

Avaliação do Programa Criança Feliz: um estudo randomizado em 30 municípios brasileiros

Evaluation of the Happy Child Program: a randomized study in 30 Brazilian municipalities

Iná S. Santos (<https://orcid.org/0000-0003-1258-9249>)¹

Tiago N. Munhoz (<https://orcid.org/0000-0003-1281-9542>)¹

Raquel Siqueira Barcelos (<https://orcid.org/0000-0001-6580-0509>)¹

Cauane Blumenberg (<https://orcid.org/0000-0002-4580-3849>)¹

Caroline Cardozo Bortolotto (<https://orcid.org/0000-0003-3318-7900>)¹

Alicia Matijasevich (<https://orcid.org/0000-0003-0060-1589>)²

Cristiane Salum (<https://orcid.org/0000-0003-0191-371X>)³

Hernane Guimarães dos Santos Júnior (<https://orcid.org/0000-0002-9998-2141>)⁴

Letícia Marques (<https://orcid.org/0000-0001-5963-2166>)⁵

Luciano Correia (<https://orcid.org/0000-0001-8948-8660>)⁶

Marta Rovey de Souza (<https://orcid.org/0000-0001-6910-843X>)⁷

Pedro Israel Cabral de Lira (<https://orcid.org/0000-0002-1534-1620>)⁸

Vitor Pereira (<https://orcid.org/0000-0002-6487-7704>)⁹

Cesar G. Victora (<https://orcid.org/0000-0002-2465-2180>)¹

Abstract *The Happy Child Program (Programa Criança Feliz - PCF, in Portuguese) reaches 1.4 million Brazilian children under three years of age with home visits aimed at promoting neuro-psychomotor development. Based on a conceptual model, PCF implementation and impact were evaluated in a randomized study in 30 municipalities. A total of 3,242 children were allocated to the intervention (IG) or control (CG) group, 80.0% of whom were prospectively followed up from late 2018 to late 2021. Development was assessed by the Ages and Stages Questionnaire (ASQ3). During the three-year study period, visits were replaced by virtual contacts for an average of 12 months due to COVID-19. At the endline survey, intent-to-treat analyses showed mean scores of 203.3 in the IG and 201.3 in the CG. Additional analyses using instrumental variables and propensity scores matching also showed no effect, since the number of contacts with the program was not associated with ASQ3 scores. No impact was observed on stimulation, responsive interactions or psychological attributes of children. The implementation study revealed low coverage in the IG, contamination of the CG, deficiencies in management and low quality of visits in many municipalities. The study did not demonstrate an impact of PCF implemented under routine conditions, but provides elements for its improvement.*

Key words *Child development, Program evaluation, Impact*

Resumo *O Programa Criança Feliz (PCF) atinge 1,4 milhão de crianças brasileiras menores de três anos com visitas domiciliares visando o desenvolvimento neuropsicomotor. Com base em modelo conceitual, avaliou-se implementação e impacto do PCF em estudo randomizado, em 30 municípios. Ao todo 3.242 crianças foram alocadas para o grupo intervenção (GI) ou controle (GC), sendo 80,0% acompanhadas prospectivamente durante três anos. O desenvolvimento foi avaliado pelo Ages and Stages Questionnaire (ASQ3). Análises por intenção de tratar mostraram escores médios de 203,3 no GI e 201,3 no GC. Análises adicionais com variáveis instrumentais e emparelhamento por escores de propensão tampouco mostraram efeito, uma vez que o número de contatos recebidos não esteve associado aos escores ASQ3. Tampouco foi observado impacto sobre estimulação, interações responsivas ou atributos psicológicos das crianças. As visitas foram interrompidas durante 12 meses devido à COVID-19, sendo substituídas por contatos virtuais. O estudo de implementação revelou baixa cobertura no GI, contaminação do GC, deficiências na gestão e baixa qualidade das visitas em muitos municípios. O estudo não demonstrou impacto do PCF implementado sob condições de rotina e fornece elementos para seu aprimoramento.*

Palavras-chave *Desenvolvimento infantil, Avaliação de programas, Impacto*

¹ Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas. R. Marechal Deodoro 1160, 3º piso, Centro. 96020-220 Pelotas RS Brasil. inasantos.epi@gmail.com

² Departamento de Medicina Preventiva, Faculdade de Medicina, Universidade de São Paulo. São Paulo SP Brasil.

³ Centro de Matemática, Computação e Cognição, Universidade Federal do ABC. Santo André SP Brasil.

⁴ Instituto de Saúde Coletiva, Universidade Federal do Oeste do Pará. Santarém PA Brasil.

⁵ Instituto de Humanidades Artes e Ciências, Universidade Federal da Bahia. Salvador BA Brasil.

⁶ Departamento de Saúde Comunitária, Universidade Federal do Ceará. Fortaleza CE Brasil.

⁷ Departamento de Saúde Coletiva, Universidade Federal de Goiás. Goiânia GO Brasil.

⁸ Departamento de Nutrição do Centro de Ciências da Saúde, Universidade Federal de Pernambuco. Recife PE Brasil.

⁹ Escola Nacional de Administração Pública. Brasília DF Brasil.

Introdução

Programas de visitas domiciliares são considerados estratégias efetivas para promoção do desenvolvimento infantil. Revisões sistemáticas de intervenções realizadas em países de renda média e baixa mostraram efeito positivo sobre os domínios cognitivo, linguagem, psicossocial e motor^{1,2}. Metanálises, que reuniram ensaios randomizados conduzidos em países de renda alta, média e baixa, também encontraram efeitos positivos sobre o desenvolvimento infantil, além de melhoras em desfechos mediadores, como conhecimento dos pais, práticas parentais e interação com a criança³⁻⁵. É importante notar que a maioria desses estudos experimentais foram realizados em amostras pequenas e restritas, sendo as intervenções implementadas pelos próprios pesquisadores.

Várias intervenções randomizadas foram implementadas em países da América Latina. Na Colômbia, em um estudo em 96 municípios com 1.420 crianças o grupo que recebeu visitas domiciliares durante 18 meses melhorou em 0,26 desvios padrão (DP) o escore cognitivo ajustado ($p=0,002$) e em 0,22 DP o escore de linguagem ajustado ($p=0,032$) em relação ao grupo não visitado⁶. No Peru, a proporção de crianças com escore acima da média para a idade, em todos os domínios, foi 12-23 pontos percentuais mais alta no grupo que recebeu visitas domiciliares⁷. No México, a integração de dois programas de larga escala em funcionamento no país (um programa de transferência de renda e outro de educação infantil) resultou em maiores índices cognitivo, de linguagem e de memória, equivalentes a 0,26-0,29 DP, em comparação ao grupo que recebia somente aporte financeiro⁸. Uma outra análise do mesmo programa mostrou efeito positivo das intervenções combinadas sobre desfechos mediadores: brincadeiras entre pais e filhos e leitura de livros infantis, os quais explicaram 32% do efeito sobre o desenvolvimento⁹.

No Brasil, uma revisão da literatura identificou uma lacuna de estudos de avaliação de impacto dos programas de visita domiciliar implementados como políticas públicas com enfoque no desenvolvimento da criança¹⁰. No âmbito das pesquisas de intervenções lideradas por universidades, um estudo de coorte com 156 crianças conduzido no interior de Pernambuco empregou visitas domiciliares e atividades em grupo realizadas por terapeutas ocupacionais¹¹. A diferença média entre os grupos intervenção e controle foi de 9,4 pontos em desenvolvimento cognitivo e 8,2 pontos em desenvolvimento psicossocial

($p<0,001$ em cada caso). Um recente estudo randomizado controlado realizado com 900 famílias na cidade de São Paulo avaliou o impacto no desenvolvimento da criança da versão adaptada do programa *Reach Up*¹², uma intervenção baseada em evidências¹². A análise por intenção de tratar não demonstrou efeitos no desenvolvimento das crianças, porém a análise por protocolo, restrita às díades mãe-criança que completaram no mínimo 10 visitas, mostrou uma melhora no desenvolvimento das crianças de 0,22 DP (IC95% 0,01-0,43)¹³. Um estudo quase-experimental do impacto do Programa Primeira Infância Melhor (PIM), em Pelotas-RS, com 601 crianças que participaram do programa e um grupo controle emparelhado individualmente não encontrou efeito do programa no desenvolvimento de crianças aos quatro anos de idade¹⁴. Porém, crianças cujas famílias foram visitadas desde a gravidez apresentaram prevalência 60% inferior de atraso no desenvolvimento, comparadas com o grupo controle ($p=0,02$), e uma forte evidência estatística de que o efeito do PIM foi maior do que quando iniciado após o nascimento (razão de prevalências=0,40; IC95% 0,18-0,89).

Em março de 2016, foi instituída a Lei 13.257, denominada Marco Legal da Primeira Infância, que estabelece que a oferta de programas e de ações de visita domiciliar de estimulação ao desenvolvimento integral na primeira infância serão consideradas como estratégias de atuação respaldadas pelas políticas públicas sociais. Após esta legislação, em outubro de 2016, por iniciativa do então Ministério do Desenvolvimento Social e Agrário (atual Ministério da Cidadania - MCid), o Brasil tornou-se o primeiro país do mundo a lançar um programa de visitas domiciliares com abrangência nacional visando promover o desenvolvimento de crianças vivendo em condições de vulnerabilidade social – o Programa Criança Feliz (PCF)¹⁵. São elegíveis para o PCF crianças pertencentes a famílias beneficiárias do Programa Bolsa Família (atual Auxílio Brasil), um amplo programa nacional de transferência de renda que cobre 40 milhões de brasileiros¹⁶. Com implementação gradativa, o PCF em junho de 2022 cobria cerca de 1,4 milhão de crianças menores de três anos.

O PCF é um programa com ações executadas de forma descentralizada e integrada entre os três entes federativos: União, Estados e Municípios¹⁷. As ações preconizadas devem ser realizadas pelos municípios de forma articulada com os serviços socioassistenciais, via Centros de Referência de Assistência Social (CRAS) e com as demais políticas públicas setoriais, objetivando a atenção in-

tegral das demandas das famílias. Com a assinatura do Termo de Aceite e Compromisso do PCF, o município decide pela adesão com a aprovação do Conselho Municipal de Assistência Social. A partir desse ponto, dentro de suas competências como gestores da política de assistência social, os municípios têm autonomia para a execução das ações, observando a legislação pertinente. Os municípios que aderem passam a receber aportes financeiros do governo federal para execução do PCF.

Especificamente, o PCF visa orientar e apoiar as gestantes e famílias, desde a preparação para o nascimento da criança, colaborando no exercício da parentalidade; fortalecer o papel das famílias em relação ao cuidado, proteção e educação das crianças; e favorecer o fortalecimento de vínculos afetivos e comunitários; além de promover ações voltadas ao desenvolvimento integral na primeira infância, estimulando o desenvolvimento de atividades lúdicas, com envolvimento de 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 crianças¹⁵. Entre as ações desenvolvidas pelo PCF, incluem-se visitas semanais para as crianças, visando a estimulação intelectual até a criança atingir 36 meses de idade, quando as visitas devem ser interrompidas.

No mesmo ano de lançamento do PCF, o Ministério encomendou a pesquisadores do Programa de Pós-Graduação em Epidemiologia da Universidade Federal de Pelotas (UFPEL) um projeto para a avaliação independente do impacto do programa sobre o desenvolvimento infantil, em diferentes estados brasileiros.

O modelo de impacto, construído pelos pesquisadores por ocasião do planejamento do estudo, postulava uma cadeia de eventos necessários para o impacto do programa sobre o desenvolvimento infantil¹⁸. O primeiro elo da cadeia seria a implantação do PCF nos municípios e o segundo, a seleção de visitadores e seu treinamento específico e estruturado, resultando no aumento de seus conhecimentos sobre estimulação e desenvolvimento infantil. O terceiro elo consistiria na realização regular de visitas às famílias selecionadas para o grupo intervenção, consistentemente com o número e periodicidade planejadas pelo programa. Seguiu-se o quarto elo da cadeia, representado pela aplicação adequada dos protocolos definidos pelo programa, em termos de conteúdo das visitas e no relacionamento apropriado entre visitadores e famílias. Com isso, seria desencadeado o quinto elo da cadeia - ou seja o aumento do conheci-

mento da mãe/cuidador sobre estimulação infantil - o qual promoveria a interação e estimulação da criança (sexto elo). Mais estimulada, a criança apresentaria melhor desenvolvimento do que seus pares em similares condições de vida, mas não beneficiários do programa. Desta forma, o PCF atingiria seu principal objetivo de promover o desenvolvimento infantil.

Durante os três anos da avaliação, cada um dos elos acima descritos foi avaliado por meio de distintos indicadores e instrumentos de coleta de informações. Por se tratar de uma cadeia, em que todas as etapas estão interligadas em uma ordem lógica e sequencial, deficiências em qualquer dos elos comprometeria os elos subsequentes^{19,20}. O objetivo deste artigo é apresentar os resultados dessa avaliação de impacto do programa.

A presente pesquisa buscou fornecer elementos para que o PCF qualifique seus esforços para aprimorar o fortalecimento parental e desenvolvimento infantil, objetivos de extrema relevância para nossas crianças em situação de vulnerabilidade e para o desenvolvimento nacional. As avaliações de impacto e de implementação podem contribuir com os gestores para o aprimoramento e fortalecimento desse e outros programas na área da primeira infância¹⁰. Além disso, a pesquisa fornece evidências científicas sobre o perfil da população atendida pelo Bolsa Família e pelo PCF, assim como sobre fatores associados ao desenvolvimento da criança. Uma publicação anterior com base nesta avaliação verificou que a cobertura vacinal no grupo intervenção foi maior do que no controle quando as crianças tinham em média 18,9 meses²¹. Também foi demonstrada uma alta prevalência (26,5%) de sintomas depressivos entre as mães participantes do estudo²², e que - além da depressão materna - os fatores associados a menores escores de desenvolvimento infantil incluíam a baixa escolaridade materna e sua percepção de não ter recebido apoio durante a gestação²³. Finalmente, a pesquisa também visa suprir a lacuna na literatura de estudos sobre o impacto de programas de visitas domiciliares de estimulação do desenvolvimento da criança vinculados a políticas públicas.

Métodos

Delineamento e elegibilidade

A avaliação longitudinal teve dois componentes principais. O primeiro ou estudo de impacto, foi um ensaio randomizado em nível in-

dividual. Menores de um ano de idade, elegíveis para o PCF, foram selecionados em 2018-2019 e acompanhadas durante três anos.

O segundo componente (estudo de implementação) baseou-se em metodologia observacional quantitativa e qualitativa. A equipe de avaliação visitou os municípios no final de 2019 e 2021 e coletou informações por telefone no final de 2020, quando não foi possível realizar visitas devido à pandemia de COVID-19.

Seleção dos municípios

Em parceria com técnicos do Ministério, foram selecionados 30 municípios, cobrindo quatro das cinco regiões do país. A região Sul (que é composta por somente três estados) não foi incluída, porque havia poucos municípios com número suficiente de crianças elegíveis e devido à presença de programas similares ao PCF, que poderiam contaminar o delineamento do estudo. Em cada uma das quatro regiões, foi selecionado um estado, exceto na região Nordeste, em que três estados foram incluídos, uma vez que esta é a região com a maior clientela do programa de transferência de renda. A Figura 1 mostra um mapa com os estados e municípios selecionados.

Em cada um dos seis estados, foram selecionados três a seis municípios, de acordo com os seguintes critérios: adesão do governo local ao programa e ao estudo de avaliação; número de habitantes ≥ 10.000 ; número de crianças elegíveis

pelo menos quatro vezes superior à capacidade instalada do programa; e capacidade para incluir pelo menos 80 crianças. A exigência de pelo menos quatro crianças elegíveis para cada vaga do programa era essencial, por motivos éticos relacionados à randomização.

Cálculos do tamanho amostral

Com base na revisão da literatura, estipulou-se que, em relação ao grupo de comparação, a participação no PCF levaria a um aumento de 0,20 DP nos resultados dos testes de desenvolvimento. Uma vez que a adesão total era improvável, adotou-se uma diferença de 0,15 DP nos cálculos do tamanho amostral. Parâmetros adicionais fixaram erro alfa (bicaudal) de 0,05, poder do estudo de 90% (erro beta de 10%) e 20% de perdas de acompanhamento, resultando em uma amostra nacional estimada de 2.880 crianças na linha de base. Este tamanho de amostra deveria ser alcançado com uma média de 96 crianças por município, isto é, 48 crianças em cada grupo.

Randomização

A base de dados do programa de transferência de renda do Ministério foi utilizada para gerar listas de crianças potencialmente elegíveis em cada município. Do final de 2018 a meados de 2019, a equipe de pesquisa visitou cada município e mapeou as áreas em que o PCF seria im-



Figura 1. Municípios incluídos no estudo.

plementado. Na maioria dos municípios, as áreas rurais e as regiões com problemas de segurança foram excluídas pelos técnicos do serviço social. A lista de elegibilidade incluía famílias com crianças menores de 12 meses ou gestantes com data provável de parto em 2018. De posse da lista das famílias elegíveis, a equipe de pesquisa realizou o estudo de linha de base e gerou uma lista de crianças cujas famílias concordaram em participar no estudo randomizado e assinaram o termo de consentimento informado. Uma vez que 26 municípios já haviam iniciado a implementação do PCF, as crianças que tivessem sido visitadas uma ou mais vezes eram inelegíveis, não sendo incluídas na avaliação.

Logo que o estudo de linha de base era concluído em um município, uma lista anonimizada das crianças elegíveis era enviada eletronicamente para a Universidade Federal de Pelotas (UFPEL), onde a randomização era imediatamente realizada. Cada criança era identificada por sua data de nascimento e pelo Número de Identificação Social (NIS), um número exclusivo atribuído pelo Ministério a todos os beneficiários do bem-estar social. Na UFPEL, todas as crianças elegíveis em cada município eram ordenadas de acordo com a data de nascimento, e blocos de 20 crianças em dez pares eram criados. Usando um dado de jogar, as crianças de cada par eram alocadas para o grupo A ou B, resultando na seguinte sequência dentro de cada grupo de 20 crianças: BABAABA-BABABABBABABA. A randomização dentro de cada grupo de 20 crianças foi realizada uma única vez, sendo utilizada nos 30 municípios. Assim que as listas de crianças elegíveis para cada município eram definidas, o último dígito do primeiro prêmio da loteria federal era usado para alocar o grupo A para receber as visitas (quando esse dígito era ímpar) ou grupo B (para dígitos pares). Todas as seções de randomização foram transmitidas ao vivo pela internet, usando a plataforma Zoom, e seguem disponíveis no Youtube²⁴. As listas eram imediatamente salvas em formato PDF e enviadas ao Ministério, que as transmitia aos 30 municípios, especificando as crianças que deveriam receber as visitas (grupo intervenção) e as que não deveriam ser visitadas (grupo controle).

Aspectos éticos

Os municípios selecionados apresentavam um número de crianças elegíveis para o PCF pelo menos quatro vezes maior do que o número potencial de vagas ofertadas, garantindo que a seleção aleatória não excluiria crianças que poderiam

se beneficiar. O Projeto foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Pelotas, filiado a Comissão Nacional de Ética em Pesquisa, sob o número de protocolo CONEP 2.148.689. O protocolo do estudo do estudo foi depositado no Registro Brasileiro de Ensaios Clínicos (protocolo REBEC RBR-4x7dny).

O protocolo do estudo garantia a confidencialidade de informações pessoais. Foi obtido consentimento informado por escrito de um dos pais ou do cuidador na entrevista de linha de base. Todos os bancos de dados são anonimizados. Todos os dados do estudo são de domínio público e poderão ser usados por outros grupos de pesquisa.

Estudo de avaliação: coleta de dados

O delineamento original do estudo de avaliação incluía quatro visitas para todas as crianças: o estudo de linha de base, no final de 2018 (designado T0 ou T zero); o estudo de acompanhamento do primeiro ano (T1), no final de 2019; o segundo acompanhamento (T2), no final de 2020; e o terceiro acompanhamento (T3), no final de 2021. A pandemia levou ao cancelamento das visitas do T2, sendo o delineamento adaptado para incluir cinco entrevistas telefônicas com as famílias em junho e setembro de 2020 e em janeiro, abril e julho de 2021.

Para as entrevistas do T0, T1 e T3, cada equipe estadual incluía um investigador principal experiente, um ou dois supervisores e 6-10 trabalhadores de campo. Todos os supervisores e trabalhadores de campo tinham pelo menos ensino médio completo. Antes das fases T0 e T1, os entrevistadores foram treinados face a face, em Brasília; e, antes da fase T3, devido à pandemia, o treinamento foi realizado via *web*. Em cada estado, os supervisores e os trabalhadores de campo realizaram as entrevistas, em cada município, durante 1-2 semanas e, se necessário, retornavam em outra ocasião para completar as entrevistas adicionais.

As medidas de controle de qualidade, em todas as fases, incluíram a padronização e teste piloto dos questionários; manuais detalhados para os trabalhadores de campo; reentrevistas a 10% de todas as famílias, realizadas pelos supervisores usando um questionário curto; e visitas a municípios selecionados, em todos os estados, por membros da equipe central (UFPEL) durante as fases de trabalho de campo. Os entrevistadores não eram informados quanto ao status de intervenção ou controle de cada criança, embora fosse

possível deduzir esta informação uma vez que os questionários incluíam variáveis relacionadas à participação no PCF.

Instrumentos

As três visitas incluíram a aplicação de questionários para as mães ou cuidadoras das crianças (doravante referidas como “mães”) e testes de desenvolvimento e saúde mental da criança. O questionário principal incluía questões amplamente usadas do *Multiple Indicator Cluster Survey* (MICS)²⁵ e dos questionários usados nos Estudos de Coorte de Nascimento de Pelotas (<http://www.epidemiologia-ufpel.org.br/site/content/studies/>).

Perguntas adicionais foram adaptadas dos manuais de treinamento do PCF. Os tópicos incluídos nos questionários abordavam informações sociodemográficas da família e dos pais, cuidados e estimulação da criança, disciplina, alimentação e uso de serviços de saúde. Perguntas sobre participação no PCF, incluindo idade de ingresso, frequência das visitas e recordatório materno sobre aconselhamentos recebidos dos visitantes foram coletados de ambos os grupos, intervenção e controle. A depressão materna foi investigada usando a versão brasileira da *Edinburgh Postnatal Depression Scale*²⁶⁻²⁸. Os questionários completos estão disponíveis no link <http://www.epidemiologia-ufpel.org.br/site/content/pesquisas/index.php?estudo=5>.

A estimulação cognitiva e sensorial foi avaliada por meio de um escore construído com base em perguntas de inquéritos populacionais MICS (UNICEF) e nas recomendações do manual de treinamento dos visitantes do PCF²⁹. De acordo com a idade da criança, perguntava-se a mãe sobre interações recomendadas, tais como contar histórias, ler livros, cantar, ensinar as partes do corpo humano, nomes dos alimentos, desenhar e participar de atividades da casa. Para crianças de 30-36 meses, o escore máximo era de 38 pontos, e para aquelas com mais de 36 meses, de 30 pontos. Para este último grupo, o escore foi multiplicado por 38/30 para harmonização com o grupo mais jovem antes das análises de regressão.

As práticas parentais de disciplina foram avaliadas usando um subconjunto de perguntas de duas escalas validadas, a *Parent-Child Conflict Tactics Scales*³⁰ e o *Juvenile Victimization Questionnaire*³¹.

O desenvolvimento infantil foi avaliado nas visitas T0, T1 e T3, por meio do *Ages and Stages*

Questionnaire (ASQ3)³². O ASQ3 inclui 30 itens de cinco domínios: cognitivo (resolução de problemas), comunicação, coordenação motora ampla e fina e pessoal social. Cada domínio contém seis questões respondidas pela mãe em relação a metas de desenvolvimento, com três opções possíveis: ainda não (0), às vezes (5 pontos) e sim (10 pontos). Os itens investigados são específicos para a idade, por exemplo, o questionário T3 incluía módulos para idades médias de 36, 42 e 48 meses. Para cada faixa de idade, o escore máximo, incluindo as 30 questões, soma 300 pontos.

A versão tradicional do ASQ3 utiliza somente o relato do respondente sobre as habilidades da criança. Esta versão foi empregada nas três fases do estudo. Na tentativa de aumentar a objetividade dos resultados, a equipe de pesquisa montou uma versão adicional do ASQ3 usando exatamente as mesmas metas de desenvolvimento da versão tradicional, mas baseada na observação, pelos entrevistadores, da realização de tarefas pela criança. Cada entrevistador recebeu objetos, como brinquedos, roupas infantis, itens de higiene pessoal, papel e lápis, e as crianças eram solicitadas a realizar cada tarefa do ASQ3 enquanto eram observadas. Quando a criança não conseguia completar a tarefa, perguntava-se ao respondente e sua informação era registrada. Isto levou a obtenção de dois tipos distintos de escore para cada criança, o tradicional (ou relatado) e o escore observado, sendo ambos analisados separadamente. O coeficiente de correlação de Pearson entre os dois escores foi 0,858 ($p < 0,001$). Nas análises de dados, foram excluídas crianças com alguma deficiência mental (conforme o respondente) e aquelas com escores abaixo de 30 pois a distribuição era bimodal, com um primeiro pico entre 0 e 30, indicando crianças que não responderam a praticamente nenhum quesito. No total, houve 16 exclusões para o teste relatado (8 de cada um dos grupos intervenção e controle) e 292 para o teste observado (145 do grupo intervenção e 147 do grupo controle). A diferença nos números está relacionada ao fato de que o teste observado foi mais exigente do que o relatado. Apenas três das crianças excluídas do grupo intervenção e uma das crianças excluídas do grupo controle haviam recebido visita do PCF nas quatro semanas anteriores à entrevista.

Uma avaliação de atributos psicológicos da criança, por meio do *Strengths and Difficulties Questionnaire* (SDQ) fez parte do questionário T3. Este instrumento de rastreamento, desenvolvido no Reino Unido e validado no Brasil^{33,34}, contém cinco escalas, cada uma com cinco itens:

sintomas emocionais, problemas de conduta, hiperatividade e desatenção, problemas de relacionamento com os pares e comportamentos pró sociais. Cada item era pontuado entre 0 e 2, e três categorias (geradas pela soma dos resultados de todas as escalas, exceto a escala de sociabilidade - pró-social) foram definidas, indicando prevalência crescente de sintomas: 0-13, 14-16 e 17 ou mais pontos. Para as análises de regressão foi utilizada a variável contínua de 0-40 pontos.

Nas rodadas T1 e T3, aproximadamente 10% das díades mãe-criança foram filmadas durante cinco minutos, enquanto realizavam duas atividades simples, visando documentar a interação responsiva entre mães e filhos. Depois de receber uma imagem impressa e vários blocos coloridos, a mãe trabalhava com a criança, para reproduzir duas imagens com os blocos. Cada vídeo foi avaliado e pontuado separadamente por duas psicólogas, em Pelotas, as quais desconheciam o status de intervenção ou controle da criança. O teste foi recentemente validado no Canadá e no Brasil^{35,36}.

Entrevistas telefônicas

Para minimizar o impacto da pandemia no delineamento da avaliação, a visita T2 foi substituída em parte por cinco entrevistas telefônicas com as famílias, durante 2020 e 2021. Utilizando os números de telefone coletados nas visitas T0 e T1, foram feitas tentativas de contato com todas as famílias. O objetivo das entrevistas telefônicas era atualizar o endereço e informações de contato e avaliar a participação no PCF, por meio de visitas domiciliares ou outros tipos de interações com os visitantes do PCF, tais como mídia social ou telefonemas.

Estudo de implementação

O estudo de implementação foi realizado paralelamente às visitas domiciliares anuais, exceto durante a fase T2, quando foi conduzido por telefone. Os supervisores da equipe estadual de avaliação realizaram a coleta de dados. Cada rodada incluía entrevistas com informantes chave da coordenação estadual e municipal do PCF, que também completavam formulários padronizados com informações sobre o programa. Os responsáveis pelo PCF nos CRAS em cada município também eram entrevistados; estes eram encarregados da implementação do programa em suas áreas de abrangência. Três tipos de atividades eram realizadas com os visitantes domiciliares de forma anônima: um teste de conhecimentos,

um questionário autoaplicado sobre suas impressões quanto a aspectos positivos e negativos do programa e um grupo focal. Em cada município, em média dez visitas domiciliares do PCF eram observadas pelo supervisor da equipe de avaliação, utilizando uma *checklist* com os conteúdos esperados de serem abordados, de acordo com a idade da criança. Finalmente, após a conclusão das entrevistas da rodada T3, uma sessão de grupo focal, via *web*, foi realizada pela equipe central do estudo, incluindo os seis supervisores estaduais da avaliação, na qual foram discutidas suas impressões sobre os aspectos fortes e as fragilidades do PCF.

Base de dados do Ministério sobre o número de visitas ou contatos

A base de dados de monitorização do PCF incluía o número de visitas ou contatos virtuais (doravante referidos como “contatos”) com cada criança, de janeiro de 2019 a dezembro de 2021. Esta informação foi emparelhada com a base de dados do estudo por meio dos Números de Identificação Social das crianças.

Processamento e análise de dados

Os dados das entrevistas domiciliares e dos telefonemas, bem como do estudo de implementação, eram digitados no campo utilizando um *tablet* programado com o *software* REDCap³⁷ e enviados diariamente por meio de internet ao servidor central de dados localizado na UFPEL. Cheques de amplitude, consistência e lógica condicional eram feitos inicialmente no campo, utilizando o *software*, e posteriormente em Pelotas, quando inconsistências que ainda persistissem eram retornadas às equipes estaduais. O programa Stata 17.0 (StataCorp®, College Station, TX, USA) foi usado na limpeza, recodificação e análises estatísticas.

Análises por intenção de tratar constituíram a estratégia primária de análise. As crianças inicialmente alocadas no grupo intervenção e controle permaneceram no mesmo grupo, independente da adesão à randomização original.

Dois tipos de análises adicionais foram realizadas. Usando análises com emparelhamento por escores de propensão (*propensity score matching*, ou PSM, em inglês), 445 crianças com 30 ou mais contatos na base de dados de monitorização do PCF foram emparelhadas a crianças que não receberam nenhum contato. As variáveis de emparelhamento incluíram idade, município,

escores do ASQ3 na linha de base, quintis de renda, características maternas (escolaridade, cor, trabalho fora de casa e depressão) e o número de crianças no domicílio.

Por fim, foram realizadas análises com a metodologia de variáveis instrumentais (*instrumental variables* ou IV, em inglês), para estimar o impacto adicional de cada visita do programa sobre o desenvolvimento infantil. Uma vez que a seleção das crianças para o grupo intervenção através de sorteio é aleatória e está associada a um aumento do número médio de contatos, a atribuição do tratamento pode ser utilizada como variável instrumental para o número de visitas recebidas. Assumindo a ausência de efeitos de transbordamento (*spillovers*, em inglês) e que a alocação da criança para o grupo intervenção não cause uma redução do número de contatos (hipótese de monotonicidade), pode-se interpretar as estimativas de IV como o efeito médio (local) de cada visita sobre os desfechos observados entre as crianças que tiveram o número de visita aumentado, ao serem sorteadas para o grupo intervenção (*Local Average Treatment Effects* ou LATE, em inglês).

Dessa forma, usou-se atribuição de tratamento do sorteio para prever o número de contatos e, em seguida, os valores preditos foram usados para estimar o impacto de cada contato adicional sobre os desfechos, após ajuste para idade e município. As estimativas de variáveis instrumentais foram calculadas inicialmente utilizando o número de contatos obtidos na base de dados de monitorização do PCF e, nas análises de sensibilidade, usando o número de contatos calculados a partir das entrevistas e telefonemas do estudo.

A amostra do estudo foi descrita por meio de frequências absolutas e relativas, para variáveis categóricas, e médias e desvios padrão (DP) para variáveis contínuas. Os grupos intervenção e controle foram comparados usando testes qui-quadrados para frequências relativas, e teste *t* de Student, para médias. Modelos de regressão linear brutos e ajustados para covariáveis foram utilizados nas análises de intenção de tratar.

Resultados

Estudo de impacto

O estudo de linha de base ou T0 incluiu 3.242 crianças de 0-11 meses de idade, sendo em média 108 por município. Dessas, 1.623 foram randomizadas para o grupo intervenção e 1.619 para

o grupo controle. A média de idade das crianças era 7,3 meses (DP=2,9). A Tabela 1 mostra as características sociodemográficas e de cuidados de saúde da amostra. A mãe era a cuidadora principal para mais de 90% das crianças. Cerca de 87% das famílias tinham água encanada dentro de casa e 78% dos domicílios tinham sanitário com descarga. Um quinto das moradias tinham mais de três pessoas por quarto de dormir. A idade média das mães era 26 a 27 anos, um pouco mais de um terço delas não viviam com o pai da criança e aproximadamente três quartos autodeclararam-se de cor parda. Cerca de 10% das mães tinham quatro anos ou menos de escolaridade. Praticamente todas as mães fizeram acompanhamento pré-natal, com oito consultas em média. Praticamente todos os partos foram hospitalares e cerca da metade, por cesariana. Não houve diferença entre os grupos quanto a presença de sintomas depressivos maternos. Cerca de 9% das crianças eram pré-termo e 5% tiveram baixo peso ao nascer.

A randomização resultou em dois grupos altamente comparáveis (Tabela 1). Dos 50 indicadores estudados, apenas um apresentou diferença estatisticamente significativa entre os dois grupos: no grupo intervenção, 2,2% foram deixados sozinhos por uma hora ou mais na semana anterior a entrevista, em comparação a 1,1% no grupo controle ($p=0,001$). Resultados detalhados da comparabilidade na fase T0 encontram-se na Tabela 1 e em publicação anterior²⁴.

O ASQ3 foi aplicado durante o estudo de linha de base. Como mostra a Tabela 2, os escores médios foram de 248,1 em ambos os grupos. Informações detalhadas sobre estes resultados estão disponíveis em publicação anterior²³. Em uma reanálise, incluindo os escores das 292 crianças excluídas porque apresentavam ASQ3 observado inferior a 30, as médias dos dois grupos (intervenção e controle) diminuíram mas permaneceram praticamente iguais e não significativas.

O primeiro acompanhamento (T1) ocorreu entre setembro de 2019 e janeiro de 2020. Um total de 3.008 crianças (92,8% dos incluídos no T0) foram localizadas, sendo 1.482 do grupo intervenção e 1.526 do grupo controle. A média de idade foi de 18,6 meses (DP=3,5). De acordo com o relato materno (Tabela 3), somente 31,3% das crianças do grupo intervenção foram visitadas no mês anterior e 24,1% das mães não lembravam nenhuma recomendação que houvessem recebido do visitador do PCF. Houve evidência de contaminação do grupo controle, com 12,9% das crianças havendo sido incluídas no PCF e

Tabela 1. Comparabilidade entre os grupos intervenção e controle na linha de base, de acordo com indicadores selecionados.

	Controle (%)	PCF (%)	P
Características do domicílio			
Tipo de construção			0,945
Tijolos/edifício	93,2	93,1	
Outros	6,8	6,9	
Água encanada			0,176
Não	8,0	7,2	
Sim, dentro de casa	87,3	86,7	
Sim, no quintal	4,8	6,1	
Privada			0,768
Não tem	2,0	2,0	
Sanitário com descarga	79,4	78,3	
Sanitário sem descarga	18,7	19,6	
Casinha/Latrina			
Número de crianças menores de 7 anos (exceto a criança índice)			0,266
0	24,4	25,5	
1	57,1	58,6	
≥2	18,5	15,9	
Aglomerado (>3 pessoas por dormitório)	22,0	23,1	0,482
Respondente do questionário			0,642
Mãe	98,6	98,8	
Pai/cuidador	1,4	1,2	
Nível socioeconômico e características paternas			
Nível socioeconômico da família			0,754
Quintil mais pobre	20,6	19,5	
Segundo quintil	20,2	19,8	
Terceiro quintil	19,8	20,2	
Quarto quintil	19,1	21,0	
Quintil mais rico	20,3	19,6	
Idade do pai (média e DP)	30,5 (8,6)	30,7 (8,8)	0,476
Cor da pele (pai)			0,206
Branca	24,2	23,9	
Preta	8,7	10,6	
Parda	65,2	64,3	
Outras	1,8	1,3	
Escolaridade do pai (média e DP)	8,4 (3,6)	8,5 (3,8)	0,621
Pai trabalha atualmente			0,601
Não	26,7	25,3	
Todos os dias	49,7	51,3	
Alguns dias	23,5	23,4	
Recebe bolsa-família	98,0	98,1	0,898

continua

Tabela 1. Comparabilidade entre os grupos intervenção e controle na linha de base, de acordo com indicadores selecionados.

	Controle (%)	PCF (%)	P
Características maternas			
Idade (média e DP)	26,7 (6,8)	26,9 (6,6)	0,466
Cor da pele			0,257
Branca	15,5	13,6	
Preta	10,1	10,0	
Parda	71,8	73,0	
Outras	2,6	3,4	
Escolaridade (anos)			0,717
0-4	8,8	9,7	
5-8	30,9	30,6	
≥ 9	60,3	59,7	
Escolaridade da mãe (média e DP)	9,3 (3,2)	9,3 (3,2)	0,829
Vive com marido ou companheiro	63,5	62,6	0,585
Trabalha fora de casa			0,599
Não	89,9	90,7	
Todos os dias	4,2	3,5	
Alguns dias	5,9	5,8	
Tabagismo	7,9	7,8	0,844
Depressão materna (Teste de Edimburgo ≥10)	25,4	27,6	0,171
Depressão materna (Teste de Edimburgo ≥13)	14,7	14,4	0,880
Características da gestação da criança			
Paridade materna (excluindo a criança índice)			0,057
0	32,7	28,6	
1	30,9	33,3	
2	20,7	22,8	
≥3	15,7	15,3	
Consultou para exame pré-natal	98,9	99,0	0,734
Número de consultas pré-natais (média e DP)	7,8 (3,0)	7,9 (3,0)	0,085
Parto hospitalar	99,7	99,5	0,580
Parto vaginal	53,1	52,9	0,972
Características dos recém-nascidos			
Sexo feminino	50,2	50,1	0,972
Prematuro (<37 semanas de idade gestacional)	8,3	8,6	0,843
Peso ao nascer (média e DP)	3.271 (509)	3.266 (513)	0,575
Baixo peso ao nascer (<2500 gramas)	6,0	6,1	0,941
Apgar no 5º minuto <7	1,1	1,2	0,699

continua

Tabela 1. Comparabilidade entre os grupos intervenção e controle na linha de base, de acordo com indicadores selecionados.

	Controle (%)	PCF (%)	p
Características das crianças			
Grupos de idade em meses			0,990
<3	9,2	9,0	
3 a 4	12,1	12,0	
5 a 6	21,0	21,5	
7 a 8	25,2	25,0	
9 a 10	23,5	23,0	
11 a 12	8,9	9,5	
Mama no peito	75,1	76,6	0,345
Características de cuidados, interação e promoção do desenvolvimento			
Principal cuidador da criança			0,559
Mãe	93,0	93,4	
Pai	0,4	0,4	
Avó/avô	3,0	3,3	
Mãe e pai	2,7	1,9	
Outro	1,0	1,0	
Criança frequente creche	1,4	1,3	1,000
Criança ficou sozinha ≥1 hora na última semana	1,1	2,2	0,012
Criança foi cuidada ≥1 hora por outra criança (≤10 anos) na última semana	0,8	0,7	0,838
Possui livro ou revistinha	8,7	10,2	0,183
Escore de estimulação e interação da mãe/cuidador (0-5 meses) (média e DP)	12,8 (2,3)	12,8 (2,3)	0,749
Escore de estimulação e interação da mãe/cuidador (6-12 meses) (média e DP)	14,9 (2,6)	14,8 (2,7)	0,926
Escore de atividades promotoras do desenvolvimento (últimos 3 dias) (média e DP)	3,4 (1,3)	3,4 (1,3)	0,950
Emprego de métodos inadequados de disciplina (último mês) (média e DP)	1,1 (0,8)	1,1 (0,8)	0,560
Concorda com a afirmativa: “para educar e criar bem uma criança, a criança não precisa apanhar”	95,1	94,3	0,308

Estimulação e interação da mãe/cuidador (0-5 meses): máximo 16 pontos. Estimulação e interação da mãe/cuidador (6-12 meses): máximo 20 pontos. Atividades promotoras do desenvolvimento (últimos 3 dias): máximo 6 pontos. Métodos inadequados de disciplina (último mês): máximo 7 pontos.

Fonte: Autores.

7,3% sido visitadas no mês anterior. Estes resultados preliminares foram reportados ao Ministério em fevereiro de 2020. Como mostra a Tabela 2, o escore médio do ASQ3 foi 2,4 pontos maior no grupo intervenção do que no controle, mas a diferença não foi estatisticamente significativa ($p=0,145$). Esta diferença equivale a cerca de 0,05 DP. Ignorando a alocação aleatória e utilizando emparelhamento por escores de propensão, uma análise adicional comparou todas as 565 crianças que haviam sido visitadas no mês anterior à entrevista com um grupo similar de 565 crianças não inscritas no programa. Nesta ocasião, o grupo visitado apresentou escore médio ligeiramente superior ao do grupo não visitado (233,8 pontos *versus* 228,5 pontos; $p=0,043$).

A Tabela 3 mostra ainda os resultados das entrevistas telefônicas sobre a participação no PCF durante a pandemia, quando as visitas domiciliares foram marcadamente reduzidas e os contatos virtuais introduzidos. Estes incluíam chamadas via internet, mensagens e vídeos, bem como telefonemas dos visitantes para as famílias. Cerca de 30% das famílias das crianças do grupo intervenção e de 10% daquelas do grupo controle relataram haver recebido tais contatos.

Com o cancelamento das entrevistas domiciliares do T2 devido à pandemia, a terceira e última visita (T3) ocorreu entre outubro de 2021 a janeiro de 2022. Um total de 2.594 entrevistas foram realizadas, com uma taxa geral de acompanhamento de 80,0% (79,9% no grupo intervenção e 80,1% no grupo controle). A média de idade das crianças foi de 42,7 meses (DP=3,5) no grupo controle e 42,8 meses (DP=3,4) no grupo intervenção ($p=0,790$).

As taxas de acompanhamento conforme características de linha de base são mostradas na Tabela 4. Embora as taxas de acompanhamento para a amostra como um todo tenham variado significativamente de acordo com a idade das crianças (maior entre as crianças mais novas), cor da pele da mãe (maior entre as mães de cor parda), condições socioeconômicas (maior nos domicílios mais pobres), escolaridade materna (maior entre aquelas com nove anos ou mais de escolaridade) e depressão materna (menor entre mães deprimidas), pelo menos 69% das crianças de todos os grupos foram acompanhadas. Estas variáveis foram incluídas como potenciais confundidores nos modelos de regressão do ASQ3. Em particular, não houve evidência estatística de diferenças entre os grupos intervenção e controle em relação às taxas de acompanhamento dentro de qualquer categoria das variáveis acima mencionadas.

Tabela 2. Resultados brutos do ASQ3 na análise por intenção de tratar na linha de base e no primeiro e terceiro estudos de acompanhamento*.

	Controle			Intervenção			P
	N	Média	DP	N	Média	DP	
ASQ-3 (relatado)							
T0	1505	248,1	40,2	1521	248,1	37,7	0,953
T1	1518	226,2	44,5	1479	228,6	44,2	0,145
T3	1245	219,8	50,6	1253	221,3	49,4	0,453
ASQ-3 (observado)							
T3	1105	201,3	58,2	1108	203,3	57,1	0,427

*Crianças com escore ASQ3 menor de 30 pontos e aquelas que apresentavam deficiência mental relatada foram excluídas; isto foi mais frequente com a escala observada do que com a relatada.

Fonte: Autores.

Tabela 3. Percentual de crianças do grupo intervenção (GI) e do grupo controle (GC) que receberam visitas ou contatos do PCF durante o mês anterior à entrevista.

Percentual de crianças	Entrevistas domiciliares (T1)				Entrevistas telefônicas						Entrevistas domiciliares (T3)			
	Setembro 2019- Janeiro 2021		Junho 2020		Setembro 2020		Janeiro 2021		Abril 2021		Julho 2021		Setembro 2019- Janeiro 2021	
	GC	GI	GC	GI	GC	GI	GC	GI	GC	GI	GC	GI	GC	GI
Visitas do PCF	7,3	31,3	1,0	3,8	6,9	28,2	4,7	17,8	4,2	11,4	2,8	9,0	0,8	6,1
Contatos virtuais do PCF	0,0	0,0	5,7	24,8	14,0	31,9	15,4	25,3	15,7	32,4	13,4	27,7	17,3	35,5
Visitas e/ou contatos	7,3	31,3	5,8	25,0	14,1	32,6	15,4	26,1	16,4	35,6	13,9	29,4	16,8	35,4
Número de crianças	1.482	1.526	1.022	1.022	897	891	894	905	770	765	820	856	1.300	1.294

Fonte: Autores.

Seis desfechos foram analisados: escores de estimulação/interação; disciplina das crianças; interações responsivas (baseadas nos vídeos); atributos psicológicos; e resultados do teste ASQ3 observado e relatado. Primeiro são apresentadas as médias e proporções dos indicadores de desfecho usando análise por intenção de tratar, sem qualquer ajuste (Tabela 5). Os escores de estimulações/interações foram muito semelhantes nos grupos intervenção e controle, para ambas as faixas de idade estudadas. Os escores médios atingiram cerca de 80% dos valores máximos da escala. Os escores de interação responsiva derivados das gravações de vídeos obtidas para uma subamostra foram relativamente baixos, com valores médios de 2,1 em um máximo de 5,0, em ambos os grupos. O uso de métodos inadequados de disci-

plina foi relatado um pouco mais frequentemente no grupo controle (média de 3,5 métodos em um total de 12), do que no grupo intervenção (3,4 métodos). As distribuições dos atributos psicológicos foram similares em ambos os grupos.

Como mostra a Tabela 2, os valores médios do ASQ3 observado e relatado na etapa T3 foram similares nos grupos intervenção e controle, com valores médios em torno de 202 em 300 pontos possíveis. Conforme esperado, os escores relatados foram mais elevados do que os observados.

Os resultados não ajustados acima apresentados foram complementados por análises baseadas em regressões por intenção de tratar e LATE (Tabela 6) que incluíram ajuste para idade da criança e município. Os coeficientes beta nas análises por intenção de tratar mostram as diferenças no

Tabela 4. Percentuais de crianças localizadas no estudo T3, conforme características medidas na linha de base do estudo (T0).

Grupo	Controle	Intervenção	Total	(Intervenção*Controle) P
Sexo da criança				
Meninos	79,1%	78,6%	78,8%	0,798
Meninas	81,1%	81,2%	81,2%	0,927
p intragrupos	0,320	0,180	0,099	
Idade da criança (T0)				
<5 meses	85,3%	83,8%	84,5%	0,603
5-6 meses	77,7%	77,9%	77,8%	0,959
7-8 meses	79,0%	80,0%	79,5%	0,717
9-10 meses	78,0%	77,7%	77,9%	0,915
11-12 meses	82,1%	81,2%	81,6%	0,841
p intragrupos	0,068	0,245	0,008	
Cor da pele materna				
Branca	76,5%	81,0%	78,6%	0,239
Preta	75,2%	69,4%	72,3%	0,247
Parda	82,2%	81,2%	81,7%	0,549
p intragrupos	0,024	0,002	0,000	
Quintis de riqueza				
Quintil mais pobre	80,4%	79,9%	80,2%	0,874
Segundo quintil	85,3%	82,7%	84,0%	0,378
Terceiro quintil	80,0%	83,4%	81,7%	0,259
Quarto quintil	77,5%	76,7%	77,1%	0,791
Quintil mais rico	77,3%	77,2%	77,2%	0,979
p intragrupos	0,076	0,098	0,006	
Escolaridade materna (anos)				
0-4	78,8%	78,1%	78,4%	0,886
5-8	73,9%	78,5%	76,2%	0,096
≥9	83,1%	80,0%	81,6%	0,085
p intragrupos	0,000	0,753	0,004	
Depressão materna				
Não	81,5%	80,4%	81,0%	0,504
Sim	75,9%	79,0%	77,5%	0,287
p intragrupos	0,016	0,526	0,032	

Fonte: Autores.

Tabela 5. Análises não ajustadas dos desfechos chave, por intenção de tratar, nos grupos controle e intervenção no acompanhamento T3.

	Escore máximo	Controle		Intervenção		P
		Média (DP)		Média (DP)		
Escore de estimulação e interação (30-36 m)	38	28,4	(4,4)	27,7	(5,0)	0,541
Escore de estimulação e interação (>36 m)	30	24,8	(3,9)	25,0	(4,1)	0,349
Escore de interação responsiva	5	2,1	(0,7)	2,1	(0,8)	0,974
Métodos inadequados de disciplina	8	1,6	(1,5)	1,5	(1,4)	0,066
Atributos psicológicos (pontos)		%		%		
0-13		57,9		60,2		0,358
14-16		15,5		13,7		
≥17		26,6		26,1		

Fonte: Autores.

Tabela 6. Análises por intenção de tratar e LATE (todas ajustadas para idade da criança e município).

Modelo	Beta	IC95%	P
Escore de estimulação/interação			
ITT	0,092	-0,215 0,399	0,558
LATE	0,004	-0,010 -0,018	0,562
Escore de interações responsivas			
ITT	-0,008	-0,180 0,160	0,926
LATE	-0,00086	-0,0084 0,0066	0,822
Escore de métodos inadequados de disciplina			
ITT	-0,060	-0,247 0,126	0,524
LATE	-0,004	-0,009 0,001	0,101
Escore de atributos psicológicos			
ITT	-0,417	-0,950 0,112	0,122
LATE	-0,019	-0,040 0,010	0,133
Escore do ASQ3 observado			
ITT*	2,62	-2,16 7,41	0,282
LATE	0,08	-0,14 0,30	0,485
Escore do ASQ3 relatado			
ITT*	1,87	-2,04 5,78	0,348
LATE	0,08	-0,10 0,26	0,373

*Ajustado para ASQ3 na linha de base e covariáveis.

Fonte: Autores.

desfecho entre os grupos intervenção e controle, enquanto nas análises LATE, os coeficientes beta correspondem à mudança no desfecho associada com cada contato adicional do PCF.

Nas análises ajustadas, os escores de estimulação/interação para os dois grupos de idade (Tabela 6) foram expressos na mesma escala (ver Métodos) e depois agrupados. Nem as análises por intenção de tratar, nem as análises LATE mostraram qualquer evidência de diferença entre os grupos intervenção e controle. Da mesma forma, não foi observada nenhuma diferença nas análises por intenção de tratar e LATE quanto aos escores de interações responsivas, métodos de disciplina ou atributos psicológicos.

As análises de regressão dos desfechos principais do estudo – os dois escores ASQ3 – foram também ajustadas para o valor do ASQ3 na linha de base e para as variáveis acima descritas, que se associaram às taxas de acompanhamento no T3. Nem as análises por intenção de tratar, nem as análises LATE mostraram diferenças estatisticamente significativas (Tabela 6). Para confirmar estes resultados, foram realizadas análises com emparelhamento por escores de propensão, comparando 445 pares de crianças com 30 ou mais contatos durante 2019-2021 (de acordo com a base de dados do PCF) com crianças similares que não receberam nenhum contato. Nestas análises,

o beta estimado para o ASQ3 observado foi -2,44 (IC95% -9,62;4,74, $p=0,506$), ou seja, o escore médio entre as crianças com 30 ou mais contatos foi levemente menor, mas não estatisticamente significativo, do que entre as crianças que não receberam nenhum contato.

Para explorar os efeitos das visitas presenciais durante 2019, os dois grupos de 565 crianças (com visitas no último mês e nunca visitadas antes da etapa T1), descritos acima, foram avaliados novamente na etapa T3. Nessa ocasião, os escores médios para o teste ASQ3 observado foram de 202,7 e 201,2 ($p=0,712$). Portanto, não houve evidência de um efeito duradouro de visitas presenciais em 2019 sobre o desempenho ao final de 2021.

Foram também realizadas análises separadas para cada um dos domínios do ASQ3 (Tabelas 7 e 8), mas não foram encontradas diferenças entre os grupos intervenção e controle em nenhuma dessas análises.

A ausência de impacto nas análises para o ASQ3 levou à investigação da correlação entre o escore observado do ASQ3 e o número de contatos durante 2019-2021 de acordo com a base de dados do PCF. A Figura 2, em que cada criança é representada por um ponto, independente da alocação por randomização, mostra que não houve associação, com um coeficiente de correlação

Tabela 7. Escores médios nos cinco domínios do teste ASQ-3 relatado nos grupos intervenção e controle, por faixa etária, excluindo 96 crianças com deficiência e 317 crianças com escore total inferior a 30 pontos:

	Controle			Intervenção			P
	N	Média	(DP)	N	Média	(DP)	
ASQ-3 Relatado	(36 meses)						
Total (0-300)	203	223,3	(49,7)	191	218,6	(52,1)	0,367
Comunicação (0-60)	203	43,1	(13,8)	191	42,6	(13,6)	0,727
Coordenação motora ampla (0-60)	203	52,4	(11,6)	191	51,0	(11,1)	0,242
Coordenação motora fina (0-60)	203	43,8	(16,0)	191	42,4	(16,5)	0,398
Resolução de problemas (0-60)	203	39,4	(15,8)	191	39,2	(15,8)	0,916
Pessoal-social (0-60)	203	44,6	(12,7)	191	43,4	(13,9)	0,351
ASQ-3 Relatado	(42 meses)						
Total (0-300)	672	219,2	(51,3)	687	222,7	(46,7)	0,190
Comunicação (0-60)	672	42,0	(15,0)	687	42,1	(14,3)	0,890
Coordenação motora ampla (0-60)	672	51,0	(11,0)	687	51,4	(11,1)	0,511
Coordenação motora fina (0-60)	672	38,1	(15,2)	687	39,1	(14,2)	0,233
Resolução de problemas (0-60)	672	39,1	(16,0)	687	40,6	(15,4)	0,077
Pessoal-social (0-60)	672	49,1	(12,5)	687	49,6	(11,0)	0,405
ASQ-3 Relatado	(48 meses)						
Total (0-300)	370	218,8	(50,0)	375	219,9	(52,8)	0,763
Comunicação (0-60)	370	46,4	(15,2)	375	47,3	(15,6)	0,432
Coordenação motora ampla (0-60)	370	51,5	(11,4)	375	52,5	(10,3)	0,187
Coordenação motora fina (0-60)	370	34,9	(17,0)	375	34,2	(17,2)	0,611
Resolução de problemas (0-60)	370	38,0	(15,0)	375	38,5	(15,5)	0,659
Pessoal-social (0-60)	370	48,1	(10,7)	375	47,4	(11,4)	0,418

Fonte: Autores.

Tabela 8. Escores médios nos cinco domínios do teste ASQ-3 observado nos grupos intervenção e controle, por faixa etária, excluindo 96 crianças com deficiência e 317 crianças com escore total inferior a 30 pontos.

	Controle			Intervenção			P
	N	Média	(DP)	N	Média	(DP)	
ASQ-3 Observado	(36 meses)						
Total (0-300)	178	205,1	(52,6)	170	197,4	(62,0)	0,208
Comunicação (0-60)	178	40,1	(14,4)	170	39,4	(15,6)	0,664
Coordenação motora ampla (0-60)	178	49,8	(13,9)	170	48,2	(14,1)	0,288
Coordenação motora fina (0-60)	178	44,0	(15,7)	170	41,2	(17,8)	0,112
Resolução de problemas (0-60)	178	36,0	(17,0)	170	34,9	(17,1)	0,538
Pessoal-social (0-60)	178	35,1	(14,8)	170	33,6	(16,1)	0,377
ASQ-3 Observado	(42 meses)						
Total (0-300)	596	204,7	(59,1)	614	207,0	(54,6)	0,494
Comunicação (0-60)	596	38,5	(16,5)	614	38,3	(15,7)	0,871
Coordenação motora ampla (0-60)	596	47,6	(13,8)	614	48,3	(13,5)	0,369
Coordenação motora fina (0-60)	596	37,2	(15,8)	614	38,0	(15,0)	0,397
Resolução de problemas (0-60)	596	37,6	(16,6)	614	38,3	(16,1)	0,431
Pessoal-social (0-60)	596	43,8	(14,7)	614	44,0	(14,1)	0,812
ASQ-3 Observado	(48 meses)						
Total (0-300)	331	193,2	(58,8)	324	199,4	(58,8)	0,177
Comunicação (0-60)	331	41,1	(17,3)	324	42,6	(17,6)	0,289
Coordenação motora ampla (0-60)	331	49,3	(13,8)	324	50,8	(11,8)	0,137
Coordenação motora fina (0-60)	331	31,8	(17,8)	324	32,0	(18,0)	0,890
Resolução de problemas (0-60)	331	32,4	(15,4)	324	34,2	(16,0)	0,147
Pessoal-social (0-60)	331	38,6	(15,3)	324	39,9	(16,3)	0,294

Fonte: Autores.

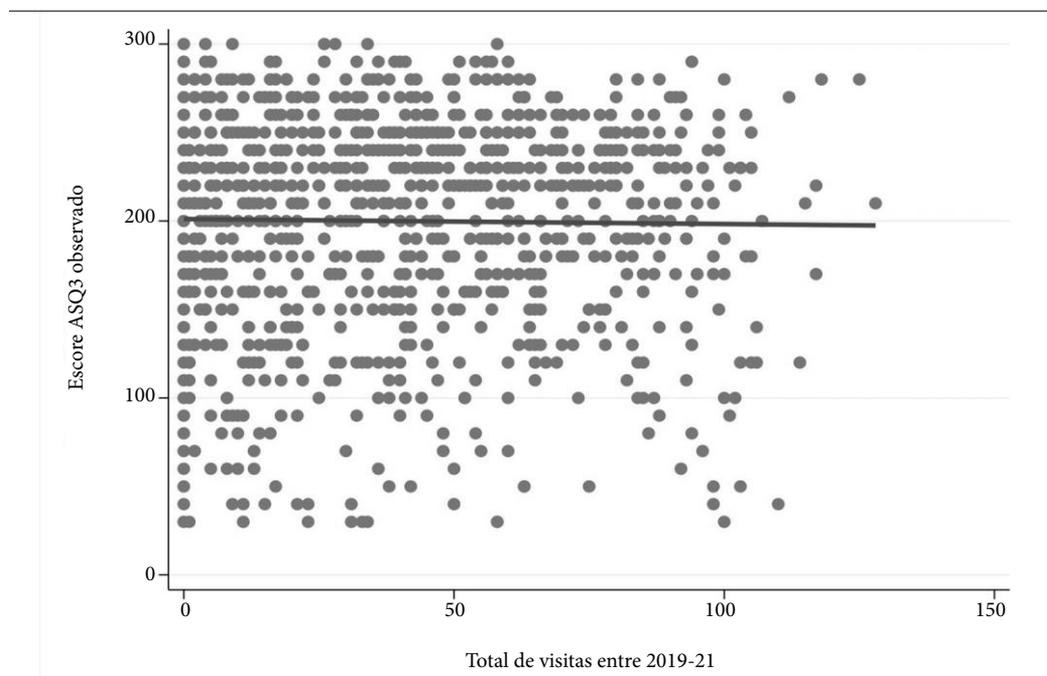


Figura 2. Diagrama de dispersão entre o escore do ASQ3 observado no estudo T3 e o número de contatos durante 2019-2021, de acordo com a base de dados do PCF.

Fonte: Autores.

de Pearson de $-0,0133$ ($p=0,436$). Uma vez que a base de dados do PCF não discrimina visitas face a face de contatos virtuais, as análises foram repetidas usando informações de 758 crianças cujas mães foram entrevistadas sete vezes: no T1, T3 e em todas as cinco rodadas de entrevistas por telefone. O coeficiente de correlação de Pearson com o ASQ3 observado foi de $0,028$ para visitas face a face e de $0,005$ para contatos virtuais.

Foram realizadas análises de sensibilidade, com restrição da amostra aos municípios com implementação forte do PCF, de acordo com o estudo de implementação; análises de subgrupos, examinando o impacto de acordo com características socioeconômicas, maternas e da criança; e análises LATE, utilizando resultados das entrevistas por telefone para estimar variáveis instrumentais. Nenhuma dessas análises mostrou impactos significativos do programa. Esses resultados são detalhados em outro documento. (<http://www.epidemiologia-ufpel.org.br/site/content/pesquisas/index.php?study=518>).

Estudo de implementação

Resultados detalhados do estudo de implementação estão disponíveis em outros documentos (<http://www.epidemiologia-ufpel.org.br/site/content/pesquisas/index.php?study=518>). Devido à pandemia, as visitas do PCF foram suspensas em todos os municípios, na maioria entre março de 2020 a cerca de junho de 2021. Durante os 36 meses da avaliação (janeiro de 2019 a dezembro de 2021), o número de meses com visitas domiciliares variou de 5 a 35 nos 30 municípios, com uma mediana de 24.

O banco de dados de monitoramento do PCF foi usado para avaliar a frequência de contatos e adesão à randomização original. A Figura 3 mostra dois círculos para cada município, sendo o círculo azul claro o número médio de visitas no grupo controle, e o círculo azul escuro o mesmo indicador para o grupo intervenção. O número de visitas foi pequeno na maioria dos municípios, com média de $6,6$ contatos no grupo controle e $28,2$ no grupo intervenção. Em média, as crianças do estudo completaram 36 meses de idade em junho de 2021, quando deveriam ser desligadas

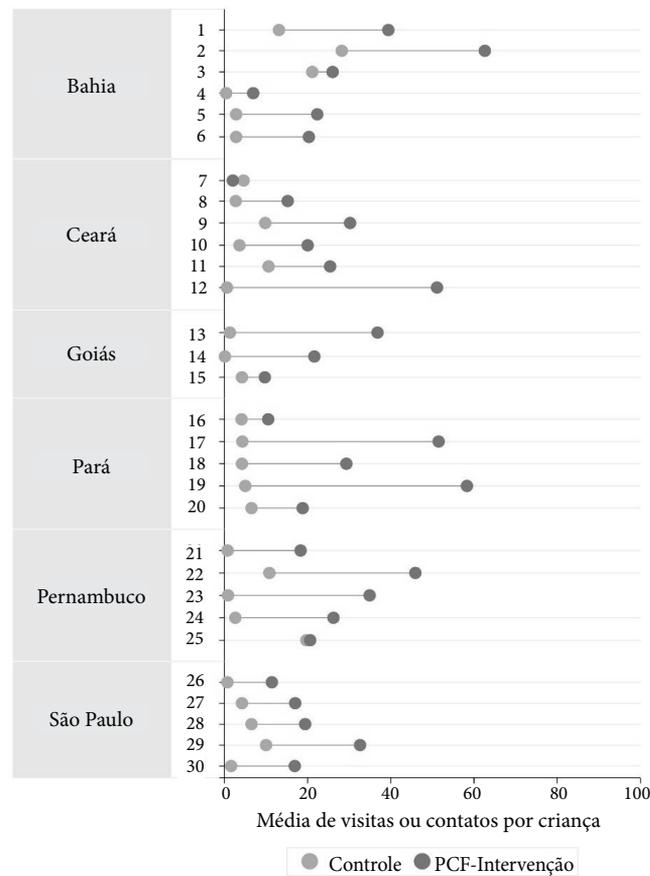


Figura 3. Número médio de visitas ou contatos do PCF com as crianças do estudo, de janeiro de 2019 a dezembro de 2020, de acordo com o grupo de alocação. Cada linha corresponde a um município.

Fonte: Autores.

do PCF. Portanto, o número médio de contatos deveria ser igual a 30, e, se os contatos fossem semanais, de 120.

Observa-se ainda que alguns municípios desconsideraram o esquema de alocação aleatória, uma vez que o número de contatos foi muito semelhante nos grupos intervenção e controle.

A parte superior da Tabela 9 mostra resultados selecionados das entrevistas com os informantes chave e da revisão documental nos seis estados e 28 municípios, uma vez que as atividades do PCF foram interrompidas em dois dos 30 municípios originalmente incluídos no estudo. Os coordenadores estaduais do PCF acreditavam que somente 30% dos municípios apresentavam implementação forte. A alta rotatividade das equipes foi um aspecto crítico, com cerca de metade dos coordenadores municipais sendo substituídos dentro do último ano. Em menos da

metade dos municípios, a maioria dos visitantes domiciliares eram os mesmos desde o início do programa. Em 60% dos municípios, os visitantes eram selecionados por processo seletivo, em vez de por indicação. Quatro ou mais métodos de contato com as famílias durante a pandemia foram empregados em 43% dos municípios, e somente 20% dispunham três ou mais veículos para o programa. O número mediano de famílias por visitante foi igual a 25, inferior ao número de 30 famílias recomendado pelo programa nacional³⁸.

Em cada município, os visitantes realizaram um teste de conhecimentos e completaram escalas de 0 a 10, descrevendo sua satisfação com o programa. Ao todo, 363 visitantes foram avaliados. Foram calculadas as médias destes escores por município, sendo que a Tabela 9 mostra a média das médias municipais. A média dos escores do teste foi de 5,9; o que deve ser interpretado

Tabela 9. Resultados principais do estudo de implementação de 2021 nos 30 municípios.

Crítérios	Percentual
Coordenador Estadual julga ser forte a implementação no município	30%
Coordenador Municipal é o mesmo do T2	47%
A maioria ou todos os visitadores são os mesmos desde o início	44%
Visitadores escolhidos por processo seletivo	60%
Pelo menos 4 meios de contato com as famílias (telefone, mensagem de voz/áudio, mensagem de texto, chamada de vídeo, outro)	43%
Pelo menos 3 veículos para uso do PCF	20%
Escores Médios (escalas de 0 a 10)	
Nota da prova dos visitadores do PCF (média)	5,9
Satisfação dos visitadores com salário	5,3
Satisfação dos visitadores com carga horária	7,3
Satisfação dos visitadores com transporte	5,2

Fonte: Autores.

à luz do fato de que as questões do teste apresentavam baixo grau de dificuldade. Os escores médios do teste por município variou de 4,4 a 7,6.

O salário médio dos visitadores nos 30 municípios era de R\$ 1.103 ou cerca de US\$ 200, variando de R\$ 400 a R\$ 1.450. Contratos temporários, sem benefícios adicionais constituíam o tipo mais prevalente de arranjo contratual (67%). O escore médio de satisfação dos visitadores com o salário foi de 5,3. Quanto à satisfação com o número de horas trabalhadas, o escore médio foi de 7,3. A indisponibilidade de transporte para as visitas foi frequentemente relatada, com um escore médio de 5,2.

A fim de avaliar a fidelidade ao programa, 182 visitas domiciliares foram observadas pelos supervisores do estudo em 22 municípios, utilizando um *checklist* com os conteúdos recomendados pelo PCF para a idade de cada criança. Foi calculado o percentual de itens abordados pelo visitador durante cada visita, em relação aos itens esperados. Em média, somente 9,8% (DP=7,4) dos itens recomendados foram abordados, com médias municipais variando de 1,0% a 71,0%. Não foi possível realizar observações em oito municípios, onde não estavam ocorrendo visitas no momento da pesquisa.

Foram realizados 28 grupos dirigidos (dois municípios estavam sem visitadores por ocasião do estudo de implementação), com três a oito visitadores por grupo. A maioria dos grupos relatou que o PCF foi impactado pela pandemia devido à perda de contato ou a contatos somente remotos com as famílias, à falta de acesso das famílias à internet ou a telefone celular e à perda de

vínculo com as famílias e/ou crianças. Na grande maioria dos municípios, as visitadoras expressaram sua insatisfação com o transporte, salários e a precariedade dos contratos de trabalho, e sua satisfação com a carga horária de trabalho. Outros pontos positivos mencionados pelas visitadoras incluíram a boa receptividade e confiança pela maioria das famílias, o fortalecimento dos vínculos entre mãe e a criança, as contribuições do PCF para o desenvolvimento infantil, os vínculos construídos entre visitadores e crianças, mães e outros membros da família.

Finalmente, foi organizado um grupo focal via Zoom, com todos os supervisores estaduais da pesquisa que haviam realizado o estudo de implementação de 2019 a 2021. Os tópicos discutidos incluíram características das visitas, do treinamento e condições de trabalho dos visitadores, rotatividade e gestão. Os supervisores destacaram fragilidades decorrentes da baixa fidelidade ao conteúdo do programa durante as visitas; observação de muitas visitas sem roteiros pré-definidos e sem levar em conta a idade da criança; visitadores atuando sem ter recebido treinamento específico em alguns municípios, com muitas crianças para atender (mediana de 26, variando de 11 a 37) e percebendo baixos salários; e alta rotatividade entre os coordenadores municipais e visitadores (principalmente em decorrência de eleições municipais, baixa remuneração e tipo de contrato de trabalho), que prejudicam a continuidade do programa. Além disso, os supervisores relataram problemas de coordenação e gestão, iniciando com a falta de fiscalização do PCF pela gestão estadual e municipal. Vários problemas

foram mencionados com respeito ao papel dos CRAS, os quais em muitos municípios estavam sobrecarregados e dissociados do PCF, sem interagir suficientemente com os visitantes. Alguns municípios não contavam com funcionários dentro dos CRAS que atuassem como responsáveis pelo PCF. Apenas em poucos municípios, o supervisor municipal do PCF comparecia regularmente aos CRAS para se reunir com os visitantes. Outros problemas identificados no grupo focal incluíram a falta de registros de dados do PCF disponíveis para a gestão e a falta de vínculo entre o PCF e o Programa Bolsa Família, o que dificulta o entendimento e a participação das famílias. Finalmente, um aspecto extremamente preocupante foi a total ausência de supervisão do trabalho dos visitantes: em nenhuma cidade os supervisores do PCF acompanham visitas domiciliares periodicamente para dar suporte e orientações aos visitantes.

Discussão

Nesta seção serão abordados o delineamento do estudo, os principais resultados dos estudos de impacto e implementação, uma comparação com a literatura existente e recomendações para políticas públicas.

O estudo empregou o delineamento mais rigoroso de avaliação de impacto de programas de saúde pública, uma vez que a randomização resultou em dois grupos de crianças altamente comparáveis e que quaisquer mudanças em variáveis contextuais afetariam ambos os grupos. O fato de que o número de crianças potencialmente elegíveis para o PCF excedia em muito o número de vagas abertas nos municípios selecionados garantiu que a alocação aleatorizada fosse eticamente justificável. O processo de randomização foi bem-sucedido, resultando em alta comparabilidade entre os grupos na linha de base (T0). O tamanho de amostra foi adequado para medir um impacto compatível com o observado na literatura, com mais de 3.000 crianças em 30 municípios de seis estados sendo estudadas. As taxas de acompanhamento foram adequadas (80% após três anos) e similares nos subgrupos de crianças e famílias, indicando um baixo risco de viés devido a perdas de acompanhamento. As análises principais foram realizadas por “intenção de tratar”, como originalmente planejado, e foram complementadas por métodos adicionais.

Outros pontos positivos incluem a natureza longitudinal do delineamento e o uso de múlti-

plos desfechos, baseados em um modelo de impacto definido *a priori*, cobrindo a avaliação de interações parentais, disciplina, atributos psicológicos e desenvolvimento infantil, esse último usando duas versões do teste ASQ3. Além de entrevistas domiciliares face a face, o estudo baseou-se em interações videogravadas entre a mãe e a criança, entrevistas com informantes chave, teste do conhecimento dos visitantes e observação de visitas do PCF.

A pandemia de COVID-19 exigiu mudanças de alguns dos aspectos do delineamento original, incluindo o cancelamento das entrevistas do segundo ano de acompanhamento (T2) e a adoção de entrevistas telefônicas para monitorar a adesão durante 2020 e 2021. Além disso, o Ministério autorizou os municípios a utilizar canais virtuais para continuarem em contato com as famílias. Uma outra mudança diz respeito ao delineamento de efetividade. Os resultados preocupantes do estudo de implementação do PCF no final de 2019 foram comunicados ao Ministério no início de 2020. Naquele momento, os pesquisadores consideraram que não seria ético privar o Ministério desta retroalimentação. Consequentemente, durante os anos de 2020 e 2021, o Ministério empreendeu uma série de medidas para fortalecer a implementação do PCF nos 30 municípios. Tais medidas afetaram o delineamento original do estudo, que consistia em uma avaliação “pura” de efetividade, isto é, sob condições de implementação de rotina. Dentro do *continuum* entre eficácia e efetividade, o reforço das atividades nos 30 municípios – mas não no restante do país – transformou o delineamento do estudo mais próximo de uma avaliação de eficácia do que era o caso do delineamento original de efetividade³⁹.

Possíveis limitações do estudo incluem o fato de que, desde que o PCF foi delineado, estava previsto que as visitas seriam interrompidas aos 36 meses de idade, e parcialmente em função da pandemia, a idade média na etapa T3 foi de 42 meses, quando muitas crianças já haveriam deixado de receber visitas. No entanto, esperava-se um impacto duradouro do programa⁴⁰, que deveria persistir após o encerramento das visitas. Outra limitação é o fato de que as observações de visitas somente puderam ser realizadas em 22 dos 30 municípios, principalmente devido ao fato de estarem as visitas suspensas nos demais oito municípios.

O modelo de impacto¹⁸, construído pelos pesquisadores na fase de planejamento do estudo, com base nas diretrizes do programa^{15,29},

postulava uma cadeia de eventos necessários para que houvesse impacto do programa sobre o desenvolvimento infantil. Essa cadeia, descrita acima na Introdução, norteou a seleção de indicadores e instrumentos da avaliação, visando documentar desde o início da implantação do programa até seu postulado impacto positivo sobre o desenvolvimento infantil e a redução geral da vulnerabilidade^{15,29}.

A combinação de estratégias para medida de processo e impacto é essencial para medir o desempenho de qualquer programa^{19,20}. Durante os três anos de avaliação, todos os elos do modelo de impacto foram avaliados por meio de várias estratégias de coleta de dados. Uma vez que o modelo é uma cadeia causal em que todos os elos estão interconectados em uma ordem lógica e sequencial, o desempenho inadequado em qualquer um deles comprometerá os elos subsequentes.

Os presentes resultados do impacto do PCF sobre os vários desfechos seguiram um processo robusto e consistente entre as diversas estratégias analíticas. As análises por intenção de tratar não mostraram nenhuma diferença significativa entre os grupos intervenção e controle em termos de escores de estimulação, interação e disciplina; interações responsivas videogravadas; atributos psicológicos (SDQ); e resultados do teste ASQ3 observado ou relatado. As análises LATE, usando variáveis instrumentais, bem como as análises com emparelhamento por escores de propensão e vários tipos de análises de sensibilidade tampouco mostraram impacto. Consistentemente com os resultados acima, não houve associação entre o número de visitas ou contatos e os desfechos estudados.

A investigação dos motivos pelos quais não houve impacto está embasada nos estudos de implementação, cujos resultados detalhados estão disponíveis em documentos adicionais (<http://www.epidemiologia.ufpel.org.br/site/content/pesquisas/index.php?study=518>). As principais razões incluem: baixa cobertura e baixa frequência das visitas no grupo intervenção; falta de adesão ao esquema de randomização, resultando em contaminação do grupo controle; interrupção das visitas face a face devido a pandemia; e desafios da implementação. Essas serão discutidas a seguir.

As análises das visitas de acompanhamento e das entrevistas telefônicas, complementadas por resultados da base de dados de monitorização do PCF, revelaram importantes desafios na cobertura do programa no grupo intervenção. De acordo com a base de dados, o número médio de

visitas, nos três anos, foi igual a 28,2, variando de 2,0 a 62,6 – isto é, bem abaixo do número esperado de contatos semanais, e também levemente inferior ao número esperado de visitas mensais, caso estas tivessem sido realizadas. Também foi evidente a contaminação em vários municípios, que aparentemente desconsideraram as listagens da randomização encaminhadas pelo Ministério em 2019 e reforçadas em 2020, após a retroalimentação fornecida pelo estudo. Este achado pode evidenciar desafios na comunicação entre os níveis federal e municipal, apesar de reiterados esforços da equipe de pesquisa para alertar sobre os problemas de implementação que se tornaram evidentes já no primeiro ano da avaliação. Outra possibilidade que não pode ser descartada é a utilização de critérios diversos aos preconizados pelo programa, como por exemplo, critérios político-partidários.

Em um programa em que interações face a face com os cuidadores e crianças é essencial, a justificada suspensão das visitas domiciliares durante a pandemia foi impactante para o programa. Embora o Ministério tenha desenvolvido materiais de aconselhamento para contatos virtuais com as famílias, e a maioria dos municípios tenham se apoiado nesses contatos em maior ou menor grau, estes são substitutos muito fracos de visitas domiciliares. No entanto, os gestores municipais relataram que foram realizadas visitas domiciliares em média durante 24 dos 36 meses de estudo, de forma que algum grau de impacto poderia ser esperado. Note-se ainda que, ao final do primeiro ano do programa – antes da pandemia, quando somente visitas domiciliares eram realizadas – os escores médios do ASQ3 no grupo intervenção foram somente 2,4 pontos ou 0,05 DP maiores no grupo intervenção do que no controle ($p=0,145$).

A baixa cobertura no grupo intervenção, contaminação do grupo controle e a pandemia podem explicar parcialmente a ausência de impacto nas análises por intenção de tratar. No entanto, as análises LATE e o emparelhamento por escores de propensão tampouco mostraram evidência de impacto entre os tratados, ou seja, mesmo entre as crianças que de fato receberam as visitas. A completa falta de associação entre o número de contatos e os escores ASQ3 (Figura 2) foi um achado inesperado que confirma a debilidade da implementação. Isto ficou evidente nos resultados do estudo da implementação, que mostraram fraca gestão do programa no nível municipal, alta rotatividade, seleção e treinamento inadequados dos visitantes, baixos salários e contratos inade-

quados de trabalho, falta de supervisão no campo e baixa fidelidade ao desenho do programa durante as visitas domiciliares.

Em relação a estudos similares na América Latina, resultados positivos foram observados para intervenções aplicadas em escala muito inferior à do PCF. Isso foi o caso dos estudos do México⁸, Peru⁷, Colômbia⁶ e no Brasil¹¹. Em todos estes estudos, houve supervisão intensiva e controle de qualidade das visitas. Dois outros estudos brasileiros avaliaram intervenções similares ao PCF. O estudo de eficácia em São Paulo¹³ não mostrou impacto nas análises por intenção de tratar, mas foi detectada melhora em crianças visitadas com maior frequência. A avaliação com delineamento observacional do Programa Primeira Infância Melhor em Pelotas¹⁴ não mostrou impacto em sua análise de todas as crianças, mas somente entre aquelas cujas mães foram recrutadas durante a gestação. Estes dois últimos estudos podem ser afetados por vieses inerentes a análises observacionais, pois tanto o maior número de visitas quanto o recrutamento precoce podem estar associados a características das famílias que favoreceriam o desenvolvimento da criança.

Nossos achados estão em acordo com a literatura internacional, a qual mostra que as avaliações de eficácia – conduzidas em estudos de pequena escala e com forte implementação, sob o monitoramento direto dos pesquisadores – geralmente revelam um impacto positivo das visitas domiciliares sobre o desenvolvimento infantil. Por outro lado, avaliações de efetividade de programas de larga escala, sob implementação de rotina, tendem a apresentar menor ou nenhum impacto. Portanto, nosso achado de falta de impacto está de acordo com a literatura⁴¹⁻⁴³.

Importante também considerar que, em seu delineamento original, o PCF era um programa extremamente ambicioso, com visitas semanais a milhões de crianças em mais de 3.000 municípios, e múltiplos conteúdos a serem trabalhados com as famílias. De fato, este é o maior programa já implementado em qualquer país, com visitas domiciliares para promoção do desenvolvimento infantil. No entanto, a pandemia afetou as características do programa, o que juntamente com os desafios de implementação, já presentes em 2019,

particularmente no que concerne à qualidade das visitas, explicam a ausência de impacto.

Deve-se considerar que as dificuldades enfrentadas pelo PCF como um programa descentralizado, em que a maioria das ações ocorrem no nível municipal, são também observadas em diversos programas brasileiros de assistência à saúde, educação e assistência social⁴⁴.

O objetivo primário das avaliações de programas é prover subsídios para melhoria de políticas e gerenciamento. Avaliações de impacto devem responder a principal pergunta: o programa está alcançando os objetivos propostos? Em caso negativo, o subestudo de avaliação de processo deve sugerir como melhorar o programa. Os resultados detalhados dos componentes de impacto e implementação estão publicamente disponíveis e foram comunicados ao Ministério em inúmeras ocasiões, a fim de auxiliar na melhora do programa.

De um ponto de vista mais amplo, o PCF teve sucesso em colocar a primeira infância e, especificamente, a promoção do desenvolvimento neuropsicomotor, no mais alto grau da agenda nacional da cidadania. O Ministério teve a visão e coragem de financiar uma avaliação randomizada prospectiva, iniciada em um estágio precoce de implementação, que oferece subsídios para o aprimoramento do programa. Poucos, se é que algum programa no Brasil ou em outros países de renda média e baixa, puderam se beneficiar de avaliações de efetividade rigorosas como a atual. Mesmo sem um impacto mensurável, o programa precisa ser valorizado por sua abordagem pioneira de implementação e avaliação. Conforme demonstrado em um estudo prévio, o grupo de intervenção do PCF apresentou uma maior cobertura vacinal do que o grupo controle²¹. Além disso, o programa pode ter impactos em outras áreas, a serem investigadas, como a identificação e notificação de violência doméstica e outras múltiplas vulnerabilidades, assim como fortalecer a intersetorialidade e a conexão das famílias com os serviços públicos. Os resultados mostram que várias medidas são necessárias para fortalecer a implementação do programa, de forma que o país possa garantir o investimento na primeira infância como uma das prioridades nacionais.

Colaboradores

IS Santos, TN Munhoz e CG Victora participaram na concepção do projeto de pesquisa. CG Victora, IS Santos, TN Munhoz, RS Barcelos, CC Bortolotto e C Blumenberg foram responsáveis pela coordenação central do estudo. A Matijasevich, C Salum, HG Santos Júnior, L Marques, L Correia, MR Souza e PIC Lira foram responsáveis pela condução da pesquisa em seus respectivos estados. C Blumenberg, TN Munhoz, RS Barcelos, CC Bortolotto e V Pereira realizaram as análises estatísticas. Todos os autores contribuíram com a interpretação dos dados. CG Victora e IS Santos redigiram a primeira versão do artigo e todos os autores participaram da revisão crítica do conteúdo intelectual. Todos os autores aprovaram a versão final a ser publicada e são responsáveis por todos os aspectos do trabalho na garantia da exatidão e integridade de qualquer parte da obra.

Agradecimentos

Este artigo foi realizado com dados do estudo “Avaliação de Impacto do Programa Criança Feliz”, conduzido pelo Programa de Pós-Graduação em Epidemiologia da Universidade Federal de Pelotas, em parceria com o Instituto Humanidades Artes e Ciências Professor Milton Santos (Bahia), Universidade Federal do ABC Paulista, Universidade Federal do Ceará, Universidade Federal de Goiás, Universidade Federal do Oeste do Pará, Universidade Federal de Pernambuco e Universidade de São Paulo.

Financiamento

O estudo foi financiado pelo Ministério da Cidadania, Fundação Maria Cecília Souto Vidigal, Itaú Social e Banco Mundial, com o apoio do Programa das Nações Unidas para o Desenvolvimento.

Conflito de interesses

Além dos potenciais conflitos de interesses relacionados aos financiamentos descritos acima, Marina Chicaro e Elisa Altafim atuam na Fundação Maria Cecília Souto Vidigal, entidade parceira do Ministério da Cidadania em ações direcionadas à primeira infância.

Referências

1. Aboud F, Yousafzai A. Global health and development in early childhood. *Annu Rev Psychol* 2015; 66:433-457.
2. Britto P, Pyle S, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, Perez-Escamilla R, Rao N, Ip P, Fernald LCH, MacMillan H, Hanson M, Wachs TD, Yao H, Yoshikawa H, Cerezo A, Leckman JF, Bhutta ZA; Early Childhood Development Interventions Review Group, for the Lancet Early Childhood Development Series Steering Committee. Nurturing care: promoting early childhood development. *Lancet* 2017; 389(10064):91-102.
3. Jeong J, Franchett E, Oliveira CR, Rehmani K, Yousafzai A. Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS Med* 2021; 18(5):e1003602.
4. Prado E, Larson L, Cox K, Bettencourt K, Kubes J, Shankar A. Do effects of early life interventions on linear growth correspond to effects on neurobehavioural development? A systematic review and meta-analysis. *Lancet Glob Health* 2019; 7(10):e1398-e1413.
5. Henwood T, Channon S, Penny H, Robling M, Waters C. Do home visiting programmes improve children's language development? A systematic review. *Int J Nurs Stud* 2020; 109:103610.
6. Attanasio O, Fernandez C, Fitzsimons E, Grantham-McGregor S, Meghir C, Rubio-Codina M. Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial *BMJ* 2014; 349:g5785.
7. Hartinger S, Lanata C, Hattendorf J, Wolf J, Gil AI, Obando MO, Noblega M, Verastegui H, Mäusezahl D. Impact of a child stimulation intervention on early child development in rural Peru: a cluster randomised trial using a reciprocal control design. *J Epidemiol Community Health* 2017; 71:217-224.
8. Fernald L, Kagawa R, Knauer H, Schnaas L, Guerra A, Neufeld L. Promoting child development through group-based parent support within a cash transfer program: Experimental effects on children's outcomes. *Dev Psychol* 2017; 53(2):222-236.
9. Knauer H, Kagawa R, Garcia-Guerra A, Schnaas L, Neufeld L, Fernald L. Pathways to improved development for children living in poverty: A randomized effectiveness trial in rural Mexico *Int J Behav Dev* 2016; 40(6):492-99.
10. Fundação Maria Cecília Souto Vidigal (FMCSV). *Visita domiciliar como estratégia de promoção do desenvolvimento e da parentalidade na primeira infância*. São Paulo: FMCSV; 2018.
11. Eickmann SH, Lima AC, Guerra MQ, Lima MC, Lira PI, Huttly SR, Ashworth A. Improved cognitive and motor development in a community-based intervention of psychosocial stimulation in northeast Brazil. *Dev Med Child Neurol* 2003; 45(8):536-541.
12. Grantham-McGregor S, Powell C, Walker S, Himes J. Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study. *Lancet* 1991; 338(8758):1-5.
13. Brentani A, Walker S, Chang-Lopez S, Grisi S, Powell C, Fink G. A home visit-based early childhood stimulation programme in Brazil—a randomized controlled trial. *Health Policy Plan* 2021; 36(3):288-297.
14. Silva EV, Hartwig F, Barros F, Murray J. Effectiveness of a large-scale home visiting programme (PIM) on early child development in Brazil: quasi-experimental study nested in a birth cohort. *BMJ Global Health* 2022; 7:e007116.
15. Brasil. Ministério do Desenvolvimento Social e Agrário (MDS). *A intersectoralidade na visita domiciliar*. Brasília: MDS; 2017.
16. Brasil. Lei nº 10.836, de 9 de janeiro de 2004. Cria o Bolsa Família. *Diário Oficial da União* 2004; 1 dez.
17. Brasil. Ministério da Cidadania. Portaria MC nº 664, de 2 de setembro de 2021. Consolida os atos normativos que regulamentam o Programa Criança Feliz/Primeira Infância no Sistema Único de Assistência Social - SUAS. *Diário Oficial da União* 2021; 6 set.
18. Bryce J, Victora C, Boerma T, Peters D, Black R. Evaluating the scale-up for maternal and child survival: a common framework. *Int Health* 2011; 3(3):139-146.
19. Habicht J, Victora C, Vaughan J. Evaluation designs for adequacy, plausibility and probability of Public Health programme performance and impact. *Int J Epidemiol* 1999; 28:10-18.
20. Donabedian A. The quality of care: how can it be assessed? *JAMA* 1988; 270:1743-1748.
21. Barcelos RS, Santos IS, Munhoz TN, Blumenberg C, Bortolotto CC, Matijasevich A, Salum C, Santos Júnior HGD, Santos LMD, Correia L, Souza MR, Lira PIC, Altafim E, Macana EC, Victora CG. Vaccination coverage in children up to 2 years old, receiving financial support from the Family Income Transfer Program, Brazil. *Epidemiol Serv Saude* 2021; 30(3):e2020983.
22. Santos IS, Munhoz TN, Blumenberg C, Barcelos R, Bortolotto CC, Matijasevich A, Santos Júnior HG, Santos LM, Correia LL, Souza MR, Lira PIC, Altafim E, Marino E, Macana EC, Silva RS, Ohana EF, Fontes MTA, Victora CG. Post-partum depression: a cross-sectional study of women enrolled in a conditional cash transfer program in 30 Brazilian cities. *J Affect Disord* 2021; 281:510-516.
23. Munhoz TN, Santos IS, Blumenberg C, Barcelos RS, Bortolotto CC, Matijasevich A, Santos Júnior HG, Santos LMD, Correia LL, Souza MR, Lira PIC, Altafim ERP, Macana EC, Victora CG. Fatores associados ao desenvolvimento infantil em crianças brasileiras: linha de base da avaliação do impacto do Programa Criança Feliz. *Cad Saude Publica* 2022; 38(2):e00316920.

24. Brasil. Ministério da Cidadania (MC). Secretaria de Avaliação e Gestão da Informação. *Estudo de Linha de Base da Avaliação de Impacto do Programa Criança Feliz. Cadernos de Estudos nº 35*. Brasília: MC; 2020.
25. United Nations Children's Fund (UNICEF). *Multiple indicator cluster surveys (MICS)* [Internet]. [acessado 2021 jan 20]. Disponível em: <https://mics.unicef.org/>.
26. Cox JGC, Murray UD, Jones P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *J Affect Disord* 1996; 39(3):185-189.
27. Santos IS, Matijasevich A, Tavares BF, Barros AJ, Botelho IP, Lapolli C, Magalhães PV, Barbosa AP, Barros FC. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study. *Cad Saude Publica* 2007; 23(11): 2577-2588.
28. Matijasevich A, Munhoz TN, Tavares BF, Barbosa AP, Silva DM, Abitante MS, Dall'Agnol TA, Santos IS. Validation of the Edinburgh Postnatal Depression Scale (EPDS) for screening of major depressive episode among adults from the general population. *BMC Psychiatry* 2014; 14:284.
29. Brasil. Ministério da Cidadania (MC). *Manual do visitador: um olhar sobre a visita domiciliar* [Internet]. 2021 [acessado 2022 jan 20]. Disponível em: <https://www.gov.br/cidadania/pt-br/acoes-e-programas/crianca-feliz/publicacoes-1/MANUALDOVISITADORVERSOFINAL.pdf>.
30. Reichenheim ME, Moraes CL. Adaptação transcultural do instrumento Parent-Child Conflict Tactics Scales (CTSPC) utilizado para identificar a violência contra a criança. *Cad Saude Publica* 2003; 19:1701-1712.
31. Tremblay RE, Nagin DS, Séguin JR, Zoccolillo M, Zelazo PD, Boivin M, Pérusse D, Japel C. Physical Aggression During Early Childhood: Trajectories and Predictors. *Pediatrics* 2004; 114(1):e43-e50.
32. Squires J, Bricker D. *Ages and Stages Questionnaires Third Edition (ASQ3)*. Baltimore: Paul H. Brookes Publishing Co.; 2009.
33. Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry* 1997; 38:581-586.
34. Fleitlich B, Cortázar P, Goodman R. Questionário de capacidades e dificuldades (SDQ). *Infanto Rev Neuropsiq Inf Adol* 2000; 8(1):44-50.
35. Prime H, Browne D, Akbari E, Wade M, Madigan S, Jenkins JM. The development of a measure of maternal cognitive sensitivity appropriate for use in primary care health settings. *J Child Psychol Psychiatry* 2015; 56:488-495.
36. Schneider A, Rodrigues M, Falenchuk O, Munhoz TN, Barros AJD, Murray J, Domingues MR, Jenkins JM. Cross-Cultural Adaptation and Validation of the Brazilian Portuguese Version of an Observational Measure for Parent-Child Responsive Caregiving. *Int J Environ Res Public Health* 2021; 18(3):1246.
37. 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. *J Biomed Inform* 2009; 42(2):377-381.
38. Brasil. Ministério de Desenvolvimento Social (MDS). *Criança Feliz: guia para visita domiciliar*. 2ª ed. Brasília: MDS; 2017.
39. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health* 2004; 94(3):400-405.
40. Heckman J. Skill Formation and the Economics of Investing in Disadvantaged Children. *Science* 2006; 312(5782):1900-1902.
41. Engle PL, Fernald LC, Alderman H, Behrman J, O'Garra C, Yousafzai A, Mello MC, Hidrobo M, Ulkuer N, Ertem I, Iltus S; Global Child Development Steering Group. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *Lancet* 2011; 378(9799):1339-1353.
42. Victora C, Barros F, Assunção M, Restrepo-Méndez M, Matijasevich A, Martorell R. Scaling up maternal nutrition programs to improve birth outcomes: A review of implementation issues. *Food Nutr Bull* 2012; 33(2 Suppl.):S6-S26.
43. List J, Suskind D, Supplee L. *The Scale-Up Effect in Early Childhood and Public Policy: Why Interventions Lose Impact at Scale and What We Can Do About It*. 1ª ed. Oxfordshire: Routledge; 2021.
44. Arretche M. Políticas Sociais no Brasil: descentralização em um Estado federativo. *Rev Bras Cien Soc* 1999; 14(40):111-141.

Artigo apresentado em 22/08/2022

Aprovado em 24/08/2022

Versão final apresentada em 26/08/2022

Editores-chefes: Romeu Gomes, Antônio Augusto Moura da Silva

