

Avaliação do Impacto do Programa Criança Feliz

Projeto de pesquisa

12 de abril de 2018

(Atualizado julho de 2019; alterações sinalizadas em itálico)

Cesar Victora

Universidade Federal de Pelotas

Colaboradores: Iná Santos e Tiago Munhoz

Agradecimentos: Fundação Maria Cecília Vidigal, Ministério do Desenvolvimento Social e Agrário, Itaú Social, Programa das Nações Unidas para o Desenvolvimento, Alessandra Schneider, Jennifer Jenkins e Vitor Azevedo Pereira.

Objetivo, antecedentes e justificativa

O objetivo do projeto é avaliar o impacto do Programa Criança Feliz (PCF) sobre a estimulação intelectual no ambiente doméstico e sobre o desenvolvimento cognitivo e psicomotor de crianças brasileiras menores de três anos.

O PCF foi lançado em outubro de 2016, com implantação gradativa nos anos de 2017 e 2018 até atingir 3,5 milhões crianças menores de três anos, cujas famílias sejam beneficiárias do Programa Bolsa Família (PBF) e do Benefício de Prestação Continuada (BPC).

Entre uma série de ações a serem desenvolvidas pelo PCF, inclui-se a realização de visitas semanais para estas crianças, visando a estimulação intelectual. Especificamente, o programa visa orientar e apoiar as gestantes e famílias desde a preparação para o nascimento da criança, colaborando no exercício da parentalidade, fortalecendo o papel das famílias em relação ao cuidado, proteção e educação de crianças na faixa etária de até seis anos de idade e favorecendo o fortalecimento de vínculos afetivos e comunitários; promover ações voltadas ao desenvolvimento integral na primeira infância, estimulando o desenvolvimento de atividades lúdicas envolvendo outros membros da família; e promover, acompanhar e monitorar o desenvolvimento infantil, inclusive mediando o acesso da família a outros serviços públicos que possam impactar positivamente no desenvolvimento das crianças.

Os impactos e benefícios esperados do Programa incluem:

- Fortalecimento dos vínculos familiares
- Prevenção de situações de negligência e violência contra a criança
- Redução da subnutrição crônica
- Melhora do desenvolvimento psicomotor, cognitivo e da capacidade de interação da criança, conforme indicadores apropriados para cada faixa etária

A avaliação aqui descrita será restrita a crianças incluídas no PCF durante seu primeiro ano de vida, no decorrer do ano de 2018, as quais serão acompanhadas prospectivamente por um período de três anos, até completarem 36 meses de idade.

Os princípios orientadores para a execução da avaliação incluíram:

- A necessidade premente de avaliar com metodologias rigorosas o impacto de programas de larga escala.
- A necessidade de planejar precocemente a avaliação do programa, permitindo a coleta de dados de linha de base antes do início da implantação do mesmo.
- A necessidade de avaliação externa por um grupo independente da equipe de gestores do programa.

- A possibilidade de realizar a alocação aleatória de crianças ou municípios para receberem o programa em diferentes períodos de tempo.
- A interação entre gestores e avaliadores para aprimorar a qualidade da avaliação e a retroalimentação de resultados.
- A integração entre o monitoramento do programa realizado pelos gestores e avaliação externa.

A opção por um delineamento com alocação aleatória é essencial para garantir o rigor metodológico da avaliação. O sorteio de crianças para receberem o programa imediata ou posteriormente é eticamente justificado pelas restrições orçamentárias do PCF, que não permitem sua implementação simultânea em todo o país, ou mesmo em todas as crianças elegíveis de um determinado município.

Outro aspecto importante da avaliação é o fato de que os entrevistadores que avaliarão os desfechos do programa não estarão informados se a criança pertence ao grupo que recebe o PCF ou ao grupo controle. Este aspecto, conhecido como “cegamento” em estudos de intervenção, evita a possibilidade de um viés de informação ao comparar ambos os grupos.

Modelo de impacto, indicadores e fontes de dados

A tabela 1 mostra o modelo de impacto esperado para o programa, assim como os indicadores a serem medidos e as diferentes fontes de dados a serem utilizadas.

Tabela 1. Modelo de impacto e indicadores

ETAPA	INDICADORES	METODOLOGIA / FONTES DE DADOS	GRUPOS
Implantação do PCF nos municípios	Características dos municípios incluídos no programa (população, renda, IDH, etc.)	Registros do programa (nacional e estadual)	Intervenção e controle
Visitadoras são selecionados e treinados	Características dos visitantes (idade, sexo, escolaridade, formação, etc)	Registros do programa (município) e entrevista com os visitantes	Intervenção
Aumenta o conhecimento das visitadoras sobre estimulação ao desenvolvimento infantil	Avaliações do conhecimento antes e depois do treinamento (Repetir 12 e 24 meses mais tarde)	Pré- e pós-testes por escrito aplicados pela equipe de avaliadores	Intervenção
Identificação de crianças vulneráveis (elegíveis para o programa)	Características das crianças arroladas no programa, e compatibilidade com os critérios de inclusão	Formulário de extração de dados de identificação de crianças cadastradas e questionário linha de base	Intervenção e controle
Visitadoras realizam as visitas regularmente	Consistência entre o número de visitas planejadas e o número efetivamente realizado	Registros do programa (cadernos das visitadoras) Questionário dos 12, 24 e 36 meses para a família (para confirmar a realização das visitas)	Intervenção
Visitadoras aplicam adequadamente os protocolos definidos pelo programa durante as visitas	Conteúdo das visitas e comportamento das visitadoras está de acordo com protocolo do programa	Observação de sub-amostra de visitas de cada visitador usando um protocolo	Intervenção

		padronizado (ex: HOVRS)	
Aumenta o conhecimento da mãe (cuidadora) sobre estimulação infantil	Indicadores de conhecimento baseados no protocolo do programa	Questionários da linha de base, 12, 24 e 36 meses	Intervenção e controle
Mãe (cuidadora) estimula mais a criança	Avaliação de indicadores de interação / estimulação	Filmagem de sub-amostra da interação entre mães (cuidadoras) e criança para posterior quantificação – linha de base, 12 e 24 meses	Intervenção e controle
Melhora o desenvolvimento infantil	Avaliação psicométrica	Escore de desenvolvimento psicomotor – linha de base, 12, 24 e 36 meses	Intervenção e controle
Redução na vulnerabilidade geral	Estado nutricional da criança Comportamentos agressivos (familiares e da criança) Outros desfechos (a decidir)	Antropometria e questionário sobre violência e saúde infantil	Intervenção e controle

Em resumo, as fontes de dados incluirão:

- revisão da documentação de implantação em nível nacional, estadual e municipal
- materiais e documentação do processo de treinamento de visitadoras
- registros com características das crianças incluídas e excluídas do programa
- pré- e pós-teste de conhecimentos das visitadoras (pós-teste repetido 12 e 24 meses após o treinamento)
- observação de visitas domiciliares (sub-amostra, 6 meses após início do programa)
- observação da interação entre pais e filhos (linha de base, 12 e 24 meses)
- mensuração do desenvolvimento psicomotor (linha de base, 12, 24 e 36 meses)
- mensuração de conhecimentos e atitudes do cuidador principal (linha de base, 12, 24 e 36 meses)
- mensuração de características sociodemográficas, ambientais e assistenciais do domicílio, família, cuidador principal e criança (linha de base, 12, 24 e 36 meses)
- antropometria das crianças ao final dos três anos de acompanhamento

Instrumentos para avaliação do desenvolvimento e da interação

A escolha dos instrumentos a serem usados na avaliação levou em conta a necessidade de dispor de testes simples, válidos, de curta duração e que – em sua maioria - não exijam profissionais altamente qualificados (por exemplo, psicólogos) para sua aplicação.

O processo de escolha dos testes envolveu vários parceiros. Os Anexos 1 e 2 incluem documentos detalhados sobre os diversos testes que foram considerados, e a justificativa para suas escolhas.

A avaliação incluirá diversos tipos de testes, listados abaixo.

Desenvolvimento neuropsicomotor

De acordo com o delineamento da avaliação, propõe-se utilizar dois tipos de teste na avaliação do desenvolvimento neuropsicomotor das crianças pertencentes aos grupos intervenção e controle. O primeiro instrumento será um teste simples e rápido, que possa ser aplicado por entrevistadoras leigas em três ocasiões: na linha de base, 12 meses mais tarde e 24 meses mais tarde. O segundo instrumento será um teste detalhado e abrangente, a ser aplicado 36 meses depois da linha de base, por psicólogas treinadas.

Para o teste simples, a ser aplicado anualmente, a equipe de pesquisa e consultores com experiência na área avaliaram quatro instrumentos:

- O teste “Ages and Stages Questionnaire”, terceira versão (ASQ-3), que contém 21 questionários adequados para distintas faixas etárias entre 1 e 66 meses.
- O teste Denver II, que contém quatro questionários aplicáveis às faixas etárias 0 a 9 meses, 9 a 24 meses, 2 a 4 anos e 4 a 6 anos.
- O questionário Caregiver Reported Early Development Index (CREDI), dividido em três grupos etários (0-11; 12-23; 25-35 meses); este é um instrumento novo, que ainda se encontra em processo de validação.
- O teste Survey of Well-being of Young Children (SWYC), um instrumento de livre acesso para o rastreamento e monitoramento do desenvolvimento de crianças de até 5 anos de idade; este também é um teste desenvolvido recentemente.

Foram considerados também os testes Peabody (linguagem e vocabulário) e MacArthur, os quais foram descartados por abordarem apenas um aspecto do desenvolvimento (no caso do primeiro) ou por serem aplicáveis somente para crianças maiores de um ano (ambos os testes).

Tendo em vista a curta duração do teste, seu amplo uso em outras avaliações similares em diversos países, seu baixo custo e a possibilidade de aplicação por entrevistadores leigos (possibilitando também seu uso para o monitoramento de rotina do programa em todos os municípios) optou-se pelo uso do ASQ-3. Propomos a realização concomitante de um estudo de validação do ASQ-3 na cidade de Pelotas, RS. Os motivos detalhados que justificam esta opção estão expostos no Anexo 1, e os itens incluídos no teste estão incorporados no questionário de linha de base (Anexo 3).

Para a avaliação final do programa, 36 meses após a linha de base (com crianças de 36-47 meses), dois instrumentos foram considerados, o teste Bayley-III e o teste de Batelle, ambos aplicados por psicólogos. O teste Bayley-III tem sido amplamente utilizado em todo o mundo, inclusive nas coortes conduzidas pela UFPEL, e, portanto, optou-se por seu uso na avaliação (ver Anexo 1).

Avaliação da saúde mental da criança

Foram considerados dois instrumentos para avaliar a saúde mental das crianças por ocasião da última visita, quando estarão com 36-47 meses: os testes Strengths and Difficulties Questionnaire (SDQ) e Ontario Child Health Study Scales (OCHS). Devido à facilidade de aplicação e à experiência acumulada com seu uso pela equipe da UFPEL, optou-se por recomendar o teste SDQ (ver Anexo 1).

Testes para depressão e inteligência maternas

Tanto a depressão materna quanto o seu nível intelectual podem influenciar no efeito do PCF. É também desejável demonstrar que, no início da intervenção, as mães dos grupos intervenção e controle apresentam níveis similares de inteligência e de morbidade psicológica.

Optou-se por avaliar a prevalência de depressão em todas as mães do estudo, através de um instrumento simples que tem sido amplamente utilizado no Brasil, inclusive pelas coortes de Pelotas: a Escala de Depressão de Edinburgo (EPDS). As perguntas que integram a escala estão incluídas no questionário de linha de base (Anexo 3).

A testagem de inteligência das mães é complexa e requer aplicação por psicólogos. Por este motivo, recomenda-se aplicar a subescala de vocabulário da Escala Wechsler Abreviada de

Inteligência (WASI) em uma subamostra de 10% das mães, durante o estudo de linha de base. Esta escala é amplamente utilizada em todo o mundo, inclusive no Brasil. O Anexo 3 (questionário de linha de base) inclui os itens desta escala.

Avaliação da interação entre visitadoras e famílias, e de interações responsivas para a aprendizagem (sensibilidade cognitiva materna)

O mecanismo central para o impacto do PCF consiste em garantir a efetiva interação entre visitadoras e as famílias, que resultem na transmissão de conhecimentos para a família e na melhoria da qualidade da estimulação da criança. Durante a avaliação, os diversos componentes que são essenciais para o sucesso do PCF serão avaliados, incluindo as seguintes etapas:

- As visitadoras apresentam conhecimentos adequados sobre estimulação
- Visitadoras realizam as visitas regularmente
- Visitadoras aplicam adequadamente os protocolos definidos pelo programa durante as visitas
- Aumenta o conhecimento da mãe (cuidadora) sobre estimulação infantil
- A mãe ou cuidador estimula mais e melhor a criança

As avaliações de cada uma destas etapas estão delineadas no Anexo 2. O passo mais complexo destas avaliações diz respeito ao último item, ou seja, a qualidade de estimulação da criança pelos pais. Foram considerados dois instrumentos para testar a interação parental. O primeiro foi o teste Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO), um instrumento observacional com 27 itens, que avalia o que os pais podem fazer para apoiar o desenvolvimento das crianças entre 10 e 47 meses de idade. O segundo – MCS ou Maternal Cognitive Sensitivity - avalia a responsividade materna em termos cognitivos, isto é, em que medida a mãe responde às necessidades e interesses da criança de 18 a 36 meses, quando ambos estão envolvidos em uma tarefa de cooperação cognitivamente desafiadora. A interação entre mães e filhos é vídeo-gravada durante 5 minutos, durante a realização de até quatro atividades estruturadas. Os vídeos são posteriormente avaliados por psicólogos treinados que desconhecem se a família pertence ao grupo intervenção ou controle. Este teste foi recentemente validado no Canadá e no Brasil (Pelotas). Pelos motivos expostos no Anexo 2, optou-se por utilizar o MCS na presente avaliação. *O termo sensibilidade cognitiva foi recentemente substituído por “interações responsivas para a aprendizagem”.*

Em resumo, os seguintes instrumentos serão aplicados na avaliação do PCF:

- Teste Ages and Stages Questionnaire (ASQ-3) para a avaliação anual do desenvolvimento neuropsicomotor;

- Teste Bayley-III para avaliação do desenvolvimento neuropsicomotor ao final dos 3 anos da intervenção;
- Escala de Edimburgo para avaliar depressão materna na linha de base;
- A Escala Weschler (WASI) para avaliar inteligência em uma subamostra de mães na linha de base;
- Questionário Strengths and Difficulties (SDQ) para avaliar a saúde mental das crianças no final da intervenção.
- Escala Maternal Cognitive Sensitivity (MCS), gravada em vídeo, para avaliar as interações cognitivas entre cuidadores e criança.

Delineamento da avaliação

Foram considerados inicialmente três opções de delineamento. Todas estas opções incluem a alocação aleatória de crianças ou áreas geográficas que irão ou não receber as visitas do PCF. A alocação aleatória é eticamente justificada pois há sobredemanda, ou seja, o número de crianças elegíveis é bastante superior à capacidade instalada do PCF de visitar regularmente suas famílias. As opções consideradas foram as seguintes:

- a) alocação aleatória em nível individual (alocação de crianças, individualmente, dentro de cada município),
- b) alocação de áreas geográficas dentro de cada município (sorteio de setores censitários onde o PCF seria implementado),
- c) implementação municipal aleatória escalonada (alocação de municípios dentro de cada estado, iniciando com um grupo de municípios no primeiro ano, um segundo grupo no ano 2, e um terceiro grupo no ano 3).

A opção “c” foi descartada após uma discussão inicial com gestores, que argumentaram ser este delineamento de difícil implementação por motivos políticos, ou seja, pela necessidade de adiar a implementação do programa em alguns municípios por até três anos.

Após extensas discussões com os parceiros envolvidos na avaliação, optou-se pela opção “b”. A opção “a” foi descartada devido à possibilidade de contaminação, ou seja, de que famílias do grupo controle recebam mensagens similares às aquelas fornecidas ao grupo intervenção. A contaminação ocorre quando vizinhos, parentes, etc., das famílias do grupo controle passam a estimular seus filhos de forma semelhante ao que ocorre no grupo que recebe o programa. A opção “b” apresenta menor risco de contaminação do que a opção “a”, pois dentro de um mesmo conglomerado (ou vizinhança) todas as crianças elegíveis receberão a intervenção. No entanto, ainda há algum risco de contaminação através de parentes ou conhecidos residentes em outros conglomerados no mesmo município. De qualquer maneira, a contaminação pode ser medida durante o programa através dos questionários a serem aplicados regularmente ao grupo controle.

A opção “b” apresenta também vantagens logísticas, com menor necessidade de deslocamento de visitadoras e avaliadores dentro de cada município, uma vez que as crianças estarão concentradas em um número pequeno de setores censitários.

É importante salientar que o delineamento aleatorizado por conglomerados (setores censitários) é extremamente rigoroso e robusto, sendo amplamente aceito para avaliação de impacto de programas.

Com a opção pelo delineamento com alocação aleatória de setores censitários, foram realizadas extensas discussões com o MDSA para a escolha de estados e municípios.

No presente documento, a amostragem proposta inclui os sete estados de acordo com o que foi acordado inicialmente. No entanto, como será discutido abaixo, há poucos municípios na região Sul e Sudeste que cumprem com os critérios de inclusão. No Apêndice, são feitas considerações sobre as vantagens e desvantagens de realizar a avaliação em um número menor de estados.

Dentro dos estados selecionados, municípios onde já existem programas similares de estimulação infantil serão excluídos, para maximizar a possibilidade de se detectar um efeito significativo do PCF.

Tendo em vista que a clientela de crianças elegíveis para o programa será via de regra bastante superior ao número de vagas a serem oferecidas – ou seja, que existirá *sobredemanda* - a proposta inicial de delineamento com 7 estados inclui as seguintes etapas:

- Escolher 4 municípios de cada estado, que apresentem um número suficiente (ver abaixo) de crianças menores de um ano elegíveis para ingresso no programa.
- Listar todas as crianças elegíveis menores de um ano em cada município a partir do cadastro do PBF.
- Entrevistar as famílias de todas as crianças elegíveis para identificar aquelas que aceitam participar do PCF, e georreferenciar seus domicílios e o setor censitário onde estão localizados.
- Identificar o número de setores necessários para completar a amostra municipal de crianças e sortear metade destes setores para constituir o grupo intervenção, sendo a outra metade designada para o grupo controle.
- Realizar as entrevistas de linha de base, seguidas de avaliações anuais durante 3 anos.

Este delineamento é bastante simples, mas depende da existência de municípios onde o número elegível de crianças seja pelo menos igual ao dobro do número de vagas oferecidas pelo programa, ou seja, onde exista *sobredemanda*.

O cálculo de tamanho amostral (ver abaixo) para o Brasil como um todo estima a necessidade de recrutar 4.320 crianças nos grupos intervenção e controle. Após diversas

simulações, optou-se por trabalhar com 30 municípios no país. Com este número de municípios, o número de crianças menores de um ano a serem recrutadas em cada município é igual a 144 (72 no grupo intervenção e 72 no grupo controle) em 24 setores (12 por grupo).

Em cada um dos 30 municípios, selecionaremos 24 setores censitários com alta densidade de crianças elegíveis. Usando dados do IBGE, estes municípios serão ordenados conforme a renda média no último censo, sendo identificados 12 pares de setores, cada par tendo renda média similar. No primeiro par, o setor mais pobre receberá o número 1 e o setor menos pobre o número 2. No segundo par, o procedimento será invertido e o município menos pobre receberá o número 3 e o mais pobre o número 4. O terceiro par seguirá a mesma lógica do primeiro par, sendo o número 5 alocado para o setor mais pobre e o número 6 para o menos pobre, e assim por diante. Serão então criados dois grupos de 12 setores, um reunindo todos aqueles com números pares e outro com os números ímpares, sendo estes grupos altamente similares em termos de renda média. De acordo com o último dígito da extração da Loteria Federal no dia do sorteio, um dos grupos de 12 setores será sorteado para receber a intervenção.

Posteriormente, para facilitar a atuação das equipes estaduais da avaliação, foi realizada uma modificação no delineamento do estudo. A idéia original para a amostragem, de pré-selecionar 24 setores censitários do IBGE por município, “reservar” estes setores (sem incluir nenhuma criança no PCF), localizar pelo menos 6 crianças elegíveis menores de um ano em cada setor, entrevistar as 144 famílias e testar as crianças, e depois sortear os setores, 12 para receberem os PCF e 12 como controles (72 crianças em cada grupo), esbarrou em três dificuldades identificadas pela equipe da Bahia ao visitar dois dos municípios da amostra. Primeiro, as crianças elegíveis não estavam georreferenciadas dentro dos setores, o que exigiria uma visita prévia para o mapeamento das residências destas crianças. Portanto, o início do trabalho de campo seria adiado, pelo menos em alguns dias/semanas, e haveria necessidade de um financiamento extra para estas pré-visitas. Segundo, alguns setores pré-selecionados tinham menos de seis crianças menores de um ano elegíveis para o PCF, conforme o Cadastro Único. Receamos que os 24 setores pré-selecionados não fossem suficientes para encontrar as 144 crianças. Terceiro, pelo menos em alguns municípios, alguns dos setores “reservados” já estariam recebendo visitas do PCF.

Assim, decidiu-se abandonar a amostragem por setores e adotar a alocação aleatória individual, como previa a opção original “a”. Antes de chegar em cada município, a equipe da pesquisa receberia do PCF uma lista com 160 crianças menores de um ano elegíveis para o PCF. Esta lista incluiria crianças residentes em bairros onde o Centro de Referência de Assistência Social (CRAS) do município tem capacidade de oferecer o programa, excluindo por exemplo áreas rurais ou vilarejos (setores urbanos isolados. A listagem será ordenada pela data de nascimento das crianças, iniciando com a mais jovem. Para ter uma margem

de folga, a equipe da pesquisa tentará entrevistar as 160 crianças mais jovens desta lista, todas menores de um ano. A listagem completa de crianças entrevistadas será enviada para Pelotas, onde a equipe central realizará a amostragem aleatória individual, a partir da lista enviada pelos municípios. A amostragem será por blocos, e dentro de cada bloco metade das crianças serão visitadas pelo PCF e a outra metade constituirá o grupo controle.

A única desvantagem do sorteio individual, relativamente ao sorteio de setores, é a possibilidade de contaminação, ou seja, que as mães das crianças-controle sejam expostas ao PCF, pois famílias vizinhas estariam sendo visitadas. Nos questionários de linha de base, serão coletadas informações sobre contaminação (vizinhos ou parentes recebendo o PCF), o que nos permitirá pelo menos medir sua magnitude e excluir estas crianças “contaminadas” nas sub-análises de impacto.

A grande vantagem da alocação individual é simplificar em muito o trabalho de campo. Além disso, não exige mapeamento prévio do município e aumenta a oferta de crianças elegíveis (por não restringir a poucos setores). Uma outra vantagem da alocação aleatória individual será diminuir ou mesmo eliminar o efeito de delineamento devido aos conglomerados (a correlação intraclasse, pois crianças vizinhas vivendo em um mesmo conglomerado tendem a ser parecidas entre si). No cálculo original do tamanho amostral, havíamos feito um acréscimo para este efeito de delineamento. A amostragem individual requer um número de crianças 33% inferior ao exigido pela amostragem por conglomerados, na qual o “efeito do delineamento” infla o tamanho amostral. Portanto, ao invés de 144 crianças por município, 96 serão suficientes com este novo delineamento.

Processo de alocação

Para cada município será gerada uma planilha em Excel contendo os números de identificação social (NIS) e as datas de nascimento das crianças, cujas famílias tenham sido entrevistadas e aceitado participar do estudo. Nesta planilha, as crianças serão ordenadas por idade, da mais jovem para a mais velha. A primeira etapa da alocação consistirá em um sorteio usando um dado previamente testado para garantir que não estava viciado. Ao ser lançado, se o dado mostrasse um número ímpar, a ordem do primeiro par de crianças seria grupo A (criança mais jovem do primeiro par) e grupo B (criança mais velha do primeiro par), e na presença de um número par, a ordem do primeiro par de crianças seria oposta, ou seja B e A. Posteriormente, como será descrito abaixo, os grupos A e B serão alocados para PCF e controle.

Para a alocação das crianças para o grupo PCF ou controle, serão realizados cinco lançamentos de dados, cada qual gerando um número e a sequência A-B ou B-A para os cinco primeiros pares de crianças (ou 10 crianças). O sorteio será realizado por videoconferência com o Ministério e coordenadores estaduais, sendo filmado para

documentação. O resultado da sequência gerada pelos lançamentos do dado encontra-se na Tabela abaixo (Tabela 1).

Tabela 1: Resultado do lançamento do dado e sequências geradas.

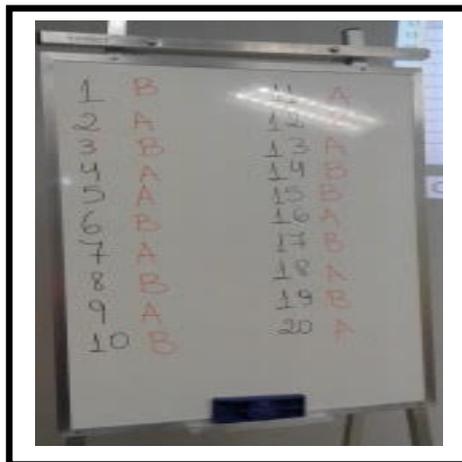
Lançamento do dado	Resultado do lançamento		Sequência
1º	2	Par	B e A
2º	2	Par	B e A
3º	5	Ímpar	A e B
4º	1	Ímpar	A e B
5º	3	Ímpar	A e B

Como, mesmo com o sorteio aleatório, seria possível que o grupo A fosse ligeiramente mais velho ou jovem do que o grupo B, para os próximos cinco pares de crianças sorteados foi invertida a ordem inicial, gerando respectivamente as sequências A-B, A-B, B-A, B-A, B-A para as crianças de número de ordem 11 a 20, totalizando assim dez pares (ou 20 crianças) (Tabela 2 e Figura 3). Essa sequência será utilizada em todos os sorteios, para todos os municípios randomizados. A alocação por blocos é amplamente usada em estudos clínicos e resultará em dois grupos altamente comparáveis em termos de idade das crianças, pois a listagem será ordenada por data do nascimento

Tabela 2: Sequência dos 20 números sorteados.

Número da criança	Grupo	Número da criança	Grupo
1	B	11	A
2	A	12	B
3	B	13	A
4	A	14	B
5	A	15	B
6	B	16	A
7	A	17	B
8	B	18	A
9	A	19	B
10	B	20	A

Figura 3: Foto do quadro no dia do primeiro sorteio.



Com os dez primeiros pares formados, as 20 crianças mais jovens da lista serão alocadas aos grupos A ou B. A mesma sequência de pares AB ou BA será copiada na planilha para as crianças de números 21 a 40, 41 a 60, etc., até completar a lista final de cada município. Para o último bloco (que provavelmente terá menos de 20 crianças) o mesmo processo se aplica, até terminar a numeração de crianças. No caso de municípios com um número ímpar de crianças, um dos grupos terá uma criança a mais. Este sorteio inicial será realizado uma única vez para toda avaliação, em 31/10/2018, e definirá os grupos A e B para todos os municípios, em todos os sorteios subsequentes.

Resta então, para cada município, determinar se o PCF será ofertado para o grupo A ou B. Para isso, serão utilizados os números premiados no último sorteio da Loteria Federal antes de cada definição da amostra. O último dígito do primeiro prêmio será usado para o primeiro município; o último dígito do segundo prêmio para o segundo município, e assim por diante. Nos sorteios em que participarem mais de cinco municípios, o primeiro prêmio será usado novamente para o sexto município, o segundo prêmio para o sétimo município, etc.

A aleatorização será realizada na UFPEL, com participação dos demais parceiros da pesquisa através de vídeo. A aleatorização será realizada mensalmente, para os municípios com linha de base concluída nos últimos 30 dias. Após a conclusão de cada sorteio, uma lista em PDF será gerada imediatamente e enviada a todos os coordenadores estaduais e a equipe do MDS. Todos os sorteios serão transmitidos por videoconferência (transmissão pela ferramenta Zoom), sendo também gravados e compartilhados por links no site Youtube, em ordem das datas dos sorteios.

Municípios escolhidos para a avaliação

A proposta inicial previa realizar a avaliação em estados das cinco regiões do país. Quando ficou disponível a listagem de municípios que aderiram ao PCF, tornou-se evidente que seria impossível realizar a avaliação na Região Sul, onde havia poucos municípios elegíveis. Optou-se então por incluir estados das quatro regiões restantes, com ênfase na Região Nordeste onde três estados seriam selecionados.

Após a listagem de municípios e estados que aderiram ao programa, foi pactuado com o MDSA como critério seleção dos estados o contingente populacional beneficiário do PBF, assim como a existência de uma equipe de pesquisa qualificada para coletar dados nos municípios. Nessa linha, foram propostos para o estudo: Pará, Ceará, Pernambuco, Bahia, São Paulo e Goiás. Nestes estados, foram listados todos os municípios com número suficiente de crianças menores de um ano a serem acompanhadas, assim como a sua distribuição por setores censitários.

Os seguintes princípios guiaram a seleção inicial de municípios:

1. Municípios que aderiram ao PCF.
2. Municípios com mais de 10.000 habitantes.
3. Para garantir uma margem de segurança, propomos selecionar apenas municípios com sobredemanda igual ou maior do que 4,0.
4. Capacidade instalada para incluir pelo menos 80 crianças menores de um ano no PCF e realizar as visitas semanalmente. Posteriormente, foi acordada com o MDSA a expansão do número de vagas oferecidas nos municípios selecionados, para garantir que as vagas necessárias para a avaliação fossem adicionais ao número já pactuado com o município (ou seja, cada município selecionado receberá vagas extras para cobrir todas as crianças incluídas no grupo intervenção).
5. Pelo menos 800 crianças menores de um ano no PBF, para garantir uma boa margem de segurança da clientela potencial do PCF.
6. Pelo menos 6 crianças elegíveis menores de 12 meses e cujas famílias aceitam participar da avaliação por setor censitário incluído no estudo – mesmo caso - (como as crianças elegíveis estarão concentradas nos setores mais pobres, a média municipal de crianças por setor pode ser um pouco inferior a 6 crianças).
7. Mais de 24 setores censitários urbanos com pelo menos 6 crianças elegíveis ao PCF menores de 12 meses.

Com o apoio do MDSA, e levando em conta os aspectos acima, assim como a localização geográfica, identificamos os municípios listados abaixo (Tabela 2)

Tabela 2. Municípios escolhidos para a avaliação. (Nota: esta ainda não é a lista definitiva de municípios pois ainda estão ocorrendo desistências antes do estudo de linha de base)

	UF	Município	Total de setores	de Setores PCF	Crianças <12 meses no PBF	Sobre-demanda	Meta de oferta de vagas no PCF
1	BA	Campo Formoso	134	68	1319	7.1	400
2	BA	Eunápolis	139	70	1278	13.8	200
3	BA	Feira de Santana	628	81	2264	4.2	1500
4	BA	Paulo Afonso	168	63	1192	6.5	500
5	BA	Santo Antônio de Jesus	146	43	964	5.0	400
6	BA	Vitória da Conquista	450	87	2364	6.2	800
7	CE	Caucaia	373	153	3049	7.4	900
8	CE	Crato	176	92	1739	8.6	500
9	CE	Itapipoca	180	88	2131	9.0	500
10	CE	Juazeiro do Norte	316	124	2527	8.0	800
11	CE	Morada Nova	107	66	1232	12.2	200
12	CE	Sobral	235	114	2262	8.7	600
13	GO	Aguas Lindas de Goiás	157	40	1276	29.2	200
14	GO	Luziania	236	92	1877	28.2	200
15	GO	Novo Gama	98	37	872	11.7	200
16	PA	Altamira	163	43	1037	13.5	200
17	PA	Bragança	106	78	2086	11.2	400
18	PA	Breu Branco	58	40	997	10.6	200
19	PA	São Miguel do Guamá	60	50	1015	10.9	200
20	PA	Tailândia	66	48	1196	16.2	200
21	PE	Abreu e Lima	114	71	1159	11.5	200
22	PE	Camargibe	197	63	1228	5.5	400
23	PE	Caruaru	405	130	2839	7.8	1000
24	PE	Olinda	466	132	2647	5.6	1000
25	PE	Serra Talhada	108	55	1195	6.3	400
26	SP	Embu	407	71	1698	4.4	800
27	SP	Francisco Morato	199	103	1768	8.3	400
28	SP	Limeira	460	52	1466	6.8	500
29	SP	Piracicaba	642	45	1388	4.7	700
30	SP	Taboão da Serra	247	75	1433	4.5	600

Mapa 1. Municípios originalmente selecionados para o estudo



No entanto, alguns dos municípios que haviam expressado interesse em aderir ao PCF e que foram incluídos na amostra original acabaram por não aderir, ou aderiram e posteriormente desistiram. Assim, foi necessário substituir quatro municípios, sendo dois na BA (Itamaraju e Santo Antônio de Jesus, que foram substituídos por Serrinha e Irecê), um em PE (Olinda, substituído por Campina Grande) e um em SP (Itapetininga por Umbu das Artes).

Além disso, as estimativas iniciais da SAGI/MDS indicavam mais de 144 crianças elegíveis em cada município. Mas, com base nas listagens disponibilizadas pela SAGI, havia vários municípios com número insuficiente de crianças, entre outras razões, porque os municípios delimitam áreas onde o PCF será implantado (ex: áreas urbanas e áreas menos violentas) e as listas do MDS não conseguem fazer esta discriminação. Em função disso, na Bahia, o município de Campo Formoso teve que ser substituído por Casa Nova e em Pernambuco, Campina Grande, por São Lourenço da Mata. Além disso, foi necessário ampliar as listagens em todos os municípios, incluindo mulheres grávidas, o que requereu visitas a todas as gestantes, pois o sistema do MDS não informa se a criança já nasceu, uma vez que essas crianças não possuem número no SIS. Os municípios efetivamente participantes do estudo encontram-se no mapa abaixo.

Mapa 2. Municípios efetivamente participantes do estudo



Cálculos do tamanho amostral

Os cálculos originais de tamanho de amostra utilizaram os seguintes parâmetros:

- Erro alfa (probabilidade de encontrar um impacto quando na verdade o programa não tem impacto) = 5%
- Erro beta (probabilidade de não detectar um impacto quando na verdade o programa tem impacto) = 10%
- Efeito do delineamento (fator que multiplica o tamanho da amostra para levar em conta o delineamento) = 1,30 para conglomerados com 4 crianças ou 1,50 para conglomerados com 6 crianças
- Adesão ao programa (continuidade de participação durante 3 anos entre as famílias que ingressaram no programa) = 60% ou 80%
- Perdas de acompanhamento por mudança de endereço ou recusa a participar das entrevistas = 20%.

O valor do erro alfa é universalmente aplicado em estudos de avaliação, e corresponde ao valor de 5% utilizado para detectar significância estatística. O valor do erro beta, na maior parte dos estudos epidemiológicos, é tradicionalmente fixado em 20%; optamos por 10% para reduzir o risco de que avaliação leve à conclusão equivocada de que o programa não teve impacto. Sendo conservadores, usamos um teste bicaudal para o poder estatístico, ou seja, pressupondo que a intervenção poderia tanto melhorar como piorar o desempenho das crianças nos testes.

Os valores do efeito do delineamento são mais difíceis de estabelecer. Em pesquisa realizada em Pelotas usando conglomerados urbanos (áreas de abrangência de postos de saúde) e medindo o quociente de inteligência em adolescentes com a escala Weschler (Iná dos Santos, dados não publicados) observou-se uma correlação intraclasse (ICC) igual a 0,02. Avaliações de programas similares ao PCF na Colômbia, Peru e Bangladesh mostraram ICCs de 0,04, 0,07 e 0,09. Sendo conservadores novamente, optamos por um ICC de 0,10 em nossos cálculos.

Em cada setor censitário selecionado, esperamos encontrar cerca de 6 crianças menores de um ano elegíveis para o PCF. Com setores incluindo 6 crianças em média e um ICC de 0,10, estima-se um efeito de delineamento de 1,50, ou seja, um aumento de 50% na amostra para compensar a alocação por setores censitários, comparada com uma amostra aleatória simples de crianças. Estimamos também a amostra considerando conglomerados de 4 crianças; nesta situação o efeito do delineamento é reduzido para 1,30.

Conforme é praxe em estudos randomizados, as análises principais deverão ser feitas conforme a “intenção de tratar” (ITT, *intent to treat* em inglês), ou seja, crianças alocadas

inicialmente para o grupo intervenção precisam ser consideradas como se fossem tratadas até o final do estudo, mesmo que suas famílias desistam ou deixem de ser visitadas por outros motivos, como mudança para outros bairros ou cidades. Isso implica em considerar o efeito do programa sobre o total de crianças no início do estudo, notando que este total inclui crianças cujas famílias deixarão de aderir ao programa durante o período de 3 anos de acompanhamento. O efeito entre as crianças que permanecem no PCF durante 3 anos também será avaliado; este é conhecido como efeito do “tratamento entre os tratados” (TOT, *treatment on the treated*).

É preciso considerar ainda que, mesmo entre as famílias que aderem ao programa, um percentual poderá ser perdido pois mudam de endereço ou se recusam a ser entrevistadas pela equipe de pesquisa. Estipulamos este percentual de perdas em 20%.

Estimamos inicialmente que 40% das crianças incluídas no programa deixarão de ser acompanhadas por mudança ou recusa da família após o início da intervenção, ou seja, que a adesão será de 60%. Portanto, se a intervenção resultar em um aumento médio de QI de 3,8 pontos (ou 0,25 desvio padrão) entre as crianças acompanhadas, o grupo intervenção como um todo terá um aumento de QI de 2,3 pontos (3,8 pontos nas 60% que recebem toda a intervenção e zero entre as 40% que desistem do programa antes dos 3 anos). Este é o cenário E na tabela 3 abaixo. Com estes parâmetros, serão necessárias 2.172 crianças no grupo intervenção e outras 2.172 no grupo controle.

Tabela 3. Cenários de tamanho de amostra para a avaliação do PCF.

CENÁRIOS	A	B	C	D	E	F	G	H
Erro alfa	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96
Erro beta	1.65	1.65	1.65	1.65	1.65	1.65	1.65	1.65
ICC	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Tamanho do conglomerado	4	4	4	4	6	6	6	6
Efeito do delineamento	1.30	1.30	1.30	1.30	1.50	1.50	1.50	1.50
Adesão ao programa	60%	60%	80%	80%	60%	60%	80%	80%
Efeito nos acompanhados (SD)	0.25	0.33	0.25	0.33	0.25	0.33	0.25	0.33
Efeito nos acompanhados (pontos)	3.8	5.0	3.8	5.0	3.8	5.0	3.8	5.0
Efeito no total do grupo intervenção (SD)	0.15	0.20	0.20	0.26	0.15	0.20	0.20	0.26
Efeito no total do grupo intervenção (pontos)	2.3	3.0	3.0	4.0	2.3	3.0	3.0	4.0
Desvio padrão do teste	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
Perdas de acompanhamento	20%	20%	20%	20%	20%	20%	20%	20%
Tamanho de amostra por grupo	1882	1080	1059	608	2172	1247	1222	701
Tamanho de amostra PCF + controle	3765	2161	2118	1215	4344	2493	2444	1402

Examinando os diversos cenários que constam da Tabela 3, propomos que uma amostra nacional de 4.320 crianças, sendo 2.160 em cada grupo, será suficiente para detectar o

impacto do programa. Este número de 4.320 foi escolhido por ser divisível por 30, que será o número de municípios conforme descrito acima. Vale notar que em todas as etapas do cálculo de tamanho amostral foram feitas estimativas conservadoras, por exemplo quanto ao poder estatístico de 90% bicaudal, o ICC, o tamanho do efeito e a taxa de adesão. Os cenários E-H apresentam a vantagem de utilizarem conglomerados com 6 ao invés de 4 crianças, o que diminui os gastos de tempo e transporte.

O mesmo tamanho de amostra de 2.160 crianças por grupo permitirá detectar como significantes as diferenças abaixo (Tabela 4) entre os dois grupos em desfechos expressos como proporções, por exemplo os resultados das avaliações da interação parental com a criança, ou na prevalência de déficit de desenvolvimento. Assim, se 20% das crianças no grupo controle apresentarem atraso no desenvolvimento será possível detectar como significativa uma redução de 6 pp (pontos percentuais), ou seja, uma prevalência de 14% no grupo intervenção.

Tabela 4. Diferenças entre proporções detectáveis com amostras de 2.000 crianças em cada grupo.

Grupo	Prevalência do desfecho (déficit de desenvolvimento)				
Grupo 1	10%	20%	30%	40%	50%
Grupo 2	6%	15%	24%	33%	43%
Diferença	4 pp	5 pp	6 pp	7 pp	7 pp

Portanto, nossos cálculos indicam que uma amostra inicial de 4.320 crianças no total do país será suficiente para detectar o impacto do programa, desde que este impacto tenha relevância em termos de magnitude.

É importante salientar que este grupo inicial de 4.320 crianças irá incluir aquelas que efetivamente permanecem no programa como também aquelas cujas famílias desistem ou se mudam. Nesse último caso, deverão ser feitos todos os esforços possíveis para encontrar e examinar as crianças uma vez por ano, mesmo que tenham deixado o programa. Este procedimento é essencial pois a análise será realizada – conforme descrito acima - por “intenção de tratar”, ou seja, mesmo as famílias e crianças que saírem do programa após o início do acompanhamento necessitam ser consideradas como se continuassem participando do programa. Sem esta premissa, o processo inicial de aleatorização fica comprometido.

A experiência das coortes de nascimentos de Pelotas é de que seja possível acompanhar pelo menos 90% das crianças por três anos, mas para sermos conservadores, estipulamos percentuais de adesão ao programa de 60% e 80% (Tabela 3).

Conforme a magnitude da não adesão e das perdas, poderão ser utilizadas variáveis instrumentais nas análises, para aumentar a comparabilidade entre os dois grupos.

Com a mudança no delineamento do estudo para alocação individual dos participantes, foi excluído o efeito do delineamento, resultando em 96 crianças por município (48 intervenção e 48 controles), totalizando 2.880 crianças no país.

Um e dois anos após o início da intervenção, subamostras dos grupos intervenção e controle serão observadas de forma sistemática e a interação entre pais e filhos será filmada, para posterior avaliação da interação entre os cuidadores e as crianças. Com base nos filmes, a interação será pontuada de acordo com a escala de Interações Responsivas para a Aprendizagem (ver a seção referente à escolha de testes, acima).

Em um estudo recente realizado pelos pesquisadores Alexandra Schneider e Tiago Munhoz, a média do teste foi de 3,2 pontos, com desvio padrão de 0,7 (Munhoz T, comunicação pessoal). Para detectar uma melhora de 0,2 ponto no escore, são necessárias 142 crianças em cada grupo, ou 284 no país. Estes cálculos pressupõem um erro alfa de 5%, poder de 90% (bicaudal) e efeito de delineamento de 1,0, uma vez que se propõe recrutar uma criança de cada conglomerado. Acrescentando 50% para não adesão e perdas, propõe-se uma amostra de 432 crianças no país, ou seja, 10% da amostra total da avaliação.

Questionários de linha de base

Na linha de base, serão administrados três questionários para todas as mães ou cuidadoras, além de um teste de inteligência para uma subamostra de 10% das mães (Anexo 3). Os questionários serão aplicados aos grupos de intervenção e controle, antes do início das visitas domiciliares do PCF. O termo de consentimento livre e esclarecido (Anexo 4) será também aplicado na mesma ocasião, antes da aplicação dos questionários.

Os questionários serão programados em *tablets*, com a incorporação de saltos (“pulos”) para omitir perguntas que não são relevantes para as distintas faixas etárias das crianças. Os questionários incluídos no Anexo 3 não se encontram no formato final, pois não serão impressos em papel, mas programados nos *tablets*.

Parte 1 (linha de base).

Este questionário incorpora perguntas padronizadas e testadas, incluídas no questionário Multiple Indicator Cluster Survey (MICS) utilizado pelo UNICEF em mais de 70 países, e pelos questionários utilizados nas Coortes de Nascimento de Pelotas, RS, nos anos 2004 e 2015. Incorpora também perguntas adaptadas dos manuais de treinamento do PCF. O questionário inclui os seguintes módulos (Tabela 5):

Tabela 5. Módulos do questionário de linha de base.

Módulo	Fonte	Conteúdos
Dados gerais	MICS	Identificação Consentimento (ver termo de consentimento livre e esclarecido no anexo 4)
Gestação e parto	Coortes de Pelotas	Atenção ao pré-natal e parto Características do recém-nascido
Cuidado e estimulação	Coortes de Pelotas e MICS	Frequência a creche Cuidados da criança Estimulação precoce (questionário simplificado do MICS)
Disciplina	MICS	Medidas utilizadas para disciplinar a criança
Alimentação	MICS	Aleitamento materno Ingesta de líquidos, semi-sólidos e sólidos (detalhada)
Imunizações	Coortes de Pelotas	Situação vacinal da criança (detalhada)
Hospitalizações	Coortes de Pelotas	Histórico de hospitalizações
Características da mãe e do pai	Coortes de Pelotas	Estrutura familiar História reprodutiva Escolaridade Raça/cor Tabagismo materno
Estimulação	PCF	Questionário do PCF para avaliar estimulação, por faixa etária da criança (obs: estas perguntas são específicas do PCF e apresentam alguma duplicação com as perguntas de estimulação do MICS, ver acima)
PCF	Perguntas novas	Conhecimentos sobre o PCF para mães do grupo intervenção e controle Participação no PCF (caso o estudo de linha de base ocorra após o início do programa)
Características ambientais e socioeconômicas	MICS ANEP Coortes de Pelotas	Escala de bens domésticos (asset index) Escala de classes sociais da ANEP Características de água, saneamento e construção Bolsa Família
Dados para contato		Endereço, telefone, etc., da família e de familiares

Parte 2: Ages and Stages Questionnaire (avaliação do desenvolvimento da criança)

O ASQ-3 é um instrumento de avaliação e monitoramento do desenvolvimento cognitivo, comunicação, psicomotor e pessoal-social de crianças entre 0 e 6 anos. Os itens do instrumento são aplicados aos pais e avaliam capacidades ou dificuldades para cada grupo etário específico. De acordo com as informações fornecidas pelo fabricante, o tempo total de aplicação do teste varia entre 10 e 15 minutos. Os resultados da avaliação são expressos em escores brutos para cada domínio. Estes escores brutos são comparados com pontos de corte em cada área, o que indica se a criança necessita de avaliação especializada, se deve ser monitorada ou se o desenvolvimento está adequado. Os resultados de teste e re-teste para a população norte-americana indicaram concordância de 92% entre as avaliações e um coeficiente de correlação intra-classe de 0,75 a 0,82, sugerindo um ótimo desempenho. A concordância entre avaliadores (pais e examinadores treinados) foi de 93%, com um coeficiente de correlação intra-classe de 0,43 a 0,69, sugerindo um desempenho satisfatório entre avaliadores. A consistência interna do questionário, avaliada pelo alpha de Cronbach, variou entre 0,60 e 0,85. A sensibilidade e especificidade do ASQ-3 foram comparadas ao *Battelle Developmental Inventory* (BDI). Os resultados indicaram sensibilidade de 86,1% e especificidade de 85,6%.(Squires 1999)

A versão brasileira foi preparada e validada pela Universidade Federal do Rio de Janeiro, e se encontra reproduzida no Anexo 3.

Parte 3. Escala de Depressão Materna

Para avaliação da saúde mental materna, será utilizada a Escala de Depressão de Edinburgo que avalia sintomas de depressão na comunidade (Santos, Matijasevich et al. 2007, Matijasevich, Munhoz et al. 2014). Esta escala contém 10 perguntas que avaliam os sintomas mais comuns para depressão. Cada pergunta é composta por quatro alternativas de respostas, com valores entre 0 e 3, onde o (a) entrevistado (a) deve escolher aquela que descreva como tem se sentido nos últimos sete dias. O somatório desses valores gera o escore total da escala. Serão considerados positivos para depressão os indivíduos as mulheres que pontuarem escore ≥ 8 (Matijasevich, Munhoz et al. 2014). A escala será aplicada verbalmente pela entrevistadora em um único bloco seguindo a ordem das perguntas no instrumento. Esta escala foi validada no Brasil para uso na população de mulheres no período pós-parto (Santos, Matijasevich et al. 2007) e na população geral (Matijasevich, Munhoz et al. 2014), tanto para homens como para mulheres. A versão a ser aplicada está reproduzida no Anexo 3.

Parte 4. Escala de inteligência materna WASI

A escala WASI é um instrumento breve de avaliação da inteligência, aplicável a crianças de 6 anos a idosos de 89 anos de idade. Fornece informações sobre os QIs Total, de Execução e Verbal, em um curto espaço de tempo, a partir de quatro subtestes: vocabulário, cubos, semelhanças e raciocínio matricial. A escala ainda fornece a possibilidade de avaliação do QI Total com apenas dois subtestes (vocabulário e raciocínio matricial). A escala é também associada à Escala de Inteligência Wechsler para Crianças – Terceira Edição e à Escala de Inteligência Wechsler para Adultos – Terceira Edição, além de fornecer tabelas para estimativa de faixas de escore de QI Total nas escalas WISC-III e WAIS-III. O WASI está validado para a população brasileira (Trentini, Yates et al. 2014) O subteste **vocabulário** avalia o desenvolvimento da linguagem e conhecimento de palavras; o subteste com **cubos** avalia a organização perceptual e visual, visualização espacial, análise do todo em suas partes componentes e formação de conceito não verbal; o subteste de **semelhanças** avalia a formação de conceito verbal e pensamento lógico abstrato; e, o **raciocínio matricial**, o raciocínio abstrato e processamento da informação visual.

A versão brasileira (ver Anexo 3) foi traduzida, testada e validada na população brasileira.

Plano de análise

O plano de análise levará em conta os seguintes aspectos:

1. **Estrutura complexa da amostra.** Será levado em conta o delineamento complexo da amostra (grupo intervenção ou controle; unidade da federação; município; setor censitário; família e criança), através de análise multinível. Os efeitos do delineamento complexo serão considerados no cálculo da variância.
2. **Análise principal por intenção de tratar.** Crianças alocadas inicialmente para o grupo intervenção serão mantidas nesse grupo nas análises primárias, independentemente do fato de haver continuado ou não a receber a intervenção durante todo o período. Famílias que desistirem ou se mudarem serão analisadas como se houvessem participado do programa. Conforme mencionado acima, este tipo de análise é conhecido na literatura como ITT (“intent to treat”).
3. **Análise adicionais.** Diversas análises adicionais serão realizadas.
 - a. **Análise por protocolo.** Nesta análise, o grupo intervenção será composto apenas pelas famílias (e crianças) que permanecerão no programa até o final. Estas serão comparadas com as crianças do grupo controle. Como descrito acima, este tipo de análise é conhecido como análise naïve. Esta análise é necessariamente enviesada, pois não respeita a alocação aleatória original, sendo possível que as crianças e famílias que completem o tratamento sejam diferentes daquelas que desistem durante o estudo. Essa análise, no entanto, nos mostra a direção do viés quando deixamos de utilizar o método de variáveis instrumentais.
 - b. **Análise de fatores relacionados à adesão** (ou “take-up”). Estas análises comparam, dentro do grupo inicialmente alocado para receber o programa, as crianças e famílias que permanecem até o final do estudo com aquelas que desistem. Isso permite entender as razões para o abandono e melhorar o programa no futuro. Se a adesão for baixa, uma estratégia complementar de análise é o uso de variáveis instrumentais para prever quais as crianças e famílias do grupo controle que provavelmente participariam do programa, e comparar estas crianças e famílias com aquelas que efetivamente aderiram ao programa no grupo intervenção.
 - c. **Análise do efeito do tratamento nos tratados (TOT):** Embora a análise principal seja sobre a intenção de tratar (ITT), é possível também calcular o efeito médio do programa sobre aqueles que decidem permanecer nele. Esse efeito é chamado de efeito sobre os tratados. O efeito sobre os tratados

é estimado através do método de variáveis instrumentais. Intuitivamente, imagine que o efeito estimado da intenção de tratar seja de 0.15 SD sobre o desenvolvimento cognitivo das crianças que moram em setores selecionados para receber o programa, e que a adesão ao programa nesses setores (take up) seja de 60%. Nesse caso, o efeito sobre cada criança efetivamente tratada (TOT) é de $0.15/0.6=0.25$ SD.

4. **Correção para diferenças de linha de base.** Apesar da alocação aleatória, há uma probabilidade de 5%, devida ao acaso, de que ocorram diferenças em cada característica das famílias e crianças. As amostras iniciais dos grupos intervenção e controle serão comparadas em termos de características medidas no início do estudo (como escolaridade dos pais, sexo da criança, peso ao nascer, tamanho da família, características da habitação, etc.), e se houver diferenças significativas entre o grupo intervenção e controle, estas características serão tratadas como fatores de confusão (ou confundimento) nas análises.
5. **Uso de testes bicaudais.** Como não se pode excluir a possibilidade de que a intervenção tenha efeitos negativos sobre o desenvolvimento psicomotor, serão utilizados testes bicaudais na análise estatística.
6. **Análise de fatores mediadores e mecanismos.** Além da análise principal que revelará os efeitos sobre o desenvolvimento psicomotor, serão realizadas análises de desfechos intermediários (ver Tabela 6) para aumentar a plausibilidade do argumento de que os efeitos sobre o desenvolvimento foram devidos efetivamente à intervenção. Serão utilizados métodos apropriados para análise de mediação, incluindo G formula e modelos generalizados de estimação.
7. **Análises de subgrupos.** Se o tamanho da amostra final assim permitir, serão realizadas análises de subgrupos definidos a priori, para identificar se o efeito da intervenção variou conforme: sexo da criança; grupos de renda familiar; número de irmãos; tipo de família (nuclear ou extensiva); escolaridade e idade maternas, níveis iniciais de desenvolvimento infantil e níveis de interação entre pais e filhos.

As variáveis a serem analisadas estão listadas na Tabela 6.

Tabela 6. Variáveis a serem incluídas nas análises.

ETAPA	INDICADORES	ANÁLISE E COMENTÁRIOS
Implantação do PCF nos municípios	Características dos municípios incluídos no programa (população, renda, IDH, etc.)	Análise descritiva de registros do programa (nacional e estadual). Comparação de municípios que aderem ou não ao programa. Importante para a validade externa do estudo de avaliação.
Visitadoras são selecionadas e treinadas	Características das visitadoras (idade, sexo, escolaridade, formação, etc)	Análise descritiva dos registros do programa (município) e entrevistas com as visitadoras. Importante para descrever a intervenção.
Aumenta o conhecimento das visitadoras sobre estimulação	Avaliações do conhecimento antes e depois do treinamento. Repetir 12 e 24 meses mais tarde.	Pré- e pós-testes por escrito aplicados pela equipe de avaliadores. Comparação dos conhecimentos antes e depois do treinamento. Com os testes realizados 12 e 24 meses mais tarde, avaliar a retenção dos conteúdos do treinamento.
Identificação de crianças vulneráveis (elegíveis para o programa)	Características das crianças arroladas no programa e compatibilidade com os critérios de inclusão	Formulário de extração de dados de identificação de crianças cadastradas e questionário linha de base. Verificação sobre se as crianças e famílias incluídas efetivamente se enquadram nos critérios de inclusão. Identificação da taxa de adesão e de fatores socioeconômicos, demográficos, etc., associados à adesão. Análises essenciais para a validade externa do programa.
Comparabilidade inicial entre os grupos intervenção e controle	Características socioeconômicas, demográficas, nutricionais, de conhecimentos e de morbidade dos dois grupos	Questionário de linha de base aplicado aos dois grupos, antes do início da intervenção. Informação sobre peso ao nascer obtida da Caderneta da Criança. Comparação entre os grupos, para identificar possíveis diferenças, que ocorrem apesar da alocação aleatória. Importante para identificar variáveis de confusão (confundimento), que deverão ser levadas em conta na análise final do impacto.

ETAPA	INDICADORES	ANÁLISE E COMENTÁRIOS
Visitadoras realizam as visitas regularmente	Consistência entre o número de visitas planejadas e o número efetivamente realizado	Registros do programa (cadernos das visitadoras). Questionário dos 12, 24 e 36 meses para a família (para confirmar a realização das visitas). Análises descritivas, essenciais para garantir que a intervenção foi efetivamente realizada.
Visitadoras aplicam adequadamente os protocolos definidos pelo programa durante as visitas	Conteúdo das visitas e comportamento das visitadoras está de acordo com protocolo do programa	Observação de sub-amostra de visitas de cada visitadora usando um protocolo padronizado. Análise descritiva da qualidade das visitas.
Aumenta o conhecimento da mãe (cuidadora) sobre estimulação infantil	Indicadores de conhecimento baseadas no protocolo do programa	Questionários da linha de base, 12, 24 e 36 meses. Análises descritivas sobre os níveis de conhecimento das mães (cuidadoras) ao longo do programa.
Mãe estimula mais a criança	Avaliação de indicadores de interação / estimulação	Filmagem da interação entre cuidadores e criança para posterior quantificação – linha de base, 12 e 24 meses. Avaliação objetiva das mudanças na qualidade da estimulação pela família.
Melhora o desenvolvimento infantil	Avaliação psicométrica nos grupos intervenção e controle.	Escore de desenvolvimento psicomotor – linha de base, 12, 24 e 36 meses. Análise principal do estudo: (a) comparação entre os dois grupos; (b) mensuração do possível progresso do grupo intervenção ao longo do tempo. Análises do tipo ITT e TOT (ver acima). Controle de diferenças na linha de base através de análises multivariáveis (se necessário).
Redução na vulnerabilidade geral	Comportamentos agressivos (familiares e da criança) Outros desfechos (a decidir)	Questionário sobre violência e saúde infantil. Comparação entre os dois grupos para identificar possíveis benefícios do programa, adicionais ao efeito principal sobre desenvolvimento.

Aspectos éticos

A alocação aleatória de crianças, conglomerados ou municípios para receberem ou não a intervenção de estimulação é eticamente justificada porque, conforme mencionado acima, o financiamento federal para o programa não é suficiente para atingir todas as crianças elegíveis. O banco de dados do MDSA mostra que, em muitos municípios, a sobredemanda será superior a 4, ou seja, a maior parte das crianças elegíveis não poderão ser absorvidas pelo programa. As crianças que forem sorteadas para o grupo controle continuarão a receber a Bolsa Família e os demais benefícios para os quais são elegíveis.

Os pais das crianças incluídas na avaliação fornecerão seu consentimento informado por escrito. Os questionários e demais testes serão confidenciais, sendo mantidos em bancos de dados anonimizados e seguros. Será dado um retorno aos pais das crianças sobre os resultados dos testes de desenvolvimento e, quando necessário, as crianças serão encaminhadas para serviços públicos especializados.

O presente projeto foi registrado na Plataforma Brasil e aprovado pela Comissão de Ética em Pesquisa da Universidade Federal de Pelotas em 30 de junho de 2017 (Parecer: 2.148.689) e em 02 de abril de 2018 (Parecer: 2.574.190), onde consta o caráter multicêntrico nacional da avaliação cobrindo seis estados. A Comissão dispensou a necessidade de aprovação pelo Conselho Nacional de Ética em Pesquisa (CONEP), ao qual a comissão está afiliada. A minuta da aprovação e o TCLE (termos de consentimento livre e esclarecido) constam do Anexo 4.

O protocolo do estudo encontra-se depositado no *site* Registros Brasileiros de Ensaios Clínicos (ReBEC) sob o número [RBR-4x7dny](#).

Disponibilidade de resultados

Resultados iniciais sobre o efeito do programa sobre a interação entre os pais e crianças estarão disponíveis cerca de um ano após o início da avaliação e resultados preliminares sobre o impacto do programa sobre o desenvolvimento estarão disponíveis cerca de 18-24 meses após a linha de base.

Os dados anonimizados serão disponibilizados para a sociedade três anos após a conclusão da coleta em campo.

Logística e equipe de pesquisa

Em cada estado, será formada uma equipe com 1 supervisora, 6 entrevistadoras e 2 motoristas. Cada equipe realizará 5 entrevistas simultaneamente, no mesmo município, com 1 entrevistadora extra para eventuais substituições e entrevistas de controle de qualidade. Estima-se que cada entrevistadora poderá realizar, com folga, 4 entrevistas por dia, totalizando 20 entrevistas para toda a equipe, ou 8 dias de campo para cada município (22 dias em GO, 36 dias em PE, SP e PA, e 43 dias na BA e CE). Acrescentando os dias de deslocamento entre municípios, folgas semanais e um dia por município para reconhecimento e listagem das crianças a serem avaliadas, estipula-se 1 mês de campo em GO, 45 dias em PE, SP e PA, e 2 meses na BA e CE.

Com base em equipes de 9 membros, em GO será necessário cobrir 270 diárias e os gastos de 2 veículos por 30 dias; em PE, SP e PA, será necessário cobrir 405 diárias e os gastos de 2 veículos por 45 dias; e, na BA e CE, será necessário cobrir 540 diárias e os gastos de 2 veículos por 60 dias. Se alguns dos municípios escolhidos forem sedes das instituições de pesquisa, o gasto de diárias será menor, pois as equipes não precisarão ser acomodadas em hotéis.

Logística prévia à coleta de dados

Para que o estudo de linha de base seja executado, será necessário atender os seguintes passos:

- Convite aos prefeitos dos 30 municípios selecionados para o estudo de avaliação do PCF
- Assinatura do Termo de Adesão pelos prefeitos municipais para participar do estudo de avaliação do PCF
- Em cada município, identificação de 24 setores censitários que atendam os critérios de disponibilidade de crianças menores de 12 meses elegíveis e de sobredemanda ≥ 4
- Seleção aleatória dos 12 setores intervenção e 12 controle, conforme definido na seção “Delineamento da avaliação” acima
- Nos setores intervenção, sorteio das crianças que receberão o PCF e, dentre essas, sorteio das que farão parte do grupo intervenção do estudo de avaliação do PCF
- Nos setores controle, sorteio das crianças que formarão o grupo controle do estudo de avaliação do PCF

Estes passos serão implementados pelos técnicos do departamento de avaliação do MDSA, juntamente com os pesquisadores responsáveis pela avaliação do PCF.

Com a mudança no delineamento do estudo para alocação individual dos participantes, em cada município foram identificadas todas as crianças menores de 12 meses e as gestantes com data provável de parto nos 11 meses anteriores ao período em que as entrevistas de linha de base seriam realizadas. A alocação subsequente das crianças para o grupo intervenção ou controle ocorreu como descrito na seção intitulada “Processo de alocação”, acima apresentada.

Logística para a coleta de dados de linha de base

Definidas as crianças que comporão a amostra do estudo de avaliação, a execução da coleta de dados envolverá as seguintes atividades:

- O Coordenador estadual faz contato com o secretário municipal de desenvolvimento social dos municípios selecionados de seu estado, para checar os dados de identificação das famílias participantes do estudo de avaliação (nome, idade e sexo da criança, nome e idade da mãe e endereço residencial)
- Deslocamento no campo para localização dos endereços e aplicação da entrevista às mães, sendo entrevistadas 20 mães (cuidadoras) por dia
- A coleta de dados em cada município deve durar cerca de 7 dias
- A entrevistadora extra entrevista 2-3 mães por dia, aplicando questionário reduzido de controle de qualidade
- A equipe de campo desloca-se para executar o campo em um próximo município somente depois de concluir as 144 entrevistas programadas para cada município

Além disso, uma sub-amostra de mães, dos grupos intervenção e controle, serão filmadas durante a execução de uma atividade intelectualmente desafiadora (teste de Interações Responsivas para a Aprendizagem), para posterior avaliação da interação entre mães (cuidadoras) e criança. Uma pessoa por estado, munida de uma câmera fotográfica e treinada para a aplicação do teste, fará a filmagem desta atividade em 10% das mães estudadas (intervenção e controle).

Logística para a coleta de dados 12 e 24 meses após o início do estudo

Preferentemente, em cada estado, a mesma equipe que realizou as entrevistas de linha de base realizará as entrevistas de seguimento, quando será aplicado um instrumento

semelhante ao de linha de base (Questionário do primeiro acompanhamento – T1), contendo questões adequadas à idade alcançada pela criança por ocasião da entrevista (Anexo 3).

Nos dois acompanhamentos, uma sub-amostra de mães (10%), dos grupos intervenção e controle, serão filmadas para avaliação de Interações Responsivas para a Aprendizagem.

Adicionalmente, para avaliar a interação entre visitadora e a família, um avaliador externo fará a observação de uma sub-amostra das visitas de cada visitadora. Para tanto, o observador utilizará um instrumento padronizado, construído com base no protocolo do PCF, especificamente para este fim, onde registrará os itens que surgirem durante a observação. O Anexo 4 contém uma cópia deste instrumento.

A força da implementação do PCF nos municípios estudados será avaliada de várias maneiras: a) aplicação de um teste escrito aos Visitadores do programa, abordando conteúdos disponíveis em seu material de treinamento; realização de discussão em grupo com 9 a 12 visitadores indicados pelo coordenador municipal; entrevista com os coordenadores estaduais e municipais do PCF; e entrevista com os responsáveis pelo PCF nos três CRAS com maior cobertura do programa no município. Os instrumentos empregados para a coleta dessas informações encontram-se no Anexo 5.

Adicionalmente, antes do início das visitas de T1 a T3, as famílias incluídas na avaliação do impacto do PCF serão contatadas por telefone, com dois objetivos: confirmar a permanência da família no endereço identificado na linha de base (e com isso otimizar o tempo de coleta de dados do acompanhamento subsequente – T1, T2 e T3) e checar o recebimento de visitas do PCF (para verificar adesão dos Visitadores às atividades do programa e identificar contaminação do grupo controle). Para os telefonemas, serão usados, inicialmente, o telefone de contato da própria mãe e, na impossibilidade de contatá-la após três tentativas, a telefonista buscará contato com as demais pessoas cujos números de telefone tenham sido fornecidos pela mãe na visita da linha de base. As telefonistas serão treinadas para padronização da abordagem das mães e pessoas de contato e aplicação de breve questionário estruturado.

Logística para a coleta de dados finais (36 meses após o início do estudo)

As visitas finais do estudo serão realizadas por psicólogas treinadas para aplicação nas crianças dos instrumentos Bayley-III e o questionário SDQ.

Equipe de pesquisa

A equipe de pesquisadores do Programa de Pós-Graduação em Epidemiologia da Universidade Federal de Pelotas (Cesar Victora e Iná Santos) possui ampla experiência em avaliação de impacto de programas de saúde, nutrição e desenvolvimento infantil, e constituirá o grupo de referência para a avaliação aqui proposta. Tiago Munhoz, também da UFPEL (Faculdade de Psicologia) coordenará a escolha, adaptação e validação dos instrumentos sobre estimulação e desenvolvimento psicomotor.

Os estados que participam da avaliação do PCF foram selecionados por conveniência, com base na disponibilidade de instituições de pesquisa com experiência na realização de estudos populacionais. Definidos os estados, os coordenadores foram selecionados por terem reconhecida experiência em pesquisas populacionais somada à recomendação de seus pares.

Em cada um dos seis estados escolhidos para a pesquisa, a entidade acadêmica selecionada será responsável pela coleta de dados nos municípios selecionados. Os coordenadores e as instituições a que estão vinculados encontram-se listados abaixo.

Devido à desistência, em abril de 2018, da equipe da Universidade Nacional de Brasília, que estaria a cargo de realizar a pesquisa em Goiás, os municípios deste estado seriam inicialmente cobertos pela equipe da Universidade Federal da Bahia. Posteriormente, com o aceite da Universidade Federal de Goiás, foi possível constituir uma equipe própria do estado.

Coordenadores	Instituições
Alicia Matijasevich amatija@hotmail.com	Universidade de São Paulo
Hernane Santos Júnior hernanegs@gmail.com	Universidade Federal do Oeste do Pará
Luciano Correia correialuciano@hotmail.com	Universidade Federal do Ceará
Letícia Marques marquesleticia@hotmail.com	Universidade Federal da Bahia
Pedro Lira lirapic@ufpe.br	Universidade Federal de Pernambuco
Marta Roverly martary@gmail.com	Universidade Federal de Goiás

Controle de qualidade

As medidas de controle de qualidade do estudo incluirão:

- Cegamento dos avaliadores dos desfechos quanto ao status de intervenção ou controle das crianças
- Padronização e pré-teste de questionários
- Treinamento para aplicação dos questionários e demais instrumentos de coleta de dados
- Manuais de instruções para todos os instrumentos a serem utilizados no estudo
- Revisão de todas as entrevistas para identificar respostas faltantes e inconsistências entre as respostas

As entrevistas de controle de qualidade terão como objetivo testar a repetibilidade das respostas das mães (cuidadoras) a algumas das perguntas do questionário principal. Para tanto, será aplicado um questionário reduzido contendo perguntas cujas respostas sejam mais estáveis no tempo, por exemplo, “quantos filhos vivos a mãe teve antes do nascimento da criança índice?” O questionário de controle de qualidade encontra-se no Anexo 6.

Cronograma simplificado

Atividades	ANOS					
	2018	2019	2020	2021	2022	2023
Treinamento das equipes estaduais	xxx	xxx	xxx	xxx		
Coleta de dados (linha de base)	xxx	x				
Alocação aleatória	xxx	xxx				
Estudo da força da implementação (1)		xxx				
Coleta de dados (T1)		xxx				
Estudo da força da implementação (2)			xxx			
Coleta de dados (T2)			xxx			
Estudo da força da implementação (3)				xxx		
Coleta de dados (avaliação final)				xxx		
Análise de dados		xxx	xxx	xxx	xxx	xxx
Redação de artigos científicos		xxx	xxx	xxx	xxx	xxx
Oficina nacional de disseminação					xxx	

Referências

Matijasevich, A., T. N. Munhoz, B. F. Tavares, A. P. Barbosa, D. M. da Silva, M. S. Abitante, T. A. Dall'Agnol and I. S. Santos (2014). "Validation of the Edinburgh Postnatal Depression Scale (EPDS) for screening of major depressive episode among adults from the general population." BMC Psychiatry **14**: 284.

Santos, I. S., A. Matijasevich, B. F. Tavares, A. J. D. Barros, I. P. Botelho, C. Lapolli, P. V. d. S. Magalhães, A. P. P. N. Barbosa and F. C. Barros (2007). "Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study." Cadernos de Saúde Pública **23**: 2577-2588.

Squires, J. (1999). Asq-3 User's Guide: Ages & Stages Questionnaires.

Trentini, C., D. Yates and V. Heck (2014). "Escala de Inteligência Wechsler Abreviada (WASI): Manual profissional." São Paulo, SP: Casa do Psicólogo.

Anexos

Anexo 1: Justificativa para o uso de ASQ-3 para avaliar o desenvolvimento infantil

Anexo 2: Teste de conhecimentos dos Visitadores do PCF com gabarito; *Checklist* para observação das visitas dos Visitadores do PCF; Teste de Interações Responsivas para a Aprendizagem

Anexo 3: Questionário de Linha de Base, ASQ-3, EPDS, WASI, Questionário do primeiro acompanhamento (T1) e Manuais de Instrução

Anexo 4: Termo de Consentimento Livre e Esclarecido

Anexo 5: Instrumentos para avaliar a implementação do PCF (Roteiro de perguntas para o Grupo dos Visitadores; Questionário para entrevista com o coordenador estadual; Questionário para entrevista com coordenador municipal; Questionário para entrevista com o responsável pelo PCF no CRAS) e respectivos manuais de instruções

Anexo 6: Questionário de controle de qualidade