 2, intitulado “*The role of questionnaire length and frequency of reminders on response rates to a web-based epidemiological study: a randomized trial*”, a ser submetido para o periódico *American Journal of Epidemiology*.
7. Comunicado para a imprensa.

1 – Projeto de Pesquisa

Universidade Federal de Pelotas
Faculdade de Medicina
Departamento de Medicina Social
Programa de Pós-Graduação em Epidemiologia

**Pesquisa epidemiológica baseada na *web*: estratégias e
características pessoais associadas à adesão e participação de adultos
jovens**

Cauane Blumenberg S.

**Projeto de pesquisa apresentado ao
Programa de Pós-Graduação em Epidemiologia
da Universidade Federal de Pelotas, como
requisito para qualificação de doutorado.**

Orientador: Prof. Dr. Aluísio J. D. Barros

Pelotas, setembro de 2016

SUMÁRIO

RESUMO	4
ARTIGOS PROPOSTOS	5
ABREVIATURAS	6
DEFINIÇÃO DE TERMOS	7
1. INTRODUÇÃO.....	9
2. JUSTIFICATIVA.....	12
3. REVISÃO DE LITERATURA	15
3.1 Taxa de resposta de questionários <i>web</i> conforme o método de recrutamento	23
3.2 Taxa de resposta de questionários <i>web</i> comparados a outros métodos de coleta.....	24
3.3 Taxa de resposta de questionários <i>web</i> conforme o fornecimento de incentivo ..	25
4. OBJETIVOS.....	26
4.1 Objetivos específicos.....	26
5. MATERIAIS E MÉTODOS	27
5.1 Coorte de nascimentos de 1993 de Pelotas	27
5.2 Logística do acompanhamento de 2015/2016.....	28
5.3 População alvo	29
5.3.1 Critérios de inclusão	29
5.4 Instrumentos	29
5.5 Logística.....	30
5.6 Desenvolvimento da plataforma <i>web</i>	31
5.7 Sistema de recompensas e incentivos.....	32
5.8 Controle de qualidade.....	32
5.9 Planos de análise	33
5.9.1 Variáveis de exposição consideradas nas análises.....	33
5.9.2 Representatividade dos elegíveis e dos participantes da plataforma <i>web</i>	35
5.9.3 Taxa de contato de acordo com o método de recrutamento	35
5.9.4 Influência da estratégia de coleta de dados e frequência de lembretes sobre a taxa de resposta.....	36
6. ASPECTOS ÉTICOS	37
7. CRONOGRAMA	38
8. DIVULGAÇÃO DOS RESULTADOS	39

9. FINANCIAMENTO.....	40
REFERÊNCIAS	41
ANEXO 1	49

RESUMO

As pesquisas epidemiológicas sempre contribuíram para importantes achados na área de saúde. Para que estes achados possam ser traduzidos em políticas públicas, é importante que as pesquisas epidemiológicas sejam conduzidas com rigor metodológico em vista de maximizar sua validade. Um fator que pode comprometer a validade desses achados é o constante declínio nas taxas de resposta, influenciado por diversos fatores, entre eles o método de coleta de dados. Com o aumento do acesso à internet, bem como do seu número de usuários, conduzir pesquisas epidemiológicas via *web* pode ser uma boa alternativa. Apesar disso, novos desafios surgem com essa nova metodologia de pesquisa, como o melhor método de recrutamento a ser utilizado, questões de representatividade das amostras, desenho dos questionários e maneiras eficientes de motivar a participação dos indivíduos nas pesquisas aplicadas via *web*. Levando isso em consideração, o objetivo deste projeto é desenvolver uma plataforma que contenha o ambiente completo para condução de pesquisas epidemiológicas via *web*, visando investigar a viabilidade de sua condução e explorar estratégias para maximizar as taxas de contato e resposta dos questionários aplicados via *web*. Serão convidados a participar da plataforma os membros da coorte de nascimentos de 1993 de Pelotas que participaram do acompanhamento dos 22/23 anos, e relataram ter acesso à internet. Para investigar a influência de diferentes metodologias sobre as taxas de contato e resposta, serão estudados diferentes métodos de recrutamento, tamanhos de questionário, frequências de lembrete e fornecimento de incentivos. Somado a isso, serão estudados fatores sócio-demográficos associados à participação dos membros da coorte na plataforma *web* a ser desenvolvida.

ARTIGOS PROPOSTOS

Artigo de Revisão: Taxas de resposta de questionários aplicados via *web* comparadas a outros métodos de coleta para a condução de estudos epidemiológicos: uma revisão da literatura.

Artigo Original 1: A influência do método de recrutamento sobre a taxa de contato e resposta a um ambiente de pesquisa epidemiológica baseada na *web*.

Artigo Original 2: Comparação de diferentes estratégias de frequência de lembretes e tamanho de questionários em relação à taxa de resposta e retenção ao longo de seis meses de questionários aplicados via *web*.

ABREVIATURAS

CEP – Comitê de Ética e Pesquisa

CSS – *Cascading Style Sheets*

CSV – *Comma Separated Values*

DAD – Discagem Aleatória de Dígitos

DAG – *Directed Acyclic Graph*

HTML – *Hypertext Markup Language*

IP – *Internet Protocol*

OMS – Organização Mundial da Saúde

PDA – *Personal Digital Assistant*

PHP – *Hypertext Processor*

PNAD – Pesquisa Nacional por Amostra de Domicílios

RR5 – *Response Rate 5*

SGBD – Sistema de Gerenciamento de Banco de Dados

SMS – *Short Message Service*

WWW – *World Wide Web*

WebSM – *Web Survey Methodology*

DEFINIÇÃO DE TERMOS

Domínio: ferramenta que funciona como uma espécie de “máscara” para o endereço IP, transformando o endereço IP em um nome mais fácil de ser lembrado, assim facilitando o acesso aos servidores da internet (1).

Endereço IP: número único que identifica o endereço de cada servidor conectado à internet (1).

Firewall: ferramenta que filtra as comunicações e transações entre a rede local e a rede pública (internet), barrando tentativas de ataque ou acesso indevido aos dados feitas através da rede pública (2).

Framework: conjunto de códigos computacionais organizados em uma determinada estrutura e que podem ser utilizados para o desenvolvimento de aplicações da internet. São aplicações pré-construídas, as quais podem ser moldadas de acordo com a necessidade do programador (3).

Gamificação: termo que se refere ao emprego de técnicas relacionadas a jogos de videogame, como *rankings*, conquistas e desafios, com o objetivo de melhorar a experiência e o engajamento dos usuários em serviços e aplicações não relacionadas a videogames (4).

Linguagem de estilo: linguagem computacional de apresentação que define um conjunto de regras que associam propriedades e estilos com estruturas de um documento de texto, definindo assim como estas estruturas devem ser apresentadas (5).

Linguagem de marcação: linguagem computacional que faz uso de etiquetas para marcar a estrutura e o funcionamento de um documento de texto. O uso dessas etiquetas facilita o processo de interpretação do documento de texto por parte do navegador *web* (1).

Linguagem de *script*: linguagens que são geralmente executadas do interior de outras linguagens de programação, e que são usadas para desempenhar tarefas específicas em tempo real (6).

Navegador *web*: *software* que auxilia os usuários da internet a obter informações de servidores conectados à internet. Com o uso de um endereço IP ou domínio, o navegador *web* é capaz de se conectar ao servidor correto e recuperar o conteúdo ali presente (1).

Página *web*: arquivo de texto hospedado em um servidor da internet que se relaciona com outros arquivos e páginas *web* que contém código computacional ou gráfico (7).

Plataforma Web: espécie de *website*, o qual é definido como um conjunto de páginas *web* interconectadas e hospedadas em um mesmo servidor da internet, e que são desenvolvidas e mantidas por uma pessoa, grupo ou organização (7).

Taxa de contato: se trata de uma métrica calculada para a plataforma *web*, sendo obtida pela divisão entre o número de indivíduos que se cadastram na plataforma *web* e o total de indivíduos elegíveis (8).

$$\text{Taxa de contato} = \frac{\text{n}^{\circ} \text{ de indivíduos cadastrados na plataforma}}{\text{n}^{\circ} \text{ de indivíduos elegíveis para cadastro}}$$

Taxa de resposta: se trata de uma métrica calculada para cada questionário, sendo obtida através da divisão entre o número de indivíduos que forneceram questionários completos e o número de indivíduos elegíveis (9). Considerando que a amostra desse projeto é conhecida, onde se tem informação sobre seu acesso à internet, será considerada a taxa de resposta *Response Rate 5* (RR5) como definida pela Associação Americana de Pesquisa para Opinião Pública (10).

$$\text{Taxa de resposta} = \frac{\text{n}^{\circ} \text{ de respondentes}}{\text{n}^{\circ} \text{ de indivíduos elegíveis a responder}}$$

Taxa de retenção: considerando que para cada participante recrutado serão aplicados no mínimo dois questionários, a taxa de retenção é calculada pelo número de participantes que completam no mínimo dois questionários, dividido pelo número de participantes que completam apenas um questionário (11).

$$\text{Taxa de retenção} = \frac{\text{n}^{\circ} \text{ de indivíduos com dois questionários completos}}{\text{n}^{\circ} \text{ de indivíduos com um questionário completo}}$$

1. INTRODUÇÃO

Ao longo dos anos, as pesquisas epidemiológicas têm contribuído com importantes achados relacionados à etiologia de doenças, ações de saúde pública, suporte à prática clínica, entre outros (12–14). Para que estes achados possam ser transformados em ações de saúde pública ou possam influenciar a prática clínica, é importante que os estudos sejam conduzidos com rigor metodológico. Para isso, todo o processo da pesquisa epidemiológica deve ser cuidadosamente pensado para minimizar barreiras que possam comprometer a qualidade dos estudos. Uma dessas barreiras, que afeta inclusive estudos epidemiológicos de alta qualidade metodológica, é o constante declínio nas taxas de resposta, o qual é estimado em cerca de um ponto percentual ao ano (15), sendo que alguns estudos relatam que esse declínio pode ser quase duas vezes maior (16). Estudos com baixas taxas de resposta são mais suscetíveis ao viés de não resposta, o qual ocorre quando os indivíduos que não participaram do estudo diferem dos participantes em termos relevantes ao objetivo do estudo (9), colocando em xeque a validade dos achados. Os motivos para o declínio nas taxas de resposta são variados, entre eles o método utilizado para coletar os dados (17).

Um dos métodos de coleta de dados epidemiológicos mais utilizados são entrevistas face-a-face, onde um entrevistador aplica a pesquisa diretamente ao participante. Por um lado, a presença de um entrevistador pode motivar os entrevistados a participar do estudo. Mas por outro, isso pode fazer com que os entrevistados não fiquem a vontade para fornecer respostas válidas para assuntos sensíveis ou que não são socialmente aceitáveis, causando viés de informação ou até mesmo se recusando a responder as perguntas, reduzindo assim as taxas de resposta (18). Outra desvantagem desse método de coleta é a necessidade de que entrevistadores e participantes estejam simultaneamente no mesmo local. Isso implica em questões logísticas e financeiras, as quais envolvem necessidade de deslocamento e disponibilidade de horário por parte dos participantes e da equipe de pesquisa. Com isso, essa metodologia de pesquisa muitas vezes é prejudicada pela restrição de cobertura imposta pelo espalhamento urbano e pela dificuldade de contato com os participantes em suas próprias residências, hoje caracterizadas por condomínios que limitam o acesso dos entrevistadores, comprometendo as taxas de resposta (19). Somado a isso, os altos custos decorrentes do salário de entrevistadores, necessidade de transporte, impressão de questionários e do longo tempo de coleta de dados podem inviabilizar a condução de estudos epidemiológicos (20).

Uma alternativa às pesquisas presenciais, que surgiu como opção para contornar problemas de cobertura e reduzir custos, foi o uso de questionários via correio. Nesse caso, um questionário impresso é enviado à residência dos indivíduos elegíveis, os quais devem

preenchê-lo e retorná-lo ao remetente. Comparado às pesquisas face-a-face, questionários via correio eliminam os custos referentes ao salário dos entrevistadores e do deslocamento, mas por outro lado agregam custos referentes à postagem. Um dos problemas dessa metodologia é a dificuldade em controlar qual dos residentes responde à pesquisa, principalmente quando são aplicados questionários sem identificação pessoal (21). Esse método também sofre com as mudanças de endereço residencial, e também com o longo período entre o envio dos questionários aos participantes e a obtenção das respostas, muitas vezes influenciada pela falta de tempo ou acesso a um serviço postal (22,23).

Outro método de coleta muito utilizado foram questionários aplicados através do telefone. Até a década de 1990, essa foi considerada uma boa metodologia de pesquisa, afinal a telefonia fixa era uma tecnologia com grande penetração populacional, capaz de gerar resultados comparáveis aos obtidos pelo método presencial e convencional de pesquisa (24). Embora as pesquisas por telefone fixo fossem restritas aos domicílios com uma linha de telefonia fixa, sua comparabilidade com as pesquisas face-a-face foi muito beneficiada pelo uso de um sistema de discagem aleatória de dígitos (DAD), o qual permite realizar uma amostragem aleatória entre os domicílios (25). Porém, nos últimos anos percebeu-se um declínio no número de linhas de telefonia fixa, muito influenciado pela migração para a telefonia móvel. E assim, a condução de pesquisas através do telefone fixo encontrou barreiras importantes, as quais impactaram as taxas de resposta dos estudos conduzidos através deste método (26,27).

Visando explorar a disseminação da telefonia móvel, algumas pesquisas passaram a ser conduzidas através de mensagens de texto, conhecidas como *Short Message Service* (SMS). O uso desse método de coleta foi motivado pela rapidez em enviar questionamentos e receber respostas, pela possibilidade de confirmar o recebimento de mensagens e pelo baixo custo do SMS (28,29). Apesar do baixo custo, o envio das respostas por parte dos entrevistados requer que eles arquem com as tarifas de envio, o que pode comprometer as taxas de resposta. Além disso, esse método de coleta possui limite de caracteres por mensagem, o qual é diferente para cada operadora de telefonia móvel mas que suporta em média 160 caracteres (30). Esse limite impossibilita a condução de questionários mais longos, e assim essa metodologia de pesquisa foi pouco empregada em estudos epidemiológicos.

Embora novos métodos de coleta de dados tenham surgido, a condução de entrevistas face-a-face utilizando questionários impressos nunca deixou de ser empregada. Porém, com a evolução da computação estes questionários começaram a ser desenvolvidos de maneira eletrônica, sendo aplicados através de assistentes pessoais digitais (PDAs), *tablets*, computadores ou outros dispositivos eletrônicos. Com isso, a qualidade do dado coletado aumentou devido a possibilidade de checar dados inconsistentes em tempo de entrevista (31).

Outra vantagem foi a redução no tempo para condução das entrevistas, afinal questões que não devem ser aplicadas são automaticamente escondidas. Tais questionários eletrônicos também podem reduzir custos referentes a impressão, porém não eliminam a necessidade de empregar um entrevistador para aplicar as perguntas, ou a necessidade de deslocamento do entrevistador ou participante para realização da entrevista (32).

Um método de coleta de dados que é capaz de se beneficiar das vantagens dos questionários eletrônicos, e que ainda é capaz de eliminar os custos referentes ao salário de entrevistadores e de deslocamento, é a utilização de pesquisas baseadas na *web*. Esta metodologia de pesquisa se torna viável devido ao aumento do número de usuários de internet, o qual está em constante crescimento, e se estima que em 2016 quase metade da população mundial terá acesso à internet (33). Outra vantagem das pesquisas baseadas na *web* é a maior flexibilidade de horário para o preenchimento dos questionários, uma vez que não é necessária a presença de um entrevistador em virtude das pesquisas serem auto-aplicadas pelo participante (34). Em contrapartida, a participação nos questionários aplicados via *web* fica limitada aos indivíduos que tem acesso à internet, o que pode acarretar em problemas de representatividade da amostra estudada (35,36). Embora a literatura internacional já descreva as vantagens e desvantagens da coleta de dados epidemiológicos via *web* (37–40), essa metodologia de pesquisa ainda é pouco empregada e estudada em países de baixa e média renda, em especial no Brasil (41,42). Em virtude disso, este projeto visa investigar a viabilidade em conduzir pesquisas epidemiológicas baseadas na *web* e explorar estratégias para maximizar as taxas de contato e resposta. Somado a isso, serão estudados fatores sócio-demográficos associados à participação de adultos jovens membros da coorte de nascimentos de 1993 de Pelotas em uma plataforma *web* desenvolvida para condução de pesquisas epidemiológicas.

2. JUSTIFICATIVA

Utilizar a internet como metodologia para condução de pesquisas epidemiológicas traz diversos benefícios, como rapidez e consistência no processo de coleta de dados, possibilidade de participação assíncrona e redução de custos (43,44). Essas vantagens têm contribuído para o aumento do número de pesquisas epidemiológicas que utilizam a internet como plataforma de pesquisa. Entre os anos de 2007 e 2008, cerca de 500 ensaios clínicos mencionavam o uso da internet como plataforma de pesquisa, em 2014 este número subiu para aproximadamente 2.000 estudos. Porém, este crescimento é percebido principalmente nos estudos conduzidos em países de alta renda, em países de baixa e média renda esta metodologia de pesquisa ainda é pouco explorada (45).

Por se tratar de um tema emergente, é possível observar na literatura resultados inconclusivos sobre a utilização da internet como plataforma de pesquisas epidemiológicas. Uma das principais discussões é com relação às taxas de resposta, onde alguns estudos relatam vantagem para os demais métodos de pesquisa (entrevistas face-a-face, por telefone, ou via correio) (46–48), outros não observam diferenças significativas (37,49,50), e outros indicam que utilizando a internet é possível atingir taxas de resposta superiores (39,51,52). Em vista disso, existe espaço para mais pesquisas nessa área, principalmente no contexto brasileiro onde poucos estudos conduzem a coleta de dados epidemiológicos via *web*.

Analisar as taxas de resposta em pesquisas epidemiológicas é muito importante, afinal a probabilidade de viés aumenta conforme o número de não respondentes (9). Isso pode afetar a validade dos estudos, logo é relevante estudar os fatores que podem influenciar as taxas de resposta de pesquisas via internet. A literatura relata que o tipo de recrutamento, a elaboração dos questionários, a frequência de contato e o fornecimento de incentivos para participação estão associados à taxa de reposta. Apesar disso, muitos pesquisadores realizam estudos pontuais, onde essas questões são estudadas individualmente (53). Embora isso seja de grande valia, se torna difícil determinar a interação entre estes fatores. Assim, torna-se desejável a implementação de uma plataforma unificada em que todos esses fatores possam ser avaliados simultaneamente.

Mesmo que o uso da internet traga diversas vantagens ao processo de pesquisa epidemiológica, utilizá-la como único método para coleta de dados pode encontrar algumas limitações. O acesso a um dispositivo conectado à internet se faz necessário para garantir a participação no estudo. Em países de alta renda, como os Estados Unidos, isso pode não ser uma barreira, afinal em 2014 cerca de 87% dos adultos declaravam ser usuários de internet (33). Por outro lado, em países de média renda, como o Brasil, o acesso à internet não é pervasivo.

Conforme mostra a Figura 1, em 2014 pouco mais da metade dos brasileiros haviam acessado a internet nos últimos 12 meses (54).

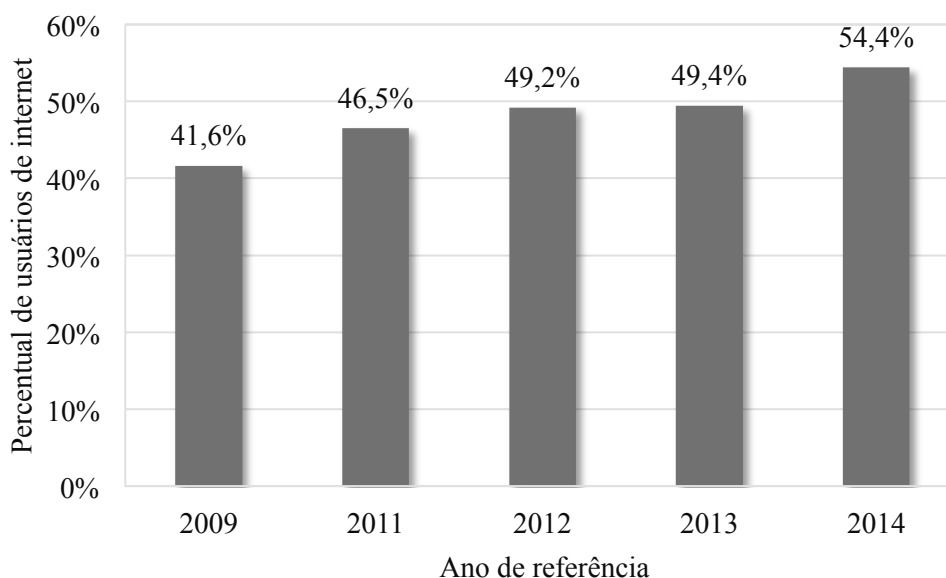


Figura 1: Percentual de brasileiros que acessaram a internet no período de um ano anterior à realização da pesquisa. Dados extraídos da Pesquisa Nacional por Amostra de Domicílio (PNAD) de 2014 (54).

Porém, quando os percentuais de usuários brasileiros de internet em 2014 são estratificados por grupo etário, percebe-se que 88% dos adultos jovens entre 16 e 24 anos se conectaram à internet no último ano (54), como mostra a Figura 2. Considerando a alta capilaridade da internet entre os adultos jovens brasileiros no ano de 2014, se espera um importante efeito de coorte, onde essas altas taxas de acesso passarão para as faixas etárias mais altas na medida em que essa parcela da população envelhece. Assim, a condução de pesquisas epidemiológicas baseadas na *web* no contexto brasileiro poderão se beneficiar de uma amostra mais heterogênea e representativa da população.

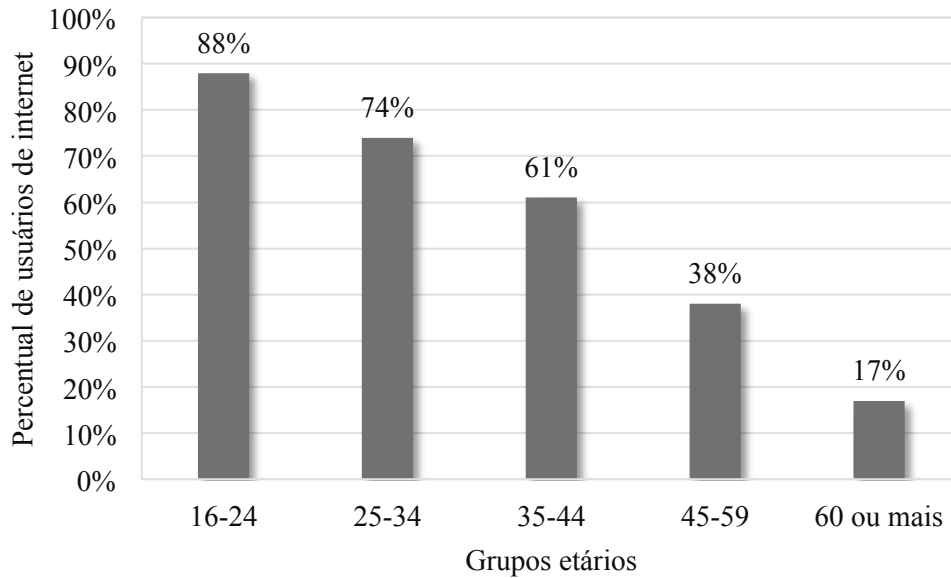


Figura 2: Percentual de brasileiros conforme grupo etário, que acessaram a internet no período de um ano anterior à realização da pesquisa de 2014. Dados extraídos da Pesquisa Nacional por Amostra de Domicílio (PNAD) de 2014 (54).

Considerando a alta capilaridade da internet no grupo de adultos jovens entre 16 e 24 anos, conduzir pesquisas epidemiológicas pela internet tendo como alvo tal população se torna viável. Durante o último acompanhamento presencial da coorte de nascimentos de 1993 de Pelotas, quando os participantes estavam com 22 ou 23 anos de idade, levantou-se que 96,5% dos 3.810 indivíduos entrevistados tinham acesso à internet pelo computador, celular ou *tablet*. Esse resultado corrobora com a estimativa nacional apresentada na Figura 2, e reforça a viabilidade de conduzir pesquisas epidemiológicas baseadas na *web* considerando essa amostra da população.

Assim, além do benefício de considerar uma população com alta prevalência de acesso à internet, conduzir pesquisas via *web* focadas na coorte de nascimentos de 1993 é uma ótima oportunidade por se tratar de uma população já conhecida. Nesse sentido, é possível fazer uso das informações coletadas em acompanhamentos presenciais prévios para identificar fatores demográficos e comportamentais associados a participação dos membros da coorte em pesquisas via *web*. Porém, a simples aplicação de questionários pode não ser suficiente para motivar a participação de uma população já muito estudada. Dessa forma, se torna necessário o desenvolvimento de uma plataforma *web* que, além da coleta de dados, possa fornecer incentivos, retorno de resultados e empregue estratégias de gamificação para maximizar as taxas de resposta e retenção desses participantes em pesquisas epidemiológicas aplicadas via *web*.

3. REVISÃO DE LITERATURA

A revisão de literatura foi feita utilizando a base PubMed (55) e teve como enfoque artigos que trataram a taxa de resposta como desfecho principal ou secundário. Para isso, foram pesquisados artigos que continham os seguintes termos em seus títulos ou resumos: (internet OR web OR online OR www) AND (response rate* OR return rate*). O único filtro utilizado foi quanto à obrigatoriedade de serem estudos focados em espécies humanas. Para serem selecionados, os estudos deveriam satisfazer aos seguintes critérios de elegibilidade:

1. A coleta deve ter sido feita através da *web*, mesmo que não exclusivamente;
2. A pesquisa deve ter sido preenchida pelo próprio participante, sem a presença de entrevistador;
3. Os estudos devem ter relatado a taxa de resposta obtida nos questionários aplicados via *web*;
4. A taxa de resposta deve ter sido considerada como um dos desfechos do estudo.

Foram excluídos estudos realizados exclusivamente com idosos, indivíduos com 60 anos de idade ou mais conforme definição da Organização Mundial da Saúde (OMS) (56), por se tratar de uma faixa etária muito distinta da que será considerada neste projeto. Além disso, estudos exclusivamente conduzidos com populações específicas, como profissionais de saúde ou outras categorias profissionais, pacientes com uma determinada condição de saúde (por exemplo, pacientes com diabetes) e pesquisadores, também não foram considerados. O motivo dessas exclusões se deve ao fato de populações específicas terem características diferentes da população geral, as quais podem influenciar sua participação nas pesquisas (57).

Nessa revisão de literatura apenas a taxa de resposta foi considerada como desfecho, embora o presente projeto irá investigar outros tipos de desfecho como taxa de contato e retenção. Para considerar estes outros desfechos, uma revisão de literatura mais aprofundada será realizada ao longo do desenvolvimento do projeto. Além disso, outras bases de dados, como Medline, SciELO e Lilacs serão consideradas. Além destas bases, será realizada uma busca no portal *Web Survey Methodology* (WebSM) (58) a fim de encontrar trabalhos que tratam sobre metodologia de pesquisa *web* aplicadas a saúde.

Até o dia 08 de agosto de 2016, a revisão de literatura com os termos acima descritos retornou 2.976 resultados. Após leitura dos títulos e resumos, foram selecionados 73 artigos para leitura na íntegra. Depois da leitura completa, 39 artigos foram considerados elegíveis de acordo com os critérios estabelecidos. A Figura 3 demonstra um fluxograma com o processo de

seleção dos artigos. Os 39 estudos selecionados após revisão da literatura estão sintetizados na Tabela 1.

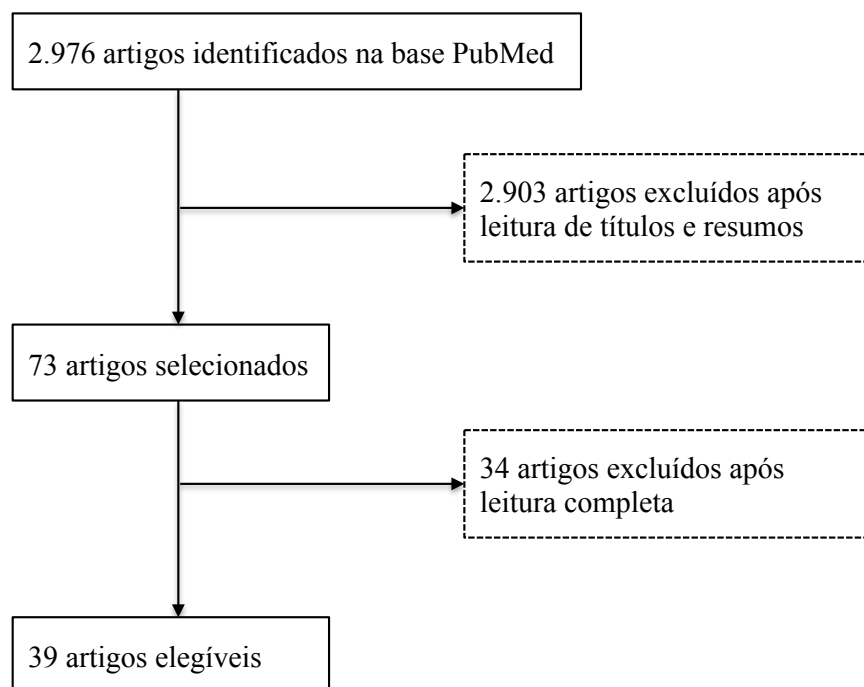


Figura 3: Esquema de busca e seleção de artigos que tratam a taxa de resposta de pesquisas em saúde aplicadas via web como desfecho principal ou secundário.

Dentre os artigos selecionados, 19 são estudos europeus, 14 norte-americanos, quatro oceânicos e dois latino-americanos, todos eles realizados entre 2001 e 2016. A localidade dos artigos reforça a falta de aplicação e investigação sobre o uso de questionários aplicados via *web* no contexto brasileiro, visto que apenas dois estudos latino-americanos foram encontrados e somente um deles foi conduzido no Brasil. Quanto aos delineamentos utilizados, 23 estudos utilizaram o delineamento transversal, 15 estudos conduziram ensaios randomizados e um estudo foi de caráter qualitativo. A mediana do tamanho de amostra dos estudos foi de 3.000 participantes, variando entre 72 e mais de 300.000 indivíduos.

A maior parte dos artigos, 17 dos 39 considerados, teve como alvo a população geral. A outra população alvo que foi bastante predominante entre os artigos foi a de estudantes, considerada em 14 artigos. A escolha de populações de estudantes se deve em virtude da alta prevalência de acesso à internet entre esses indivíduos, visto que grande parte das instituições de ensino fornecem o acesso aos alunos matriculados. Dessa forma, os pesquisadores são capazes de determinar a elegibilidade de cada um dos estudantes convidados a participar dos questionários aplicados via *web*. Quando a população geral é considerada, alguns dos indivíduos convidados a responder os questionários via *web* podem não ser elegíveis já no início do estudo,

o que pode comprometer as taxas de resposta. Apesar disso, saber a prevalência de acesso à internet do local de onde a amostra será retirada pode indicar a viabilidade dessa metodologia de coleta de dados. Os 17 artigos que consideraram a população geral foram conduzidos em países com prevalência de acesso à internet superior a 80%, conforme dados de 2013 (59).

Tabela 1: Descrição dos estudos selecionados após revisão de literatura.

Autor (ano)	País	Delineamento (n)	População alvo	Tipo de recrutamento	Incentivo	Taxa de resposta
Pealer et al. (60) (2001)	Estados Unidos	Ensaio Randomizado (600)	Estudantes	Correio E-mail	US\$2	58,3%
Andersson et al. (61) (2002)	Suécia	Transversal (1.167)	População geral	Banner na <i>web</i> Correio		51,9%
Goldman et al. (62) (2004)	Canadá	Ensaio Randomizado (198)	Pacientes	E-mail		53,0%
McCabe (63) (2004)	Estados Unidos	Ensaio Randomizado (7.000)	Estudantes	E-mail	US\$10	63,0%
McCabe et al. (64) (2004)	Estados Unidos	Transversal (1.723)	Estudantes	Correio		89,0%
Kypri et al. (65) (2004)	Nova Zelândia	Transversal (1.910)	Estudantes	Correio E-mail Telefone	Caneta	81,9%
Bälter et al. (49) (2005)	Suécia	Transversal (875)	População geral	Correio		50,0%
Mangunkusumo et al. (66) (2005)	Holanda	Ensaio Randomizado (591)	Estudantes	Correio		96,5%
Satia et al. (67) (2005)	Estados Unidos	Transversal (5.000)	População geral negra	Correio	US\$4	2,0%

Autor (ano)	País	Delineamento (n)	População alvo	Tipo de recrutamento	Incentivo	Taxa de resposta
McCabe et al. (68) (2006)	Estados Unidos	Transversal (6.500)	Estudantes	Correio E-mail	Sorteio de US\$500	34,3%
Ekman et al. (39) (2006)	Suécia	Transversal (47.859)	Mulheres participantes de uma coorte	Correio		41,0%
Rodriguez et al. (69) (2006)	Estados Unidos	Ensaio Randomizado (22.152)	Pacientes	Correio		18,4%
Kongsved et al. (70) (2007)	Dinamarca	Ensaio Randomizado (533)	Pacientes	Correio		64,2%
Ekman et al. (71) (2007)	Suécia	Transversal (47.859)	Mulheres participantes de uma coorte	Correio E-mail		41,8%
Brøgger et al. (72) (2007)	Noruega	Ensaio Randomizado (4.213)	População geral	Correio		8,3%
Dinitto et al. (73) (2008)	Estados Unidos	Transversal (1.610)	População geral	Telefone	Cartão presente US\$20	1,9%
Blank et al. (74) (2009)	Brasil	Transversal (4.718)	Estudantes	E-mail		34,5%
Balajti et al. (75) (2010)	Hungria	Transversal (1.544)	Estudantes	E-mail	Sorteio de viagem	58,8%

Autor (ano)	País	Delineamento (n)	População alvo	Tipo de recrutamento	Incentivo	Taxa de resposta
Lim et al. (76) (2010)	Austrália	Ensaio Randomizado (72)	População geral	Correio E-mail SMS	Pré-pago: \$3 Pós-pago: \$25	91,7%
Atherton et al. (77) (2010)	Inglaterra	Transversal (1.329)	Estudantes	E-mail		43,0%
Scott-Johnson et al. (78) (2010)	Estados Unidos	Transversal (357)	Estudantes	Correio	US\$25	31,7%
Sinadinovic et al. (79) (2011)	Suécia	Transversal (5.000)	População geral	Correio		38,1%
Whitehead et al. (80) (2011)	Estados Unidos	Ensaio Randomizados (2.000)	Estudantes	E-mail		54,4%
Sinclair et al. (48) (2012)	Austrália	Transversal (17.000)	População geral	Correio	Bilhete de loteria	3,5%
Lagerros et al. (81) (2012)	Suécia	Transversal (637)	População geral	Correio E-mail		79,8%
Gajic et al. (8) (2012)	Canadá	Transversal (3.000)	População geral	Correio	Pré-pago: \$2 Loteria: \$25 Loteria: \$250	9,8%
Dokkum et al. (82) (2012)	Holanda	Transversal (300.000)	População geral	Correio		20,0%

Autor (ano)	País	Delineamento (n)	População alvo	Tipo de recrutamento	Incentivo	Taxa de resposta
Hohwü et al. (37) (2013)	Dinamarca	Ensaio Randomizado (3.148)	População geral	Correio	Sorteio de um <i>tablet</i>	53,0%
Hunter et al. (38) (2013)	Austrália	Qualitativo (4.315)	Pacientes	Correio E-mail		7,4%
Bailey et al. (83) (2013)	Reino Unido	Ensaio Randomizado (2.006)	População geral	Correio E-mail Facebook Pessoal	Voucher de £10 ou £20	60,2%
Elliott et al. (84) (2013)	Estados Unidos	Transversal (26.223)	Pacientes	Correio		12,0%
Patrick et al. (85) (2013)	Estados Unidos	Ensaio Randomizado (3.000)	Estudantes	Correio E-mail	Pré-pago: US\$10 Pós-pago: US\$10	44,0%
Flüß et al. (86) (2014)	Escócia	Ensaio Randomizado (4.600)	População geral	Correio		7,3%
Johnson et al. (87) (2014)	Inglaterra	Transversal (2.376)	População geral	Correio		5,0%
Griffin et al. (88) (2014)	Estados Unidos	Ensaio Randomizado (300)	Estudantes	Telefone	US\$25	37,2%
Wallander et al. (16) (2015)	Suécia	Transversal (6.000)	População geral	Correio E-mail		25,0%

Autor (ano)	País	Delineamento (n)	População alvo	Tipo de recrutamento	Incentivo	Taxa de resposta
Wallander et al. (16) (2015)	Suécia	Transversal (24.000)	População geral	Correio E-mail		24,0%
Conde et al. (89) (2015)	Argentina	Transversal (120)	Estudantes	E-mail		33,0%
Couturier et al. (90) (2015)	França	Ensaio Randomizado (1.141)	Pacientes			39,1%
Galante et al. (91) (2016)	Inglaterra	Transversal (331.013)	População geral	E-mail		53,0%

3.1 Taxa de resposta de questionários *web* conforme o método de recrutamento

Toda pesquisa que envolve a aplicação de questionários inicia com a etapa de contato e recrutamento dos participantes, e isso não é diferente nas pesquisas aplicadas via *web*. Existe uma grande variedade de métodos de recrutamento para pesquisas baseadas na *web*, como convite pessoal, propagandas, telefonemas, cartas enviadas pelo correio, e-mail, redes sociais e outros. Independente do método de recrutamento utilizado, é importante que ele seja simples de ser compreendido pelos convidados, seja breve e objetivo para descrever os interesses da pesquisa, tenha cobertura suficiente para alcançar os indivíduos elegíveis e seja interessante a ponto de evitar que o convite para participação seja descartado (19). Para atingir esses objetivos, alguns estudos empregam mais de um método de recrutamento (16,38,60,61,65,68,71,76,81,83,85). Por um lado, isso pode aumentar a cobertura do recrutamento por não ser restrito a um único meio de comunicação. Mas, por outro, utilizar múltiplos métodos pode se tornar logisticamente complicado e mais caro.

Um dos principais métodos de recrutamento utilizados em pesquisas *web*, empregado em 28 estudos considerados nesta revisão de literatura, são convites via correio. Nesse tipo de recrutamento uma carta endereçada à residência do participante é enviada contendo uma breve descrição do estudo, um domínio para acesso a página *web* da pesquisa e, geralmente, um nome de usuário e senha para acesso aos questionários (19). Essa metodologia de recrutamento é simples, pois o acesso ao endereço residencial dos participantes é facilmente obtido através de registros sociais. Porém, para acessar a página *web* contendo a pesquisa o participante deve digitar o domínio do estudo em um navegador. Se este domínio não for simples o suficiente, o participante pode se negar ou desistir de participar do estudo.

Outro método bastante utilizado para recrutamento de pesquisas via *web* é o envio de e-mails, o qual foi empregado em 18 artigos selecionados na revisão. Neste método as mesmas informações presentes na carta utilizada no recrutamento via correio são enviadas através de uma mensagem eletrônica, com a facilidade do participante poder clicar no domínio e ser redirecionado automaticamente para a página *web* da pesquisa. Embora este método de recrutamento pareça ser mais adequado para pesquisas via *web*, as taxas de resposta obtidas através de convites via e-mail podem ser tão baixas quanto 7,4% (38) ou tão altas quanto 91,7% (76). Esta variabilidade nas taxas de resposta é similar a encontrada através do recrutamento via correio, que variou entre 2,0% (67) e 96,5% (66).

Apenas dois artigos selecionados na revisão de literatura compararam as taxas de resposta de acordo com o método de recrutamento utilizado. Um deles comparou recrutamentos feitos

através de contato pessoal, e-mail e da rede social Facebook (83). O outro estudo investigou se o uso de convites personalizados enviados via correio são capazes de obter maiores taxas de resposta para questionários aplicados via *web* do que convites sem personalização (48). Os achados do primeiro estudo demonstram que a maior parte dos participantes (84,0%) foi recrutado através do Facebook, o que pode ter sido motivado pela faixa etária jovem considerada no estudo, de 16 a 20 anos (83). Já o segundo estudo mostrou que o uso de personalização na fase de recrutamento pode gerar taxas de resposta 53,2% maiores do que os convites sem personalização (48). Dessa forma, é importante considerar as redes sociais como método de recrutamento para populações de adultos jovens, bem como o uso de convites personalizados para buscar maximizar as taxas de resposta de questionários aplicados via *web*.

3.2 Taxa de resposta de questionários *web* comparados a outros métodos de coleta

Com base nos artigos revisados, a taxa de resposta média dos questionários aplicados via *web* foi 40,9% (desvio padrão de 26,3 pontos percentuais). Dentre os 39 estudos, 22 compararam as taxas de resposta conforme o método de coleta. Além dos questionários aplicados via *web*, os estudos consideraram questionários aplicados em papel e caneta, telefone, SMS e correio. Entre os quatro estudos que compararam as taxas de resposta de questionários *web* com a metodologia tradicional (papel e caneta), dois relataram maiores taxas para o método de coleta tradicional (70,89), um não encontrou diferença entre as taxas de resposta (37) e o outro relatou taxa de resposta superior para os questionários aplicados via *web* (66). Essa mesma variação foi relatada pelos artigos que compararam os questionários aplicados via *web* com os aplicados via correio. Cinco artigos encontraram taxas de resposta superiores para o método de coleta via *web* (39,75,77,80,81), quatro encontraram o resultado inverso (48,61,72,84) e dois não encontraram diferença (60,76).

A taxa de resposta obtida por questionários aplicados através de telefone fixo foi comparado aos questionários *web* por seis artigos. Desses, cinco relataram superioridade por parte dos questionários aplicados pelo telefone (48,62,73,84,90), e apenas um relatou o inverso (79). O uso de telefones celulares também foi comparado aos questionários aplicados via *web* por dois estudos. Um dos estudos não encontrou diferença entre as taxas de resposta (76) e o outro relatou maior taxa de resposta para os questionários aplicados pela *web* (81). A grande variação das taxas de resposta obtidas, as quais ora mostram vantagem para os questionários aplicados via *web*, ora mostram desvantagem desse método comparado a outros, não apontam a melhor maneira para conduzir investigações epidemiológicas. Apesar disso, com o aumento do número de usuários de internet, e também com a melhora no grau de instrução para o uso da internet, se espera que a metodologia de pesquisa via *web* ganhe ainda mais força e seja capaz

de obter taxas de resposta equivalentes, e até superiores, às obtidas pelos demais métodos de coleta de dados (19,54).

3.3 Taxa de resposta de questionários *web* conforme o fornecimento de incentivo

Quinze dos artigos revisados forneceram algum tipo de incentivo visando impulsionar as taxas de resposta dos questionários aplicados via *web*. O principal incentivo fornecido foi monetário (66,7%), seguido de bilhetes para participação de sorteios (13,3%), bilhetes de loteria (13,3%) e fornecimento de brinde (6,7%). Os valores dos incentivos monetários, pagos previamente (pré-pagos) ou somente após a participação dos indivíduos (pós-pagos), variou entre US\$2,00 e US\$25,00. Dois dos artigos revisados compararam as taxas de resposta entre o fornecimento de incentivos monetários pré-pagos de menor valor, com as taxas obtidas através de incentivos monetários pós-pagos de maior valor. Ambos relataram maiores taxas de resposta entre o grupo de incentivo pré-pago e de menor valor (8,85). Em contraponto a estes achados, no estudo de Bailey e colaboradores foi conduzido um ensaio aleatorizado para investigar a influência do valor de um vale presente sobre as taxas de resposta de questionários aplicados via *web*. Foram comparados dois valores de vales pós-pagos, £10,00 e £20,00, e a maior taxa de resposta foi obtida pelo fornecimento do vale presente de maior valor (83).

Embora a literatura demonstre que o fornecimento de incentivos monetários é capaz de aumentar as taxas de resposta, isso faz com que os custos das pesquisas conduzidas via *web* aumentem conforme aumenta o número de participantes (8,83,85,92,93). Por isso, alguns autores forneceram incentivos mais custo efetivos, como brindes (65), entradas para sorteios (37,75) e bilhetes de loteria (48,68), os quais também foram capazes de motivar a participação dos entrevistados. Ainda mais custo efetivo do que estes, pode ser a utilização do retorno de resultados como forma de incentivo para participação de pesquisas aplicadas via *web* (94). Porém, nenhum dos artigos revisados estudou a possível influência desse tipo de incentivo sobre as taxas de resposta, e assim existe espaço para tal investigação.

4. OBJETIVOS

Este projeto tem como objetivo investigar a viabilidade de conduzir pesquisas epidemiológicas via *web*, explorando estratégias de metodologia de pesquisa com o objetivo de maximizar as taxas de contato, resposta e retenção dos usuários. Além disso, o projeto irá estudar os fatores sócio-demográficos associados à participação de adultos jovens membros da coorte de nascimentos de 1993 de Pelotas nos questionários aplicados através da plataforma *web*.

4.1 Objetivos específicos

- Desenvolver uma plataforma *web* com o ambiente completo para a condução de pesquisas epidemiológicas, compreendendo recrutamento de participantes, controle de acesso, coleta de dados epidemiológicos via *web*, retorno de resultados epidemiológicos, convite para participação em pesquisas e fornecimento de recompensas.
- Avaliar a representatividade dos indivíduos elegíveis com relação à população original da coorte de nascimentos de 1993 de Pelotas.
- Avaliar a representatividade dos indivíduos cadastrados na plataforma *web* com relação aos indivíduos elegíveis.
- Avaliar a influência de diferentes fatores demográficos e socioeconômicos na taxa de contato da plataforma *web*.
- Estudar determinantes demográficos e socioeconômicos que influenciam a taxa de resposta dos questionários aplicados via *web*.
- Identificar variáveis demográficas e socioeconômicas que influenciam na taxa de retenção.
- Aplicar diferentes estratégias de recrutamento para convidar os indivíduos a se cadastrar na plataforma *web*, como correspondência, e-mail e redes sociais. O objetivo é estimar a influência desses diferentes métodos na taxa de contato.
- Comparar diferentes tamanhos de questionários (questionários longos vs. curtos) e frequência de lembretes (frequência alta vs. baixa) em relação à taxa de resposta dos questionários aplicados via *web*.

5. MATERIAIS E MÉTODOS

5.1 Coorte de nascimentos de 1993 de Pelotas

A coorte de 1993 é a segunda de quatro coortes de nascimento conduzidas na cidade de Pelotas, as quais captaram os nascidos vivos na zona urbana do município nos anos de 1982 (95), 1993 (96), 2004 (97) e 2015. A larga experiência na condução desses estudos longitudinais, bem como os excelentes resultados obtidos, tornaram a cidade de Pelotas referência nacional e mundial em estudos de coorte.

A coorte de nascimentos de 1993 considerou em seu estudo perinatal 5.265 nascimentos ocorridos em Pelotas entre os dias 1º de janeiro e 31 de dezembro daquele ano. Do total de nascidos vivos, foram considerados como membros da coorte somente aqueles nascidos na zona urbana da cidade e que aceitaram participar do estudo, representando 5.249 indivíduos (96). A partir do estudo perinatal, aconteceram diversos acompanhamentos com os nascidos em 1993, conforme mostra a Figura 4. Aos 11, 15 e 18 anos foram recrutados todos os membros da coorte, e nos demais acompanhamentos uma sub-amostra foi estudada. Durante parte dos anos de 2015 e 2016, quando os membros da coorte tinham 22 ou 23 anos, ocorreu um novo acompanhamento com todos os participantes do estudo. Neste acompanhamento 3.810 participantes foram entrevistados e, somados aos óbitos, geraram um percentual de acompanhamento de 76,3%.

Durante estes acompanhamentos, variáveis cobrindo múltiplos propósitos são coletadas, incluindo informações socioeconômicas, demográficas, comportamentais, nutricionais e sobre doenças crônicas. Além disso, testes e exames para coleta de material biológico, avaliação da composição corporal e da saúde mental também foram aplicados. Maiores detalhes metodológicos sobre a coorte de nascimentos de 1993 estão presentes na publicação de Gonçalves e colaboradores (96).

Acompanhamentos principais e percentuais de acompanhamento Coorte de Nascimentos de 1993



Figura 4: Descrição dos acompanhamentos realizados na Coorte de Nascimentos de 1993, com exceção do acompanhamento de 2015/2016. Fonte: Gonçalves et al. (2014) (96).

5.2 Logística do acompanhamento de 2015/2016

O acompanhamento dos 22/23 anos da coorte de nascimentos de 1993, o qual ocorreu durante os anos de 2015 e 2016, contemplou entrevistas domiciliares e conduzidas dentro da clínica do Centro de Pesquisas Epidemiológicas da Universidade Federal de Pelotas. Previamente ao início do estudo, as informações cadastrais dos membros da coorte foram atualizadas de forma a facilitar o contato para realização do agendamento para este acompanhamento. Posteriormente, uma equipe contactava os membros da coorte através de ligações telefônicas para agendar uma data e um horário para que comparecessem à clínica.

Ao chegar na clínica, os membros da coorte se dirigiam à recepção onde suas informações cadastrais eram mais uma vez atualizadas. Um crachá contendo o nome e o número de identificação era fornecido para cada participante, bem como uma pulseira com o mesmo número de identificação em formato de código de barras. Este código de barras era utilizado para acessar cada uma das estações de coleta de dados, a fim de evitar erros de digitação.

Durante a visita foram aplicados questionários para investigação de diversos interesses, como saúde mental, consumo alimentar, uso de medicamentos, questões comportamentais e também sobre uso de internet, computador e mídias sociais. Além disso, foram realizadas antropometria, exames para avaliação da composição corporal e coleta de material biológico.

Aqueles membros da coorte que tinham filhos também foram convidados a levá-los até a clínica de pesquisa. Além destes, o pai ou mãe biológicos(as) destes filhos, que não eram membros da coorte, também foram convidados a participar do acompanhamento. Os filhos dos membros da coorte realizavam exames físicos e tinham material biológico coletado. Já os pais ou mães que não eram membros da coorte, além de exames de composição corporal e coleta de material biológico, também respondiam um questionário com relação a cada filho. Todo o processo de coleta de dados, considerando entrevistas, exames e coleta de material biológico durava em média aproximadamente 4 horas.

5.3 População alvo

A população alvo deste projeto consiste nos nascidos vivos na zona urbana do município de Pelotas no ano de 1993, pertencentes à coorte de nascimentos de 1993.

5.3.1 Critérios de inclusão

1. Ter participado do acompanhamento dos 22/23 anos da coorte de nascimentos de 1993, o qual ocorreu em 2015 e 2016.
2. Ter acesso à internet disponível durante 24 horas por dia, independentemente do dispositivo com o qual este acesso é realizado (computador ou dispositivo móvel).
3. Ter respondido aos questionários do acompanhamento dos 22/23 anos sem o auxílio de outra pessoa.

5.4 Instrumentos

O presente projeto tem como objetivo o desenvolvimento de uma plataforma que contenha o ambiente completo para condução de pesquisas epidemiológicas aplicadas via *web*. Para a coleta de dados epidemiológicos, a intenção é usar questionários padronizados que já são aplicados durante as pesquisas presenciais da coorte de nascimentos de 1993, os quais investigam questões como tabagismo, consumo de álcool, frequência alimentar, prática de atividade física, doenças crônicas, entre outras. Os questionários aplicados nos acompanhamentos presenciais da coorte estão disponíveis em www.epidemiio-ufpel.org.br.

Visto que um dos objetivos do projeto é testar a associação do tamanho dos questionários com as taxas de resposta e retenção, duas extensões de questionário serão adotadas: longos e

curtos. De acordo com a literatura, questionários curtos são aqueles com duração média entre 5 e 10 minutos. Já questionários longos têm duração média entre 20 e 30 minutos (50). Um estudo piloto será conduzido para determinar o tempo médio de resposta de cada questionário, visando classificá-los entre questionários curtos ou longos. Este estudo piloto, além de determinar a duração dos questionários, também servirá para testar a aceitação e a experiência de uso da plataforma *web* considerando uma amostra por conveniência, diferente da população alvo do projeto.

Para que seja possível traçar um perfil dos participantes, serão utilizadas variáveis demográficas e socioeconômicas coletadas no acompanhamento presencial dos 22/23 anos da coorte 1993. Adicionalmente, serão consideradas variáveis sobre utilização de internet, computador, dispositivos móveis e redes sociais, também coletadas durante o último acompanhamento presencial. As variáveis a serem utilizadas estão descritas na Tabela 2.

Tabela 2: Variáveis demográficas, socioeconômicas e sobre uso de dispositivos e internet coletadas durante o acompanhamento dos 22/23 anos da coorte 1993 e que serão utilizadas no projeto.

Demográficas e Socioeconômicas	Dispositivos e Internet
Sexo	Posse de computador, celular e <i>tablet</i>
Escolaridade	Acesso à internet fixa e móvel
Índice de bens	Tempo de uso do computador, celular e <i>tablet</i>
Situação conjugal	Comunicação com amigos e familiares
Cor da pele	Tipo de internet móvel
Número de filhos	Aplicativos usados no celular ou <i>tablet</i>
	Frequência de uso por redes sociais

5.5 Logística

Com base nas informações coletadas durante o acompanhamento presencial dos 22/23 anos da coorte de nascimentos de 1993, os indivíduos elegíveis para participação no estudo serão identificados conforme os critérios descritos na Seção 5.3.1. No mês de setembro de 2017 será iniciado o processo de recrutamento dos participantes para a plataforma *web*. Nesse momento os indivíduos serão convidados para se cadastrar e criar um perfil com nome de usuário e senha, os quais servirão como chave para futuros acessos à plataforma *web*.

O processo de recrutamento de indivíduos ainda não cadastrados, e de coleta de dados entre os indivíduos já cadastrados terá duração de seis meses. Dessa forma, convites para cadastramento de novos usuários e lembretes para preenchimento de novos questionários serão sistematicamente enviados aos indivíduos elegíveis ao longo desse período. Depois dos seis meses, o recrutamento e o envio de lembretes serão encerrados, os questionários serão removidos e a plataforma *web* ficará disponível exclusivamente para que os participantes possam acessar seus perfis e verificar os resultados dos questionários aplicativos através da plataforma *web* e dos exames realizados em acompanhamentos presenciais da coorte.

5.6 Desenvolvimento da plataforma *web*

A plataforma *web* será hospedada em um computador servidor rodando o sistema operacional Linux com a distribuição Ubuntu Server (98), o qual será configurado pelo próprio doutorando. Para facilitar o acesso à plataforma *web* e evitar que os participantes do estudo tenham que acessar o endereço do *Internet Protocol* (IP) do servidor, um domínio será adquirido com o seguinte nome: www.coortesnaweb.club. Para evitar que a extensão desse domínio possa afastar os participantes, também será adquirido o domínio www.coortesnaweb.com.

A plataforma *web* será desenvolvida utilizando a linguagem de marcação *Hypertext Markup Language* (HTML) (99) e a linguagem de *script Hypertext Processor* (PHP) (100), as quais são amplamente utilizadas no desenvolvimento dos *websites* hoje presentes na internet. Os recursos gráficos serão implementados utilizando a linguagem de estilo *Cascading Style Sheets* (CSS) (99), a qual também é extensamente adotada na internet. Utilizar tais linguagens para o desenvolvimento da plataforma *web* vai permitir com que esta seja compatível com qualquer dispositivo capaz de acessar a internet através de um navegador *web*. Com o objetivo de simplificar e agilizar o processo de desenvolvimento, serão utilizados os *frameworks* PHP Symfony (101) e CSS Bootstrap (102). Estas ferramentas fornecem ferramentas pré-construídas, que podem simplificar, por exemplo, a interação entre os recursos visuais da plataforma *web* e o conteúdo presente no banco de dados.

Visto que o acesso à plataforma *web* poderá ser realizado através de qualquer dispositivo que suporte um navegador *web*, é importante garantir sua compatibilidade tanto com computadores, quanto com dispositivos móveis. Fatores importantes a serem considerados são as diferentes capacidades de processamento, de tamanho de tela e tipos de conexão à internet. Para garantir a compatibilidade com dispositivos de baixa capacidade de processamento, bem como com dispositivos que tenham conexões mais lentas, recursos gráficos avançados serão poupados. Para lidar com dispositivos de diferentes tamanhos de tela, a plataforma será

desenvolvida de maneira responsiva, onde seu conteúdo será redimensionado e reposicionado de maneira adequada para qualquer tamanho de tela.

O processo de coleta de dados se dará exclusivamente através da plataforma *web*. Os questionários serão desenvolvidos utilizando formulários HTML, onde os participantes responderão a questões textuais abertas e de múltipla escolha. Todos os questionários serão desenvolvidos utilizando lógica condicional. Este recurso é capaz de definir, em tempo real, quais perguntas devem ser aplicadas a partir das respostas fornecidas em perguntas anteriores. Todas as informações obtidas através dos questionários serão armazenadas e criptografadas em um banco de dados, o qual será gerenciado através do sistema de gerenciamento de banco de dados (SGBD) MySQL (103). Todas as ferramentas utilizadas para o desenvolvimento da plataforma *web* são gratuitas, com exceção do domínio que deverá ser adquirido.

5.7 Sistema de recompensas e incentivos

Outra funcionalidade que será implementada na plataforma *web* é um sistema de retorno de resultados relacionados a desfechos em saúde. Dessa forma, cada usuário poderá acessar uma página *web* segura e pessoal onde seus resultados individuais poderão ser relacionados aos resultados da população da coorte de nascimentos de 1993 de Pelotas. Os resultados serão liberados conforme os usuários respondem aos questionários, e assim esse retorno servirá como recompensa e também como um método de incentivo para a participação nas demais pesquisas.

Além do retorno de resultados, será implementado um sistema de gamificação em formato de *ranking*, conquista de insígnias e evolução de níveis. O *ranking* será baseado no nível dos participantes, sendo que a evolução para cada nível será definida a partir de pontuações garantidas após completar cada questionário. Quanto maior for a pontuação acumulada, maior será o nível e o *ranking* do participante. Somado a isso, os participantes também poderão obter conquistas em forma de insígnias. Por exemplo, aquele participante que ao final do estudo tiver respondido a todos os questionários, será presenteado com uma insígnia de “participante assíduo”. Os sistemas de níveis, *ranking* e conquista de insígnias serão utilizados para liberar novas funcionalidades na plataforma *web*. Por exemplo, o participante deverá ter um determinado nível para obter a liberação de resultados extra.

5.8 Controle de qualidade

Todos os questionários serão desenvolvidos utilizando um sistema de validação de dados em tempo real. Dessa maneira, as respostas fornecidas pelos entrevistados serão verificadas para checar que estão de acordo com o tipo de dado esperado. Por exemplo, um campo que requer uma resposta numérica não será validado se uma letra for inserida. Caso isso ocorra, o

participante será avisado e terá a oportunidade de corrigir o erro. Nas questões que requerem a inserção de valores numéricos, será definido um intervalo de valores aceitáveis. Dessa forma, os participantes serão avisados caso respondam com valores extremos.

Questões com categorias pré-definidas e que devam aceitar apenas uma resposta válida serão implementadas dessa maneira para evitar possíveis inconsistências. Nas questões que necessitam de um formato especial, como campos de data, máscaras serão aplicadas para que os dados já sejam coletados com a estrutura correta. Por exemplo, as barras que separam dia/mês/ano serão automaticamente inseridos.

Visando reduzir o número de erros devido a não compreensão por parte dos entrevistados, todas as questões terão instruções de preenchimento que poderão ser acessadas pelos participantes quando necessário. Após coletadas, as informações serão transferidas para bancos de dados do programa estatístico Stata versão 13.1 (104), onde *scripts* para checagem de inconsistências serão implementados para verificar possíveis erros que possam não ter sido evitados durante o processo de coleta de dados.

5.9 Planos de análise

5.9.1 Variáveis de exposição consideradas nas análises

Visto que um dos objetivos deste projeto é estudar os fatores sociais e demográficos associados à participação dos membros da coorte na plataforma *web*, serão calculadas as taxas de contato, resposta e retenção conforme as variáveis descritas na Tabela 2. As variáveis a serem consideradas serão operacionalizadas da seguinte maneira:

- Sexo: variável dicotômica coletada no acompanhamento perinatal da coorte de nascimentos de 1993 de Pelotas. Categorias: feminino e masculino.
- Escolaridade: variável numérica discreta, medida em anos de estudo coletada no acompanhamento presencial dos 22/23 anos.
- Índice de bens: variável numérica contínua que indica o poder de compra do indivíduo de acordo com a posse de determinados bens de consumo (105) coletada durante o acompanhamento presencial dos 22/23 anos.
- Situação conjugal: variável categórica politômica coletada no acompanhamento presencial dos 22/23 anos. Categorias: casado(a), divorciado(a), solteiro(a) e viúvo(a).
- Cor da pele: variável categórica politômica auto-relatada pelo participante no acompanhamento presencial dos 15 anos. Categorias: branca, preta, parda, amarela e indígena.

- Número de filhos: variável numérica discreta coletada no acompanhamento presencial dos 22/23 anos.

Além de traçar o perfil social e demográfico dos participantes da plataforma web, estes também serão descritos conforme o perfil de acesso à internet e a posse de dispositivos que se conectam à internet. Essas variáveis foram obtidas através de 20 perguntas presentes no questionário do acompanhamento presencial dos 22/23 anos da coorte de 1993. Abaixo é descrita a operacionalização de cada uma das variáveis, fazendo referência à numeração das perguntas do questionário que podem ser consultadas no ANEXO 1:

- Posse de computador: variável dicotômica (sim/não) coletada através da questão D038.
- Acesso à internet pelo computador: variável dicotômica (sim/não) coletada através da questão D038b.
- Principal meio de comunicação com amigos e familiares: variável categórica politômica coletada através das questões L144 e L145.
- Aplicativos usados no celular ou *tablet*: variável categórica politômica coletada através das questões L147 e L147a.
- Uso de redes sociais: variável dicotômica (sim/não) a ser construída através das informações coletadas nas questões L144, L145, L147, L147a e L148.
- Frequência de uso por redes sociais: variável numérica discreta, medida em dias, que foi coletada através das questões L150 até L150n.
- Dispositivo para acesso às redes sociais: variável categórica politômica coletada através da questão L151.
- Uso de celular e *tablet* para acesso à internet: variável dicotômica coletada através das questões D223 e D224.
- Tempo de uso internet no celular e *tablet*: variável numérica contínua, medida em minutos, coletada através das questões D223a e D224a.
- Tipo de internet do celular e *tablet*: variável categórica politômica coletada através das questões L152 e L153.
- Principal dispositivo para acesso à internet: variável categórica politômica coletada através da questão L146.
- Tempo de uso do computador: variável numérica contínua, medida em minutos, coletada através da questão D225a.

5.9.2 Representatividade dos elegíveis e dos participantes da plataforma web

Um dos principais questionamentos com relação a validade da condução de pesquisas aplicadas via *web* é quanto a representatividade de suas amostras, afinal apenas indivíduos conectados à internet têm a possibilidade de participar desse tipo de estudo. Com o objetivo de investigar essa questão, serão realizadas duas análises de representatividade:

1. Análise da representatividade dos indivíduos elegíveis em comparação à população original da coorte 1993.
2. Análise da representatividade dos indivíduos cadastrados na plataforma *web* em comparação aos indivíduos elegíveis.

A análise de representatividade entre os grupos será realizada através da comparação entre as frequências e médias das variáveis categóricas e numéricas. Apenas as variáveis demográficas e socioeconômicas, descritas na Tabela 2, serão consideradas. As variáveis sobre acesso à internet e utilização de dispositivos não serão consideradas por terem sido coletadas apenas no último acompanhamento presencial da coorte de nascimentos de 1993. Para a comparação das frequências de variáveis categorias, serão conduzidos testes qui-quadrado. Já a frequência de variáveis numéricas serão comparadas através de testes t de Student.

5.9.3 Taxa de contato de acordo com o método de recrutamento

Com o objetivo de determinar a influência de diferentes métodos de recrutamento sobre a taxa de contato, os indivíduos elegíveis serão randomizados para determinar a partir de qual método serão convidados a se cadastrar na plataforma *web*. Para garantir que os braços desse ensaio randomizado sejam equilibrados com relação ao número de indivíduos, um processo de randomização aleatória por blocos será empregado. A análise será realizada através da comparação das taxas de contato (variável numérica contínua) conforme cada um dos métodos de recrutamento (variável categórica).

Considerando uma randomização feita de forma adequada e um estudo que contenha um número elevado de participantes, se espera que as características dos indivíduos do estudo sejam comparáveis entre os braços de um ensaio randomizado (106). Levando em conta que este projeto conta com 3.677 indivíduos elegíveis, é previsto que haja tal comparabilidade entre os grupos do estudo. Assim, não se torna necessária a aplicação de análises estatísticas ajustadas para fatores de confusão. Um modelo de regressão linear será ajustado para estudar a influência do método de recrutamento sobre a taxa de contato.

5.9.4 Influência da estratégia de coleta de dados e frequência de lembretes sobre a taxa de resposta

Serão aplicadas duas extensões de questionário através da plataforma *web*, questionários curtos e longos. O objetivo é avaliar a influência do tamanho dos questionários sobre a taxa de resposta. Para os indivíduos que não responderem aos questionários após o primeiro convite, lembretes serão enviados para requisitar sua participação. Dessa forma, juntamente com o tamanho dos questionários, será testada a influência de diferentes frequências de lembretes sobre a taxa de resposta. Para que os dois fatores (tamanho de questionário e frequência de lembretes) possam ser avaliados simultaneamente, um estudo fatorial 2x2 será conduzido. Dessa forma, os indivíduos já cadastrados na plataforma *web* serão randomizados para ocupar um dos seguintes grupos: (1) questionário longo e frequência alta, (2) questionário longo e frequência baixa, (3) questionário curto e frequência alta, e (4) questionário curto e frequência baixa.

Com o desenho fatorial, é possível analisar a influência do tamanho dos questionários e da frequência de lembretes de forma separada. Conforme mostra a Figura 5, caso seja necessário avaliar a influência do tamanho dos questionários sem levar em consideração a frequência de lembretes, as taxas de resposta das células A e C, bem como das células B e D podem ser comparadas. Assim, o fator frequência de lembretes é considerado “fixo”, e o único fator variável se trata do tamanho dos questionários. Será ajustado um modelo linear para avaliar as taxas de resposta, visto que o desfecho será tratado de forma contínua.

		Frequência de lembretes	
		Alta	Baixa
Tamanho de questionário	Curto	A	B
	Longo	C	D

Figura 5: Organização dos quatro grupos do desenho fatorial 2x2 para testar a influência do tamanho dos questionários e da frequência de lembretes sobre a taxa de resposta.

6. ASPECTOS ÉTICOS

O acompanhamento dos 22/23 anos da coorte de nascimentos de 1993 de Pelotas foi aprovado pelo Comitê de Ética e Pesquisa da Faculdade de Medicina da Universidade Federal de Pelotas. Todos os membros da coorte que participaram deste acompanhamento, bem como os responsáveis pelos filhos dos membros, assinaram um Termo de Consentimento Livre e Esclarecido. Aqueles membros que aceitaram levar seus filhos até a clínica, consentiram em sua participação. Para o presente projeto, uma descrição detalhada sobre o processo de coleta de dados via *web* será elaborada para ser submetida à apreciação do Comitê de Ética e Pesquisa (CEP) da Universidade Federal de Pelotas. Além disso, o termo de uso que deverá ser aceito por todos os participantes da plataforma *web* também será encaminhado para análise do CEP.

Para garantir a segurança da plataforma *web* a ser desenvolvida, algumas soluções serão implementadas. O acesso ao perfil de cada usuário exigirá um nome de usuário e uma senha definida pelo próprio participante, e nem mesmo o administrador do projeto terá acesso ao conteúdo dessa senha. Com o objetivo de evitar o acesso indevido aos dados presentes no servidor e o vazamento de informações, a máquina que irá hospedar a plataforma *web* contará com um *firewall*: ferramenta que filtra as comunicações e transações entre a rede local e a rede pública (internet), e que barra tentativas de ataque ou acesso indevido aos dados feitas através da rede pública (2). Para garantir o anonimato dos participantes, as informações de contato e identificação armazenadas no banco de dados serão criptografadas. Apenas o administrador do projeto terá acesso a estes dados. Todos os participantes, no ato do cadastro na plataforma *web*, devem estar de acordo com os termos de uso. O cadastro só será confirmado mediante aceitação dos termos por parte dos participantes.

7. CRONOGRAMA

Abaixo estão descritas as atividades a serem desenvolvidas durante o doutorado. Todas as atividades foram organizadas por trimestre, conforme descrito na Tabela 3.

- A) Revisão da literatura;
- B) Escrita do projeto;
- C) Prova de qualificação;
- D) Defesa do projeto;
- E) Desenvolvimento da plataforma *web*;
- F) Estágio sanduíche;
- G) Coleta dos dados;
- H) Análise dos dados;
- I) Escrita dos artigos;
- J) Escrita da tese;
- K) Defesa da tese.

Tabela 3: Cronograma de atividades conforme trimestres de cada ano de duração do doutorado.

Atividade	2015				2016				2017				2018			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
A	x	x	x	x	x	x	x	x	x	x	x	x	x			
B					x	x										
C						x										
D							x									
E							x	x	x	x						
F									x	x						
G											x	x				
H													x	x		
I													x	x	x	x
J													x	x	x	
K																x

8. DIVULGAÇÃO DOS RESULTADOS

Ao longo do desenvolvimento deste estudo, serão escritos artigos científicos para publicação em periódicos nacionais e/ou internacionais indexados. Embora os resultados acadêmicos sejam importantes, o acesso a este tipo de informação é limitado ao meio científico. Portanto, visando possibilitar que todos os participantes do estudo possam ter acesso fácil aos seus resultados, uma página *web* segura contendo resultados personalizados será desenvolvida na plataforma *web* a ser desenvolvida. Além do fornecimento de resultados individualizados, esta página *web* fará uma comparação dos resultados do participante com relação aos resultados da população da coorte de nascimentos de 1993 de Pelotas. Outros resultados relacionados à perfis de risco também serão fornecidos, os quais serão baseado em evidências de saúde bem estabelecidas no meio científico. A comparação dos resultados servirá, exclusivamente, para fins de informação e não como critério diagnóstico. Mesmo após o término da aplicação dos questionários, a página contendo os resultados dos participantes ficará disponível para consulta.

9. FINANCIAMENTO

A coorte de nascimentos de 1993 de Pelotas foi financiada através de recursos da *Wellcome Trust*. A União Europeia, o Programa de Apoio a Núcleos de Excelência (PRONEX), o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) e o Ministério da Saúde também financiaram algumas fases e acompanhamentos da coorte 1993. O doutorando Cauane Blumenberg Silva recebe bolsa do CNPq e também da Associação Brasileira de Saúde Coletiva (ABRASCO).

REFERÊNCIAS

1. Wang PS, Katila S. An introduction to Web design and programming. Belmont, CA: Thomson/Brooks/Cole; 2004. 572 p.
2. Coley CD, Wesinger RE. Firewall system for protecting network elements connected to a public network [Internet]. 1998. Available at: <https://www.google.com/patents/US5826014>
3. Porebski B, Przystalski K, Nowak L. Building PHP applications with Symfony, CakePHP, and Zend Framework. Hoboken, N.J. : Chichester: Wiley ; John Wiley [distributor]; 2011. 542 p. (Wrox programmer to programmer).
4. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. using game-design elements in non-gaming contexts. In ACM Press; 2011 [citado 5 de agosto de 2016]. p. 2425. Available at: <http://portal.acm.org/citation.cfm?doid=1979742.1979575>
5. Lie HW, Bos B. Cascading style sheets: designing for the Web. Harlow, England ; Reading, Mass: Addison Wesley Longman; 1997. 279 p.
6. Williams HE, Lane D. Web database applications with PHP and MySQL [Internet]. Sebastopol: O'Reilly; 2004 [citado 23 de julho de 2016]. Available at: <http://public.ebib.com/choice/publicfullrecord.aspx?p=540545>
7. American Heritage Dictionaries. American heritage dictionary of the english language. 5^o ed. [Place of publication not identified]: Houghton Mifflin Harcourt; 2015. 2112 p.
8. Gajic A, Cameron D, Hurley J. The cost-effectiveness of cash versus lottery incentives for a web-based, stated-preference community survey. Eur J Health Econ HEPAC Health Econ Prev Care. dezembro de 2012;13(6):789–99.
9. Groves RM. Survey errors and survey costs. Hoboken, N.J: Wiley; 2004. 590 p. (Wiley series in survey methodology).
10. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 8^o ed. AAPOR; 2015.
11. Goritz AS. Incentives in Web Studies: Methodological Issues and a Review. Int J Internet Sci. março de 2006;1(1):58–70.
12. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet Lond Engl. 26 de janeiro de 2002;359(9303):341–5.
13. Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet Lond Engl. 12 de janeiro de 2002;359(9301):145–9.
14. Centers for Disease Control and Prevention (CDC). Ten great public health achievements--worldwide, 2001-2010. MMWR Morb Mortal Wkly Rep. 24 de junho de 2011;60(24):814–8.

15. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol.* 1 de fevereiro de 2006;163(3):197–203.
16. Wallander L, Tikkanen RH, Mannheimer LN, Östergren P-O, Plantin L. The problem of non-response in population surveys on the topic of HIV and sexuality: a comparative study. *Eur J Public Health.* fevereiro de 2015;25(1):172–7.
17. Galea S, Tracy M. Participation Rates in Epidemiologic Studies. *Ann Epidemiol.* setembro de 2007;17(9):643–53.
18. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health.* 1 de setembro de 2005;27(3):281–91.
19. Dillman DA, Smyth JD, Christian LM. *Internet, phone, mail, and mixed-mode surveys: the tailored design method.* 4. ed. Hoboken, NJ: Wiley; 2014. 509 p.
20. Blinson K, Dignan M, Michielutte R, Wells HB. The cost of conducting face-to-face household interviews in a rural, Native American population. The North Carolina Native American Cervical Cancer Prevention Project. *Cancer.* 1 de outubro de 1996;78(7 Suppl):1587–91.
21. Bryman A. *Social research methods.* Fifth edition. Oxford: Oxford University Press; 2016. 747 p.
22. Akl EA, Maroun N, Klocke RA, Montori V, Schünemann HJ. Electronic mail was not better than postal mail for surveying residents and faculty. *J Clin Epidemiol.* abril de 2005;58(4):425–9.
23. Kroth PJ, McPherson L, Leverence R, Pace W, Daniels E, Rhyne RL, et al. Combining web-based and mail surveys improves response rates: a PBRN study from PRIME Net. *Ann Fam Med.* junho de 2009;7(3):245–8.
24. Groves RM. Three Eras of Survey Research. *Public Opin Q.* 1 de dezembro de 2011;75(5):861–71.
25. Johnson TP, organizador. *Handbook of health survey methods.* Hoboken, New Jersey: John Wiley & Sons, Inc; 2015. 816 p. (Wiley handbooks in survey methodology).
26. Hu SS, Balluz L, Battaglia MP, Frankel MR. The impact of cell phones on public health surveillance. *Bull World Health Organ.* 1 de novembro de 2010;88(11):799–799.
27. Blumberg SJ, Ganesh N, Luke JV, Gonzales G. Wireless substitution: state-level estimates from the National Health Interview Survey, 2012. *Natl Health Stat Rep.* 18 de dezembro de 2013;(70):1–16.
28. Lim MSC, Hocking JS, Hellard ME, Aitken CK. SMS STI: a review of the uses of mobile phone text messaging in sexual health. *Int J STD AIDS.* 1 de maio de 2008;19(5):287–90.
29. Kuntsche E, Robert B. Short message service (SMS) technology in alcohol research--a feasibility study. *Alcohol Alcohol Oxf Oxf.* agosto de 2009;44(4):423–8.

30. Steeh C, Buskirk TD, Callegaro M. Using Text Messages in U.S. Mobile Phone Surveys. *Field Methods*. 1 de fevereiro de 2007;19(1):59–75.
31. Hesse BW. Public Health Informatics. In: Gibbons MC, organizador. *eHealth Solutions for Healthcare Disparities* [Internet]. New York, NY: Springer New York; 2008 [citado 20 de maio de 2016]. p. 109–26. Available at: http://link.springer.com/10.1007/978-0-387-72815-5_11
32. Blumenberg C, Barros AJD. Electronic data collection in epidemiological research: The use of REDCap in the Pelotas birth cohorts. *Appl Clin Inform*. 2016;7(3):672–81.
33. Internet users (per 100 people) [Internet]. The World Bank. 2016 [citado 20 de julho de 2016]. Available at: <http://data.worldbank.org/indicator/IT.NET.USER.P2?end=2014&start=2014&view=bar>
34. Sax LJ, Gilmartin SK, Bryant AN. Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys. *Res High Educ*. 44(4):409–32.
35. Pizzi C, De Stavola BL, Pearce N, Lazzarato F, Ghiotti P, Merletti F, et al. Selection bias and patterns of confounding in cohort studies: the case of the NINFEA web-based birth cohort. *J Epidemiol Community Health*. novembro de 2012;66(11):976–81.
36. Hatch EE, Hahn KA, Wise LA, Mikkelsen EM, Kumar R, Fox MP, et al. Evaluation of Selection Bias in an Internet-based Study of Pregnancy Planners: *Epidemiology*. janeiro de 2016;27(1):98–104.
37. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-Based Versus Traditional Paper Questionnaires: A Mixed-Mode Survey With a Nordic Perspective. *J Med Internet Res*. 26 de agosto de 2013;15(8):e173.
38. Hunter J, Corcoran K, Leeder S, Phelps K. Is it time to abandon paper? The use of emails and the Internet for health services research—a cost-effectiveness and qualitative study. *J Eval Clin Pract*. outubro de 2013;19(5):855–61.
39. Ekman A, Dickman PW, Klint Å, Weiderpass E, Litton J-E. Feasibility of Using Web-based Questionnaires in Large Population-based Epidemiological Studies. *Eur J Epidemiol*. fevereiro de 2006;21(2):103–11.
40. Richiardi L, Baussano I, Vizzini L, Douwes J, Pearce N, Merletti F, et al. Feasibility of recruiting a birth cohort through the Internet: the experience of the NINFEA cohort. *Eur J Epidemiol*. 2007;22(12):831–7.
41. Apostolico MR, Egry EY. Uso da internet na coleta de dados primários na pesquisa em Enfermagem. *Rev Bras Enferm*. dezembro de 2013;66(6):949–55.
42. Dainesi SM, Goldbaum M. E-survey with researchers, members of ethics committees and sponsors of clinical research in Brazil: an emerging methodology for scientific research. *Rev Bras Epidemiol*. dezembro de 2012;15(4):705–13.
43. Schleyer TK, Forrest JL. Methods for the design and administration of web-based surveys. *J Am Med Inform Assoc JAMIA*. agosto de 2000;7(4):416–25.

44. Wyatt JC. When to use web-based surveys. *J Am Med Inform Assoc JAMIA*. agosto de 2000;7(4):426–9.
45. Brice A, Price A, Burls A. Creating a database of internet-based clinical trials to support a public-led research programme: A descriptive analysis. *Digit Health [Internet]*. 20 de novembro de 2015 [citado 10 de abril de 2016];1(0). Available at: <http://dhj.sagepub.com/lookup/doi/10.1177/2055207615617854>
46. Martins Y, Lederman RI, Lowenstein CL, Joffe S, Neville BA, Hastings BT, et al. Increasing response rates from physicians in oncology research: a structured literature review and data from a recent physician survey. *Br J Cancer*. 13 de março de 2012;106(6):1021–6.
47. van Gelder MMHJ, Bretveld RW, Roeleveld N. Web-based Questionnaires: The Future in Epidemiology? *Am J Epidemiol*. 1 de dezembro de 2010;172(11):1292–8.
48. Sinclair M, O’Toole J, Malawaraarachchi M, Leder K. Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Med Res Methodol*. 2012;12:132.
49. Bälter KA, Bälter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiol Camb Mass*. julho de 2005;16(4):577–9.
50. Clark M, Rogers M, Foster A, Dvorchak F, Saadeh F, Weaver J, et al. A randomized trial of the impact of survey design characteristics on response rates among nursing home providers. *Eval Health Prof*. dezembro de 2011;34(4):464–86.
51. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study “Snart-Gravid”. *Eur J Epidemiol*. maio de 2010;25(5):297–304.
52. McCabe SE, Boyd CJ, Couper MP, Crawford S, D’Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. *J Stud Alcohol*. novembro de 2002;63(6):755–61.
53. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. *Comput Hum Behav*. março de 2010;26(2):132–9.
54. Instituto Brasileiro de Geografia e Estatística, organizador. Acesso à Internet e à televisão e posse de telefone móvel celular para uso pessoal, 2014: Pesquisa Nacional por Amostra de Domicílios. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística - IBGE; 2016. 84 p.
55. US National Library of Medicine National Institutes of Health [Internet]. PubMed - NCBI. [citado 21 de fevereiro de 2016]. Available at: <http://www.ncbi.nlm.nih.gov/pubmed>
56. Definition of an older or elderly person [Internet]. World Health Organization. [citado 28 de julho de 2016]. Available at: <http://www.who.int/healthinfo/survey/ageingdefnolder/en/>

57. Shih T-H, Xitao Fan. Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods*. 3 de março de 2008;20(3):249–71.
58. Web Survey Methodology [Internet]. WebSM.org. [citado 21 de fevereiro de 2016]. Available at: <http://www.websm.org>
59. Top 50 countries with the highest internet penetration rates - 2013 [Internet]. Internet World Stats. [citado 18 de agosto de 2016]. Available at: <http://www.internetworldstats.com/top25.htm>
60. Pealer LN, Weiler RM, Pigg RM, Miller D, Dorman SM. The feasibility of a web-based surveillance system to collect health risk behavior data from college students. *Health Educ Behav Off Publ Soc Public Health Educ*. outubro de 2001;28(5):547–59.
61. Andersson G, Lindvall N, Hursti T, Carlbring P. Hypersensitivity to sound (hyperacusis): a prevalence study conducted via the Internet and post. *Int J Audiol*. dezembro de 2002;41(8):545–54.
62. Goldman RD, Mehrotra S, Pinto TR, Mounstephen W. Follow-up after a pediatric emergency department visit: telephone versus e-mail? *Pediatrics*. outubro de 2004;114(4):988–91.
63. McCabe SE. Comparison of web and mail surveys in collecting illicit drug use data: a randomized experiment. *J Drug Educ*. 2004;34(1):61–72.
64. McCabe SE, Boyd CJ, Young A, Crawford S. Feasibility study for collecting alcohol and other drug use data among secondary school students: a web-based survey approach. *J Drug Educ*. 2004;34(4):373–83.
65. Kypri K, Stephenson S, Langley J. Assessment of nonresponse bias in an internet survey of alcohol use. *Alcohol Clin Exp Res*. abril de 2004;28(4):630–4.
66. Mangunkusumo RT, Moorman PW, Van Den Berg-de Ruiters AE, Van Der Lei J, De Koning HJ, Raat H. Internet-administered adolescent health questionnaires compared with a paper version in a randomized study. *J Adolesc Health Off Publ Soc Adolesc Med*. janeiro de 2005;36(1):70.e1–6.
67. Satia JA, Galanko JA, Rimer BK. Methods and strategies to recruit African Americans into cancer prevention surveillance studies. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. março de 2005;14(3):718–21.
68. McCabe SE, Diez A, Boyd CJ, Nelson TF, Weitzman ER. Comparing web and mail responses in a mixed mode survey in college alcohol use research. *Addict Behav*. setembro de 2006;31(9):1619–27.
69. Rodriguez HP, von Glahn T, Rogers WH, Chang H, Fanjiang G, Safran DG. Evaluating patients' experiences with individual physicians: a randomized trial of mail, internet, and interactive voice response telephone administration of surveys. *Med Care*. fevereiro de 2006;44(2):167–74.

70. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response Rate and Completeness of Questionnaires: A Randomized Study of Internet Versus Paper-and-Pencil Versions. *J Med Internet Res.* 30 de setembro de 2007;9(3):e25.
71. Ekman A, Klint A, Dickman PW, Adami H-O, Litton J-E. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. *Eur J Epidemiol.* 2007;22(5):293–300.
72. Brøgger J, Nystad W, Cappelen I, Bakke P. No increase in response rate by adding a web response option to a postal population survey: a randomized trial. *J Med Internet Res.* 2007;9(5):e40.
73. Dinitto DM, Bridget Busch-Armendariz N, Bender K, Hyeyoung Woo null, Tackett-Gibson M, Dyer J. Testing telephone and web surveys for studying men's sexual assault perpetration behaviors. *J Interpers Violence.* outubro de 2008;23(10):1483–93.
74. Blank D, Hohgraefe Neto G, Grando E, Siqueira PZ, Lunkes RP, Pietroboli JL, et al. A Web-based survey on students' conceptions of "accident". *Inform Health Soc Care.* dezembro de 2009;34(4):189–208.
75. Balajti I, Daragó L, Adány R, Kósa K. College students' response rate to an incentivized combination of postal and web-based health survey. *Eval Health Prof.* junho de 2010;33(2):164–76.
76. Lim MSC, Sacks-Davis R, Aitken CK, Hocking JS, Hellard ME. Randomised controlled trial of paper, online and SMS diaries for collecting sexual behaviour information from young people. *J Epidemiol Community Health.* outubro de 2010;64(10):885–9.
77. Atherton H, Oakeshott P, Aghaizu A, Hay P, Kerry S. Use of an online questionnaire for follow-up of young female students recruited to a randomised controlled trial of chlamydia screening. *J Epidemiol Community Health.* julho de 2010;64(7):580–4.
78. Scott-Johnson PE, Gross SM, Browne DC. Web-based data collection: an effective strategy for increasing African Americans' participation in health-related research. *Ethn Dis.* 2010;20(1 Suppl 1):S1–201 – 6.
79. Sinadinovic K, Wennberg P, Berman AH. Population screening of risky alcohol and drug use via Internet and Interactive Voice Response (IVR): a feasibility and psychometric study in a random sample. *Drug Alcohol Depend.* 1 de março de 2011;114(1):55–60.
80. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. *J Med Internet Res.* 2011;13(4):e109.
81. Lagerros YT, Sandin S, Bexelius C, Litton J-E, Löf M. Estimating physical activity using a cell phone questionnaire sent by means of short message service (SMS): a randomized population-based study. *Eur J Epidemiol.* julho de 2012;27(7):561–6.
82. Dokkum NFB, Koekenbier RH, van den Broek IVF, van Bergen JEAM, Brouwers EEHG, Fennema JSA, et al. Keeping participants on board: increasing uptake by

automated respondent reminders in an Internet-based chlamydia screening in the Netherlands. *BMC Public Health*. 2012;12:176.

83. Bailey JV, Pavlou M, Copas A, McCarthy O, Carswell K, Rait G, et al. The Sexunzipped trial: optimizing the design of online randomized controlled trials. *J Med Internet Res*. 2013;15(12):e278.

84. Elliott MN, Brown JA, Lehrman WG, Beckett MK, Hambarsoomian K, Giordano LA, et al. A randomized experiment investigating the suitability of speech-enabled IVR and Web modes for publicly reported surveys of patients' experience of hospital care. *Med Care Res Rev MCCR*. abril de 2013;70(2):165–84.

85. Patrick ME, Singer E, Boyd CJ, Cranford JA, McCabe SE. Incentives for college student participation in web-based substance use surveys. *Addict Behav*. março de 2013;38(3):1710–4.

86. Flüb E, Bond CM, Jones GT, Macfarlane GJ. The effect of an internet option and single-sided printing format to increase the response rate to a population-based study: a randomized controlled trial. *BMC Med Res Methodol*. 2014;14:104.

87. Johnson S, Seaton SE, Manktelow BN, Smith LK, Field D, Draper ES, et al. Telephone interviews and online questionnaires can be used to improve neurodevelopmental follow-up rates. *BMC Res Notes*. 2014;7:219.

88. Griffin J, Patrick ME. Nonresponse bias in a longitudinal measurement design examining substance use across the transition out of high school. *Drug Alcohol Depend*. 1 de outubro de 2014;143:232–8.

89. Conde K, Cremonte M. [Data quality in surveys on alcohol consumption among university students]. *Cad Saúde Pública*. janeiro de 2015;31(1):39–47.

90. Couturier B, Carrat F, Hejblum G, SENTIPAT Study Group. Comparing Patients' Opinions on the Hospital Discharge Process Collected With a Self-Reported Questionnaire Completed Via the Internet or Through a Telephone Survey: An Ancillary Study of the SENTIPAT Randomized Controlled Trial. *J Med Internet Res*. 2015;17(6):e158.

91. Galante J, Adamska L, Young A, Young H, Littlejohns TJ, Gallacher J, et al. The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *Br J Nutr*. fevereiro de 2016;115(04):681–6.

92. Young JM, O'Halloran A, McAulay C, Pirotta M, Forsdike K, Stacey I, et al. Unconditional and conditional incentives differentially improved general practitioners' participation in an online survey: randomized controlled trial. *J Clin Epidemiol*. junho de 2015;68(6):693–7.

93. Wilson PM, Petticrew M, Calnan M, Nazareth I. Effects of a financial incentive on health researchers' response to an online survey: a randomized controlled trial. *J Med Internet Res*. 2010;12(2):e13.

94. Edwards PJ, Roberts I, Clarke MJ, DiGiuseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. In: The Cochrane Collaboration, organizador. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2009 [citado 21 de fevereiro de 2016]. Available at: <http://doi.wiley.com/10.1002/14651858.MR000008.pub4>
95. Horta BL, Gigante DP, Gonçalves H, dos Santos Motta J, Loret de Mola C, Oliveira IO, et al. Cohort Profile Update: The 1982 Pelotas (Brazil) Birth Cohort Study. *Int J Epidemiol.* abril de 2015;44(2):441, 441a – 441e.
96. Gonçalves H, Assunção MC, Wehrmeister FC, Oliveira IO, Barros FC, Victora CG, et al. Cohort profile update: The 1993 Pelotas (Brazil) birth cohort follow-up visits in adolescence. *Int J Epidemiol.* agosto de 2014;43(4):1082–8.
97. Santos IS, Barros AJD, Matijasevich A, Zanini R, Chrestani Cesar MA, Camargo-Figuera FA, et al. Cohort profile update: 2004 Pelotas (Brazil) Birth Cohort Study. Body composition, mental health and genetic assessment at the 6 years follow-up. *Int J Epidemiol.* outubro de 2014;43(5):1437–1437a – f.
98. Ubuntu Server [Internet]. Ubuntu. 2016 [citado 21 de julho de 2016]. Available at: <http://www.ubuntu.com/server>
99. Duckett J. HTML & CSS: design and build websites. Indianapolis, IN: Wiley; 2011. 490 p.
100. PHP: Hypertext Processor [Internet]. PHP: Hypertext Processor. 2016 [citado 21 de julho de 2016]. Available at: <http://php.net>
101. Symfony [Internet]. Symfony. 2016 [citado 21 de julho de 2016]. Available at: <http://symfony.com>
102. Bootstrap [Internet]. Bootstrap. 2016 [citado 21 de julho de 2016]. Available at: <http://getbootstrap.com>
103. MySQL [Internet]. MySQL. 2016 [citado 21 de julho de 2016]. Available at: <https://www.mysql.com>
104. Stata Statistical Software: Release 13. College Station, TX: StataCorp; 2013.
105. Barros AJD, Victora CG. Indicador econômico para o Brasil baseado no censo demográfico de 2000. *Rev Saúde Pública.* agosto de 2005;39(4):523–9.
106. Gordis L. *Epidemiology.* Fifth edition. Philadelphia, PA: Elsevier/Saunders; 2014. 392 p.

ANEXO 1

Questões aplicadas no questionário do acompanhamento presencial dos 22/23 anos da coorte de nascimentos de 1993 para avaliação do perfil de uso de internet e dispositivos que se conectam à internet.

<p>D038. Vocês têm computador de mesa ou notebook?</p> <p>(0) Não (1) Sim</p> <p><i>SE SIM:</i></p> <p>D038b. Este(s) computador(es) tem acesso a internet 24 horas?</p> <p>(0) Não (1) Sim</p>
<p>L144. Pensando na maneira como tu conversas com teus amigos quando eles não estão por perto. Qual o <u>principal</u> meio de comunicação que tu usas? <i>Ler opções</i></p> <p>(1) Chamada de telefone fixo (2) Chamada de telefone celular (6) E-mail (3) SMS/torpedo (5) Facebook (4) WhatsApp</p>
<p>L145. Pensando na maneira como tu conversas com teus familiares quando eles não estão por perto. Qual o <u>principal</u> meio de comunicação que tu usas? <i>Ler opções</i></p> <p>(1) Chamada de telefone fixo (2) Chamada de telefone celular (6) E-mail (3) SMS/torpedo (5) Facebook (4) WhatsApp</p>
<p>MARCAR O QUE O ENTREVISTADO RELATAR E PERGUNTAR OS DEMAIS:</p> <p>L147. Dos aplicativos para celular ou <i>tablet</i> que vou te ler, quais são aqueles que tu usas mais..? <i>Ler opções/Múltipla escolha</i></p> <p>(1) WhatsApp (2) Facebook (3) Instagram (4) Twitter (5) Viber (6) Youtube (7) Snapchat (8) Gmail/e-mail</p>

<p>(9) Tinder (10) Outro (0) Não usa aplicativos</p> <p><i>SE OUTRO:</i> L147a. Qual? _____</p>
<p><i>SE NÃO RELATOU NENHUMA REDE SOCIAL NAS QUESTÕES L144, L145, L147:</i> L148. Tu usas redes sociais como Facebook, Twitter ou Instagram? (0) Não (1) Sim</p>
<p><i>SE SIM NA L148 OU SE RELATOU ALGUMA REDE SOCIAL NAS QUESTÕES L144, L145, L147:</i> <i>MARCAR O QUE O ENTREVISTADO RELATAR E PERGUNTAR OS DEMAIS:</i> L149. Quais são as redes sociais que tu usas? <i>Ler opções/Múltipla escolha</i> (1) Facebook (2) Twitter (3) Instagram (4) Google Plus (5) LinkedIn (6) Badoo (7) Tinder (8) Whatsapp (10) Snapchat (9) Outra</p> <p><i>SE OUTRA:</i> L149a. Qual? _____</p> <p><i>REPETIR PARA CADA UMA DAS REDES SOCIAIS MARCADAS:</i> L150-L150n. Quantos dias por semana tu acessas o <REDE SOCIAL 1 até 10>? __</p>
<p><i>SE SIM NA L148 OU SE RELATOU ALGUMA REDE SOCIAL NAS QUESTÕES L144, L145, L147:</i> L151. Na maioria das vezes, tu acessas as redes sociais no..? <i>Ler opções</i> (1) Computador (2) Celular (3) Tablet</p>
<p>D223. Tu usas o celular para acessar a internet ou jogar no teu tempo livre? (0) Não → vá para D224 (1) Sim</p>

<p><i>SE SIM:</i></p> <p>D223a. Agora, eu quero que tu não consideres as olhadas rápidas no celular e me digas quanto tempo no total, em um dia normal, tu usas o celular para acessar a internet ou jogar de segunda a sexta-feira?</p> <p>__ __ horas __ __ minutos</p> <p><i>SE SIM:</i></p> <p>L152. A internet que tu usas no celular é? Ler opções</p> <p>(1) Somente WiFi (2) Somente 3G/4G (3) WiFi e 3G/4G (4) Não usa internet</p>
<p>D224. Tu usas o <i>tablet</i> para acessar a internet ou jogar no teu tempo livre?</p> <p>(0) Não → vá para D225 (1) Sim</p>
<p><i>SE SIM:</i></p> <p>D224a. Quanto tempo tu usas o <i>tablet</i> para internet ou jogar em um dia de semana sem ser sábado e domingo? __ __ horas __ __ minutos</p> <p><i>SE SIM:</i></p> <p>L153. A internet que tu usas no <i>tablet</i> é? Ler opções</p> <p>(1) Somente WiFi (2) Somente 3G/4G (3) WiFi e 3G/4G (4) Não usa internet</p>
<p>L146. Na maior parte do tempo, tu usas a internet no...? Ler opções</p> <p>(1) Computador (2) Celular (3) Tablet (0) Não usa a internet</p>
<p>D225. Tu usas computador em casa?</p> <p>(0) Não → vá para D226 (1) Sim</p>
<p><i>SE SIM:</i></p> <p>D225a. Quanto tempo tu ficas no computador em um dia de semana sem ser sábado e domingo? __ __ horas __ __ minutos</p>

2 – Ajustes ao projeto durante o curso do trabalho

Ajustes referentes à logística

No projeto original estava prevista a utilização de diversos meios de recrutamento para convidar os indivíduos elegíveis a se cadastrar na plataforma *coortesnaweb*, inclusive o uso de correspondências enviadas por correio. Visto que este projeto não contou com nenhum tipo de financiamento, tivemos que optar por métodos de recrutamento gratuitos. Dessa forma, o envio de correspondências não foi empregado, e convites através de mensagens enviadas por e-mail, Whatsapp e Facebook foram empregados.

O recrutamento, que tinha início previsto para setembro de 2017, teve início no mês de janeiro de 2018 e duração de seis meses. O atraso no início do recrutamento se deu em virtude da data do período de doutoramento sanduíche, que foi realizado durante o segundo e terceiro trimestres de 2017. Assim como o recrutamento, a coleta de dados teve início no mês de janeiro de 2018. Ao invés de remover os questionários da plataforma, conforme era previsto no projeto original, optamos por manter os questionários ativos ao longo de todo o período do estudo. Essa decisão se deu com o objetivo de maximizar as taxas de resposta dos questionários, possibilitando a participação dos participantes mesmo após o fim do envio de lembretes aos não respondentes.

Ajustes referentes à análise

No projeto original um dos desfechos a ser analisado era a taxa de contato. Após revisão de literatura, optamos por chamar esse desfecho de taxa de recrutamento. Ambos termos são usados para calcular o percentual de indivíduos cadastrados em uma plataforma *web*, dividindo-se o número de indivíduos cadastrados pelo total de indivíduos convidados a se cadastrar. Optamos pelo uso da nomenclatura taxa de recrutamento, porque esta foi definida pela Associação Americana para Pesquisas de Opinião Pública (1), mesma instituição que definiu os demais desfechos que utilizamos nessa tese de doutorado.

Ao contrário da proposta original, os desfechos não foram analisados conforme as variáveis de uso de dispositivos e internet, mas somente conforme as características demográficas e socioeconômicas dos participantes. Isso foi feito para simplificar os resultados apresentados nos artigos produzidos.

Adicionamos um novo critério de elegibilidade às análises da taxa de recrutamento conforme o método utilizado para o recrutamento. Apenas os indivíduos que tinham fornecido um endereço de e-mail, um número de Whatsapp e um nome de perfil no Facebook (n = 2394) foram considerados elegíveis para esta análise. Apesar disso, todos os demais participantes que se encaixavam nos demais critérios de elegibilidade foram convidados a se registrar, porém não fizeram parte das análises referentes à taxa de recrutamento.

Referência

1. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9^o ed. AAPOR; 2016.

3 – Trabalho de campo

Participação em trabalhos de campo das coortes de Pelotas

Durante os anos de 2015 e 2016, fiz parte da equipe que planejou e conduziu os acompanhamentos dos 11 e dos 22 anos das Coortes de Nascimentos de Pelotas de 2004 e 1993, respectivamente. Meu principal papel foi na elaboração dos instrumentos eletrônicos de coleta de dados por meio do *software* REDCap (1), e também no monitoramento e manutenção dos bancos de dados durante os acompanhamentos. A experiência que adquiri no acompanhamento da Coorte 2004 teve como produto a publicação de um artigo que narra as vantagens obtidas ao se utilizar o *software* REDCap para a condução de pesquisas epidemiológicas (2).

Além da minha participação em dois acompanhamentos de coorte, fui responsável pela gerência dos bancos de dados da Coorte de Nascimentos de Pelotas de 2004. A gerência de tais dados envolveu diversas etapas, incluindo organização, limpeza, unificação, documentação, anonimização e distribuição dos dados.

A primeira etapa envolveu a organização e armazenamento dos dados na nuvem, os quais foram organizados de acordo com o ano de acompanhamento. Em seguida, realizei a limpeza dos bancos, eliminando arquivos duplicados. O processo de unificação consistiu em organizar todos os dados referentes a um acompanhamento da coorte em um mesmo banco de dados, visando facilitar a análise e o processo de distribuição. Uma vez que os bancos estavam organizados, limpos e unificados, passou-se para a etapa de documentação. Com o auxílio da equipe da Coorte 2004, construí dicionários de variáveis para serem disponibilizados aos pesquisadores. Nesses dicionários existem informações sobre o nome e rótulo das variáveis geradas a partir de questionários, exames de composição corporal e variáveis secundárias. Esses dicionários têm como principal objetivo facilitar a distribuição dos dados para fins de análise, pois facilitam o processo de escolha, separação e envio de variáveis para serem analisadas pelos pesquisadores.

Por fim, desenvolvi um *script* para anonimização dos dados de pesquisa em parceria com o professor Aluísio J. D. Barros. O principal objetivo desse *script* é gerar um número de identificação aleatório para as observações dos bancos de dados distribuídos para análise. Isso impossibilita a identificação dos participantes e também evita o reuso do mesmo banco de dados para a realização de análises que não foram previstas anteriormente.

Período de doutorado sanduíche, Universidade de Turim, Itália

Entre abril e outubro de 2017, trabalhei com o grupo de pesquisa do projeto *Nascita e Infanzia* (NINFEA), coordenado pelo professor Lorenzo Richiardi da Universidade de Turim (Itália). A NINFEA é uma coorte de nascimentos italiana conduzida inteiramente pela internet, e que conta com uma amostra de mais de 7.000 pessoas. Meu contato com a equipe do projeto NINFEA serviu para elaborar estratégias de pesquisa via *web* com o objetivo de atingir altas taxas de

participação. Dentre as estratégias elaboradas e que foram aplicadas à plataforma *coortesnaweb* estão os desenhos dos questionários, o controle de acesso à plataforma e o modelo de contato com os participantes.

Durante o período de doutoramento sanduíche, também tive a oportunidade de trabalhar com os dados do projeto NINFEA. Como o principal foco desta tese de doutorado é na análise de fatores relacionados à metodologia de pesquisa via *web*, liderei um estudo seguindo essa mesma linha. Analisei a influência de fatores demográficos e socioeconômicos sobre a taxa de desistência e não resposta aos itens do primeiro questionário respondido pelos participantes do projeto NINFEA. Publicamos esse artigo no Journal of Medical Internet Research (3).

Condução do acompanhamento através da plataforma coortesnaweb

Aqui estão descritas todas as atividades referentes ao desenvolvimento, recrutamento e coleta de dados utilizando a plataforma *coortesnaweb*.

Desenvolvimento da plataforma coortesnaweb

Entre novembro de 2016 e dezembro de 2017, foi desenvolvida a plataforma *coortesnaweb*. Para isso foram utilizadas ferramentas gratuitas: ambiente LAMP (Linux, Apache, MySQL e PHP), o *framework* Symfony e CSS Bootstrap. O visual da plataforma foi elaborado em parceria com a designer digital Cíntia Borges, e os demais recursos da plataforma foram desenvolvidos por mim utilizando a linguagem PHP através do *framework* Symfony. Optou-se por utilizar um *framework* para que o desenvolvimento da plataforma se desse de maneira mais ágil, haja visto que existem diversas ferramentas desenvolvidas por terceiros que podem ser aproveitadas.

Uma dessas ferramentas é o pacote *FOSUserBundle*, o qual foi utilizado para controlar o registro e o acesso dos participantes à plataforma. Utilizando esse pacote, construímos um formulário de registro que coletava o e-mail, o nome completo, a data de nascimento, o CPF e uma senha definida pelo participante.

A plataforma *coortesnaweb* foi desenvolvida contendo quatro módulos principais: responder questionários, verificar resultados, consultar as conquistas e o *ranking*, e atualizar de dados (Figura 1).

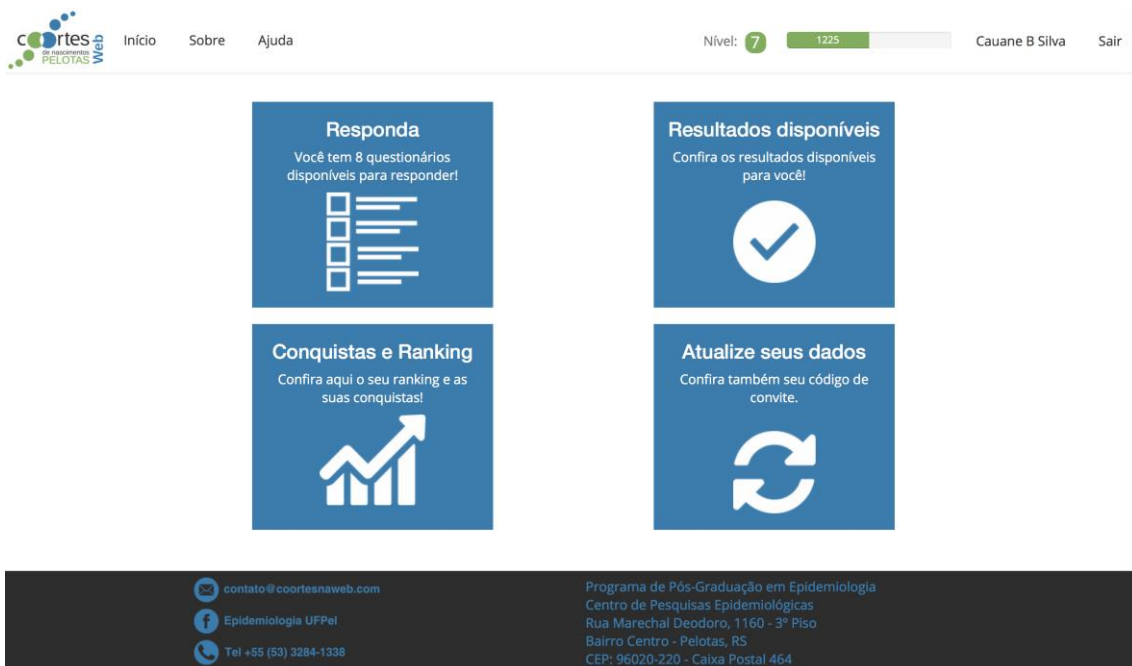


Figura 1. Tela inicial da plataforma *coortesnaweb* contendo seus quatro módulos principais: responder questionários, verificar resultados, consultar conquistas e *ranking* e atualizar dados.

Conforme mostrado na Figura 2, o módulo para responder questionários contém a lista de questionários disponíveis para serem respondidos pelos participantes. O assunto de cada questionário é exibido na tela, juntamente com o tempo estimado para responder e a quantidade de pontos que é conquistada ao finalizar seu preenchimento. Os questionários foram desenvolvidos utilizando a função de pesquisa do REDCap. Os questionários desenvolvidos foram hospedados em um servidor diferente do utilizado para hospedar a plataforma *coortesnaweb*. Dessa forma, as informações de cadastro na plataforma *coortesnaweb* e as informações de saúde coletas pelos questionários são armazenadas de maneira descentralizada. Assim que um questionário é respondido, ele é automaticamente removido da lista de questionários disponíveis para serem respondidos.

Pontos	Pesquisa	Duração
6 pontos	Medidas coporais	1 minutos
45 pontos	Consumo de álcool	8 minutos
48 pontos	Atividade física	8 minutos
39 pontos	Uso da Internet	7 minutos
51 pontos	Agressão e violência	9 minutos
36 pontos	Consumo alimentar	6 minutos
60 pontos	Qualidade do sono	10 minutos
33 pontos	Desejo de fumar	6 minutos

Figura 2. Módulo contendo a lista de questionários disponíveis para serem respondidos.

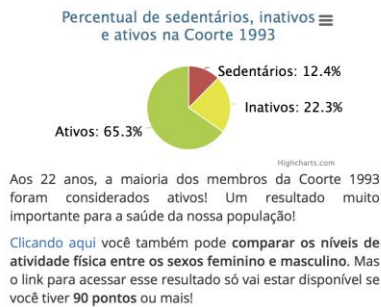
O módulo para verificação de resultados foi desenvolvido com um desenho similar à lista de questionários (Figura 3). A consulta aos resultados é dependente de pelo menos uma das seguintes condições: i) quantidade de pontos acumulados pelo usuário; e ii) ter respondido a um questionário específico. Os resultados disponíveis para consulta são marcados em verde. Aqueles ainda bloqueados, porque uma das condições não foi satisfeita, são marcados em cinza. Ao clicar em um dos resultados disponíveis para consulta, é apresentado para o participante seu resultado individual (considerando as respostas fornecidas para os questionários respondidos). Por exemplo, na Figura 4 é apresentado o resultado referente ao nível de atividade física do participante. Além disso são exibidas estimativas agregadas da Coorte 1993 e também informações que explicam a importância da prática de atividade física.

Aqui você pode conferir resultados personalizados sobre a sua saúde. Aqueles marcados em verde estão disponíveis, os demais você deve conquistar a pontuação necessária e responder às pesquisas necessárias para liberá-los!

Pontos necessários	Pesquisa necessária	Resultado
25 pontos	Medidas corporais	Status de IMC
70 pontos	Atividade física	Nível de atividade física
170 pontos	Uso da Internet	Perfil de uso da Internet
10 pontos	-	Exame de sangue aos 22 anos
300 pontos	Qualidade do sono	Qualidade do sono

Figura 3. Módulo para verificação dos resultados. Resultados disponíveis são marcados em verde, já os bloqueados são marcados em cinza.

Confira o seu nível de atividade física



Atividade física recomendada

A Organização Mundial da Saúde (OMS) recomenda que adultos entre 18 e 64 anos devem fazer no mínimo **150 minutos de atividade física por semana**. Isso inclui tanto as atividades físicas praticadas por lazer (por exemplo: futebol, corrida, natação, tênis, etc.), quanto aquelas atividades realizadas no trabalho, em casa ou para ir de um lugar ao outro. A classificação da OMS é feita da seguinte forma:

- **Ativos:** pessoas que praticam atividade física por 150 minutos ou mais por semana.
- **Inativos:** pessoas que fazem menos de 150 minutos semanais de atividade física.
- **Sedentárias:** pessoas que não fazem nenhuma atividade física.

Uma pessoa considerada ativa pode reduzir a probabilidade de ter doenças do coração, diabetes, câncer, fraturas e muitos outros problemas de saúde!

*Fonte: Organização Mundial da Saúde

Seu nível de atividade física

De acordo com as suas respostas, você realiza em média **165 minutos** de atividade física por semana. Com esse resultado, você é considerado **ativo**.



Figura 4. Tela descrevendo o resultado individual do nível de atividade física do participante, baseado nas respostas fornecidas nos questionários respondidos.

No módulo para consulta das conquistas e *ranking*, é apresentado o ranking na parte esquerda e as conquistas na parte direita da tela (Figura 5). O *ranking* foi desenvolvido com base na pontuação total do participante. O participante com o maior número de pontos estará a frente de 100% dos usuários cadastrados. Diversas conquistas foram adicionadas à plataforma coortesweb. Aquelas marcadas em azul já foram obtidas, enquanto as marcadas em cinza ainda não foram obtidas. Abaixo de cada conquista é descrita a condição que o participante deve

atingir para obter a conquista. A obtenção das conquistas garante uma pontuação extra para os usuários

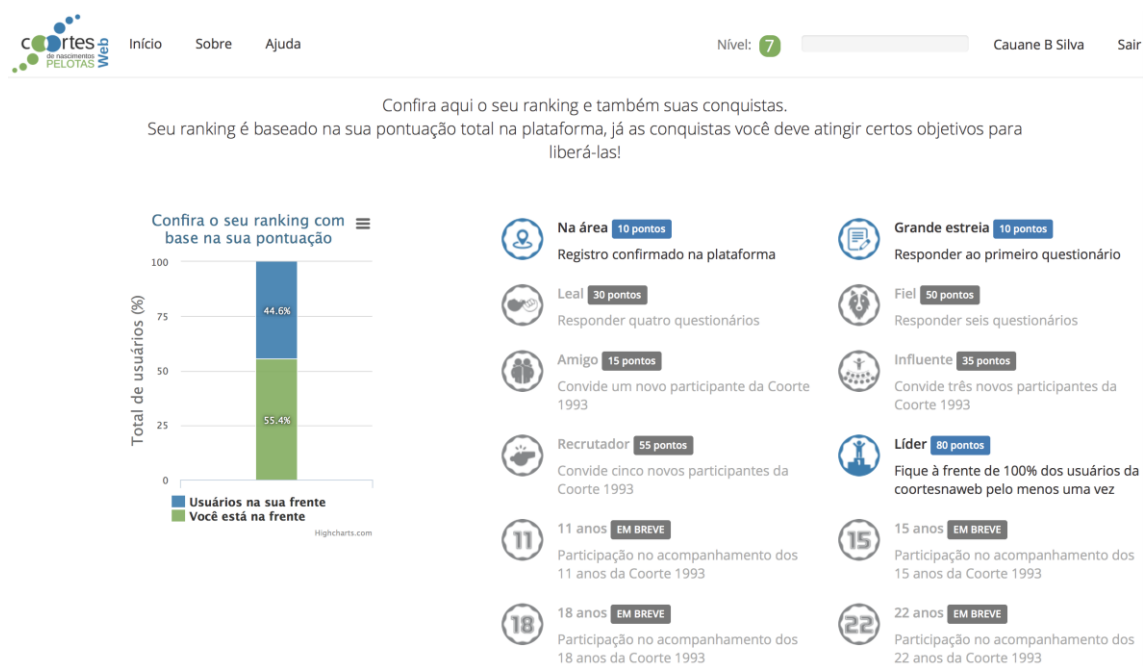


Figura 5. Módulo para consulta de conquistas e *ranking*. Conquista obtidas são marcadas em azul, conquistas bloqueadas são marcadas em cinza.

O último módulo desenvolvido é utilizado para atualizar os dados dos participantes cadastrados. Conforme mostrado na Figura 6, os participantes podem atualizar informações básicas de cadastro e contato.

Ao convidar mais membros da Coorte 1993 para se cadastrar você ganha pontos, e quem você convidou também ganha pontos!

Seu código de convite é: **999999**

Atualize seu cadastro e digite sua senha para confirmar as alterações

Nome	Sobrenome
Cauane	B Silva
Data de nascimento	
09/09/1989	
E-mail	Whatsapp
cauane.epi@gmail.com	(53) 98117-5711

Insira a sua senha e clique em 'Atualizar' para confirmar as modificações

Senha atual

Atualizar

Figura 6. Módulo de atualização cadastral.

Estudo piloto

Um estudo piloto foi conduzido para identificar possíveis problemas na plataforma *coortesnaweb*, bem como para estimar o tempo médio gasto para responder a cada questionário. O estudo piloto teve duração de 2 meses e 4 dias (início no dia 4 de Novembro de 2017 e fim no dia 8 de Janeiro de 2018). A amostra do estudo piloto foi escolhida por conveniência, e os participantes foram recrutados através de mensagens de Whatsapp, Facebook e e-mail. O estudo piloto contou com 40 participantes.

Os mesmos questionários que seriam usados na população alvo do estudo foram aplicados à amostra do estudo piloto. Também foi aplicado um questionário extra, que perguntava sobre possíveis problemas e sugestões para a plataforma. Este questionário serviu como base para melhorar o desempenho da *coortesnaweb* antes do seu lançamento oficial.

Recrutamento dos participantes

Previamente ao início do processo de recrutamento dos participantes, todos os membros da Coorte 1993 considerados elegíveis para serem convidados a se registrar na plataforma *coortesnaweb* ($n = 3.537$) foram adicionados à uma planilha do Google Drive. Além disso, o subgrupo dos membros elegíveis que tinham informações de contato de e-mail, Facebook e Whatsapp também foram identificados nessa mesma planilha ($n = 2.394$). A identificação desse subgrupo foi necessária para sinalizar quais membros seriam elegíveis para o ensaio randomizado que investigou a influência do método de recrutamento sobre a taxa de recrutamento.

O envio das mensagens de recrutamento iniciou em janeiro de 2018, e foi conduzido por uma equipe composta por dois integrantes (Cauane Blumenberg e Patrícia Simões). Cada indivíduo poderia receber até seis mensagens de recrutamento. Uma nova mensagem de recrutamento era enviada 10 dias após o envio da mensagem anterior. Caso o registro na plataforma *coortesnaweb* fosse realizado antes do recebimento da sexta mensagem, o envio de novas mensagens era interrompido. A mensagem era escrita em português e tinha o seguinte conteúdo:

Conheça a nova plataforma da Coorte 1993 na internet

Olá [nome do indivíduo], tudo bem?

*Gostaríamos de te convidar para participar da **coortesnaweb.com**, a nova plataforma de pesquisa da **Coorte 1993** na internet!*

Nessa plataforma você vai poder responder questionários, ganhar conquistas e acumular pontos! Com esses pontos você poderá desbloquear resultados sobre a sua saúde e comparar os seus resultados com os demais participantes da Coorte 1993!

*Faça o seu cadastro em **coortesnaweb.com** e comece agora mesmo!*

*Equipe **coortesnaweb***

Foram utilizados três tipos de método de recrutamento: e-mails (utilizando o recurso de mala direta), mensagens pessoais enviadas por Whatsapp (manualmente) e mensagens pessoais enviadas por Facebook (manualmente). A cada nova mensagem de recrutamento enviada, o método de recrutamento era alterado – respeitando a ordem de métodos descrita anteriormente. Assim, a primeira mensagem era enviada por e-mail, a segunda por Whatsapp e a terceira por Facebook, repetindo essa ordem para as mensagens de recrutamento subsequentes. Caso alguma informação de contato não estivesse disponível, o próximo método de recrutamento (respeitando a ordem original) era utilizado. No caso dos indivíduos elegíveis para o ensaio randomizado que investigou a influência dos métodos de recrutamento, as três primeiras mensagens de recrutamento foram enviadas utilizando o método para o qual o indivíduo havia sido alocado. A quarta mensagem de recrutamento (assim como as subsequentes) era enviada através de um novo método de recrutamento, e ao recebê-la o indivíduo tinha sua participação finalizada no ensaio randomizado e era considerado como “não registrado”.

Registro e coleta de dados

Para se registrar na plataforma coortesnaweb, os indivíduos devem acessar o site do estudo (<https://www.coortesnaweb.com>) e realizar um cadastro simples fornecendo e-mail, nome completo, data de nascimento, CPF e senha definida pelo próprio indivíduo. No momento do registro o indivíduo também pode fornecer um código de convite caso tenha sido convidado a se registrar por outro participante já cadastrado. O fornecimento de um código de convite garante

pontuação extra tanto para quem convidou, quanto para quem foi convidado. Somado a isso, no momento do cadastro é apresentado aos participantes o termo de consentimento livre e esclarecido, o qual deve ser lido e assinado pelos participantes clicando na opção “de acordo”.

Após a realização do cadastro, o perfil do indivíduo fica suspenso (Figura 7) para que os coordenadores do estudo possam verificar se o indivíduo cadastrado é um membro da Coorte 1993. A verificação é feita através da comparação do CPF fornecido no momento do registro com o CPF presente nos bancos de dados da Coorte 1993. Em alguns casos o número de CPF pode não estar presente nos bancos de dados da Coorte 1993, então as informações de nome completo e data de nascimento são utilizadas para verificar a elegibilidade do indivíduo.

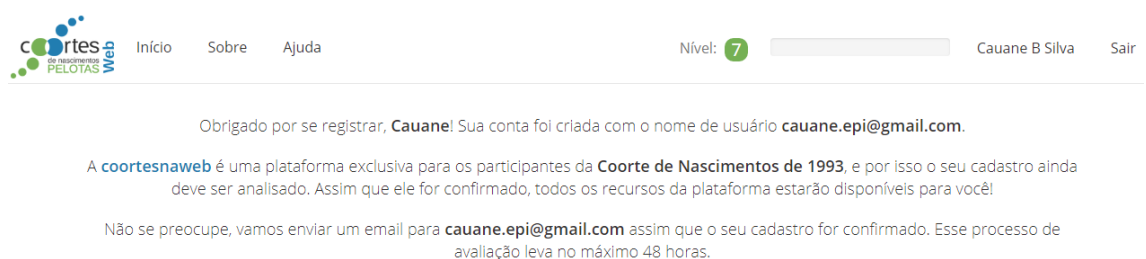


Figura 7. Mensagem de registro suspenso até que a elegibilidade do indivíduo seja verificada.

Uma vez verificada a elegibilidade, é enviado um e-mail comunicando o participante que seu perfil foi liberado na plataforma *coortesnaweb*. Além disso, o participante é aleatoriamente alocado para um dos seguintes braços do estudo: i) questionários longos e baixa frequência de lembretes; ii) questionários longos e alta frequência de lembretes; iii) questionários curtos e baixa frequência de lembretes; ou iv) questionários curtos e alta frequência de lembretes. O participante é cegado quanto sua alocação.

Um total de oito questionários foram publicados na plataforma, três no primeiro mês de estudo e um em cada mês subsequente. Cinco destes questionários foram desenvolvidos em duas extensões, uma versão longa (contendo entre 21 e 33 itens, e duração média de 14 minutos) e uma versão curta (contendo entre 11 e 17 itens, e duração média de quatro minutos). De acordo com o grupo para o qual o participante é alocado, a extensão do questionário a ser respondido é automaticamente determinada. Os questionários com apenas uma extensão são liberados para todos os participantes, independente do grupo de alocação. Uma listagem dos questionários aplicados está presente na Tabela 1, e seu conteúdo presente nos anexos.

Tabela 1. Lista de questionários aplicados através da plataforma *coortesnaweb*.

Questionário	Extensão	Número de itens	Pontuação
(Q1) Medidas corporais	Única	3	6
(Q2) Consumo de álcool	Curta	15	42
(Q2) Consumo de álcool	Longa	26	75
(Q3) Prática de atividade física	Curta	16	45
(Q3) Prática de atividade física	Longa	31	90
(Q4) Uso abusivo da internet	Curta	13	36
(Q4) Uso abusivo da internet	Longa	21	60
(Q5) Perpetração de crimes e agressão	Curta	16	45
(Q5) Perpetração de crimes e agressão	Longa	31	90
(Q6) Alimentos com glúten e lactose	Única	13	36
(Q7) Qualidade do sono	Única	20	57
(Q8) Hábito de fumar	Curta	15	42
(Q8) Hábito de fumar	Longa	32	93

Lembretes são enviados para os participantes que não respondem aos questionários, utilizando os mesmos métodos de contato do recrutamento: e-mail, Whatsapp e Facebook. São enviados até dois lembretes para cada questionário, sendo que participantes alocados para o grupo de alta frequência recebem lembretes a cada 15 dias, enquanto os alocados para o grupo de baixa frequência recebem lembretes a cada 30 dias. Caso o questionário seja respondido pelo participante, o envio de lembretes referente àquele questionário é interrompido. A mensagem de lembrete enviada é escrita em português e tem o seguinte conteúdo:

Você tem questionários para responder!

Olá, [nome do indivíduo]! Tudo bem?

*Estamos enviando esta mensagem para lembrar que existem novos questionários disponíveis para você responder na plataforma **coortesnaweb**. Para responder é bem fácil, basta entrar no site **www.coortesnaweb.com** e entrar com o seu e-mail e senha.*

Acesse agora mesmo e continue acumulando pontos!

*Equipe **coortesnaweb***

Aspectos éticos

O projeto coortesnaweb foi submetido e aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Medicina da Universidade Federal de Pelotas (CAAE: 79124917.5.0000.5317 / parecer número: 2.382.790). Uma cópia do parecer está anexada a este relatório de trabalho de campo, bem como uma cópia do termo de consentimento livre e esclarecido que deve ser lido e assinado pelos participantes do estudo.

Os ensaios randomizados para testar a eficiência de diferentes métodos sobre a taxa de recrutamento, e a influência do da extensão de questionários e frequência de lembretes sobre a taxa de resposta, foram registrados no Registro Brasileiro de Ensaios Clínicos (Número do registro: RBR-3dv7gc; página do registro: [http://www.ensaiosclinicos.gov.br/rg/RBR-3dv7gc/](http://www ensaiosclinicos.gov.br/rg/RBR-3dv7gc/)).

Referências

1. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*. abril de 2009;42(2):377–81.
2. Blumenberg C, Barros A. Electronic data collection in epidemiological research: The use of REDCap in the Pelotas birth cohorts. *Applied Clinical Informatics*. julho de 2016;07(03):672–81.
3. Blumenberg C, Zugna D, Popovic M, Pizzi C, Barros AJ, Richiardi L. Questionnaire breakoff and item nonresponse in web-based questionnaires: a multilevel analysis of individual and item-related factors in a birth cohort (Preprint). *Journal of Medical Internet Research* [Internet]. 14 de maio de 2018 [citado 7 de setembro de 2018]; Disponível em: <http://preprints.jmir.org/preprint/11046/accepted>

4 – Artigo 1

Publicado no International Journal of Public Health em 24 de Abril de 2018



REVIEW

Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature

Cauane Blumenberg¹  · Aluísio J. D. Barros¹Received: 3 November 2017 / Revised: 10 April 2018 / Accepted: 19 April 2018 / Published online: 24 April 2018
© Swiss School of Public Health (SSPH+) 2018

Abstract

Objectives To systematically review the literature and compare response rates (RRs) of web surveys to alternative data collection methods in the context of epidemiologic and public health studies.

Methods We reviewed the literature using PubMed, LILACS, SciELO, WebSM, and Google Scholar databases. We selected epidemiologic and public health studies that considered the general population and used two parallel data collection methods, being one web-based. RR differences were analyzed using two-sample test of proportions, and pooled using random effects. We investigated agreement using Bland-and-Altman, and correlation using Pearson's coefficient.

Results We selected 19 studies (nine randomized trials). The RR of the web-based data collection was 12.9 percentage points (p.p.) lower (95% CI = - 19.0, - 6.8) than the alternative methods, and 15.7 p.p. lower (95% CI = - 24.2, - 7.3) considering only randomized trials. Monetary incentives did not reduce the RR differences. A strong positive correlation ($r = 0.83$) between the RRs was observed.

Conclusions Web-based data collection present lower RRs compared to alternative methods. However, it is not recommended to interpret this as a meta-analytical evidence due to the high heterogeneity of the studies.

Keywords Response rate · Web surveys · Public health · E-epidemiology · Epidemiology · Survey methodology

Introduction

Much of the epidemiologic and public health findings are dependent on interviewing individuals. The main objective of the interviews is to produce a combined estimate that can be generalized to a broader population. To achieve this goal, it is very important to reduce the total survey error (TSE) (Biemer 2010). One of the main components of the TSE is the nonresponse error, which is the failure to obtain data from all the members of a sample (Groves 2004). If the non-respondents are different from the respondents on characteristics that are important to the study objectives,

this could skew the results and limit the generalization of the findings (Beebe et al. 2012; Keiding and Louis 2016). The nonresponse error is related to the survey response rate (RR), as low RRs increase the probability of occurring this error (Groves 2004). Recently, epidemiologic and public health studies are facing a constant reduction on their RRs, which lies between one (Morton et al. 2006) and two percentage points (p.p.) per year (Wallander et al. 2015). Several aspects can influence the RR, including the data collection method (Galea and Tracy 2007).

Face-to-face interviews are commonly used to collect data; however, the complicated logistics and the elevated costs are turning this methodology prohibitive (Dillman et al. 2014). Questionnaires sent via mail are also widely used among high-income populations, where the mailing system works and the costs related to printing and sending questionnaires via traditional mail are not a problem (Hardigan et al. 2012). Another data collection method is the telephone, which involves simple logistics but is compromised by the massive reduction of landline phones (Blumberg et al. 2013). The Internet is a natural candidate

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00038-018-1108-4>) contains supplementary material, which is available to authorized users.

✉ Cauane Blumenberg
cauane.epi@gmail.com

¹ Postgraduate Programme in Epidemiology, Federal University of Pelotas, Pelotas, Brazil

to overcome the disadvantages of the alternative data collection methods, as its coverage is in constant growth worldwide (The World Bank 2018). Web surveys are defined as instruments located in a server in the World Wide Web (WWW), where the questionnaire lies in a web page and is accessed using a web browser (Bethlehem and Biffignandi 2012). This method can facilitate the logistics (Sax et al. 2003) and reduce costs, since there is only the survey setup cost, which does not increase with the number of persons interviewed (Hardigan et al. 2012).

The number of epidemiologic and public health surveys using the web to collect data is growing. From 2007 to 2014, there was a 400% increase on the number of randomized trials using the web (Brice et al. 2015). However, the web still is a novel approach to collect epidemiologic data; thus, it is important to compare the quality of web surveys to alternative data collection methods. There are several survey quality factors that constitute the TSE, but in this study, we will focus on the nonresponse component and use the RR as its indicator.

The goal of this study was to systematically review the epidemiologic and public health literature to compare RRs of web surveys and alternative data collection methods applied to the general population, and check whether there is enough evidence about how the different approaches differ in relation to this indicator.

Methods

The literature review was based on PubMed, LILACS, SciELO, WebSM, and Google Scholar databases. The search ended in March-2017 and was limited to human populations. No date limit or language restriction was used. The list of references of all selected articles was checked to find potential studies of interest not yet identified through our search.

The expression used to search the databases was constructed as follows: [(Internet OR web OR online OR www) AND (response OR return OR completion) AND (rate OR proportion)]. The search for the terms was limited to the title and abstract of the studies. To search the WebSM database, we used the following filters: "Mode comparison", "Web Surveys", "Mixed-Mode Surveys", "Journal article", "Conference proceedings", "Book", "Edited book", "Book section", and "Thesis/diplomas".

To be included, a study had to report the RRs of at least two data collection methods, being one of them a self-administered web survey. To be considered a web survey, it should be a questionnaire located in a web page accessible using a web-browser (Bethlehem and Biffignandi 2012). The web-based and the alternative data collection methods had to be concurrently administered. The studies

had to be about health-related topics, based on representative samples, not including, for example, case reports or studies based on convenience samples. Studies that targeted specific populations such as professional groups, students, and persons with a given disease or any other specific group were excluded. These specific groups were not included in the review, since they are more likely to respond compared to the general population (Lozar Manfreda et al. 2008).

Every search result was added to a Zotero 4.0 (Zotero, Virginia, USA) database, where duplicates were removed. If different studies were based on the same survey or inquiry, only one study was selected. This exclusion occurred to avoid duplicates in terms of RRs, since different studies that used the same survey or inquiry have the same estimates. In this case, the earliest study, according to the publication year, was selected. Following the exclusion of duplicates, the titles and abstracts of the studies were checked according to the selection criteria. Finally, the full text of the articles was read to confirm their eligibility, to check the reference list, and to extract the results of the studies. If any relevant detail was not described in the full text, an email was sent to the first author of the study to achieve such information.

Descriptive variables

Among the 19 selected studies, two types of study designs were identified: cross-sectional and randomized trials. To be considered a randomized trial, the participants had to be randomly assigned to one of the different data collection methods compared by the study. The web recruitment method describes the mean through which the participants were invited to respond the web survey. E-mails had to contain a link pointing to the web survey, letters had to be sent via traditional mail, and the other category included virtual banners and phone calls. A study was considered to provide monetary incentives if any money amount or monetary gift cards were given. Lottery entrance, donations, or points were considered non-monetary. The use of reminders consisted on any attempt to contact the non-respondents after the initial recruitment. Finally, the type of RR calculation was analyzed according to the Standard Definitions of the American Association for Public Opinion Research (AAPOR) (The American Association for Public Opinion Research 2016). If the study did not provide the RR considered, the type of calculation was assessed using the formulas of the AAPOR publication.

To account for the differences in Internet coverage and use that might occur across the years, the mean RRs considering two different publication periods were analyzed. The first period was from 2002 to 2009, while the second was from 2010 to 2015. These periods were chosen,

because they are from two different decades, and because the Internet coverage suffered important changes. In 2002, only 10.6% of the world population had Internet access, reaching 25.5% in 2009. While in 2010, 28.9% of the population had Internet access, reaching 43.2% of the population in 2015 (The World Bank 2018). The whole decade period (e.g., 2000–2009 and 2010–2020) was not considered, since no studies were published before 2002, neither after 2015.

Statistical analyses

The selected studies were organized in an Excel 2010 (Microsoft Corp., Washington, USA) spreadsheet, where the relevant methodological aspects of the studies were extracted. Statistical analyses were performed using Stata 13.1 (Stata Corp., Texas, USA). The difference between the RR estimates of the web-based and the alternative data collection methods was assessed via two-sample test of proportions. The pooled RR differences were estimated using random effects, and the percentage of variation among the studies due to heterogeneity was analyzed using the i^2 statistic (Higgins and Thompson 2002). The agreement between RR estimates was analyzed using the Bland and Altman (1999) method, and the correlation was analyzed using Pearson's correlation coefficient. It was considered a significance level equal to 5%.

Results

Selection process

The search among the five databases returned 4707 studies. All titles and abstracts were checked, and 4689 studies were excluded: 46.2% did not target the general population, 23.7% were not health-related, 21.4% did not report RRs of at least two data collection methods, 4.2% were based on the same survey with identical RRs, 3.6% did not use web surveys, and 0.9% did not use concurrent data collection methods. A total of 18 eligible studies were selected according to the selection criteria. After reading their full text, one additional study was selected via the reference list, totalizing 19 eligible studies, as shown in Fig. 1.

Description of the selected studies

A complete list of the selected studies is provided in Online Resource 1. Some of the studies used more than one alternative method to the web-based data collection method, and also more than one web-based approach compared to the alternative methods. This is the reason

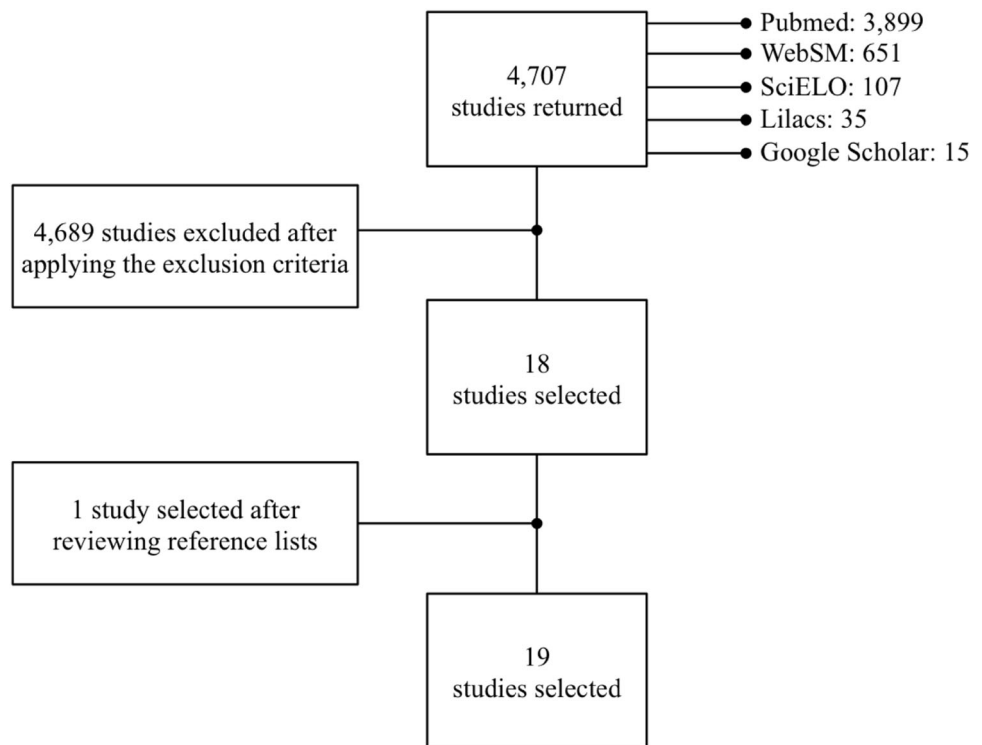
why we describe the studies more than once, adding up to 28 groups assessed. All the selected studies were conducted in high-income countries, almost half of them (nine studies) being based in the United States, one conducted in South Korea, and the remaining nine studies were performed in four European countries (one in Denmark, one in Finland, two in The Netherlands, and five in Sweden). Mail questionnaires were the most used alternative to the web-based method, being employed 11 times.

In Table 1, we describe the methodological characteristics of the selected studies. Only nine studies formally compared the RRs by randomizing the participants between the different data collection methods. Only six studies explicitly provided the type of RR calculation according to the AAPOR Standard Definitions (The American Association for Public Opinion Research 2016). Almost half of the studies calculated the RRs assuming that there were no eligible cases among the cases of unknown eligibility (RR5), and two other studies adopted the same methodology but also considered partial respondents as complete respondents (RR6). Two studies also considered the partial respondents as complete ones, but one included the real number of cases of unknown eligibility in the calculation (RR2), and the other performed an estimate of this number (RR4). Six studies calculated the minimum RR (RR1). Although there were differences between the RR calculations used by the studies, we directly compared the estimates, since there was not enough information in the manuscript, or provided by the authors, to standardize the types of RR for all the studies. In Table 2, we present the mean and median sample sizes and RRs for the web-based and the alternative data collection methods. The different number of groups assessed is to avoid repeated RR and sample size estimates for studies that compared more than two alternative or web-based approaches. The studies had a wide range of sample sizes, regardless of the data collection method used. The difference between the mean and median indicate the asymmetry of the distribution of sample sizes for the web-based and the alternative data collection methods. This difference was not as large for the RR values, but still assumed an asymmetric distribution. However, to simplify the interpretation and to enable comparisons with other studies, we adopt the mean to perform the RR analyses.

Response rate analyses

The difference between the RRs of the web-based and the alternative data collection methods is presented in Fig. 2. The RR was higher for the web-based data collection method only in four comparisons, two times compared to telephone interviews, one time compared to cell phone interviews, and one time compared to IVR interviews. The

Fig. 1 Flow chart of the study selection process (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)



pooled mean RR difference considering all 28 groups was 12.9 p.p. (95% CI = - 19.0, - 6.8; $i^2 = 99.6\%$) lower for the web-based method compared to the alternative data collection methods. A similar result was found when pooling only the estimates of the randomized trials, 15.7 p.p. (95% CI = - 24.2, - 7.3; $i^2 = 98.5\%$) lower compared to the alternative methods. We also analyzed the pooled RR differences comparing web-based and the alternative methods according to the reward of monetary incentives and reminders (calculated as: web RR - alternative methods RR). Providing monetary incentives did not influence the RR differences, as they were - 11.1 p.p. (95% CI = - 26.8, 4.5) for studies that provided incentives, and - 13.2 p.p. (95% CI = - 19.6, - 6.8) for those not providing. However, the pooled RR difference was lower among studies that did not send reminders (RRdiff. = - 2.0; 95% CI = - 19.0, 9.5) compared to the ones that sent reminders for the non-respondents (RRdiff. = - 14.2; 95% CI = - 20.7, - 7.7). As shown in Table 3, the mean RRs increased from the 2002–2009 period to the 2010–2015 period. The absolute RR difference comparing web-based to alternative data collection methods remained constant. The agreement analysis is presented in Fig. 3. The majority of the points lie below the line that indicates RR equality ($y = 0$), demonstrating the superiority of the RRs of the alternative methods compared to the web-based data collection. The points are well distributed along the x -axis, illustrating that the RR differences are not dependent on the study mean RR. Thus,

independently if the study achieves a high or a low RR, the difference between the web-based and the alternative data collection methods (regardless of the data collection method used) is, on average, similar. The correlation between the web-based and the alternative methods RR was $r = 0.83$ (95% CI = 0.66, 0.92).

Discussion

Our systematic literature review found that RRs of web surveys are systematically lower compared to RRs of alternative data collection methods. A similar difference was found when we considered only randomized trials. However, it is not possible to consider these results meta-analytical evidences due to the high degree of heterogeneity of the studies: $i^2 = 99.6\%$ considering all the studies, and $i^2 = 98.5\%$ for randomized trial only. Our study limited the analysis of survey quality to a single component of the TSE (nonresponse error), and used the RR as its indicator. This was a methodological choice to achieve the study goal and also to simplify the analyses, since the RR is usually available on published articles. However, we identified five different RR formulas among the selected studies, and few that explicitly described the type of RR calculation used. To standardize the estimates and enable better comparisons, the use of the AAPOR Standard Definitions (The American Association for Public Opinion Research 2016) is strongly recommended.

Table 1 Description of the methodological characteristics of the studies (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)

Characteristic	Number of studies (%)
Study design	
Cross-sectional	10 (52.6)
Randomized trial	9 (47.4)
Web recruitment method	
Email	10 (52.7)
Letter	7 (36.8)
Other	2 (10.5)
Monetary incentive	
Yes	3 (15.8)
No	16 (84.2)
Reminder	
Yes	16 (84.2)
No	3 (15.8)
Response rate type ^a	
RR1	6 (31.6)
RR2	1 (5.3)
RR4	1 (5.3)
RR5	9 (47.3)
RR6	2 (10.5)
Total	19 (100.0)

RR response rate

^aResponse rate codes according to the American Association for Public Opinion Research Standard Definition report (The American Association for Public Opinion Research 2016)

Other literature reviews and meta-analyses comparing RRs estimates are already published in the literature, but our study differs from those, since it focused exclusively on epidemiologic and public health studies that had the general population as their target. Considering the general population is important, since the variability of the Internet access and coverage in this group is higher compared to other specific groups (Pew Research Center 2016). For this

reason, it is important to understand if web surveys are feasible to be used as the main data collection method in the general population. Focusing on public health and epidemiologic studies is important, since the use of the web as mean of data collection is fast increasing in these areas (Brice et al. 2015). Furthermore, health surveys usually require more sensitive and personal information than other survey subjects (i.e., market research and opinion polls); hence, the cooperation and RRs should be studied. A meta-analysis published by Shih and Fan (2008) found a mean RR 11.0 p.p. lower for the web-based data collection compared to mail surveys and a similar heterogeneity $i^2 = 98.9\%$. However, besides comparing the web only to mail data collection, the authors included specific groups such as students and professional parties, which are known to have higher RRs than the general population (Lozar Manfreda et al. 2008; Shih and Fan 2008). Another meta-analysis conducted by Lozar Manfreda et al. (2008) investigated the RR differences between concurrently administered web surveys and alternative data collection methods. Among the general population, the RR difference was 19.0 p.p. lower for the web surveys. Different from our study, Lozar Manfreda and colleagues did not limit their review to studies that investigated health-related subjects, what might explain the differences between the pooled estimations.

Although there were methodological variations between ours and the other reviews, they were consistent in concluding that the RR of the web-based data collection method is lower than the alternative methods. The lower RRs of web-based methods held from 2002 to 2015, even when we considered the publication year as a proxy to account for the differences in Internet coverage and use across the years. We also found that the mean difference between the web-based and the alternative data collection methods does not vary according to the RR level (high or low). For this reason, it is clear that other methodological aspects, such as the type of recruitment method (Brown et al. 2015; Akmatov et al. 2015), giving incentives (Gajic et al. 2012; Holland et al. 2015), and sending reminders for

Table 2 Mean and median sample size and response rate for web-based and alternative data collection methods (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)

Data collection	Mean	Median	IQR	Min	Max	Groups assessed
Web-based						22
Sample size	9861.4	1659.5	293.0–7423.0	37	70,932	
Response rate (%)	40.5	38.3	20.2–53.4	1.7	93.4	
Alternative						25
Sample size	4043.0	1918.0	735.0–3009.0	34	28,407	
Response rate (%)	56.3	59.7	43.0–70.7	8.0	97.1	

IQR interquartile range, Max maximum, Min minimum

Fig. 2 Response rate difference in percentage points between web-based and alternative data collection methods. Randomized trials are marked with an asterisk (*) (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)

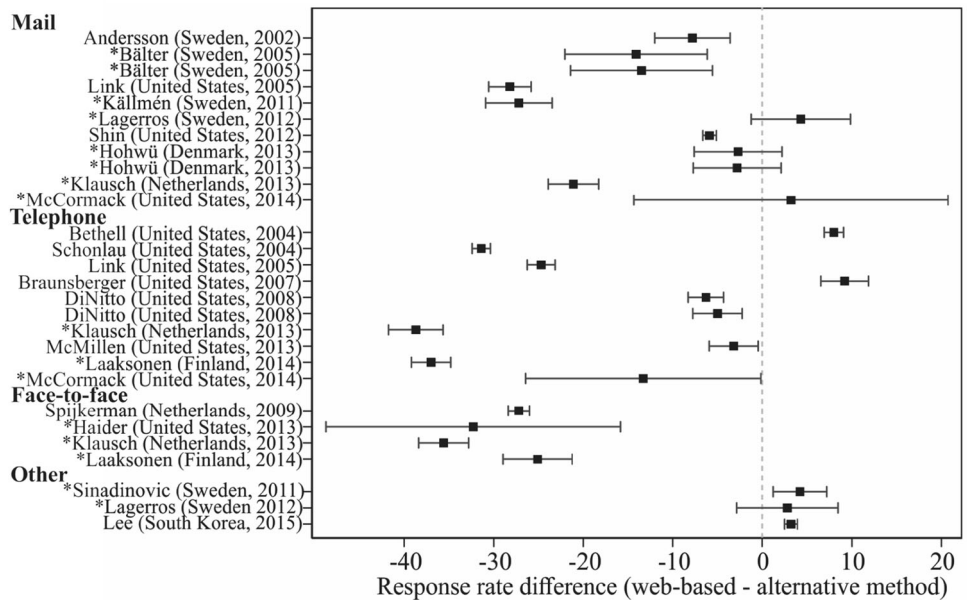
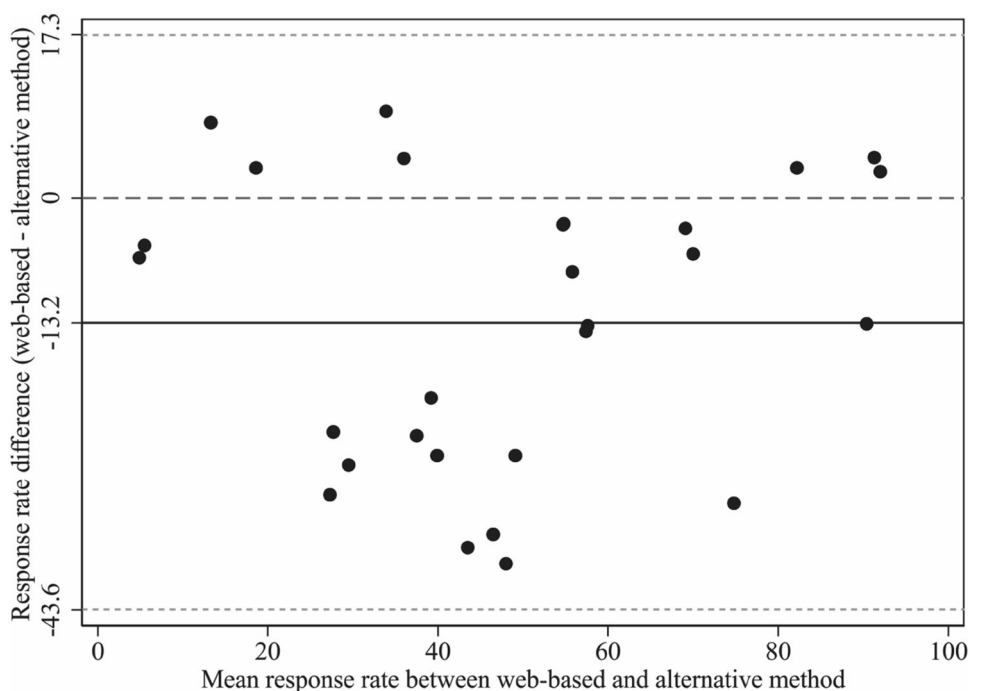


Table 3 Mean response rate for web-based and alternative data collection methods considering two time periods (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)

Data collection	2002–2009		2010–2015		Groups assessed
	Mean RR (%)	95% CI	Mean RR (%)	95% CI	
Web-based	27.6	13.3–42.0	51.3	36.0–66.6	22
Alternative	40.0	23.7–56.3	65.4	53.7–77.1	25
Difference	– 12.4 p.p.	– 54.7–29.9	– 14.1 p.p.	– 50.7–22.6	

CI confidence interval, p.p. percentage points, RR response rate

Fig. 3 Bland and Altman plot between response rates estimates of the web-based and the alternative data collection methods (Denmark, 2013; Finland, 2014; The Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014)



the non-respondents (Van Mol 2016; Sebo et al. 2017), play an important role to define the RR of both web-based and alternative data collection methods.

Edwards et al. (2009) described the failure to increase web survey RRs by giving monetary incentives, but identified an increase on the traditional mail RR estimates. This is in line to our findings, showing that giving monetary incentives fail to reduce the RR gap between the data collection methods, and hence, other types of incentives and rewards should be explored (Edwards et al. 2009). Our results also showed that the use of reminders may be less effective for web-based data collection methods than for alternative methods, since the gap between RRs is higher when a reminder is sent. We did not analyze the recruitment methods, since they were very different between the web-based and the alternative data collection methods.

Although achieving high RRs reduces the probability of nonresponse error, it is important to note that lower RRs do not invalidate survey results (Groves 2004). If non-respondents are not different from the respondents regarding to aspects that are relevant to the study objectives, this will not bias the results, but could only cause loss of precision (Dillman et al. 2014). In addition, Groves and Peytcheva (2008) describe that nonresponse bias can affect studies with high or low RRs. However, the gap between RRs of web-based and alternative methods tends to be reduced. With more users looking for medical information through web (Tonsaker et al. 2014), the increase of Internet banking use (Yuen et al. 2010), and more sales being performed online (United Problem Solvers (UPS) 2016), it is expected a more confident relation between respondents and web surveys. To further increase the confidence of participants on web surveys, it is encouraged to adopt the web as main, or part of a mixed-mode survey, to disseminate its use as a data collection method. In addition, web surveys bring further advantages like cheaper data collection (Greenlaw and Brown-Welty 2009; Huybrechts et al. 2010), results of better quality (Barrios et al. 2011; Shin et al. 2012), the possibility to use audio and visual resources, asynchronous participation, and the possibility to provide instant feedback to the participant. In addition, with the recent increase of Internet coverage around the globe (The World Bank 2018), this is a good opportunity for low- and middle-income countries to collect health-related data using web-based surveys.

Our study presents some limitations. Using binary variables to indicate the use of reminders and monetary incentives made impossible for us to analyze the differences between the number of reminders sent and the amount of the incentive. This could not be done, since this information was not available for some studies. Finally, our study compared different types of RRs without standardization. We could not standardize the estimates due to the

lack of information provided in the manuscript, or by authors, of the selected studies. However, our study also presents some strengths. We adopted a comprehensive approach by including both observational and experimental studies. In addition, this is one of the first studies to analyze web survey RRs in the context of health-related studies that focused on the general population.

Conclusions

This review fills a gap in the epidemiologic and public health literature in the sense that it compared the RRs of web surveys to alternative data collection methods. We show that web RRs are systematically lower than the ones achieved using traditional data collection methods and that there is a high variability of the RR estimates between the studies reviewed. Although the web fails to achieve a comparable RR to traditional methods, it involves lower costs and other advantages like validation during data collection, structured data collection, and more efficient data consistency leading to better quality results. Considering these advantages and also the constant growth of the Internet use worldwide, it might be beneficial to use the web other than alternative data collection methods for epidemiologic research.

Funding No funding was received to execute this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Akmatov MK, Rübsamen N, Schultze A et al (2015) Diverse recruitment strategies result in different participation percentages in a web-based study, but in similar compliance. *Int J Public Health* 60:937–943. <https://doi.org/10.1007/s00038-015-0737-0>
- Andersson G, Lindvall N, Hursti T, Carlbring P (2002) Hypersensitivity to sound (hyperacusis): a prevalence study conducted via the Internet and post. *Int J Audiol* 41:545–554
- Bälter KA, Bälter O, Fondell E, Lagerros YT (2005) Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiol Camb Mass* 16:577–579
- Barrios M, Villarroja A, Borrego A, Olle C (2011) Response rates and data quality in web and mail surveys administered to PhD holders. *Soc Sci Comput Rev* 29:208–220. <https://doi.org/10.1177/0894439310368031>
- Beebe TJ, McAlpine DD, Ziegenfuss JY et al (2012) Deployment of a mixed-mode data collection strategy does not reduce

- nonresponse bias in a general population health survey. *Health Serv Res* 47:1739–1754. <https://doi.org/10.1111/j.1475-6773.2011.01369.x>
- Bethell C, Fiorillo J, Lansky D et al (2004) Online consumer surveys as a methodology for assessing the quality of the united states health care system. *J Med Internet Res* 6:e2. <https://doi.org/10.2196/jmir.6.1.e2>
- Bethlehem JG, Biffignandi S (2012) *Handbook of web surveys*. Wiley, Hoboken
- Biemer PP (2010) Total survey error: design, implementation, and evaluation. *Public Opin Q* 74:817–848. <https://doi.org/10.1093/poq/nfq058>
- Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* 8:135–160
- Blumberg SJ, Ganesh N, Luke JV, Gonzales G (2013) Wireless substitution: state-level estimates from the National Health Interview Survey, 2012. *Natl Health Stat Rep* 70:1–16
- Braunberger K, Wybenga H, Gates R (2007) A comparison of reliability between telephone and web-based surveys. *J Bus Res* 60:758–764. <https://doi.org/10.1016/j.jbusres.2007.02.015>
- Brice A, Price A, Burls A (2015) Creating a database of Internet-based clinical trials to support a public-led research programme: a descriptive analysis. *Digit Health*. <https://doi.org/10.1177/2055207615617854>
- Brown O, Quick V, Colby S et al (2015) Recruitment lessons learned from a tailored web-based health intervention Project Y.E.A.H. (Young Adults Eating and Active for Health). *Health Educ* 115:470–479. <https://doi.org/10.1108/he-06-2014-0075>
- Dillman DA, Smyth JD, Christian LM (2014) *Internet, phone, mail, and mixed-mode surveys: the tailored design method*, 4th edn. Wiley, Hoboken
- Dinitto DM, Bridget Busch-Armendariz N, Bender K et al (2008) Testing telephone and web surveys for studying men's sexual assault perpetration behaviors. *J Interpers Violence* 23:1483–1493. <https://doi.org/10.1177/0886260508314341>
- Edwards PJ, Roberts I, Clarke MJ et al (2009) Methods to increase response to postal and electronic questionnaires. In: *The Cochrane Collaboration (ed) Cochrane database of systematic reviews*. Wiley, Chichester
- Gajic A, Cameron D, Hurley J (2012) The cost-effectiveness of cash versus lottery incentives for a web-based, stated-preference community survey. *Eur J Health Econ HEPAC Health Econ Prev Care* 13:789–799. <https://doi.org/10.1007/s10198-011-0332-0>
- Galea S, Tracy M (2007) Participation rates in epidemiologic studies. *Ann Epidemiol* 17:643–653. <https://doi.org/10.1016/j.annepidem.2007.03.013>
- Greenlaw C, Brown-Welty S (2009) A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. *Eval Rev* 33:464–480. <https://doi.org/10.1177/0193841x09340214>
- Groves RM (2004) *Survey errors and survey costs*. Wiley, Hoboken
- Groves RM, Peytcheva E (2008) The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opin Q* 72:167–189. <https://doi.org/10.1093/poq/nfn011>
- Haider S, Dodge LE, Brown BA et al (2013) Evaluation of e-mail contact to conduct follow-up among adolescent women participating in a longitudinal cohort study of contraceptive use. *Contraception* 88:18–23. <https://doi.org/10.1016/j.contraception.2012.11.016>
- Hardigan PC, Succar CT, Fleisher JM (2012) An analysis of response rate and economic costs between mail and web-based surveys among practicing dentists: a randomized trial. *J Community Health* 37:383–394. <https://doi.org/10.1007/s10900-011-9455-6>
- Higgins JPT, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–1558. <https://doi.org/10.1002/sim.1186>
- Hohwü L, Lyshol H, Gissler M et al (2013) Web-based versus traditional paper questionnaires: a mixed-mode survey with a nordic perspective. *J Med Internet Res* 15:e173. <https://doi.org/10.2196/jmir.2595>
- Holland CM, Ritchie ND, Du Bois SN (2015) iTunes song-gifting is a low-cost, efficient recruitment tool to engage high-risk MSM in Internet research. *AIDS Behav* 19:1914–1918. <https://doi.org/10.1007/s10461-015-1130-y>
- Huybrechts KF, Mikkelsen EM, Christensen T et al (2010) A successful implementation of e-epidemiology: the Danish pregnancy planning study “Snart-Gravid”. *Eur J Epidemiol* 25:297–304. <https://doi.org/10.1007/s10654-010-9431-y>
- Källmén H, Sinadinovic K, Berman A, Wennberg P (2011) Risky drinking of alcohol in Sweden: a randomized population survey comparing web- and paper-based self-reports. *Nord Stud Alcohol Drugs*. <https://doi.org/10.2478/v10199-011-0013-4>
- Keiding N, Louis TA (2016) Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A (Statistics in Society)* 179:319–376. <https://doi.org/10.1111/rssa.12136>
- Klausch T, Hox JJ, Schouten B (2013) Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociol Methods Res* 42:227–263. <https://doi.org/10.1177/0049124113500480>
- Laaksonen S, Heiskanen M (2014) Comparison of three modes for a crime victimization survey. *J Surv Stat Methodol* 2:459–483. <https://doi.org/10.1093/jssam/smu018>
- Lagerros YT, Sandin S, Bexelius C et al (2012) Estimating physical activity using a cell phone questionnaire sent by means of short message service (SMS): a randomized population-based study. *Eur J Epidemiol* 27:561–566. <https://doi.org/10.1007/s10654-012-9708-4>
- Lee C-K, Back K-J, Williams RJ, Ahn S-S (2015) Comparison of telephone RDD and online panel survey modes on CPGI scores and co-morbidities. *Int Gambl Stud* 15:435–449. <https://doi.org/10.1080/14459795.2015.1068353>
- Link MW, Mokdad AH (2005) Alternative modes for health surveillance surveys: an experiment with web, mail, and telephone. *Epidemiol Camb Mass* 16:701–704
- Lozar Manfreda K, Bosnjak M, Berzelak J et al (2008) Web surveys versus other survey modes: a meta-analysis comparing response rates. *Int J Mark Res* 50:79–104
- McCormack LA, Friedrich C, Fahrenwald N, Specker B (2014) Feasibility and acceptability of alternate methods of postnatal data collection. *Matern Child Health J* 18:852–857. <https://doi.org/10.1007/s10995-013-1310-1>
- McMillen RC, Winickoff JP, Wilson K et al (2013) A dual-frame sampling methodology to address landline replacement in tobacco control research. *Tob Control* 24:7–10. <https://doi.org/10.1136/tobaccocontrol-2012-050727>
- Morton LM, Cahill J, Hartge P (2006) Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol* 163:197–203. <https://doi.org/10.1093/aje/kwj036>
- Pew Research Center (2016) *Smartphone ownership and internet usage continues to climb in emerging economies*. Pew Research Center
- Sax LJ, Gilmartin SK, Bryant AN (2003) Assessing response rates and nonresponse bias in web and paper surveys. *Res High Educ* 44:409–432. <https://doi.org/10.1023/a:1024232915870>
- Schonlau M, Zapert K, Simon LP et al (2004) A comparison between responses from a propensity-weighted web survey and an identical RDD survey. *Soc Sci Comput Rev* 22:128–138. <https://doi.org/10.1177/0894439303256551>
- Sebo P, Maisonneuve H, Cerutti B et al (2017) Rates, delays, and completeness of general practitioners' responses to a postal

- versus web-based survey: a randomized trial. *J Med Internet Res* 19:e83. <https://doi.org/10.2196/jmir.6308>
- Shih T-H, Fan Xitao (2008) Comparing response rates from web and mail surveys: a meta-analysis. *Field Methods* 20:249–271. <https://doi.org/10.1177/1525822x08317085>
- Shin E, Johnson TP, Rao K (2012) Survey mode effects on data quality: comparison of web and mail modes in a U.S. National Panel Survey. *Soc Sci Comput Rev* 30:212–228. <https://doi.org/10.1177/0894439311404508>
- Sinadinovic K, Wennberg P, Berman AH (2011) Population screening of risky alcohol and drug use via Internet and interactive voice response (IVR): a feasibility and psychometric study in a random sample. *Drug Alcohol Depend* 114:55–60. <https://doi.org/10.1016/j.drugalcdep.2010.09.004>
- Spijkerman R, Knibbe R, Knoop K et al (2009) The utility of online panel surveys versus computer-assisted interviews in obtaining substance-use prevalence estimates in the Netherlands. *Addiction* 104:1641–1645. <https://doi.org/10.1111/j.1360-0443.2009.02642.x>
- The American Association for Public Opinion Research (2016) Standard definitions: final dispositions of case codes and outcome rates for surveys, 9th edn. AAPOR
- The World Bank (2018) Internet users (per 100 people). World Bank. <http://data.worldbank.org/indicator/IT.NET.USER.P2?end=2014&start=2014&view=bar>. Accessed 21 Feb 2018
- Tonsaker T, Bartlett G, Trpkov C (2014) Health information on the internet: gold mine or minefield? *Can Fam Physician Med Fam Can* 60:407–408
- United Problem Solvers (UPS) (2016) UPS pulse of the online shopper: tech-savvy shoppers transforming retail—a UPS white paper. United Problem Solvers (UPS)
- Van Mol C (2016) Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response. *Int J Soc Res Methodol*. <https://doi.org/10.1080/13645579.2016.1185255>
- Wallander L, Tikkanen RH, Mannheimer LN et al (2015) The problem of non-response in population surveys on the topic of HIV and sexuality: a comparative study. *Eur J Public Health* 25:172–177. <https://doi.org/10.1093/eurpub/cku154>
- Yuen YY, Yeow PHP, Lim N, Saylani N (2010) Internet banking adoption: comparing developed and developing countries. *J Comput Inf Syst* 51:52–61. <https://doi.org/10.1080/08874417.2010.11645449>

International Journal of Public Health

Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature

Cauane Blumenberg¹, Aluísio J D Barros¹

¹Postgraduate Programme in Epidemiology, Federal University of Pelotas,
Pelotas, Brazil

Cauane Blumenberg: cauane.epi@gmail.com (corresponding author)

Online Resource 1 Description of the selected studies organized by alternative methods to the web-based data collection method. (Denmark, 2013; Finland, 2014; Netherlands, 2009, 2013; South Korea, 2015; Sweden, 2002, 2005, 2011, 2012; United States, 2004, 2005, 2007, 2008, 2012, 2013, 2014).

Alternative method	Author (year)	Location	Web-based method		Alternative method	
			Sample size	RR (%)	Sample size	RR (%)
Mail						
	Andersson et al (2002) (Andersson et al. 2002)	Sweden	1,147	51.9	987	59.7
	Bälter et al (2005) (Bälter et al. 2005)	Sweden	290	50.3	292	64.4
	Bälter et al (2005) (Bälter et al. 2005)	Sweden	293	50.9	292	64.4

Alternative method	Author (year)	Location	Web-based method		Alternative method	
			Sample size	RR (%)	Sample size	RR (%)
			Link and Mokdad (2005) (Link and Mokdad 2005)	United States	7,423	15.4
Källmén et al (2011) (Källmén et al. 2011)	Sweden	1,232	26.3	1,240	53.5	
Lagerros et al (2012) (Lagerros et al. 2012)	Sweden	182	93.4	211	89.1	
Shin et al (2012) (Shin et al. 2012)	United States	27,514	67.0	28,407	72.9	
Hohwü et al (2013) (Hohwü et al. 2013)	Denmark	787	53.4	789	56.1	
Hohwü et al (2013) (Hohwü et al. 2013)	Denmark	786	53.3	789	56.1	
Klausch et al (2013) (Klausch et al. 2013)	Netherlands	2,199	28.7	2,182	49.8	
McCormack et al (2014) (McCormack et al. 2014)	United States	37	83.8	36	80.6	

Telephone

Alternative method	Author (year)	Location	Web-based method		Alternative method	
			Sample size	RR (%)	Sample size	RR (%)
	Bethell et al (2004)	United States	13,434	17.3	4,302	9.3
	Schonlau et al (2004)	United States	70,932	11.6	9,510	43.0
	Link and Mokdad (2005)	United States	7,423	15.4	5,168	40.1
	Braunsberger et al (2007)	United States	2,217	38.5	2,733	29.3
	DiNitto et al (2008)	United States	1,343	1.7	821	8.0
	DiNitto et al (2008)	United States	267	3.0	821	8.0
	Klausch et al (2013)	Netherlands	2,199	28.7	1,474	67.4
	McMillen et al (2013)	United States	2,272	67.5	2,128	70.7

Alternative method	Author (year)	Location	Web-based method		Alternative method	
			Sample size	RR (%)	Sample size	RR (%)
	Laaksonen and Heiskanen (2014) (Laaksonen and Heiskanen 2014)	Finland	3,993	25.0	3,009	62.0
	McCormack et al (2014) (McCormack et al. 2014)	United States	37	83.8	34	97.1
Face-to-face						
	Spijkerman et al (2009) (Spijkerman et al. 2009)	Netherlands	57,125	35.5	7,204	62.7
	Haider et al (2013) (Haider et al. 2013)	United States	46	58.7	46	91.0
	Klausch et al (2013) (Klausch et al. 2013)	Netherlands	2,199	28.7	2,081	64.3
	Laaksonen and Heiskanen (2014) (Laaksonen and Heiskanen 2014)	Finland	3,993	25.0	735	50.1
Other						
	Sinadinovic et al (2011) (Sinadinovic et al. 2011)	Sweden	1,976	38.1	1,960	33.9

Alternative method	Author (year)	Location	Web-based method		Alternative method	
			Sample size	RR (%)	Sample size	RR (%)
			Lagerros et al (2012) (Lagerros et al. 2012)	Sweden	182	93.4
Lee et al (2015) (Lee et al. 2015)	South Korea	21,456	20.2	23,636	17.0	

RR – response rate

5 – Artigo 2

A ser submetido para o *International Journal of Epidemiology*

How different online recruitment methods impact on recruitment rates for the web-based *coortesnaweb* project: a randomized trial

Cauane Blumenberg¹, Ana M. Menezes¹, Helen Gonçalves¹, Maria Cecília Formoso Assunção¹, Fernando C. Wehrmeister¹, Aluísio J. D. Barros¹

¹ Post-Graduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil.

ABSTRACT

The number of web-based E-epidemiologic studies using online recruitment methods is increasing in the literature. There is no optimal online recruitment method in terms of maximizing recruitment rates. Our aim was to compare the recruitment rates of three online recruitment methods. A total of 2,394 members of the 1993 Pelotas birth cohort that provided an e-mail address, a Facebook name, and a Whatsapp number during a face-to-face follow-up were randomly allocated to be recruited by e-mail, Facebook or Whatsapp (798 individuals per method). Between January and February 2018, we sent messages inviting them to register into the web-based *coortesnaweb* platform. Recruitment rates were calculated for each method, and according to gender, schooling, skin color, and socioeconomic position. We also analyzed absolute and relative inequalities on recruitment according to schooling and socioeconomic position. The overall recruitment rate was 26.8%. Facebook was the most effective recruitment method (recruitment rate = 30.6%). Recruitment rates of e-mail and Whatsapp were similar. The highest inequalities on recruitment were observed using e-mail. The effectiveness of online recruitment methods depend on the characteristics of the sample. It is important to know the profile of the target sample in order to decide which online recruitment method to use.

Keywords: E-epidemiology; Epidemiology; Research Subject Recruitment; Survey methodology; Web surveys

INTRODUCTION

The increase on the number of web-based E-epidemiologic studies is influenced by several factors, including the ever increasing number of people with access to broadband internet (1), the reduced study costs (2), and similar validity of the data compared to traditional data collection methods (3,4). A very important aspect to consider when conducting web-based surveys is the way the participants are invited to participate, as the recruitment method can influence the recruitment rates (RECR) (proportion of all invited individuals that register to participate on the survey) (5–7), and consequently the response rates (6).

Some web-based epidemiologic studies compared online and offline recruitment methods, showing lower costs (5,8,9) and higher recruitment rates (5,10) for the online methods.

Additionally, using online recruitment methods is logistically simpler compared to offline methods, since it can be done by placing ads on websites or sending automatic messages over the internet (11,12).

In the literature, it is described that recruitment rates of online methods can range from 6% to over 50% (5,6,13). Until today, there is no optimum online recruitment method in terms of maximizing recruitment rates (9). However, studies fail to formally compare the effectiveness of these methods. For this reason, the objective of this study was to compare the recruitment rates of three online recruitment methods in the context of the web-based *coortesnaweb* project.

METHODS

The 1993 Pelotas birth cohort recruited 5,249 participants who were born in that year from mothers living in the urban area of Pelotas, a medium-sized Southern Brazilian city. The original cohort sample was representative of all births occurring in the city (14). In 2015, when the members of the birth cohort were 22 years of age, they were invited to participate in a face-to-face follow-up assessment, and 3,810 individuals (76.3% of the original cohort, including those who died as being followed-up) were interviewed (15). Several health-related and life-style subjects were investigated, including internet access and the use of online social networks. We asked the participants whether they had access to broadband internet, where and how often to quantify how many participants would be able to participate in a web-based project called *coortesnaweb*.

Coortesnaweb is an experimental web-based platform developed to collect epidemiological data from the members of the 1993 Pelotas birth cohort. The data is collected by means of self-administered online surveys, designed using REDCap (16), which were integrated to the *coortesnaweb* platform. In order to encourage the participation and to reduce attrition, gamification strategies were developed. After responding to the questionnaires, the participants earn points and these points are used to unlock personal results about their health (e.g. level of physical activity, internet addiction). Additionally, the participants can earn badges after achieving some goals.

To be eligible to the *coortesnaweb*, the members of the 1993 Pelotas birth cohort had to have: (i) participated in the 2015 follow-up, (ii) confirmed that they had internet access at home or in their mobile device, and (iii) responded to the 2015 face-to-face interview without help of a third individual. A total of 3,537 cohort members (67.4% of the original cohort) were eligible to participate of the *coortesnaweb* project. However, since this study focuses only on the effectiveness of online recruitment methods, only the 2,394 cohort members that provided an e-mail address, a Facebook name, and a Whatsapp number during the 2015 face-to-face assessment were considered eligible for this study.

This was a parallel randomized trial with equal groups. The 2,394 eligible cohort members were randomly assigned to be exclusively recruited by one of the three online recruitment methods: e-mail, Facebook message or Whatsapp message (Facebook Inc. 2018. Menlo Park, CA, United States). A block randomization (block size = 3) was applied to guarantee equal sample sizes on each group (798 individuals per group). Each participant received at most three invitations to register into the *coortesnaweb* platform, using the same recruitment method. The invitation messages were sent within a 10-day interval, between 11 January and 19 February 2018, on different times and days of the week. If the participant had registered into the platform before receiving the third invitation message, no more messages were sent. The invitation messages were tailored to each individual (using their first name), were written in Portuguese and had the same content, independent of the recruitment method used. The only exception was the e-mail message that also had a subject line, since the structure of e-mail messages require a subject (Supplemental Material 1).

E-mail messages were sent to the group allocated to be recruited by e-mail using mail merge, sending 300 e-mails per day to avoid the messages to be considered as spam. The Whatsapp numbers of the individuals allocated to be recruited by Whatsapp were added to the contact list of the *coortesnaweb* mobile phone, and the messages were sent manually. Facebook profiles of the individuals allocated to be recruited by Facebook were searched by two project members, and the invitation messages were also sent manually.

The main outcome analyzed was the recruitment rate (17), which was computed as the ratio between the number of individuals who had registered and the total individuals invited (irrespective of a successful contact or not, since e-mails could return, Whatsapp numbers could be wrong and Facebook profiles could not be found). The recruitment rate for each group was calculated in order to compare their effectiveness.

The recruitment rate was also calculated according to sex, schooling, skin color, and socioeconomic position (SEP). Skin color was self-reported and categorized as white, brown, black, and other. This is commonly used as a proxy for social disparities in the Brazilian population (18). Schooling was measured as total completed years of schooling until 2015, and categorized into 0-8, 9-11, and 12 or more years. Socioeconomic position was estimated through a wealth index based on the assets of the household in 2015. The index was calculated using principal components analysis, and individuals divided into five equal sized groups (quintiles).

Relative and absolute inequalities in recruitment rate were calculated according to schooling and socioeconomic position using the concentration index (CIX) and the slope index of inequality (SII), respectively. The CIX is calculated by ranking the individuals according to groups of

schooling or socioeconomic position in ascending order. If the CIX is zero, the recruitment rate would be equal across individuals, positive CIX values indicate higher recruitment rates among richer or better educated individuals, while negative values indicate higher recruitment rates among the poorer and less educated. The SII was calculated by a logistic regression, using the recruitment rate as outcome and the schooling or socioeconomic position as exposure variables. The slope index can be interpreted as the difference between the recruitment rates (expressed in percentage points) of the top and the bottom groups of schooling and socioeconomic position. Detailed information about these measures can be obtained elsewhere (19). All the results presented in tables were obtained using intention to treat analyses. Per-protocol analyses, considering only the contacted individuals in the analyses, are provided as supplemental material. All analyses were conducted using the Stata (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC).

The *coortesnaweb* project was approved by the Research Ethics Committee of the School of Medicine of the Federal University of Pelotas (protocol number: 79124917.5.0000.5317). This randomized trial was approved and registered by the Brazilian Registry of Clinical Trials (register number: RBR-3dv7gc). All participants had to agree to the consent form in order to register into the platform.

RESULTS

Compared to the cohort participants seen in 2015, at 22 years of age, the eligible participants for the *coortesnaweb* study were more educated, wealthier, and more likely to be white. There was no statistical difference according to sex (Table 1). There were no statistical differences according to sex, skin color and any of the socioeconomic characteristics between the three recruitment methods, after the randomization process (Table 2).

The flowchart presented in Figure 1 shows that 85.0%, 74.3% and 84.0% of the individuals assigned to be recruited by e-mail, Whatsapp and Facebook, respectively, could be contacted (e-mail did not return, Whatsapp number existed and Facebook profile could be found). There were three individuals that explicitly refused to participate, two from the Whatsapp and one from the Facebook group. There were two losses in the Facebook group, because they wrongly received invitations by Whatsapp. Two deaths were identified, one because a family member answered the Whatsapp message, and another because the Facebook profile was memorialized.

A total of 642 individuals registered into the platform, representing an overall recruitment rate of 26.8% (95% CI = 25.1, 28.6). Of those, 491 individuals (76.5%) went on to complete the first questionnaire. Considering all the recruitment methods together, 31.0% (95% CI = 25.5, 33.6) of the invited females registered into the *coortesnaweb* platform compared to 21.8% (95% CI = 19.4, 24.3) of the males.

In Figure 2A it is shown that the most effective recruitment method was the Facebook, with 30.6% (95%CI = 27.5, 33.9) of recruitment. The recruitment rates of Whatsapp and e-mail were the same (24.9%; 95%CI = 22.1, 28.1 for both methods). Sending Facebook messages was also the most effective method to recruit the poorest individuals (Figure 2D). Moreover, it was seen that the recruitment rates obtained by using Facebook were similar across socioeconomic positions. In contrast, for the e-mail and Whatsapp recruitment methods, the recruitment rates increased with the increase of the socioeconomic position. The e-mail, Whatsapp and Facebook groups achieved similar recruitment rates according to skin color groups (Figure 2C), and higher recruitment rates the higher the schooling (Figure 2E). E-mail was the most effective recruitment method to invite highly educated and wealthier participants.

The median number of days between the first invitation message and the day of registration taking into account only those who registered was also computed (data not shown in tables). Overall, it took a median of four days for participants to register into the *coortesnaweb* platform. The Whatsapp was the recruitment method that took the least time from the first invitation to registration, median of only one day (ranging from 0 to 31 days). For the Facebook method the median number of days was four (range: 0-45 days), while for e-mail it was six (range: 0-44 days).

The slope and concentration indexes presented in Table 3 showed that the highest inequalities on recruitment were observed in the e-mail group, for both schooling categories and socioeconomic position. The lowest inequalities according to socioeconomic position were observed in the Facebook group; being the slope index approximately a fifth compared to e-mail and Whatsapp methods. Overall, both absolute and relative inequalities on recruitment rates are higher according to schooling compared to the socioeconomic position.

DISCUSSION

Our findings showed that using free online recruitment methods, we were able to recruit 26.8% of the eligible birth cohort members to participate in a web-based epidemiologic study. The most effective method was the Facebook. We also identified that e-mails were the most effective method to attract highly educated and wealthier individuals, and that the Facebook was the method that attracted the most equitable sample according to socioeconomic position. In our sample, the recruitment rate inequalities were higher according to schooling than according to socioeconomic position.

The overall recruitment rate of our web-based study was lower compared to the previous face-to-face follow-ups of the 1993 Pelotas birth cohort (15), and this might have occurred due to two main factors. First, in the face-to-face follow-ups we perform an extensive recruitment process, trying several times to contact the participant and using both online and offline

recruitment methods, such as phone calls and home visiting. Second, in the face-to-face follow-ups we offer a monetary incentive for the participants, whereas in the web-based study we only offered non-monetary incentives (access to personal results). Offering monetary incentives in web-based surveys can increase participation rates (20), but due to the lack of funding, we were not able to employ such approach.

Using Facebook ads to recruit participants to epidemiologic web-based studies is very common in the literature (5,7,9,21–23). However, when a study is not supposed to have open registration, but is focused on a list of eligible individuals, using targeted Facebook ads may become very expensive and not feasible. We chose to send individual Facebook messages as we could confirm the eligibility of the individuals by checking their Facebook profile information (mainly using name, education, place and date of birth). Only one study, conducted in the United States, employed a similar approach as ours and achieved 24.6% of recruitment rate (13). In our study, the Facebook recruitment rate was higher (30.6%).

The effectiveness of online recruitment methods depend on the characteristics of the sample (9). Some studies state that sending e-mails is more effective to recruit older participants, and that Facebook would be more effective to recruit a sample of young adults (21,24). This may be one of the reasons why the Facebook was the most effective recruitment method in our study, since individuals from 24 to 25 years of age composed our sample. Another reason that might explain the higher effectiveness of Facebook compared to the other recruitment methods is that errors in the Facebook contact information provided during the face-to-face follow-up could be solved. Even if the participant had provided an incorrect Facebook profile name, in some cases we could identify the correct profile by searching the name of the individual into the social network and checking its personal information against the Facebook profile. In contrast, if there were errors in the Whatsapp number or in the e-mail address, it was not possible to identify or correct the error (apart from minor typing errors in the e-mail address).

To the best of our knowledge, this was the first study that used Whatsapp to recruit participants to an epidemiologic study. Although the Whatsapp achieved a similar recruitment rate as sending e-mails, it was the method that presented the highest percentage of unreachable individuals – 25.6% versus around 15.0% for Facebook and e-mail. The higher percentage can be attributed to two factors: i) the impossibility to correct errors in Whatsapp numbers, and ii) the frequent changes in mobile phone numbers. In Brazil, the mobile carriers launch new plans that are usually financially better than the existing ones, encouraging clients to change plans and, sometimes, their phone number. If the unreachable individuals were not considered, the Whatsapp recruitment rate would reach 33.4% and would be similar to Facebook's (see per-protocol analyses in Supplemental Material 3).

Unlike Whatsapp, sending e-mails as recruitment method to epidemiologic research is more common in the literature. Similar to our finding, Buckingham and colleagues found e-mails to be less effective than recruitment via a social network (21). In contrast, two other studies described that the recruitment rates of e-mails were higher than Facebook's (7,24). The higher effectiveness of e-mails compared to Facebook can be attributed to the sample composition of these two studies, which were older than ours.

In our study, compared to Facebook, the recruitment via e-mail was related to higher absolute and relative inequalities according to schooling and socioeconomic position. A similar finding was described in a study that attracted better educated and richer participants using e-mails compared to Facebook recruitment (7). Our hypothesis to explain this is that e-mails are mainly used for workplace and university communication, hence attracting individuals with formal jobs, with better education and from a higher socioeconomic position.

One notable finding is the marked sex differences on recruitment rates independent of the method used, which were almost 10 percentage points higher for females compared to males. This finding is consistent to three other web-based studies (7,22,24). Another epidemiologic study, which also found higher participation among females, identified that the reasons for males not to participate are due to lack of interest and time constraints (25). We did not investigate the reasons for not registering into the *coortesnaweb* platform, but we hypothesize that males could be less interested in the study as the participation of females was always higher than males' in the previous face-to-face follow-ups of the 1993 Pelotas birth cohort (15).

A limitation that could have affected our study is the contamination between participants. For instance, a participant could have mentioned the study to another eligible individual before this individual read its recruitment message. This individual would become aware of the study by word of mouth, but in our analyses he would be considered recruited by the method for which he/she was originally allocated. We could not quantify the total amount of contamination that could have affected our results, but two individuals allocated to receive e-mails were unreachable (did not receive any e-mail), but still registered into the *coortesnaweb* platform. Another limitation is that we did not use a mixed approach of recruitment methods, what could have increased the recruitment rates, as the number of not contacted individuals would decrease. It is important to note that our findings are inserted in the context of the 1993 Pelotas birth cohort, a known sample that were already aware of the study and had participated of previous face-to-face follow-ups. In other situations, such as an unknown target population, other online recruitment methods could be employed (e.g. online ads) and their effectiveness could be different. Our study also presents some strengths: i) this was the first study that formally compared, using a randomized trial design, the effectiveness of online recruitment methods for

epidemiologic research; ii) this was also the first web-based epidemiological study fully conducted online in the context of a middle-income country; and iii) we could compare the effectiveness of the recruitment methods using a standardized metric (the recruitment rate) (17), what usually does not happen in web-based epidemiologic research as the number of eligible individuals is not known.

We were able to recruit members of a birth cohort to a web-based epidemiologic study, using free online recruitment methods in the context of a middle-income country (Brazil). The effectiveness of the online recruitment methods is dependent on the individual characteristics of the target sample. Overall, the Facebook showed to be the most effective method to recruit young adults, also achieving the most equitable sample according to schooling and socioeconomic position. In contrast, the use of e-mails as a recruitment method should be employed when the sample is composed by better educated or richer individuals. Other online recruitment methods, such as sending Whatsapp messages, should be exploited by e-epidemiological studies as they are cost effective recruitment options.

ACKNOWLEDGEMENTS

This article is based on data from the study "Pelotas Birth Cohort, 1993" conducted by Postgraduate Program in Epidemiology at Universidade Federal de Pelotas with the collaboration of the Brazilian Public Health Association (ABRASCO). From 2004 to 2013, the Wellcome Trust supported the 1993 birth cohort study. The European Union, National Support Program for Centers of Excellence (PRONEX), the Brazilian National Research Council (CNPq), and the Brazilian Ministry of Health supported previous phases of the study. The 22-year follow-up was supported by the Science and Technology Department / Brazilian Ministry of Health, with resources transferred through the Brazilian National Council for Scientific and Technological Development (CNPq), grant 400943/2013-1. Capes foundation sponsored the exchange period of the first author of this work (grant n. 88881.133234/2018-01).

REFERENCES

1. The World Bank. Internet users (per 100 people). *World Bank*. 2018;(http://data.worldbank.org/indicator/IT.NET.USER.P2?end=2014&start=2014&view=bar). (Accessed April 30, 2018)
2. McMaster HS, LeardMann CA, Speigle S, et al. An experimental comparison of web-push vs. paper-only survey procedures for conducting an in-depth health survey of military spouses. *BMC Med. Res. Methodol.* [electronic article]. 2017;17(1). (http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0337-1). (Accessed April 27, 2017)
3. Krogh A-B, Larsson B, Salvesen Ø, et al. A comparison between prospective Internet-based and paper diary recordings of headache among adolescents in the general population. *Cephalalgia Int. J. Headache*. 2016;36(4):335–345.
4. Christensen AI, Ekholm O, Glumer C, et al. Effect of survey mode on response patterns: comparison of face-to-face and self-administered modes in health surveys. *Eur. J. Public Health*. 2014;24(2):327–332.
5. Christensen T, Riis AH, Hatch EE, et al. Costs and Efficiency of Online and Offline Recruitment Methods: A Web-Based Cohort Study. *J. Med. Internet Res*. 2017;19(3):e58.
6. Brown O, Quick V, Colby S, et al. Recruitment lessons learned from a tailored web-based health intervention Project Y.E.A.H. (Young Adults Eating and Active for Health). *Health Educ*. 2015;115(5):470–479.
7. Dworkin J, Hessel H, Gliske K, et al. A Comparison of Three Online Recruitment Strategies for Engaging Parents: Online Recruitment. *Fam. Relat*. 2016;65(4):550–561.
8. Harris ML, Loxton D, Wigginton B, et al. Recruiting Online: Lessons From a Longitudinal Survey of Contraception and Pregnancy Intentions of Young Australian Women. *Am. J. Epidemiol*. 2015;181(10):737–746.
9. Lane TS, Armin J, Gordon JS. Online Recruitment Methods for Web-Based and Mobile Health Studies: A Review of the Literature. *J. Med. Internet Res*. 2015;17(7):e183.
10. Wise LA, Rothman KJ, Mikkelsen EM, et al. Design and Conduct of an Internet-Based Preconception Cohort Study in North America: Pregnancy Study Online. *Paediatr. Perinat. Epidemiol*. 2015;29(4):360–371.
11. Trespalacios JH, Perkins RA. Effects of Personalization and Invitation Email Length on Web-Based Survey Response Rates. *TechTrends*. 2016;60(4):330–335.

12. Loxton D, Powers J, Anderson AE, et al. Online and Offline Recruitment of Young Women for a Longitudinal Health Survey: Findings From the Australian Longitudinal Study on Women's Health 1989-95 Cohort. *J. Med. Internet Res.* 2015;17(5):e109.
13. Jones L, Saksvig BI, Grieser M, et al. Recruiting adolescent girls into a follow-up study: Benefits of using a social networking website. *Contemp. Clin. Trials.* 2012;33(2):268–272.
14. Victora CG, Barros FC, Halpern R, et al. Longitudinal study of the mother and child population in an urban region of southern Brazil, 1993: methodological aspects and preliminary results. *Rev. Saude Publica.* 1996;30(1):34–45.
15. Gonçalves H, Wehrmeister FC, Assunção MCF, et al. Cohort Profile Update: The 1993 Pelotas (Brazil) Birth Cohort follow-up at 22 years. *Int. J. Epidemiol.* 2017;
16. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* 2009;42(2):377–381.
17. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th ed. AAPOR; 2016.
18. Travassos C, Williams DR. The concept and measurement of race and their relationship to public health: a review focused on Brazil and the United States. *Cad. Saúde Pública.* 2004;20(3):660–678.
19. Barros AJD, Victora CG. Measuring Coverage in MNCH: Determining and Interpreting Inequalities in Coverage of Maternal, Newborn, and Child Health Interventions. *PLoS Med.* 2013;10(5):e1001390.
20. David MC, Ware RS. Meta-analysis of randomized controlled trials supports the use of incentives for inducing response to electronic health surveys. *J. Clin. Epidemiol.* 2014;67(11):1210–1221.
21. Buckingham L, Becher J, Voytek CD, et al. Going social: Success in online recruitment of men who have sex with men for prevention HIV vaccine research. *Vaccine.* 2017;35(27):3498–3505.
22. Carter-Harris L, Bartlett Ellis R, Warrick A, et al. Beyond Traditional Newspaper Advertisement: Leveraging Facebook-Targeted Advertisement to Recruit Long-Term Smokers for Research. *J. Med. Internet Res.* 2016;18(6):e117.
23. Juraschek SP, Plante TB, Charleston J, et al. Use of online recruitment strategies in a randomized trial of cancer survivors. *Clin. Trials.* 2018;15(2):130–138.

24. Nolte MT, Shauver MJ, Chung KC. Analysis of four recruitment methods for obtaining normative data through a Web-based questionnaire: a pilot study. *HAND*. 2015;10(3):529–534.
25. Markanday S, Brennan SL, Gould H, et al. Sex-differences in reasons for non-participation at recruitment: Geelong Osteoporosis Study. *BMC Res. Notes*. 2013;6(1):104.

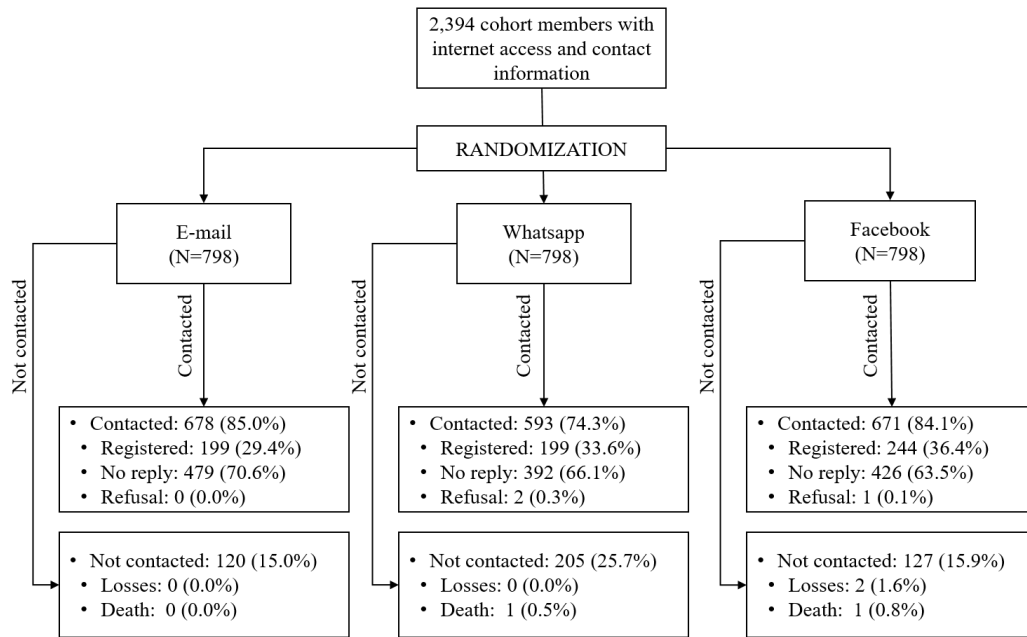


Figure 1. Flowchart Depicting the Logistic and Design of the Study, and the Number of Individuals Contacted and not Contacted by Each Recruitment Method. Pelotas, Brazil, 2018.

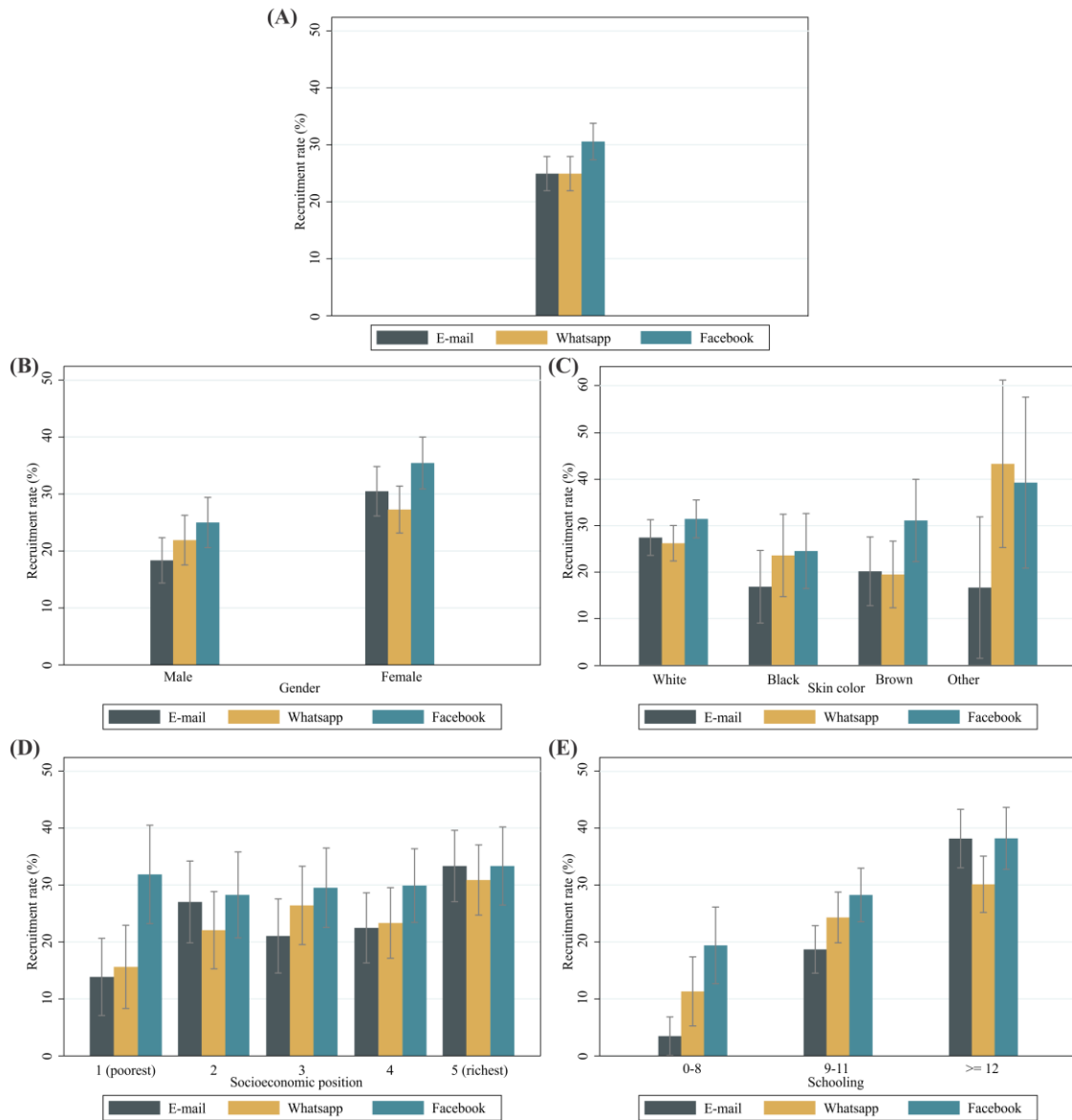


Figure 2. (A) Overall Recruitment Rate According to Recruitment Method, and Also According to (B) Sex, (C) Skin Color, (D) Socioeconomic Position, and (E) Schooling. Pelotas, Brazil, 2018.

Table 1. Comparison of the Sample Followed-Up at 22 Years of Age and the Group Eligible to the Recruitment Study. Pelotas, Brazil, 2018.

	Followed-up at 22 years of age			Eligible group			P value
	N	%	95%CI	N	%	95%CI	
Total	3,810			2,394			
Sex							0.243
Female	2,027	53.2	51.6, 54.8	1,310	54.7	52.7, 56.7	
Male	1,783	46.8	45.2, 48.4	1,084	45.3	43.3, 47.3	
Schooling (years)							< 0.001
0-8	1,124	29.5	28.1, 31.0	355	14.9	13.5, 16.3	
9-11	1,560	41.0	39.4, 42.6	1,049	43.8	41.9, 45.9	
12+	1,121	29.5	28.0, 30.9	987	41.3	39.3, 43.3	
Skin color							0.001
White	2,262	63.3	61.7, 64.9	1,531	68.3	66.4, 70.3	
Brown	637	17.8	16.6, 19.1	338	15.1	13.7, 16.6	
Black	538	15.1	13.9, 16.3	288	12.9	11.5, 14.3	
Other	137	3.8	3.3, 4.5	82	3.7	3.0, 4.5	
Socioeconomic position							< 0.001
1st (poorest)	761	20.0	18.8, 21.3	310	13.0	11.7, 14.4	
2nd	761	20.0	18.8, 21.3	431	18.0	16.5, 19.6	
3rd	761	20.0	18.8, 21.3	477	20.0	18.4, 21.6	
4th	761	20.0	18.8, 21.3	552	23.1	21.5, 24.8	
5th (richest)	760	20.0	18.7, 21.3	619	25.9	24.2, 27.7	

CI – confidence interval

Table 2. Characteristics of the Sample According to Randomization Group. Pelotas, Brazil, 2018.

	E-mail		Whatsapp		Facebook		P value
	N	% (95%CI)	N	% (95%CI)	N	% (95%CI)	
Total	798		798		798		
Sex							0.431
Female	433	54.3 (50.8, 57.7)	451	56.5 (53.1, 59.9)	426	53.4 (49.9, 56.8)	
Male	365	45.7 (42.3, 49.2)	347	43.5 (40.1, 47.0)	372	46.6 (43.2, 50.1)	
Schooling (years)							0.179
0-8	115	14.4 (12.1, 17.0)	106	13.3 (11.1, 15.9)	134	16.8 (14.4, 19.6)	
9-11	337	42.2 (38.8, 45.7)	358	45.0 (41.5, 48.5)	354	44.4 (41.0, 47.9)	
12+	346	43.4 (40.0, 46.8)	332	41.7 (38.3, 45.2)	309	38.8 (35.4, 42.2)	
Skin color							0.561
White	521	69.7 (66.3, 72.8)	511	68.3 (64.9, 71.6)	499	67.1 (63.7, 70.4)	
Brown	114	15.2 (12.8, 18.0)	118	15.8 (13.3, 18.6)	106	14.3 (11.9, 17.0)	
Black	89	11.9 (9.8, 14.4)	89	11.9 (9.8, 14.4)	110	14.8 (12.4, 17.5)	
Other	24	3.2 (2.2, 4.7)	30	4.0 (2.6, 5.4)	28	3.8 (2.6, 5.4)	
Socioeconomic position							0.510
1st (poorest)	101	12.7 (10.5, 15.2)	96	12.1 (10.0, 14.5)	113	14.2 (12.0, 16.8)	
2nd	148	18.6 (16.0, 21.4)	145	18.2 (15.7, 21.0)	138	17.4 (14.9, 20.2)	
3rd	152	19.1 (16.5, 21.9)	159	20.0 (17.3, 22.9)	166	20.9 (18.2, 23.9)	
4th	178	22.3 (19.5, 25.3)	180	22.6 (19.8, 25.6)	194	24.4 (21.6, 27.5)	
5th (richest)	219	27.4 (24.5, 30.6)	217	27.2 (24.3, 30.4)	183	23.1 (20.2, 26.1)	

CI – confidence interval

Table 3. Overall Inequalities on Recruitment Rates Comparing Schooling Categories and Socioeconomic Position, and According to Randomization Group. Pelotas, Brazil, 2018.

	Overall		E-mail		Whatsapp		Facebook	
	SII (95%CI)	CIX (95%CI)	SII (95%CI)	CIX (95%CI)	SII (95%CI)	CIX (95%CI)	SII (95%CI)	CIX (95%CI)
Schooling	29.9 (23.8, 36.0)	9.3 (5.6, 13.0)	46.5 (37.1, 56.0)	24.4 (18.1, 30.7)	20.0 (9.4, 30.6)	1.7 (-5.1, 8.5)	24.5 (13.2, 35.7)	4.1 (-1.9, 10.0)
Socioeconomic position	11.3 (5.1, 17.4)	3.0 (-0.8, 6.8)	16.6 (6.2, 27.0)	8.7 (1.8, 15.6)	14.9 (4.6, 25.2)	4.7 (-2.2, 11.6)	3.1 (-8.1, 14.3)	-2.5 (-8.5, 3.6)

CI – confidence interval; CIX – concentration index; SII – slope index of inequality

SUPPLEMENTAL MATERIAL

E-mail heading: Meet the new 1993 cohort platform on the Internet

Hello [name], how are you?

We would like to invite you to participate of the coortesnaweb.com, the new research platform of the 1993 Cohort on the Internet! In this platform you are able to respond to questionnaires, earn points and badges! With these points you will be able to unlock personal results about your health and compare you results to the other participants from the 1993 Cohort!

Register at coortesnaweb.com and start right now!

coortesnaweb team

Supplemental Material 1. Invitation Message Content Used in all Recruitment Methods. Heading was Only Used in the E-Mail Messages. Pelotas, Brazil, 2018.

Supplemental Material 2. Recruitment Rate According to Randomization Group and Individual Characteristics Using Intention to treat Analysis. Pelotas, Brazil, 2018.

	Overall		E-mail		Whatsapp		Facebook	
	RECR (%)	95%CI	RECR (%)	95%CI	RECR (%)	95%CI	RECR (%)	95%CI
Overall	26.8	25.1, 28.6	24.9	22.1, 28.1	24.9	22.1, 28.1	30.6	27.5, 33.9
Sex								
Female	31.0	28.5, 33.6	30.5	26.3, 35.0	27.3	23.4, 31.6	35.5	31.0, 40.1
Male	21.8	19.4, 24.3	18.4	14.7, 22.7	21.9	17.9, 26.6	25.0	20.9, 29.7
Schooling (years)								
0-8	11.8	8.9, 15.6	3.5	1.3, 8.9	11.3	6.5, 18.9	19.4	13.5, 27.0
9-11	23.8	21.3, 26.5	18.7	14.9, 23.2	24.3	20.1, 29.0	28.3	23.8, 33.2
12+	35.5	32.5, 38.5	38.2	33.2, 43.4	30.1	25.4, 35.3	38.2	32.9, 43.7
Skin color								
White	28.4	26.1, 30.7	27.5	23.8, 31.5	26.2	22.6, 30.2	31.5	27.5, 35.7
Brown	23.4	19.2, 28.2	20.2	13.8, 28.6	19.5	13.3, 27.7	31.1	23.0, 40.6
Black	21.9	17.5, 27.0	16.9	10.4, 26.2	23.6	15.9, 33.6	24.5	17.4, 33.5
Other	34.2	24.7, 45.1	16.7	6.3, 37.5	43.3	26.8, 61.5	39.3	23.0, 58.4
Socioeconomic position								
1st (poorest)	21.0	17.9, 25.3	13.9	8.4, 22.1	15.6	9.6, 24.4	31.9	23.9, 41.0
2nd	25.8	21.6, 29.4	27.0	20.5, 34.8	22.1	16.0, 29.6	28.3	21.4, 36.4
3rd	25.8	23.2, 31.2	21.1	15.3, 28.3	26.4	20.1, 33.8	29.5	23.1, 36.9
4th	25.4	23.6, 31.6	22.5	16.9, 29.2	23.3	17.7, 30.1	29.9	23.9, 36.7
5th (richest)	32.5	28.8, 37.3	33.3	27.4, 39.9	30.9	25.1, 37.4	33.3	26.9, 40.5

CI – confidence interval; RECR – recruitment rate

Supplemental Material 3. Recruitment Rate According to Randomization Group and Individual Characteristics Using Per-Protocol Analysis. Pelotas, Brazil, 2018.

	Overall		E-mail		Whatsapp		Facebook	
	RECR (%)	95%CI	RECR (%)	95%CI	RECR (%)	95%CI	RECR (%)	95%CI
Overall	32.9	30.8, 35.0	29.1	25.8, 32.7	33.4	29.7, 37.3	36.1	32.6, 39.9
Sex								
Female	37.3	34.4, 40.2	34.9	30.2, 39.8	35.4	30.5, 40.6	41.5	36.5, 46.6
Male	27.4	24.5, 30.5	22.1	17.8, 27.1	30.6	25.2, 36.7	29.9	25.1, 35.2
Schooling (years)								
0-8	17.1	12.9, 22.4	5.3	2.0, 13.5	19.7	11.5, 31.6	23.9	16.7, 32.8
9-11	30.2	27.1, 33.4	22.8	18.2, 28.2	33.3	27.9, 39.3	34.3	29.0, 39.9
12+	40.0	36.8, 43.3	40.0	34.8, 45.4	36.7	31.1, 42.6	43.2	37.5, 49.2
Skin color								
White	33.6	31.1, 36.3	30.7	26.7, 35.1	34.0	29.5, 38.9	36.4	32.0, 41.1
Brown	31.2	25.8, 37.2	25.3	17.4, 35.2	29.1	20.1, 40.1	39.8	29.8, 50.7
Black	29.4	23.7, 35.9	24.2	15.1, 36.5	31.3	21.3, 43.4	31.8	22.7, 42.4
Other	40.0	29.2, 51.9	19.1	7.2, 41.8	29.1	37.7, 77.5	40.7	23.9, 60.1
Socioeconomic position								
1st (poorest)	28.9	23.3, 35.2	17.3	10.5, 27.2	28.3	17.8, 41.9	39.6	30.0, 50.0
2nd	33.0	28.2, 38.3	32.8	25.0, 41.6	32.3	23.8, 42.2	33.9	25.8, 43.1
3rd	31.7	27.2, 36.5	26.2	19.2, 34.8	32.8	25.1, 41.5	35.5	27.9, 43.9
4th	30.5	26.5, 34.9	25.5	19.2, 33.0	30.4	23.3, 38.6	35.4	28.4, 43.0
5th (richest)	37.0	33.1, 41.2	36.4	29.9, 43.3	37.9	31.0, 45.2	37.0	29.9, 44.6

CI – confidence interval; RECR – recruitment rate

6 – Artigo 3

A ser submetido para o *American Journal of Epidemiology*.

The role of questionnaire length and frequency of reminders on response rates to a web-based epidemiological study: a randomized trial

Cauane Blumenberg¹, Ana M. Menezes¹, Helen Gonçalves¹, Maria Cecília Formoso Assunção¹, Fernando C. Wehrmeister¹, Fernando C. Barros², Aluísio J. D. Barros¹

¹ Post-Graduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil.

² Post-Graduate Program in Health and Behavior, Catholic University of Pelotas, Pelotas, Brazil.

ABSTRACT

Introduction: Web-based epidemiologic surveys are being widely used in the literature, but they still present lower response rates compared to traditional methods. Sending reminders respondents and the design of the questionnaires can influence the response rates of surveys.

Objective: Analyze the influence of questionnaire length and frequency of reminders, and the interaction between them, on the response rates of five web-based questionnaires.

Methods: This is a 2x2 factorial design. Members of the 1993 Pelotas birth cohort, registered into the *coortesnaweb* platform (n=1,277), were randomly allocated to respond to short or long questionnaire, and to receive high or low frequency of reminders. We analyzed the influence of these factors on the response rates of five web-based questionnaires, and also the relative risk of responding to an additional questionnaire according to socio demographic characteristics of the sample.

Results: The mean response rate considering the five questionnaires together was 54.3%. Sending reminders more frequently was positively associated to the response rates for the first two questionnaires, but there was no difference according questionnaire length. We found no interaction between questionnaire length and frequency of reminders on the response rates. Women and participants that studied 12 or more years had, respectively, 13% (95%CI=1.03, 1.23) and 28% (95%CI=1.09, 1.50) increased probability of responding an additional questionnaire compared to the reference.

Conclusions: We obtained high response rates for the first questionnaires applied, but these rates reduced through time. Strategies to keep participants motivated should be exploited in longitudinal web-based studies, and reminders should be sent even more frequently. Using long questionnaires in web-based epidemiologic studies do not jeopardize response rates.

Keywords: E-epidemiology; Epidemiology; Response Rate; Survey methodology; Web surveys

INTRODUCTION

Epidemiologic surveys are facing a constant decrease on their response rates, estimated to vary between one and two percentage points per year (1,2). Several factors influence the decrease on response rates, including the way the data is collected. Traditional epidemiologic surveys – conducted by mail, face-to-face or telephone interviews – are facing difficulties with the inefficiency of the mailing system, the decrease on the coverage of landline telephone, and with barriers to contact hard-to-reach populations. The internet is a technology that can overcome these difficulties, since its coverage is increasing worldwide and it does not involve complicated logistics. However, response rates of web-based studies still present lower response rates compared to studies using traditional data collection methods (3), but this can be changed if internet-based studies are well designed (4).

Edwards and colleagues (5) conducted a comprehensive literature review on methods to increase response rates of electronic questionnaires, including web surveys. The effects of incentives, content and appearance of the questionnaire, and origin and way of communication were widely evaluated. In contrast, the influence of questionnaire length and frequency of reminders were not enough covered. Only two studies analyzed questionnaire length, and found increased response rate for shorter questionnaires (5), agreeing with the result of another systematic review published one year later (6). However, there are mixed results about this association. A study found increased response rate using longer questionnaires (7), and other studies found no difference on response rates according to questionnaire length (8–10).

One aspect that is a consent in the literature is that sending reminders to non-responders positively affect the response rate (5,6,11,12). However, studies fail to understand how frequent these reminders should be sent in order to increase response rate and avoid being a burden. Some studies analyzed the number of reminders, but did no distinction about the frequency that they were sent (11,12). The studies that analyzed the influence of reminders and questionnaire length on response rates performed cross-sectional analyses, and did not study the influence of these exposures in a longitudinal basis.

The objective of this study was to analyze the influence of questionnaire length and frequency of reminders, and the interaction between them, on the response rates of five web-based questionnaires applied using the *coortesnaweb* platform.

METHODS

This study was conducted using the *coortesnaweb* platform, a gamified web-based platform developed to collect data from the members of the 1993 Pelotas birth cohort. After registering into the platform, cohort members respond to questionnaires and earn virtual points, used to

unlock personal results about their health. To be eligible to participate of the *coortesnaweb*, the members of the birth cohort had to have: i) participated in the 2015 face-to-face follow-up (when they were 22 years old); ii) internet access (either at home, work, or mobile device); and iii) responded to the 2015 interview without help of a third individual. Out of the 3,810 cohort members interviewed in the 2015 face-to-face follow-up, 3,537 were eligible (67.4% of the original cohort) for the *coortesnaweb* project.

Recruitment of the eligible individuals lasted seven months (from January to July-2018). During this period, we sent messages briefly presenting the *coortesnaweb* platform with an invitation to register into the platform. The messages were sent first using e-mails, followed by Whatsapp messages, and then by Facebook messages (Facebook Inc. 2018. Menlo Park, CA, United States). Each individual received at most six recruitment messages, if the registration into the *coortesnaweb* platform was performed before the sixth invitation, no further recruitment messages were sent. Although the recruitment process had finished in July-2018, registration into the platform is still open (but without active invitation). Until September-2018, 1,277 individuals (36.1% of the eligible individuals) had registered into the platform and are used as analytical sample in this study.

In order to assess the influence of questionnaire length and frequency of reminders on response rates, this study followed a 2x2 factorial design, applying five questionnaires in a longitudinal manner during almost nine months of follow-up. At registration, individuals were randomly allocated to one of the following arms: i) short questionnaires and low frequency of reminders; ii) short questionnaires and high frequency of reminders; iii) long questionnaires and low frequency of reminders; and iv) long questionnaires and high frequency of reminders. We created a list using block randomization (block size = 4) to guarantee the balance between the four arms. Registered individuals were blinded and did not know to which group they were allocated.

Our main outcome is the response rate, calculated for each questionnaire using the RR2 formula defined by the American Association for Public Opinion Research (13). This formula is the number of complete plus partial interviews divided by the total number of individuals eligible to respond. We considered an interview complete if more than 80% of the items of the questionnaire were responded, a partial interview if between 50% and 80% of the items were responded, and not responded if less than 50% of the items of the questionnaire were responded. Items left blank due to missing by questionnaire design were not considered in the calculus. Eligible individuals were all those registered into the *coortesnaweb* platform. A secondary outcome was the total number of questionnaires responded by each participant, which ranged from zero to five questionnaires.

We applied five questionnaires using two lengths, a short and a long version. The first questionnaire was about alcohol consumption, the second about physical activity, the third about internet use, the fourth about violence and the fifth about smoking. Only one version (short or long) of the questionnaire was displayed to each individual, depending on the arm to which he was allocated. Short questionnaires ranged from 11 to 17 items, while long questionnaires ranged from 21 to 33 items. According to results from a pilot study, on average, short questionnaires took four minutes to be completed, while long questionnaires took 14 minutes.

Reminders were sent to non-respondents using the same methods used for recruitment: e-mails, Whatsapp messages, and Facebook messages. We sent, at most, two reminders for each questionnaire. If the registered individual had responded to that questionnaire, no further reminders were sent until a new questionnaire was published into the platform. Individuals allocated to the high frequency group received reminders every 15 days, while those in the low frequency group received reminders every 30 days.

All results are presented using intention to treat analyses. We describe the crude response rates for each questionnaire, and according to the arms and main effects of the factorial design. We calculated the relative risk of responding each questionnaire according to questionnaire length and frequency of reminders using Poisson regression. As the time available to respond each questionnaire was different, we calculated the proportion of the total time that each questionnaire was online and used it as offset. Also using Poisson regression, we analyzed the relative risk of responding an additional questionnaire according to the socio demographic characteristics of the sample. In this case, we used as offset the proportion of the total time that each participant was registered into the *coortesnaweb* platform.

The socio demographic characteristics analyzed were sex (female, male), schooling in years of study (0-8, 9-11, and 12 or more), skin color, and socioeconomic position. Skin color was self-reported as white, brown, black and other. Schooling was calculated as the total years of study until the participants were 22 years of age. Socioeconomic position was constructed using a self-reported list of assets from the household, and calculated using principal component analysis and divided into tertiles. We performed the analyses using Stata 15.1 software (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC).

The Research Ethics Committee of the School of Medicine of the Federal University of Pelotas approved the *coortesnaweb* project (protocol number: 79124917.5.0000.5317). This study is registered into the Brazilian Registry of Clinical Trials (register number: RBR-3dv7gc). All participants had to agree to a consent form in order to register into the platform.

RESULTS

From January to September 2018, 1,277 individuals registered into the *coortesnaweb* platform. As described in Table 1, 319 participants composed each group of this 2x2 factorial design study, except for the group that received long questionnaires and low frequency of reminders, which was composed by 320 participants. There was no difference between the randomization groups regarding to sex, schooling, skin color or socioeconomic position.

The mean response rate, considering all five questionnaires, was 54.3%. Considering the questionnaires individually, the crude response rates for the first two were near to 70.0%, reducing for the next questionnaires and reaching 31.2% in the last questionnaire (Table 2).

In Figure 1A we present the response rate for each questionnaire considering the four randomization groups. For the first four questionnaires, the groups that received short questionnaires and high frequency of reminders, or long questionnaires and high frequency of reminders had higher response rates compared to the other groups, but without significance. There was no interaction between questionnaire length and frequency of reminders in any of the questionnaires (data not shown). Analyzing the frequency of reminders alone (Figure 1B), the prevalence of response was 9.0% (95%CI=1.02, 1.17) and 11.0% (95%CI=1.03, 1.20) higher for the high frequency group compared to the low frequency, considering the first two questionnaires (refer to Supplemental Material 1). There was no response rate difference analyzing the questionnaire length alone (Figure 1C).

Analyzing the total number of questionnaires responded, it is seen in Table 3 that women have 13% (95%CI=1.03, 1.23) increased probability of responding an additional questionnaire compared to men. This difference was also present in the short questionnaire and in the low frequency of reminders groups, where the probability was 22% higher (95%CI=1.08, 1.38 and 95%CI=1.07, 1.40, respectively). Overall, participants that studied for 12 years or more had 28% (95%CI=1.09, 1.50) higher probability of responding an additional questionnaire compared to those that studied from zero to eight years. A similar finding was seen in the short and long questionnaire groups, and in the low frequency group. Regarding socioeconomic position, participants allocated for the low frequency group that were from the richest category had 20% (95%CI=1.03, 1.41) higher probability of responding an additional questionnaire compared to the poorest, but this association was not perceived in the overall estimate.

DISCUSSION

Our study obtained high response rates for the first questionnaires applied, reaching almost 70.0% for the first two questionnaires, but 31.2% for the last one. Sending reminders more often was associated with higher response rates, and there was no difference on response rates

according to questionnaire length. Women and individuals from the higher schooling category responded to more questionnaires compared to men and individuals with lower schooling.

Although the response rate alone is not an evidence of study quality and validity (14), our study obtained very high response rates for the first two questionnaires, an intermediate response rate for the third questionnaire, and decreasing response rates for the following questionnaires. Due to the decreasing response rates during the course of the study, it is important to employ additional retention methods to keep participants motivated in studies with longitudinal designs. One alternative would be to give monetary incentives, which is known to increase response and retention rates (5,15). However, these rates are not exclusively influenced by methodological aspects, but also by the characteristics of the target sample. A Danish study with women willing to get pregnant, reported response rates of 87.5% for a follow-up questionnaire, showing that the interest on the study subject and the composition of the target population is also an important factor to be considered (16).

In our study, we found that women and participants with higher schooling responded to more questionnaires compared to men and individuals that studied from zero to eight years, respectively. The higher participation of these groups was already described by other web-based epidemiologic studies in the literature (17,18). Hence, web-based studies that focus on producing prevalence estimates should be interpreted in light of these differences to avoid the generalization of biased findings.

We found no influence of the questionnaire length on response rates for any of the five questionnaires. Thus, in the context of our study, using longer questionnaires do not jeopardized the response rates. Few studies analyzed this association using repeated questionnaires, but two of them reported that the questionnaire length had no influence on response rates of the follow-up questionnaires (16,19). Other randomized trials, that applied a single questionnaire, also found no association between questionnaire length and response rates (8–10). However, there are some studies that found a positive association between shorter questionnaires and response rates (5,20,21), but also a negative association (7). The mixed findings are due to the different settings of the surveys (e.g. location, target population), and to the different ways the length of the questionnaires is operationalized, for instance number of items, number of screens, or estimated time to complete (6). In order to standardize the findings, we encourage authors to use the total number of items in a questionnaire as the length measure. This would be the most comparable metric between different studies, since the number of screens in a web-based survey depends on the number of items per screen, and the estimated time to complete is closely dependent on the computer literacy and socio demographic characteristics of the sample (6,16).

Regarding reminders, some studies in the literature describe that sending more reminders is better for the response rate (7,22–24). However, Cho and colleagues found that sending two reminders was better for the response rate compared to three or more reminders (25). This shows that it is important to consider the possible overburden of the participants. In our study, we can discard overburden, as sending reminders every 15 days was more efficient to increase response rates compared to sending reminders every 30 days. Hence, sending reminders more frequently, for instance every week or 10 days, could reflect in even higher response rates (24).

One limitation that could have affected our study is that we used only three online methods to send reminders (e-mails, Facebook and Whatsapp messages). We chose these methods because they are largely used in Brazil and because they were free to use. Other types of offline methods, such as phone calls or short message services could be adopted in order to increase the response rates, but their use would involve additional costs. A second limitation is that the two questionnaire lengths were similar between short (from 11 to 17 items; estimated four minutes to complete) and long questionnaires (from 21 to 33 items; estimated 14 minutes to complete). This could explain the similar response rates according to questionnaire length. Our definition was an intermediate length based on the findings of a cross-sectional study, which identified 10-minute surveys as being the ideal length, and 20-minute surveys as the maximum acceptable length. However, as stated before, measuring questionnaire length through its estimated time to complete is not the ideal metric. Our study also present some strengths, such as the 2x2 factorial design that enabled the analysis of the main effects of questionnaire length and frequency of reminders, including their interaction, on the response rates. Another strength was the longitudinal design used to apply the questionnaires, being possible to understand the response behavior of the participants in a longitudinal web-based epidemiologic survey.

CONCLUSIONS

Our study obtained high response rates in the first three questionnaires, but it reduced for the following questionnaires. Strategies to keep participants motivated (e.g. giving monetary incentives and covering different survey topics) should be considered in longitudinal web-based epidemiologic studies. We found no influence of the questionnaire length on the response rate of the questionnaires, indicating that the use of longer questionnaires in epidemiologic web-based studies do not jeopardize response rates. The higher response rates among the group that received reminders more frequently encourages the use of an even higher frequency of reminders, but the overburden of the participants should be considered.

REFERENCES

1. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. *Am. J. Epidemiol.* 2006;163(3):197–203.
2. Wallander L, Tikkanen RH, Mannheimer LN, et al. The problem of non-response in population surveys on the topic of HIV and sexuality: a comparative study. *The European Journal of Public Health.* 2015;25(1):172–177.
3. Blumenberg C, Barros AJD. Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature. *Int J Public Health.* 2018;63(6):765–773.
4. Dillman DA, Smyth JD, Christian LM. Internet, phone, mail, and mixed-mode surveys: the tailored design method. 4. ed. Hoboken, NJ: Wiley; 2014 509 p.
5. Edwards PJ, Roberts I, Clarke MJ, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database of Systematic Reviews* [electronic article]. 2009;(http://doi.wiley.com/10.1002/14651858.MR000008.pub4). (Accessed April 23, 2018)
6. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior.* 2010;26(2):132–139.
7. Koitsalu M, Eklund M, Adolfsson J, et al. Effects of pre-notification, invitation length, questionnaire length and reminder on participation rate: a quasi-randomised controlled trial. *BMC Medical Research Methodology* [electronic article]. 2018;18(1). (https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0467-5). (Accessed March 2, 2018)
8. Yetter G, Capaccioli K. Differences in responses to Web and paper surveys among school professionals. *Behavior Research Methods.* 2010;42(1):266–272.
9. Hardigan PC, Popovici I, Carvajal MJ. Response rate, response time, and economic costs of survey research: A randomized trial of practicing pharmacists. *Research in Social and Administrative Pharmacy.* 2016;12(1):141–148.
10. Tai X, Smith AM, McGeer AJ, et al. Comparison of response rates on invitation mode of a web-based survey on influenza vaccine adverse events among healthcare workers: a pilot study. *BMC Medical Research Methodology* [electronic article]. 2018;18(1). (https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-018-0524-8). (Accessed July 4, 2018)

11. Van Mol C. Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology*. 2016;1–11.
12. Sebo P, Maisonneuve H, Cerutti B, et al. Rates, Delays, and Completeness of General Practitioners' Responses to a Postal Versus Web-Based Survey: A Randomized Trial. *Journal of Medical Internet Research*. 2017;19(3):e83.
13. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th ed. AAPOR; 2016.
14. Morton SMB, Bandara DK, Robinson EM, et al. In the 21st Century, what is an acceptable response rate? *Australian and New Zealand Journal of Public Health*. 2012;36(2):106–108.
15. David MC, Ware RS. Meta-analysis of randomized controlled trials supports the use of incentives for inducing response to electronic health surveys. *Journal of Clinical Epidemiology*. 2014;67(11):1210–1221.
16. Mikkelsen EM, Hatch EE, Wise LA, et al. Cohort Profile: The Danish Web-based Pregnancy Planning Study--'Snart-Gravid'. *International Journal of Epidemiology*. 2009;38(4):938–943.
17. Ebert JF, Huibers L, Christensen B, et al. Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost. *Journal of Medical Internet Research*. 2018;20(1):e24.
18. RübSamen N, Akmatov MK, Castell S, et al. Comparison of response patterns in different survey designs: a longitudinal panel with mixed-mode and online-only design. *Emerging Themes in Epidemiology* [electronic article]. 2017;14(1). (<http://e-tonline.biomedcentral.com/articles/10.1186/s12982-017-0058-2>). (Accessed October 2, 2018)
19. McCambridge J, Kalaitzaki E, White IR, et al. Impact of Length or Relevance of Questionnaires on Attrition in Online Trials: Randomized Controlled Trial. *Journal of Medical Internet Research*. 2011;13(4):e96.
20. Galesic M, Bosnjak M. Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*. 2009;73(2):349–360.
21. Guo Y, Kopec JA, Cibere J, et al. Population Survey Features and Response Rates: A Randomized Experiment. *Am J Public Health*. 2016;106(8):1422–1426.

22. Aerny-Perreten N, Domínguez-Berjón MF, Esteban-Vasallo MD, et al. Participation and factors associated with late or non-response to an online survey in primary care: Late and non-response to online surveys. *Journal of Evaluation in Clinical Practice*. 2015;21(4):688–693.
23. Toledo D, Aerny N, Soldevila N, et al. Managing an Online Survey about Influenza Vaccination in Primary Healthcare Workers. *International Journal of Environmental Research and Public Health*. 2015;12(1):541–553.
24. Svensson M, Svensson T, Hansen AW, et al. The effect of reminders in a web-based intervention study. *European Journal of Epidemiology*. 2012;27(5):333–340.
25. Cho YI, Johnson TP, VanGeest JB. Enhancing Surveys of Health Care Professionals: A Meta-Analysis of Techniques to Improve Response. *Evaluation & the Health Professions*. 2013;36(3):382–407.

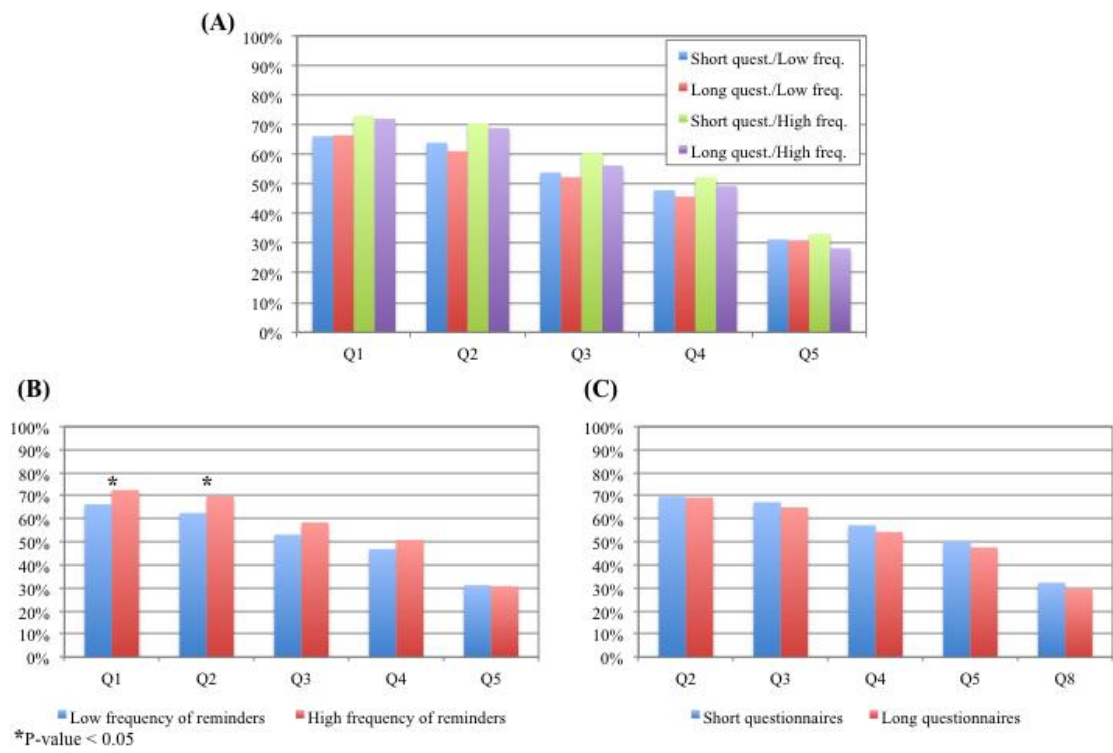


Figure 1. Response rate for each questionnaire considering (A) the four randomization groups and the individual effects of (B) frequency of reminders and (C) questionnaire length. Pelotas, Brazil, 2018.

Table 1. Characteristics of the sample according to randomization group. Pelotas, Brazil, 2018.

	Short quest.	Long quest.	Short quest.	Long quest.	P-value
	Low freq.	Low freq.	High freq.	High freq.	
	N (%)	N (%)	N (%)	N (%)	
Total	319	320	319	319	
Sex					0.706
Male	119 (37.3)	114 (35.6)	125 (39.2)	112 (35.1)	
Female	200 (62.7)	206 (64.4)	194 (60.8)	207 (64.9)	
Schooling (years)					0.869
0-8	40 (12.6)	46 (14.4)	41 (12.9)	47 (14.8)	
9-11	133 (42.0)	141 (44.0)	130 (40.9)	137 (43.2)	
12+	144 (45.4)	133 (41.6)	147 (46.2)	133 (42.0)	
Skin color					0.864
White	209 (69.9)	212 (70.1)	207 (68.0)	210 (68.5)	
Brown	49 (16.4)	40 (13.3)	44 (14.5)	43 (14.1)	
Black	29 (9.7)	40 (13.3)	37 (12.2)	40 (13.1)	
Other	12 (4.0)	10 (3.3)	16 (5.3)	13 (4.3)	
Socioeconomic position (tertiles)					0.090
1st (poorest)	96 (30.3)	113 (35.3)	105 (33.1)	110 (34.8)	
2nd	95 (30.0)	118 (36.9)	109 (34.4)	101 (32.0)	
3rd (richest)	126 (39.7)	89 (27.8)	103 (32.5)	105 (33.2)	

Freq. – frequency; Quest. – questionnaire

Table 2. Crude response rates per questionnaire. Pelotas, Brazil, 2018.

	Crude response rate	Time online
	%	Months
Q1	69.4	8.7
Q2	66.1	8.7
Q3	55.8	7.5
Q4	48.9	6.4
Q5	31.2	3.2
Mean	54.3	

Q1-Q5 – questionnaire 1 to 5

Table 3. Relative risk of number of questionnaires responded according questionnaire length, frequency of reminders and socio demographic characteristics. Pelotas, Brazil, 2018.

	Overall	Long quest.	Short quest.	Low freq.	High freq.
	RR (95%CI)	RR (95%CI)	RR (95%CI)	RR (95%CI)	RR (95%CI)
Sex					
Male	1.00	1.00	1.00	1.00	1.00
Female	1.13 (1.03, 1.23)	1.05 (0.92, 1.18)	1.22 (1.08, 1.38)	1.22 (1.07, 1.40)	1.06 (0.95, 1.18)
Schooling (years)					
0-8	1.00	1.00	1.00	1.00	1.00
9-11	1.11 (0.94, 1.31)	1.09 (0.86, 1.38)	1.13 (0.90, 1.43)	1.23 (0.95, 1.59)	1.03 (0.83, 1.27)
12+	1.28 (1.09, 1.50)	1.29 (1.02, 1.62)	1.26 (1.01, 1.58)	1.43 (1.11, 1.83)	1.16 (0.95, 1.42)
Skin color					
White	1.00	1.00	1.00	1.00	1.00
Brown	0.93 (0.82, 1.06)	1.00 (0.83, 1.20)	0.88 (0.73, 1.05)	0.89 (0.73, 1.09)	0.97 (0.82, 1.15)
Black	0.94 (0.82, 1.07)	0.97 (0.81, 1.17)	0.91 (0.74, 1.11)	1.01 (0.82, 1.24)	0.88 (0.74, 1.05)
Other	0.81 (0.64, 1.02)	0.97 (0.72, 1.31)	0.68 (0.48, 0.97)	0.67 (0.43, 1.04)	0.92 (0.71, 1.18)
Socioeconomic position (tertiles)					
1st (poorest)	1.00	1.00	1.00	1.00	1.00
2nd	1.11 (1.00, 1.23)	1.06 (0.91, 1.23)	1.16 (1.00, 1.35)	1.16 (0.99, 1.37)	1.06 (0.93, 1.21)
3rd (richest)	1.09 (0.99, 1.21)	1.04 (0.90, 1.21)	1.14 (0.98, 1.31)	1.20 (1.03, 1.41)	1.00 (0.87, 1.14)

CI – confidence interval; Freq. – frequency; Quest. – questionnaire; RR – relative risk

SUPPLEMENTAL MATERIAL

Supplemental Material 1. Relative risk of responding to each questionnaire according to questionnaire length and frequency of reminders. Pelotas, Brazil, 2018.

	Q1	Q2	Q3	Q4	Q5
	RR	RR	RR	RR	RR
	(95%CI)	(95%CI)	(95%CI)	(95%CI)	(95%CI)
Questionnaire length					
Long	1.00	1.00	1.00	1.00	1.00
Short	1.01 (0.94, 1.08)	1.03 (0.96, 1.11)	1.05 (0.95, 1.16)	1.05 (0.94, 1.18)	1.10 (0.93, 1.29)
Frequency of reminders					
Low	1.00	1.00	1.00	1.00	1.00
High	1.09 (1.02, 1.17)	1.11 (1.03, 1.20)	1.10 (1.00, 1.21)	1.08 (0.97, 1.21)	0.99 (0.83, 1.16)

CI – confidence interval; Q1-Q5 – questionnaire 1 to 5; RR – relative risk

7 – Nota de imprensa

Coorte 1993 cria plataforma na internet para coletar dados e testar nova metodologia de pesquisa

A Coorte de Nascimentos de Pelotas de 1993 vem acompanhando seus participantes desde o nascimento através de entrevistas presenciais. As primeiras foram realizadas nos hospitais e nas casas dos participantes, enquanto as mais recentes foram realizadas no Centro de Pesquisas Amílcar Gigante. Em 2018, um estudo coordenado pelos pesquisadores Cauane Blumenberg e Aluísio J. D. Barros, do Programa de Pós-graduação em Epidemiologia da Universidade Federal de Pelotas, testou a possibilidade de conduzir pesquisas utilizando a internet.

Para isso, foi desenvolvida a plataforma *coortesnaweb*, que teve como principal objetivo coletar dados de pesquisa dos participantes da Coorte 1993 através da internet, e testar diferentes metodologias de pesquisas utilizando a internet. Ao se registrar na *coortesnaweb*, os participantes respondem a questionários e conquistam pontos virtuais. Esses pontos podem ser utilizados para desbloquear novas funções dentro da plataforma, como por exemplo os resultados pessoais sobre a saúde. Além disso, os participantes podem obter conquistas e medalhas para acumular ainda mais pontos virtuais.

Algumas das metodologias testadas pelos coordenadores do estudo foram a maneira como os participantes eram convidados a participar do estudo, a frequência de envio de lembretes e o tamanho dos questionários. Os achados do estudo mostram que o método de recrutamento mais eficiente foi o Facebook, o qual foi capaz de atrair mais participantes para a plataforma quando comparado a e-mails e mensagens de Whatsapp. Além disso, a participação daqueles que receberam lembretes mais frequentes foi maior. Por outro lado, a participação dos indivíduos que responderam questionários longos ou curtos não teve diferença.

“Nosso projeto foi um dos primeiros a explorar a coleta de dados através da internet no Brasil, e mostramos que isso é possível. Esperamos que nosso estudo possa servir de base para que outros grupos de pesquisa adotem a internet para conduzir pesquisas epidemiológicas.”, comentou o autor.

8 – Anexos

Anexo 1. Termo de consentimento livre e esclarecido exibido no momento do registro do participante na plataforma *coortesnaweb*.

Caro(a) participante,

O Centro de Pesquisas em Saúde, da Universidade Federal de Pelotas (UFPEL), vem acompanhando de forma presencial os nascidos em 1993, na cidade de Pelotas. Agora, queremos acompanhar você também pela internet. Por isso estamos criando essa plataforma, chamada *coortesnaweb*, onde vamos enviar para você questionários que vão avaliar as suas condições de saúde e outros aspectos como, trabalho, renda, composição familiar e escolaridade.

Para que possamos avaliar corretamente esses aspectos, alguns questionários serão necessários - semelhante ao que fizemos nos acompanhamentos presenciais, mas agora com questionários muito menores e que podem ser respondidos em qualquer hora ou lugar. Você mesmo vai decidir como responder aos questionários, afinal qualquer tipo de dispositivo conectado à internet poderá ser usado. Alguns questionários podem conter perguntas mais íntimas, mas nenhum deles terá o seu nome – apenas usaremos um código. Além disso, todos os servidores onde os dados serão gravados contêm um elaborado sistema de segurança.

Participar desse estudo poderá trazer a você benefícios diretos ou indiretos, pois as informações coletadas serão usadas para ajudar na prevenção de doenças comuns como diabetes, doenças do coração, hipertensão, entre outras. Além disso, compreender como está a saúde de nossos jovens faz com que o sistema de saúde possa planejar ações para ajudá-los, como por exemplo, campanhas com informações de prevenção. Ao responder os questionários, você poderá ter acesso a resultados personalizados sobre a sua saúde, e ainda poderá compará-los com a população da Coorte 1993.

É importante esclarecer que todas as suas informações são confidenciais, guardadas em segurança, identificadas por um código numérico e serão utilizadas exclusivamente para esta pesquisa. Só terão acesso a elas os pesquisadores do estudo, mas sem sua identificação. A identificação (nome e telefone) só é utilizada quando vamos convidá-lo(a) a participar do estudo e está separada das suas respostas.

A sua participação deve ser inteiramente voluntária e você poderá se recusar ou deixar de fazer parte da pesquisa a qualquer momento, sem qualquer problema, prejuízo ou discriminação no futuro. Não há nenhum gasto/despesas, nem qualquer outra obrigação para participar deste estudo. Apenas pedimos que você nos responda às perguntas com muita sinceridade. Se você está de acordo com o conteúdo que acabou de ler, por favor marque o campo abaixo.

Estou ciente dos detalhes do projeto e aceito participar.

Anexo 2. Parecer emitido pelo Comitê de Ética em Pesquisa da Faculdade de Medicina da Universidade Federal de Pelotas.

UFPEL - FACULDADE DE
MEDICINA DA UNIVERSIDADE
FEDERAL DE PELOTAS



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: PESQUISA EPIDEMIOLOGICA BASEADA NA WEB: ESTRATEGIAS E CARACTERISTICAS PESSOAIS ASSOCIADAS A ADESAO E PARTICIPACAO DE ADULTOS JOVENS

Pesquisador: Aluisio Jardim Dornellas de Barros

Área Temática:

Versão: 1

CAAE: 79124917.5.0000.5317

Instituição Proponente: Faculdade de Medicina da Universidade Federal de Pelotas

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.382.790

Apresentação do Projeto:

As pesquisas epidemiológicas sempre contribuíram para importantes achados na área de saúde. Para que estes achados possam ser traduzidos em políticas públicas, é importante que as pesquisas epidemiológicas sejam conduzidas com rigor metodológico em vista de maximizar sua validade. Um fator que pode comprometer a validade desses achados é o constante declínio nas taxas de resposta, influenciado por diversos fatores, entre

eles o método de coleta de dados. Com o aumento do acesso à internet, bem como do seu número de usuários, conduzir pesquisas epidemiológicas via web pode ser uma boa alternativa. Apesar disso, novos desafios surgem com essa nova metodologia de pesquisa, como o melhor método de recrutamento a ser utilizado, questões de representatividade das amostras, desenho dos questionários e maneiras eficientes de motivar a participação dos indivíduos nas pesquisas aplicadas via web.

Objetivo da Pesquisa:

Objetivo Primário:

Investigar a viabilidade de conduzir pesquisas epidemiológicas via web, explorando estratégias de metodologia de pesquisa que visam maximizar as taxas de contato, resposta e retenção. Além disso, o projeto irá estudar os fatores sócio-demográficos associados à participação de adultos jovens membros da coorte de nascimentos de 1993 de Pelotas nos questionários aplicados através

Endereço: Av Duque de Caxias 250

Bairro: Fragata

CEP: 96.030-001

UF: RS

Município: PELOTAS

Telefone: (53)3284-4960

Fax: (53)3221-3554

E-mail: cep.famed@gmail.com

Continuação do Parecer: 2.382.790

da plataforma web.

Objetivo Secundário:

- Desenvolver uma plataforma web com o ambiente completo para a condução de pesquisas epidemiológicas, compreendendo recrutamento de participantes, controle de acesso, coleta de dados epidemiológicos via web, retorno de resultados, convite para participação em pesquisas e fornecimento de recompensas.
- Avaliar a representatividade dos indivíduos elegíveis com relação à população original da coorte de nascimentos de 1993 de Pelotas, e dos indivíduos cadastrados na plataforma web com relação aos elegíveis.
- Avaliar a influência de diferentes fatores demográficos e socioeconômicos nas taxas de contato, resposta e retenção dos questionários aplicados através da plataforma web.
- Aplicar diferentes estratégias de recrutamento para convidar os indivíduos a se cadastrar na plataforma web, como correspondência, e-mail e redes sociais. O objetivo é estimar a influência desses diferentes métodos na taxa de contato.
- Comparar diferentes tamanhos de questionários (questionários longos vs. curtos) e frequência de lembretes (frequência alta vs. baixa) em relação à taxa de resposta dos questionários aplicados via web.

Avaliação dos Riscos e Benefícios:

Riscos:

o estudo apresenta riscos mínimos quando da aplicação de questionário, pois pode causar algum constrangimento.

Benefícios:

Os benefícios aos participantes podem ser diretos ou indiretos, afinal as informações coletadas através dos questionários aplicados são usadas para ajudar na prevenção de doenças comuns como diabetes, doenças do coração, hipertensão, entre outras. Além disso, compreender como está a saúde dos jovens pode auxiliar os serviços de saúde na elaboração de ações de prevenção para a população. Somado a isso, os participantes após responderem os questionários poderão ter acesso a resultados personalizados sobre a sua saúde, e assim poderão comparar seus resultados com os resultados agregados da população da coorte 1993.

Comentários e Considerações sobre a Pesquisa:

Pesquisa relevante e adequada para o conhecimento de estratégias de investigação de problemas de saúde através da web. Trata-se de um projeto de pesquisa de doutorado do PPGE da UFPel. Serão convidados a participar da plataforma os membros da coorte de nascimentos de 1993 de

Endereço: Av Duque de Caxias 250
Bairro: Fragata **CEP:** 96.030-001
UF: RS **Município:** PELOTAS
Telefone: (53)3284-4960 **Fax:** (53)3221-3554 **E-mail:** cep.famed@gmail.com

Continuação do Parecer: 2.382.790

Pelotas que participaram do acompanhamento dos 22/23 anos, e relataram ter acesso à internet. Para investigar a influência de diferentes metodologias sobre as taxas de contato e resposta, serão estudados diferentes métodos de recrutamento, tamanhos de questionário, frequências de lembrete e fornecimento de incentivos. Somado a isso, serão estudados fatores sociodemográficos associados à participação dos membros da coorte na plataforma web a ser desenvolvida.

Considerações sobre os Termos de apresentação obrigatória:

Projeto Adequado.

Recomendações:

OK

Conclusões ou Pendências e Lista de Inadequações:

OK

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Outros	PROJETO_OK.docx	15/11/2017 18:51:49	Patricia Abrantes Duval	Aceito
Outros	PROJETO_OK.docx	15/11/2017 18:51:49	Patricia Abrantes Duval	Aceito
Outros	FR.pdf	11/11/2017 09:14:28	Patricia Abrantes Duval	Aceito
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1007060.pdf	03/10/2017 12:14:39		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	tcle_coortesnaweb.pdf	03/10/2017 12:10:51	Aluisio Jardim Dornellas de Barros	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoPhDparaCEP.docx	03/10/2017 12:10:06	Aluisio Jardim Dornellas de Barros	Aceito
Folha de Rosto	folhaDeRosto.pdf	03/10/2017 12:07:40	Aluisio Jardim Dornellas de Barros	Aceito

Situação do Parecer:

Aprovado

Endereço: Av Duque de Caxias 250
Bairro: Fragata **CEP:** 96.030-001
UF: RS **Município:** PELOTAS
Telefone: (53)3284-4960 **Fax:** (53)3221-3554 **E-mail:** cep.famed@gmail.com

UFPEL - FACULDADE DE
MEDICINA DA UNIVERSIDADE
FEDERAL DE PELOTAS



Continuação do Parecer: 2.382.790

Necessita Apreciação da CONEP:
Não

PELOTAS, 15 de Novembro de 2017

Assinado por:
Patricia Abrantes Duval
(Coordenador)

Endereço: Av Duque de Caxias 250
Bairro: Fragata **CEP:** 96.030-001
UF: RS **Município:** PELOTAS
Telefone: (53)3284-4960 **Fax:** (53)3221-3554 **E-mail:** cep.famed@gmail.com

Anexo 3. Questionário sobre medidas corporais aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 1

Medidas corporais

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 1 minuto para completá-la, e isso vai render 6 pontos!!

Este questionário é sobre suas medidas corporais. As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

1) Quantos quilos você pesa?

_____ (Digite o peso em quilos, usando ponto para separar os números. Por exemplo: 73.8)

2) Qual a sua altura?

_____ (Digite a altura em metros, usando ponto para separar os números. Por exemplo: 1.67)

3) IMC @HIDDEN

Anexo 4. Questionário sobre consumo de álcool (versão curta) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 2

Consumo de álcool

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 7 minutos para completá-la, e isso vai render 56 pontos!!

Este questionário é sobre consumo de álcool. As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Você bebe alguma bebida de álcool?

- Sim
- Não

Você se considera uma pessoa que bebe normalmente?	<input type="radio"/> Sim <input type="radio"/> Não (Beber normalmente seria beber menos ou a mesma quantidade do que as outras pessoas.)
Seu(ua) companheiro(a), seus pais ou parentes se preocupam ou reclamam sobre o quanto você bebe?	<input type="radio"/> Sim <input type="radio"/> Não
Você já se sentiu culpado por ter bebido?	<input type="radio"/> Sim <input type="radio"/> Não
Seus amigos ou parentes acham que você é uma pessoa que bebe normalmente?	<input type="radio"/> Sim <input type="radio"/> Não (Beber normalmente seria beber menos ou a mesma quantidade do que as outras pessoas.)
Você consegue parar de beber na hora que quer?	<input type="radio"/> Sim <input type="radio"/> Não
Você já participou de uma reunião dos Alcoólicos Anônimos (AA)?	<input type="radio"/> Sim <input type="radio"/> Não
A bebida já criou problemas com seu(ua) companheiro(a), pais ou parentes próximos?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você teve problemas no trabalho por causa da bebida?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você deixou de cumprir com suas obrigações no trabalho, ou com a família, por dois ou mais dias seguidos porque você estava bebendo?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você pediu ajuda para alguém com relação à bebida?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você foi parar no hospital por causa da bebida?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você foi levado para a delegacia, mesmo que por poucas horas, por estar bêbado(a)?	<input type="radio"/> Sim <input type="radio"/> Não
Alguma vez você foi preso por dirigir alcoolizado ou dirigir após ter bebido?	<input type="radio"/> Sim <input type="radio"/> Não
Total: @HIDDEN (0-1 pontos: não alcoolismo/2 pontos: risco de alcoolismo/3+ pontos: alcoolismo)	_____
Clique em "Enviar" para finalizar a pesquisa	

Anexo 5. Questionário sobre consumo de álcool (versão longa) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 3

Consumo de álcool

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 13 minutos para completá-la, e isso vai render 75 pontos!!

Este questionário é sobre consumo de álcool. As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Você bebe alguma bebida de álcool?

- Sim
- Não

-
-
- Você se considera uma pessoa que bebe normalmente? Sim
 Não
(Beber normalmente seria beber menos ou a mesma quantidade do que as outras pessoas.)
- Alguma vez você acordou de manhã, depois de ter bebido na véspera, e se deu conta de que não se lembrava de algumas coisas que aconteceram na noite anterior? Sim
 Não
- Seu(ua) companheiro(a), seus pais ou parentes se preocupam ou reclamam sobre o quanto você bebe? Sim
 Não
- Você consegue parar de beber, sem dificuldade, após um ou dois copos? Sim
 Não
- Você já se sentiu culpado por ter bebido? Sim
 Não
- Seus amigos ou parentes acham que você é uma pessoa que bebe normalmente? Sim
 Não
(Beber normalmente seria beber menos ou a mesma quantidade do que as outras pessoas.)
- Você consegue parar de beber na hora que quer? Sim
 Não
- Você já participou de uma reunião dos Alcoólicos Anônimos (AA)? Sim
 Não
- Você já se envolveu em uma briga enquanto estava bebendo? Sim
 Não
- A bebida já criou problemas com seu(ua) companheiro(a), pais ou parentes próximos? Sim
 Não
- Seu(ua) companheiro(a), ou outro membro da família, já procurou alguém para ajudar com relação ao seu hábito de beber? Sim
 Não
- Alguma vez você perdeu um amigo ou o(a) namorado(a) por causa da bebida? Sim
 Não
- Alguma vez você teve problemas no trabalho por causa da bebida? Sim
 Não
- Alguma vez você perdeu o emprego por causa da bebida? Sim
 Não
- Alguma vez você deixou de cumprir com suas obrigações no trabalho, ou com a família, por dois ou mais dias seguidos porque você estava bebendo? Sim
 Não
- Você costuma beber antes do almoço? Sim
 Não
- Alguma vez lhe disseram que você tem problemas no fígado ou cirrose? Sim
 Não

Alguma vez você delirou, tremeu forte, ouviu vozes ou viu coisas que não existiam após ter bebido muito?

- Sim
 Não

Alguma vez você pediu ajuda para alguém com relação à bebida?

- Sim
 Não

Alguma vez você foi parar no hospital por causa da bebida?

- Sim
 Não

Alguma vez você foi internado em um hospital psiquiátrico, ou em uma ala psiquiátrica de um hospital, pelo menos em parte por conta da bebida?

- Sim
 Não

Alguma vez você foi consultar em uma clínica psiquiátrica, médico, assistente social, padre ou pastor para buscar ajuda para um problema emocional em que a bebida era parte do problema?

- Sim
 Não

Alguma vez você foi levado para a delegacia, mesmo que por poucas horas, por estar bêbado(a)?

- Sim
 Não

Alguma vez você foi preso por dirigir alcoolizado ou dirigir após ter bebido?

- Sim
 Não

Total: @HIDDEN (0-4 pontos: não alcoolismo/5-6 pontos: risco de alcoolismo/7+ pontos: alcoolismo)

Clique em "Enviar" para finalizar a pesquisa

Anexo 6. Questionário sobre atividade física (versão curta) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 3

Atividade física - IPAQ

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 6 minutos para completá-la, e isso vai render 36 pontos!!

Nós estamos interessados em saber que tipos de atividade física as pessoas fazem como parte do seu dia-a-dia. As perguntas estão relacionadas ao tempo que você gasta fazendo atividade física em uma semana normal. As perguntas incluem as atividades que você faz no trabalho, para ir de um lugar a outro, por lazer, por esporte, por exercício ou como parte das suas atividades em casa ou no jardim. Por favor responda cada questão mesmo que considere que não seja ativo.

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Em quantos dias de uma semana normal, você realiza atividades vigorosas por pelo menos 10 minutos contínuos, como por exemplo correr, fazer ginástica aeróbica, jogar futebol, pedalar rápido na bicicleta, jogar basquete, fazer serviços domésticos pesados em casa, no quintal ou no jardim, carregar pesos elevados ou qualquer atividade que faça você suar bastante ou aumentem muito sua respiração ou batimentos do coração.

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades vigorosas são aquelas que precisam de um grande esforço físico e fazem respirar muito mais forte que o normal.)

Nos dias em que você faz essas atividades vigorosas por pelo menos 10 minutos contínuos, quanto tempo no total você gasta fazendo essas atividades por dia?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal, você realiza atividades moderadas por pelo menos 10 minutos contínuos, como por exemplo pedalar leve na bicicleta, nadar, dançar, fazer ginástica aeróbica leve, jogar vôlei recreativo, carregar pesos leves, fazer serviços domésticos na casa, no quintal ou no jardim como varrer, aspirar, cuidar do jardim, ou qualquer atividade que faça você suar leve ou aumentem moderadamente sua respiração ou batimentos do coração. Por favor, não inclua caminhada.

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades moderadas são aquelas que precisam de algum esforço físico e fazem respirar um pouco mais forte que o normal.)

Nos dias em que você faz essas atividades moderadas por pelo menos 10 minutos contínuos, quanto tempo no total você gasta fazendo essas atividades por dia?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você caminha por pelo menos 10 minutos contínuos em casa ou no trabalho, como forma de transporte para ir de um lugar para outro, por lazer, por prazer ou como forma de exercício?

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Nos dias em que você caminha por pelo menos 10 minutos contínuos quanto tempo no total você gasta caminhando por dia?

Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Estas últimas questões são sobre o tempo que você permanece sentado todo dia, no trabalho, na escola ou faculdade, em casa e durante seu tempo livre. Isto inclui o tempo sentado estudando, sentado enquanto descansa, fazendo lição de casa, lendo ou deitado assistindo TV. Não inclua o tempo gasto sentado durante o transporte em ônibus, trem, metrô ou carro.

Quanto tempo no total você gasta sentado durante um dia de semana?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Quanto tempo no total você gasta sentado durante um dia de final de semana?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Anexo 7. Questionário sobre atividade física (versão longa) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 9

Atividade física - IPAQ

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 13 minutos para completá-la, e isso vai render 75 pontos!!

Nós estamos interessados em saber que tipos de atividade física as pessoas fazem como parte do seu dia-a-dia. As perguntas estão relacionadas ao tempo que você gasta fazendo atividade física em uma semana normal. As perguntas incluem as atividades que você faz no trabalho, para ir de um lugar a outro, por lazer, por esporte, por exercício ou como parte das suas atividades em casa ou no jardim. Por favor responda cada questão mesmo que considere que não seja ativo.

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Vamos falar sobre atividades que você faz no seu serviço, que incluem trabalho remunerado ou voluntário, as atividades na escola ou faculdade e outro tipo de trabalho não remunerado fora da sua casa. Não incluir trabalho não remunerado que você faz na sua casa como tarefas domésticas, cuidar do jardim e da casa ou tomar conta da sua família.

Atualmente você trabalha ou faz trabalho voluntário fora de sua casa?

- Sim
 Não

As próximas questões são com relação a toda a atividade física que você faz em uma semana normal como parte do seu trabalho remunerado ou não remunerado. Não inclua o transporte para o trabalho. Pense unicamente nas atividades que você faz por pelo menos 10 minutos contínuos:

Em quantos dias de uma semana normal você faz atividades vigorosas por pelo menos 10 minutos contínuos, como trabalho de construção pesada, carregar grandes pesos, trabalhar com enxada, escavar ou subir escadas como parte do seu trabalho:

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades vigorosas são aquelas que precisam de um grande esforço físico e fazem respirar muito mais forte que o normal.)

Quanto tempo no total você usualmente gasta por dia fazendo atividades físicas vigorosas como parte do seu trabalho?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você faz atividades moderadas por pelo menos 10 minutos contínuos, como carregar pesos leves como parte do seu trabalho:

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades moderadas são aquelas que precisam de algum esforço físico e fazem respirar um pouco mais forte que o normal.)

Quanto tempo no total você usualmente gasta por dia fazendo atividades físicas moderadas como parte do seu trabalho?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você anda, durante pelo menos 10 minutos contínuos, como parte do seu trabalho? Por favor não inclua o andar como forma de transporte para ir ou voltar do trabalho.

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Quanto tempo no total você usualmente gasta por dia caminhando como parte do seu trabalho?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Estas questões se referem a forma típica como você se desloca de um lugar para outro, incluindo seu trabalho, escola, cinema, lojas e outros.

Em quantos dias de uma semana normal você anda de carro, ônibus, metrô ou trem?

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Quanta tempo no total você usualmente gasta por dia andando de carro, ônibus, metrô ou trem?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você anda de bicicleta por pelo menos 10 minutos contínuos para ir de um lugar para outro? Por favor, não inclua o pedalar por lazer ou exercício.

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Nos dias que você pedala, quanto tempo no total você pedala por dia para ir de um lugar para outro?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você caminha por pelo menos 10 minutos contínuos para ir de um lugar para outro? Por favor, não inclua caminhadas por lazer ou exercício.

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Quando você caminha para ir de um lugar para outro, quanto tempo por dia você gasta? Por favor, não inclua as caminhadas por lazer ou exercício.
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Vamos falar sobre atividades físicas que você faz em uma semana normal na sua casa e ao redor da sua casa, por exemplo trabalho em casa, cuidar do jardim, cuidar do quintal, trabalho de manutenção da casa ou para cuidar da sua família. Novamente, pense somente naquelas atividades físicas que voce faz por pelo menos 10 minutos contínuos.

Em quantos dias de uma semana normal você faz atividades físicas vigorosas no jardim ou quintal por pelo menos 10 minutos como capinar, lavar o quintal, esfregar o chão:

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

(Atividades vigorosas são aquelas que precisam de um grande esforço físico e fazem respirar muito mais forte que o normal.)

Nos dias que você faz este tipo de atividades vigorosas no quintal ou jardim quanto tempo no total voce gasta por dia?

Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você faz atividades físicas moderadas no jardim ou quintal por pelo menos 10 minutos como carregar pesos leves, limpar vidros, varrer:

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

(Atividades moderadas são aquelas que precisam de algum esforço físico e fazem respirar um pouco mais forte que o normal.)

Nos dias que você faz este tipo de atividades quanto tempo no total voce gasta por dia fazendo essas atividades moderadas no jardim ou no quintal?

Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal você faz atividades moderadas por pelo menos 10 minutos como carregar pesos leves, limpar vidros, varrer ou limpar o chão dentro da sua casa.

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

(Atividades moderadas são aquelas que precisam de algum esforço físico e fazem respirar um pouco mais forte que o normal.)

Nos dias que você faz este tipo de atividades moderadas dentro da sua casa quanto tempo no total você gasta por dia?

Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Agora vamos falar sobre as atividades físicas que você faz em uma semana normal unicamente por recreação, esporte, exercício ou lazer. Novamente pense somente nas atividades físicas que faz por pelo menos 10 minutos contínuos. Por favor não inclua atividades que você já tenha citado.

Sem contar qualquer caminhada que você tenha citado anteriormente, em quantos dias de uma semana normal, você caminha por pelo menos 10 minutos contínuos no seu tempo livre?

- 1 dia por semana
- 2 dias por semana
- 3 dias por semana
- 4 dias por semana
- 5 dias por semana
- 6 dias por semana
- 7 dias por semana
- Nenhum dia na semana

Nos dias em que você caminha no seu tempo livre, quanto tempo no total você gasta por dia?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal, você faz atividades vigorosas no seu tempo livre por pelo menos 10 minutos, como correr, fazer aeróbicos, nadar rápido, pedalar rápido ou fazer jogging:

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades vigorosas são aquelas que precisam de um grande esforço físico e fazem respirar muito mais forte que o normal.)

Nos dias em que você faz estas atividades vigorosas no seu tempo livre, quanto tempo no total você gasta por dia?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Em quantos dias de uma semana normal, você faz atividades moderadas no seu tempo livre por pelo menos 10 minutos, como pedalar ou nadar a velocidade regular, jogar bola, vôlei, basquete, tênis:

- 1 dia por semana
 - 2 dias por semana
 - 3 dias por semana
 - 4 dias por semana
 - 5 dias por semana
 - 6 dias por semana
 - 7 dias por semana
 - Nenhum dia na semana
- (Atividades moderadas são aquelas que precisam de algum esforço físico e fazem respirar um pouco mais forte que o normal.)

Nos dias em que você faz estas atividades moderadas no seu tempo livre, quanto tempo no total você gasta por dia?
Quantidade de horas:

_____ (Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Estas últimas questões são sobre o tempo que você permanece sentado todo dia, no trabalho, na escola ou faculdade, em casa e durante seu tempo livre. Isto inclui o tempo sentado estudando, sentado enquanto descansa, fazendo lição de casa, lendo ou deitado assistindo TV. Não inclua o tempo gasto sentado durante o transporte em ônibus, trem, metrô ou carro.

Quanto tempo no total você gasta sentado durante um dia de semana?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Quanto tempo no total você gasta sentado durante um dia de final de semana?
Quantidade de horas:

(Se menos de 1 hora, preencher com 0.)

Quantidade de minutos:

Anexo 8. Questionário sobre uso abusivo da internet (versão curta) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 2

Uso da Internet

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 6 minutos para completá-la, e isso vai render 36 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Essas perguntas vão avaliar os seus hábitos em relação ao uso da Internet. Por isso, ao responder as perguntas você deve considerar somente o tempo que você passa online por lazer ou diversão, e não o tempo gasto na Internet para o trabalho ou estudo.

Com que frequência você fica online mais tempo do que pretendia?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você deixa de fazer as tarefas em casa para poder ficar mais tempo online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência as suas notas ou trabalhos da faculdade ou escola são prejudicados devido à quantidade de tempo que você passa online?

- Nunca
 - Raramente
 - De vez em quando
 - Várias vezes
 - Sempre
 - Não se aplica
- (Se você não estuda, marque "Não se aplica".)

Com que frequência você se torna defensivo(a) ou guarda segredo quando alguém pergunta o que você faz online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você grita ou fica irritado(a) quando alguém incomoda quando você está online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você perde o sono por estar online até tarde durante a noite?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você se sente preocupado(a) com a Internet quando está desconectado(a) ou fica imaginando estar online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você diz "só mais alguns minutos" quando está online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você tenta reduzir a quantidade de tempo que você passa online e não consegue?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você tenta esconder a quantidade de tempo que você passa online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você prefere ficar mais tempo online do que sair com outras pessoas?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você se sente deprimido(a), mal-humorado(a) ou nervoso(a) quando está desconectado(a) e, deixa de estar assim quando está online novamente?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Anexo 9. Questionário sobre uso abusivo da internet (versão longa) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 3

Uso da Internet

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 10 minutos para completá-la, e isso vai render 60 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Essas perguntas vão avaliar os seus hábitos em relação ao uso da Internet. Por isso, ao responder as perguntas você deve considerar somente o tempo que você passa online por lazer ou diversão, e não o tempo gasto na Internet para o trabalho ou estudo.

Com que frequência você fica online mais tempo do que pretendia?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você deixa de fazer as tarefas em casa para poder ficar mais tempo online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você prefere a Internet do que a intimidade com o teu(tua) namorado(a) ou ficante?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você cria novas relações com outras pessoas online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência as outras pessoas se queixam em relação à quantidade de tempo que você passa online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência as suas notas ou trabalhos da faculdade ou escola são prejudicados devido à quantidade de tempo que você passa online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre
- Não se aplica
(Se você não estuda, marque "Não se aplica".)

Com que frequência você verifica o e-mail (ou sites como Facebook, Whatsapp, Instagram) antes de fazer qualquer outra coisa que precisa?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência o seu desempenho ou produtividade no trabalho são prejudicados por causa da Internet?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre
 Não se aplica
(Se você não trabalha marque "Não se aplica".)

Com que frequência você se torna defensivo(a) ou guarda segredo quando alguém pergunta o que você faz online?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você bloqueia pensamentos perturbadores sobre a sua vida com pensamentos calmantes da Internet?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você pensa sobre quando vai usar a Internet novamente?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você pensa que a vida sem Internet seria chata, vazia e sem graça?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você grita ou fica irritado(a) quando alguém incomoda quando você está online?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você perde o sono por estar online até tarde durante a noite?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você se sente preocupado(a) com a Internet quando está desconectado(a) ou fica imaginando estar online?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você diz "só mais alguns minutos" quando está online?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você tenta reduzir a quantidade de tempo que você passa online e não consegue?

- Nunca
 Raramente
 De vez em quando
 Várias vezes
 Sempre

Com que frequência você tenta esconder a quantidade de tempo que você passa online?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você prefere ficar mais tempo online do que sair com outras pessoas?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Com que frequência você se sente deprimido(a), mal-humorado(a) ou nervoso(a) quando está desconectado(a) e, deixa de estar assim quando está online novamente?

- Nunca
- Raramente
- De vez em quando
- Várias vezes
- Sempre

Anexo 10. Questionário sobre perpetração de crimes e agressão (versão curta) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 2

Perpetração de crimes

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 8 minutos para completá-la, e isso vai render 51 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

As próximas perguntas são sobre brigas e outros comportamentos. Você deve preencher o número (0, 1, 2, 3...) de vezes, se não aconteceu você deve preencher com 0.

Nos últimos doze meses, quantas vezes você roubou lojas, supermercados, vendas? _____

Nos últimos doze meses, quantas vezes você estragou ou destruiu propriedades dos outros intencionalmente (por exemplo: janelas, carros, iluminação pública)? _____

Nos últimos doze meses, quantas vezes você arrombou carros para tentar roubar objetos? _____

Nos últimos doze meses, quantas vezes você roubou carros ou motos? _____

Nos últimos doze meses, quantas vezes você vendeu drogas ilegais a alguém? _____

Nos últimos doze meses, quantas vezes você arrombou casas ou prédios para roubar? _____

Nos últimos doze meses, quantas vezes você bateu em outras pessoas com a intenção de machucá-las? _____
(Não inclui irmãos, irmãs, nem brincadeiras de luta e chutes em jogos.)

Nos últimos doze meses, na briga mais séria que você teve, quais machucados você causou?

- Roxoes ou olho roxo
- Arranhos ou pequenos cortes
- Cortes profundos ou graves
- Machucados na cabeça ou na face
- Osso quebrado
- Outros machucados
- Nenhum machucado
- Não briguei ano passado

(Se for o caso, marque mais de uma resposta.)

Nos últimos doze meses, quantas vezes você vendeu objetos que pertenciam a outras pessoas ou que você sabia que eram roubados? _____

Nos últimos doze meses, quantas vezes você roubou dinheiro ou objetos que alguém estava carregando ou usando? _____

Neste(s) roubo(s) de dinheiro ou outros objetos, você fez ameaças ou usou força e violência contra outra pessoa?

Não
 Sim

Nos últimos doze meses, quantas vezes você colocou fogo ou tentou incendiar objetos intencionalmente (por exemplo: uma escola, uma parada de ônibus, uma casa, etc.)?

Nos últimos doze meses, quantas vezes você carregou uma faca ou outra arma para se proteger ou brigar?

Nos últimos doze meses, quantas vezes você usou alguma arma contra outra pessoa?

Nos últimos doze meses, quantas vezes você fez sexo com alguém à força, contra a vontade da pessoa?

Nos últimos doze meses, você participou de alguma quadrilha, facção ou gangue?

- Nao
- Sim

Alguma vez na vida você foi detido(a) ou preso(a)?

- Nao
- Sim

Anexo 11. Questionário sobre perpetração de crimes e agressão (versão longa) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 4

Perpetração de crimes e agressão

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 10 minutos para completá-la, e isso vai render 90 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

As próximas perguntas são sobre agressão contra ti, feita por alguém da sua família. Você deve preencher o número (0, 1, 2, 3...) de vezes, se não aconteceu você deve preencher com 0.

Nos últimos doze meses, quantas vezes alguém da tua família ameaçou seriamente te machucar fisicamente?

Nos últimos doze meses, quantas vezes alguém da tua família bateu em ti, te empurrou, te chutou ou te agrediu fisicamente sem arma?

Nos últimos doze meses, quantas vezes alguém da tua família te atacou com uma faca, arma de fogo ou outra arma?

Nos últimos doze meses, quantas vezes alguém da tua família roubou um objeto teu, com uso de violência ou de ameaça?

Nos últimos doze meses, quantas vezes alguém da tua família te agarrou, te tocou ou agrediu as tuas partes sexuais íntimas contra a tua vontade?

As próximas perguntas são sobre agressão contra ti, feita por alguém que não é da tua família

Nos últimos doze meses, quantas vezes alguém que não é da tua família ameaçou seriamente te machucar fisicamente?

Nos últimos doze meses, quantas vezes alguém que não é da tua família te bateu, te empurrou, te chutou ou te agrediu fisicamente sem arma?

Nos últimos doze meses, quantas vezes alguém que não é da tua família te atacou com uma faca, arma de fogo ou outra arma?

Nos últimos doze meses, quantas vezes alguém que não é da tua família roubou um objeto teu, com o uso de violência ou ameaça?

Nos últimos doze meses, quantas vezes alguém que não é da tua família te agarrou, te tocou ou agrediu as tuas partes sexuais íntimas contra a tua vontade?

As próximas perguntas são sobre brigas e outros comportamentos. Você deve preencher o número (0, 1, 2, 3...) de vezes, se não aconteceu você deve preencher com 0.

Nos últimos doze meses, quantas vezes você roubou lojas, supermercados, vendas? _____

Nos últimos doze meses, quantas vezes você estragou ou destruiu propriedades dos outros intencionalmente (por exemplo: janelas, carros, iluminação pública)? _____

Nos últimos doze meses, quantas vezes você arrombou carros para tentar roubar objetos? _____

Nos últimos doze meses, quantas vezes você roubou carros ou motos? _____

Nos últimos doze meses, quantas vezes você vendeu drogas ilegais a alguém? _____

Nos últimos doze meses, quantas vezes você arrombou casas ou prédios para roubar? _____

Nos últimos doze meses, quantas vezes você bateu em outras pessoas com a intenção de machucá-las? _____
(Não inclui irmãos, irmãs, nem brincadeiras de luta e chutes em jogos.)

Nos últimos doze meses, na briga mais séria que você teve, quais machucados você causou?

Roxoes ou olho roxo
 Arranhões ou pequenos cortes
 Cortes profundos ou graves
 Machucados na cabeça ou na face
 Osso quebrado
 Outros machucados
 Nenhum machucado
 Não briguei ano passado
 (Se for o caso, marque mais de uma resposta.)

Qual(is)? _____
(Descreva brevemente.)

Nos últimos doze meses, quantas vezes você vendeu objetos que pertenciam a outras pessoas ou que você sabia que eram roubados? _____

Nos últimos doze meses, quantas vezes você roubou dinheiro ou objetos que alguém estava carregando ou usando? _____

Neste(s) roubo(s) de dinheiro ou outros objetos, você fez ameaças ou usou força e violência contra outra pessoa?

Não
 Sim

Nos últimos doze meses, quantas vezes você colocou fogo ou tentou incendiar objetos intencionalmente (por exemplo: uma escola, uma parada de ônibus, uma casa, etc.)? _____

Nos últimos doze meses, quantas vezes você carregou uma faca ou outra arma para se proteger ou brigar? _____

Nos últimos doze meses, quantas vezes você usou alguma arma contra outra pessoa? _____

Qual(is) armas(s)?

(Descreva brevemente.)

Nos últimos doze meses, quantas vezes você fez sexo com alguém à força, contra a vontade da pessoa?

Nos últimos doze meses, você participou de alguma quadrilha, facção ou gangue?

- Não
- Sim

Alguma vez na vida você foi detido(a) ou preso(a)?

- Não
- Sim

Você foi:

- Preso(a)
- Detido(a)
- Preso(a) e detido(a)

Anexo 12. Questionário sobre alimentos com glúten e lactose aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 1

Consumo alimentar

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 3 minutos para completá-la, e isso vai render 36 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

- Pensando na última semana, você bebeu cerveja, comeu pão, biscoito, bolo ou pizza? Sim Não
- Pensando nesses alimentos que você comeu na última semana, eles eram sem glúten? Sim Não
- Você evita beber ou comer alimentos que contenham glúten? Sim Não
- O quanto você evita os alimentos com glúten? Não come ou bebe nada com glúten Quase nunca come ou bebe nada com glúten
- Por que você evita os alimentos com glúten? É saudável Para emagrecer Tem intolerância Tem um familiar intolerante ao glúten
- Você visitou um médico para confirmar que tinha intolerância ao glúten? Sim Não
- Pensando na última semana, você comeu queijo, bebeu leite, iogurte ou outro alimento derivado do leite? Sim Não
- Pensando nesses alimentos que você comeu na última semana, eles eram sem lactose? Sim Não
- Você evita comer ou beber alimentos que contenham lactose? Sim Não
- O quanto você evita os alimentos com lactose? Não come ou bebe nada com lactose Quase nunca come ou bebe nada com lactose
- Por que você evita os alimentos com lactose? É saudável Para emagrecer Tem intolerância Tem um familiar intolerante a lactose
- Você visitou um médico para confirmar que tinha intolerância a lactose? Sim Não

Anexo 13. Questionário sobre qualidade do sono aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 2

Qualidade do sono

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 7 minutos para completá-la, e isso vai render 60 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

As seguintes perguntas são relativas aos seus hábitos usuais de sono durante o último mês somente. Suas respostas devem indicar a lembrança mais exata da maioria dos dias e noites no último mês. Por favor, responda a todas as perguntas.

Durante o último mês, quando você geralmente foi para a cama a noite? _____
(Hora usual de deitar)

Durante o último mês, quanto tempo (em minutos) você geralmente levou para dormir à noite? _____
(Número de minutos)

Durante o último mês, quando você geralmente levantou de manhã? _____
(Hora usual de levantar)

Durante o último mês, quantas horas de sono você teve por noite? (Este pode ser diferente do número de horas que você ficou na cama). _____
(Número de horas)

Para cada uma das questões restantes, marque a melhor resposta. Por favor, responda a todas as questões.

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você não conseguiu adormecer em até 30 minutos?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você acordou no meio da noite ou de manhã cedo?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você precisou levantar para ir ao banheiro?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você não conseguiu respirar confortavelmente?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você tossiu ou roncou forte?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você sentiu muito frio?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você sentiu muito calor?

Nenhuma no último mês
 Menos de 1 vez/semana
 1 ou 2 vezes/semana
 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você teve sonhos ruins?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

Durante o último mês, com que frequência você teve dificuldade de dormir, porque você teve dor?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

Durante o último mês, você teve dificuldade de dormir por outras razão(ões)?

- Não
- Sim

Por favor, descreva a razão:

Durante o último mês, com que frequência você teve dificuldade de dormir, por causa da razão que você descreveu?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

Durante o último mês, como você classificaria a qualidade do seu sono de uma maneira geral?

- Muito boa
- Boa
- Ruim
- Muito ruim

Durante o último mês, com que frequência você tomou medicamento (prescrito ou "por conta própria") para lhe ajudar a dormir?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

No último mês, com que frequência você teve dificuldade de ficar acordado enquanto dirigia, comia ou participava de uma atividade social (festa, reunião de amigos, trabalho, estudo)?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

Durante o último mês, quão problemático foi para você manter o entusiasmo (ânimo) para fazer as coisas (suas atividades habituais)?

- Nenhuma no último mês
- Menos de 1 vez/semana
- 1 ou 2 vezes/semana
- 3 ou mais vezes/semana

Score componente 1: @HIDDEN

Score componente 2 (parte 1): @HIDDEN

Score componente 2 (parte 2): @HIDDEN

Score componente 3: @HIDDEN

Score componente 5: @HIDDEN

Score componente 6: @HIDDEN

Score componente 7: @HIDDEN

Score total: (scores > 5 indicam baixa qualidade de sono) @HIDDEN

PSQI result para coortesnaweb: @HIDDEN

Anexo 14. Questionário sobre hábito de fumar (versão curta) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 3

Fumo

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 3 minutos para completá-la, e isso vai render 33 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Você fuma?

- Não
- Sim

Desejo fumar um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu1b: @HIDDEN

Nada seria melhor do que fumar um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

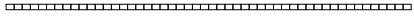
Calculation qsu2b: @HIDDEN

Se fosse possível, eu provavelmente fumaria agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu3b: @HIDDEN

Eu controlaria melhor as coisas, neste momento, se eu pudesse fumar.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu4b: @HIDDEN

Tudo o que eu quero agora é fumar um cigarro.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

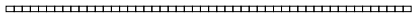
Calculation qsu5b: @HIDDEN

Eu tenho necessidade de um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

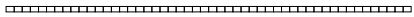
Calculation qsu6b: @HIDDEN

Fumar um cigarro seria gostoso nesse momento.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu7b: @HIDDEN

Eu faria praticamente qualquer coisa por um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

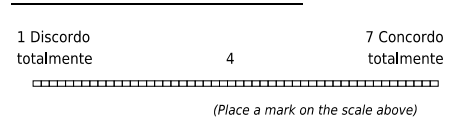
Calculation qsu8b: @HIDDEN

Fumar me faria ficar menos deprimido.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu9b: @HIDDEN

Eu vou fumar assim que for possível.



Calculation qsu10b: @HIDDEN

Anexo 15. Questionário sobre hábito de fumar (versão longa) aplicado através da plataforma *coortesnaweb*.

Confidential

Page 1 of 5

Desejo de fumar

Obrigado por aceitar responder esta pesquisa! Você deve demorar no máximo 8 minutos para completá-la, e isso vai render 99 pontos!!

As respostas são confidenciais, dessa forma você pode responder da maneira mais honesta possível. Vamos começar?

Indique o quanto você concorda ou discorda das afirmações abaixo. Quanto mais perto estiver sua marca de um dos lados, mais você estará concordando ou discordando. Queremos saber o que você está pensando e sentindo agora, enquanto preenche o questionário.

Você fuma?

- Não
 Sim

Fumar faria eu me sentir muito bem agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu1: @HIDDEN

Eu estaria menos irritado agora se pudesse fumar.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu2: @HIDDEN

Nada seria melhor do que fumar um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu3: @HIDDEN

Eu não estou sentindo necessidade de fumar neste momento.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu4: @HIDDEN

Fumarei assim que tiver uma chance.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu5: @HIDDEN

Eu não quero fumar neste momento.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu6: @HIDDEN

Fumar faria eu me sentir menos deprimido.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu7: @HIDDEN

Fumar não ajudaria a me acalmar agora.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu8: @HIDDEN

Se me oferecessem um cigarro agora, eu o fumaria imediatamente.

1 Discordo totalmente	4	7 Concordo totalmente
		
<i>(Place a mark on the scale above)</i>		

Calculation qsu9: @HIDDEN

A partir deste momento, eu poderia ficar sem fumar por um longo período de tempo.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu10: @HIDDEN

Fumar um cigarro não seria agradável.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu11: @HIDDEN

Se eu estivesse fumando agora, me sentiria menos entediado.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu12: @HIDDEN

Tudo o que eu quero, neste momento, é um cigarro

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu13: @HIDDEN

Fumar, neste momento, me deixaria menos cansado.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu14: @HIDDEN

Fumar me faria mais feliz agora.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu15: @HIDDEN

Mesmo que fosse possível, eu provavelmente não fumaria agora.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu16: @HIDDEN

Eu não tenho vontade de fumar um cigarro agora.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

Calculation qsu17: @HIDDEN

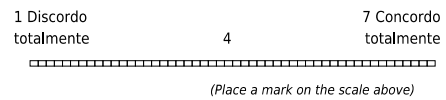
Meu desejo por fumar parece mais forte do que tudo.

1 Discordo totalmente	4	7 Concordo totalmente
-----------------------	---	-----------------------

(Place a mark on the scale above)

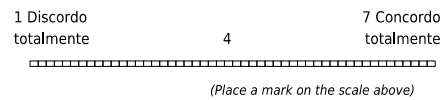
Calculation qsu18: @HIDDEN

Fumar agora faria tudo parecer perfeito.



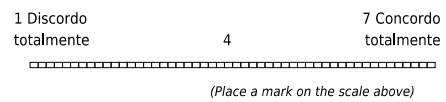
Calculation qsu19: @HIDDEN

Eu desejo intensamente um cigarro agora.



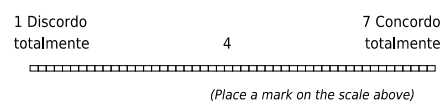
Calculation qsu20: @HIDDEN

Eu não apreciaria um cigarro agora.



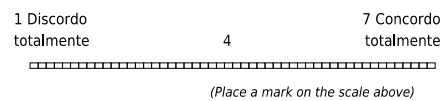
Calculation qsu21: @HIDDEN

Fumar um cigarro não seria gostoso agora.



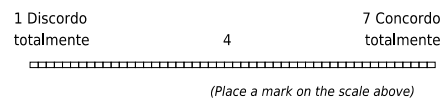
Calculation qsu22: @HIDDEN

Eu desejo um cigarro.



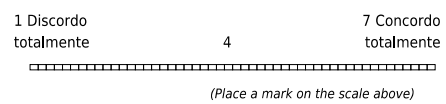
Calculation qsu23: @HIDDEN

Eu poderia controlar as coisas melhor agora, se pudesse fumar um cigarro.



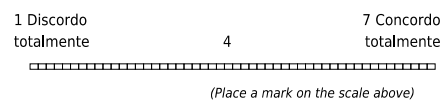
Calculation qsu24: @HIDDEN

Vou fumar assim que for possível.



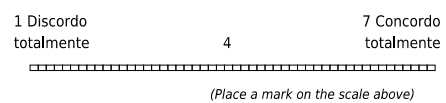
Calculation qsu25: @HIDDEN

Eu não me sentiria melhor fisicamente se estivesse fumando.



Calculation qsu26: @HIDDEN

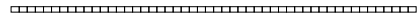
Um cigarro não seria muito satisfatório agora.



Calculation qsu27: @HIDDEN

Se eu tivesse um cigarro aceso em minha mão agora, eu não o fumaria.

1 Discordo totalmente 4 7 Concordo totalmente

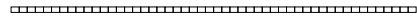


(Place a mark on the scale above)

Calculation qsu28: @HIDDEN

Se eu estivesse fumando agora, eu poderia pensar com mais clareza.

1 Discordo totalmente 4 7 Concordo totalmente

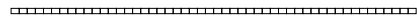


(Place a mark on the scale above)

Calculation qsu29: @HIDDEN

Eu faria praticamente qualquer coisa por um cigarro agora.

1 Discordo totalmente 4 7 Concordo totalmente

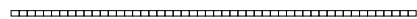


(Place a mark on the scale above)

Calculation qsu30: @HIDDEN

Eu preciso fumar agora.

1 Discordo totalmente 4 7 Concordo totalmente

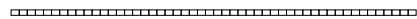


(Place a mark on the scale above)

Calculation qsu31: @HIDDEN

Neste momento, eu não estou planejando fumar.

1 Discordo totalmente 4 7 Concordo totalmente



(Place a mark on the scale above)

Calculation qsu32: @HIDDEN