

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE MÉTODOS ESTATÍSTICOS

JOÃO FRANCISCO FORTES ANDRÉ

**IMPACTOS DE EVENTOS EXTREMOS DE CALOR SOBRE A PREMATURIDADE:
UMA ABORDAGEM ESTATÍSTICA APLICADA À SAÚDE PÚBLICA**

RIO DE JANEIRO-RJ

2026

JOÃO FRANCISCO FORTES ANDRÉ

**IMPACTOS DE EVENTOS EXTREMOS DE CALOR SOBRE A PREMATURIDADE:
UMA ABORDAGEM ESTATÍSTICA APLICADA À SAÚDE PÚBLICA**

Trabalho de Conclusão de Curso apresentado
ao Instituto de Matemática da Universidade
Federal do Rio de Janeiro (UFRJ), como parte
dos requisitos necessários para a obtenção do
título de Bacharel em Estatística.

Orientadora:

Profa. Dra. Flávia Maria Pinto Ferreira
Landim.

Coorientador:

Prof. Dr. Djacinto Aparecido Monteiro dos
Santos Junior.

RIO DE JANEIRO-RJ

2026

CIP - Catalogação na Publicação

F738i Fortes André, João Francisco
Impactos de eventos extremos de calor sobre a prematuridade: uma abordagem estatística aplicada à saúde pública / João Francisco Fortes André. -- Rio de Janeiro, 2026.
57 f.

Orientadora: Flávia Maria Pinto Ferreira Landim.
Coorientador: Djacinto Aparecido Monteiro dos Santos Junior.

Trabalho de conclusão de curso (graduação) - Universidade Federal do Rio de Janeiro, Instituto de Matemática, Bacharel em Estatística, 2026.

1. Regressão Logística. 2. Ondas de Calor. 3. Saúde Perinatal. 4. Estatística. 5. Nascimento Prematuro. I. Pinto Ferreira Landim, Flávia Maria, orient. II. Aparecido Monteiro dos Santos Junior, Djacinto, coorient. III. Título.

JOÃO FRANCISCO FORTES ANDRÉ

**IMPACTOS DE EVENTOS EXTREMOS DE CALOR SOBRE A PREMATURIDADE:
UMA ABORDAGEM ESTATÍSTICA APLICADA À SAÚDE PÚBLICA**

Trabalho de Conclusão de Curso apresentado
ao Instituto de Matemática da Universidade
Federal do Rio de Janeiro (UFRJ), como parte
dos requisitos necessários para a obtenção do
título de Bacharel em Estatística.

Aprovado em 26 de Maio de 2026.

BANCA EXAMINADORA:

Profa. Dra. Andreza Pereira Rodrigues

Prof. Dr. João Batista de Moraes Pereira

Prof. Dr. Rafael Souza dos Santos

RIO DE JANEIRO-RJ

2026

AGRADECIMENTOS

Em primeiro lugar, agradeço aos meus pais, Kátia Rejane e João Luiz, por todo o amor, apoio incondicional e por serem meus pilares em cada etapa dessa caminhada. Aos meus irmãos Paulo Estevão e Pedro Gabriel, agradeço pela companhia, pelo incentivo e apoio ao longo de toda minha graduação.

À minha noiva, Raphaela Corrêa, com quem tenho o privilégio de compartilhar a vida, meu profundo agradecimento por ser minha companheira e minha inspiração ao longo de toda minha graduação. Sua companhia tem tornado possível cada conquista em minha caminhada.

Agradeço à minha orientadora, Profa. Dra. Flávia Maria Pinto Ferreira Landim, e ao meu coorientador, Prof. Dr. Djacinto Aparecido Monteiro dos Santos Junior, pela confiança, disponibilidade, orientação cuidadosa ao longo de todo o desenvolvimento deste trabalho e pelo constante incentivo à minha pesquisa e ao meu esforço.

Ao Laboratório de Aplicações de Satélites Ambientais (LASA/UFRJ), agradeço por ter sido um espaço de aprendizado, troca e crescimento acadêmico essencial para a realização deste estudo, bem como pela infraestrutura e pelo ambiente colaborativo que tanto enriqueceram minha formação. Agradeço, em especial, à Profa. Dra. Andreza Pereira Rodrigues, à Profa. Dra. Renata Libonati dos Santos e à Profa. Dra. Fernanda Pinheiro Aguiar, pelo apoio, contribuições e pela convivência enriquecedora. Agradeço também a cada membro da equipe do LASA pela contribuição a este estudo e pelo ambiente agradável proporcionado.

Aos professores do Instituto de Matemática da UFRJ, em especial aos docentes do Departamento de Métodos Estatísticos, agradeço pelo conhecimento compartilhado, pela dedicação ao ensino e pela sólida base que me possibilitou adquirir os conhecimentos necessários para a realização deste projeto e da minha formação.

A todos os familiares e amigos que, direta ou indiretamente, contribuíram para que este trabalho fosse possível, deixo registrada minha sincera gratidão pelo apoio e pelas palavras de incentivo.

Por fim, dedico este trabalho com carinho e saudade à minha querida cachorrinha Tária, minha fiel companheira, falecida no dia 22 de dezembro de 2024, que esteve presente em inúmeros momentos difíceis da minha jornada.

RESUMO

Este estudo investiga a relação estatística entre a exposição a eventos de ondas de calor e a ocorrência de nascimentos prematuros no município do Rio de Janeiro, no período de 2000 a 2023. A metodologia utilizada integra dados meteorológicos do Departamento de Controle do Espaço Aéreo (DECEA - Aeroporto Santos Dumont) e dados de nascimentos, incluindo prematuridade provenientes do Sistema de Informações sobre Nascidos Vivos (SINASC). A identificação e caracterização dos eventos de onda de calor foi realizado por meio do método Excess Heat Factor (EHF). A análise desenvolvida neste estudo utiliza modelos de regressão logística do tipo logito, bem como testes estatísticos não paramétricos, como os testes de Kruskal-Wallis e Mann-Whitney, com o objetivo de avaliar a associação entre a taxa de exposição gestacional a ondas de calor e a ocorrência de prematuridade, tanto no cenário geral quanto no contexto específico de gestações que atravessam integralmente o verão. Os resultados obtidos com esses modelos indicam uma associação positiva e estatisticamente significativa em ambos os contextos, porém com magnitudes distintas. Enquanto no modelo geral a taxa de exposição a ondas de calor está associada a um aumento moderado nas chances de prematuridade, esse efeito mostra-se mais intenso no modelo restrito ao verão. Este trabalho ilustra a importância da análise estratificada por sazonalidade e fornece evidências estatísticas que reforçam a relevância da análise climática no planejamento da saúde perinatal.

Palavras-Chave: Ondas de Calor; Saúde Perinatal; Nascimento Prematuro; Excess Heat Factor; Regressão Logística; Estatística.

ABSTRACT

This study investigates the statistical association between exposure to heatwave events and the occurrence of preterm birth in the city of Rio de Janeiro from 2000 to 2023. The methodology integrates meteorological data from the Department of Airspace Control (DECEA – Santos Dumont Airport) and birth data, including prematurity, obtained from the Live Birth Information System (SINASC). The identification and characterization of heatwave events were performed using the Excess Heat Factor (EHF) method. The analysis conducted in this study employs logistic regression models (logit) as well as nonparametric statistical tests, such as the Kruskal–Wallis and Mann–Whitney tests, in order to investigate the association between the rate of gestational exposure to heatwaves and the occurrence of prematurity, both in the overall scenario and in the specific context of pregnancies that fully overlap the summer season. The results obtained with the models indicate a positive and statistically significant association in both contexts, although with distinct magnitudes. While in the general model exposure to heat waves is associated with a moderate increase in the odds of preterm birth, this effect is more pronounced in the model restricted to summer. This study illustrates the importance of seasonal stratification analysis and provides statistical evidence that reinforces the relevance of climatic analysis in perinatal health planning.

Keywords: Heatwaves; Perinatal Health; Preterm Birth; Excess Heat Factor; Logistic Regression; Statistics.

SUMÁRIO

1	INTRODUÇÃO	7
2	CONTEXTO CLIMATOLÓGICO E DA PREMATURIDADE.....	10
2.1	Descrição dos Materiais e Bases de Dados	10
2.2	Dados Sobre Prematuridade	11
2.3	Identificação e Classificação das Ondas de Calor.....	15
3	APLICAÇÃO DE MÉTODOS ESTATÍSTICOS E ANÁLISE DOS RESULTADOS ...	22
3.1	Relação entre Eventos de Onda de Calor e Prematuridade	22
3.1.1	Análise Exploratória.....	23
3.1.2	Modelo de Regressão Logística	26
3.2	Associação Sazonal entre Prematuridade e Ondas de Calor	35
3.2.1	Análise das Estações do Ano	35
3.2.2	Análise Exploratória das Gestações no Verão	37
3.2.3	Análise Inferencial das Gestações no Verão – Testes Não Paramétricos	40
3.2.4	Modelo de Regressão Logística para Gestações no Verão	45
4	CONSIDERAÇÕES FINAIS.....	50
	REFERÊNCIAS	54

1 INTRODUÇÃO

A ocorrência de eventos extremos de ondas de calor (OC) no Brasil, conforme evidenciado por Monteiro dos Santos et al. (2024), tem se tornado cada vez mais frequente nas últimas décadas, seguindo tendências observadas em escala global, como explicitado no artigo de Fischer, Sippel e Knutti (2021), que demonstra que eventos de calor extremo têm superado recordes, tanto em frequência quanto em intensidade, desde o final do século XX. De acordo com o artigo de Chersich et al. (2020), o qual aborda 70 estudos realizados em 27 países distintos, sem inclusão de estudos no Brasil, identificou-se que o risco de parto prematuro aumenta 16% durante os períodos de OC. Tendo em vista este estudo, é possível que, em regiões urbanas, como a cidade do Rio de Janeiro, caracterizada por clima tropical e alta densidade populacional, tais fenômenos de calor extremo representem potenciais riscos à saúde, particularmente para grupos vulneráveis, como gestantes e recém-nascidos.

Além dos estudos conduzidos por Chersich et al. (2020), a associação entre eventos de OC e prematuridade também foi investigada por Darrow et al. (2024) ao analisarem cinquenta regiões metropolitanas dos Estados Unidos. Nesse estudo, foi identificada associação estatisticamente significativa entre a exposição a OC e o aumento das taxas de nascimentos prematuros. De forma semelhante, Fitch et al. (2025b) conduziram uma análise realizada em oito estados norte-americanos e verificaram um aumento da razão de chances de ocorrência de prematuridade após períodos de calor extremo. Os resultados obtidos indicaram que a exposição a temperaturas elevadas constitui, potencialmente, um fator de risco gestacional, especialmente em áreas urbanas com elevada concentração populacional.

No contexto brasileiro, embora o aumento da frequência e intensidade de eventos de calor extremo esteja documentado na literatura (FISCHER; SIPPEL; KNUTTI, 2021; MONTEIRO DOS SANTOS et al., 2024), ainda são escassos os estudos nacionais que investigam especificamente a associação entre OC e prematuridade. Nesse sentido, estudos que avaliem estatisticamente a relação entre eventos de OC e prematuridade são necessários para ampliar a compreensão dos efeitos das condições climáticas extremas sobre a saúde perinatal no Brasil.

A prematuridade, cuja relação com eventos de OC é alvo deste estudo, trata-se de uma condição relevante no contexto da saúde perinatal. De acordo com um estudo da Organização Mundial de Saúde (2023), estima-se que, em contexto mundial, um em cada dez bebês nasce prematuro, ou seja, com menos de 37 semanas de gestação, de forma que essa condição representa a principal causa de óbito em crianças menores de cinco anos. Ainda de acordo

com tal estudo, a prematuridade é responsável por mais de um terço de todas as mortes de recém-nascidos no mundo (OMS, 2023).

Assim como no cenário mundial, a prematuridade também representa um desafio para a saúde pública nacional. De acordo com o Boletim Epidemiológico emitido pelo Ministério da Saúde (BRASIL, 2024), foram registrados, no período de 2012 a 2022, aproximadamente 3,5 milhões de nascimentos prematuros no Brasil. Tendo em vista estes números, o presente estudo tem como objetivo verificar, no período de 2000 a 2023, a existência de associação significativa entre a exposição a eventos de OC e a ocorrência de nascimentos prematuros na cidade do Rio de Janeiro, especialmente na última década em análise. Para avaliar essa associação, realizou-se uma análise cruzada entre as bases de dados climáticos e de nascimentos prematuros, de modo a verificar, através de testes estatísticos, se tal associação entre essas duas variáveis é estatisticamente significativa.

De forma a tornar possível essa análise, duas bases de dados oficiais distintas, as quais cobrem todo o período de 2000 a 2023, foram utilizadas: a primeira diz respeito aos registros de prematuridade provenientes do Sistema de Informações sobre Nascidos Vivos (SINASC), obtidos por meio do Departamento de Informática do Sistema Único de Saúde (DATASUS); já a segunda diz respeito aos dados meteorológicos fornecidos pelo Departamento de Controle do Espaço Aéreo (DECEA).

Nesse contexto, a pesquisa apresentada neste trabalho busca avaliar o impacto da presença de eventos de OC, cujos parâmetros determinísticos estão descritos na Seção 2.3, sobre o desfecho de prematuridade. O principal parâmetro utilizado para tal análise é representado pelo número de dias ao qual a gestante esteve exposta a OC. Além da análise da influência da variável climática sobre a prematuridade, presente na Seção 3.1, também foi avaliado seu impacto de acordo com as estações do ano, com o objetivo de identificar possíveis efeitos de sazonalidade sobre essa associação. Ao longo da seção 3.2 deste estudo, este objetivo foi explorado através da análise das gestações que transcorrem todo o período do verão carioca, no período de 2000 a 2012 e de 2013 a 2023.

Ao realizar a análise exploratória dos dados, bem como aplicar testes estatísticos de comparação entre diferentes categorias de prematuridade, como Mann-Whitney e Kruskal-Wallis, torna-se possível elucidar essas questões. Por fim, deseja-se obter modelos de regressão logística cujos coeficientes e razões de chances permitam estimar o aumento percentual da probabilidade de ocorrência de prematuridade relativa ao aumento da taxa de exposição a OC, cuja metodologia encontra-se desenvolvida nas Subseções 3.1.2 e 3.2.4.

A investigação acerca dos possíveis riscos da exposição ao calor extremo sobre a prematuridade torna possível a obtenção de bases estatísticas significativas que estimulem, por parte das autoridades competentes, a adoção de políticas públicas nacionais e locais de adaptação climática no contexto da saúde perinatal.

Este trabalho encontra-se estruturado em quatro capítulos: o primeiro corresponde à introdução; o segundo capítulo aborda o contexto climatológico e de prematuridade que motivam o objeto de estudo deste projeto; o terceiro apresenta todas as metodologias estatísticas aplicadas na análise da associação entre OC e prematuridade, tanto no cenário geral, como no contexto sazonal, bem como os resultados das análises estatísticas aplicadas. Por fim, o quarto capítulo traz as considerações finais e discussões acerca dos resultados encontrados.

2 CONTEXTO CLIMATOLÓGICO E DA PREMATURIDADE

Neste capítulo, encontram-se descritos os contextos e origens dos dados relativos às ondas de calor (OC) e à prematuridade, bem como os parâmetros, metodologias e indicadores utilizados no processamento desses dados. Complementarmente a essa análise, dissertamos, ao longo deste capítulo, sobre os materiais e recursos computacionais utilizados para a construção das análises realizadas neste estudo.

Este capítulo está dividido entre as seguintes subseções: “Descrição dos Materiais e Bases de Dados”, “Dados sobre Prematuridade” e “Identificação e Classificação das Ondas de Calor”.

2.1 Descrição dos Materiais e Bases de Dados

Para a realização deste estudo foi utilizado o software de programação *R* (R Core Team, 2025), o qual tornou possível o processamento dos dados, bem como a aplicação das funções e algoritmos computacionais necessários para a geração de gráficos e dos testes estatísticos utilizados ao longo de toda a pesquisa.

O software *R* apresenta uma biblioteca de pacotes computacionais essenciais para a obtenção da análise estatística dos dados trabalhados neste estudo, facilitando a escrita dos códigos e a exploração de parâmetros gráficos. Os pacotes *tidyverse* (WICKHAM et al., 2019), utilizado para manipulação e organização dos dados, e o *ggplot2* (WICKHAM, 2016), utilizado para a construção dos gráficos, foram alguns dos pacotes mais empregados ao longo deste estudo. Portanto, através da programação em *R*, obtém-se a análise dos parâmetros climáticos e de saúde perinatal, explorando visualmente tais indicadores através da geração de mapas de calor, bem como histogramas, diagramas de dispersão e boxplots que relacionam prematuridade com exposição térmica extrema.

Antes de explorar a associação entre eventos de OC e desfechos de prematuridade, tornou-se necessário o carregamento dos dados meteorológicos, obtidos junto ao Departamento de Controle do Espaço Aéreo (DECEA), mediante solicitação formal via ofício. Os dados disponibilizados trazem as séries históricas de temperatura máxima e mínima diária registradas entre o período de 1970 a 2023, no Aeroporto Santos Dumont (SBRJ), localizado na cidade do Rio de Janeiro.

Após a exploração e análise dos dados meteorológicos, foram carregados os dados de saúde pública referentes à prematuridade. Tais informações estão disponíveis no site oficial

do DATASUS, contendo parâmetros referentes às gestações ocorridas no município do Rio de Janeiro, no período de 2000 a 2023 (DATASUS, 2025). Devido à grande quantidade de variáveis perinatais presentes nessa base de dados, tornou-se necessária a filtragem e processamento dela, de forma que informações como data de nascimento e duração total da gestação (em semanas) foram selecionadas, afinal elas representam o alvo desta análise.

2.2 Dados Sobre Prematuridade

Os dados utilizados correspondem aos partos de nascidos vivos ocorridos no município do Rio de Janeiro e de mães residentes no mesmo município compreendidos entre o período de 1º de janeiro de 2000 e 31 de dezembro de 2023. Tal base de dados, disponibilizada pelo DATASUS, apresenta uma particularidade que, em um primeiro momento, representa um desafio para a realização da análise de dados. Esse desafio remete à identificação do número exato de semanas de gestação e, conseqüentemente, à determinação precisa do dia de concepção, visto que, entre os anos de 2000 e 2010, os valores referentes a essa variável não são fornecidos de forma direta, mas sim através de intervalos categóricos. A abordagem utilizada para lidar com essa particularidade é o principal foco desta seção.

Conforme definições da Organização Mundial da Saúde e documentos institucionais nacionais (OMS, 2023; BRASIL; EBSEH, 2023), a classificação dos nascimentos quanto à idade gestacional segue os seguintes critérios: partos com 42 semanas ou mais são classificados como “Pós-termo”; entre 37 e 41 semanas, como “A-termo”; entre 32 e 36 semanas, como “Prematuros moderados”; entre 28 e 31 semanas, como “Muito prematuros”; e entre 22 e 27 semanas, como “Prematuros extremos”, sendo que os raros casos de partos com menos de 22 semanas de gestação (apenas 27 ocorrências) também são incluídos na categoria de prematuros extremos.

Após a importação dos dados para o software *R*, verificou-se a existência de duas variáveis que indicam o número de semanas de gestação em cada registro de nascido vivo. A primeira é uma variável quantitativa discreta que indica o número exato de semanas completas de gestação, enquanto a segunda corresponde a uma variável categórica que representa a duração da gestação em seis intervalos distintos. As classes da variável categórica coincidem com a classificação definida pela Organização Mundial da Saúde explicitada no parágrafo anterior.

A princípio, a variável mais adequada para verificar a exposição materna a OC durante a gestação é a variável que apresenta o número exato de semanas de gestação. No entanto,

conforme mencionado no início desta seção, entre os anos de 2000 e 2010, os valores referentes a essa variável não são fornecidos, enquanto que as categorias de duração de gestação encontram-se registradas ao longo de todo o período de 2000 a 2023. Foi observado, ainda, que os dados da variável quantitativa passaram a ser reportados de forma mais consistente apenas a partir de 2013.

Devido a essa circunstância, a variável categórica foi utilizada para determinar a exposição dos partos a OC. Entretanto, como essa variável não informa o número exato de semanas de gestação, foi necessária a adoção de um valor padrão representativo para cada categoria de duração das gestações. Dessa forma, pode-se explorar os dados ao longo de todo o período de estudo, e não somente entre 2013 e 2023. Nesse contexto, destacam-se, dentre outras, três possíveis abordagens distintas: considerar a primeira semana do intervalo; a semana intermediária; ou adotar a última semana do intervalo.

Ao longo do processo de escolha da melhor abordagem possível, foi definido como referência o período de 2013 a 2023, visto que este intervalo temporal apresenta valores disponíveis tanto para a variável quantitativa, quanto para a categórica. Desse modo, foram realizados testes estatísticos e análises de correlação entre a taxa de exposição a OC obtida a partir da variável quantitativa discreta, a qual indica o número exato de semanas de gestação, e as taxas de exposição calculadas segundo as três distintas abordagens que utilizam os intervalos de gestação.

A taxa de exposição a OC durante a gestação foi um parâmetro utilizado ao longo deste estudo para avaliar os possíveis impactos de eventos de OC na ocorrência de nascimentos prematuros. Essa taxa de exposição foi calculada por meio da razão entre o número de dias de exposição a OC até a 36ª semana de gestação e o número total de dias do período, definido por $\text{Mínimo}(N, 252)$, em que N representa a duração total da gestação, em dias, e 252 corresponde ao número de dias equivalentes a 36 semanas de gestação. Caso o parto ocorra antes da 36ª semana, considera-se a taxa de exposição como a razão entre o número de dias de exposição até o término da gravidez e sua duração total.

$$\text{Taxa de Exposição} = \frac{\text{Dias de Exposição a Ondas de Calor até a 36ª Semana de Gestação}}{\text{Mínimo}(N, 252)}.$$

A estimativa do número de dias de gestação, utilizado no cálculo da taxa de exposição, foi obtida a partir do número de semanas da gestação. Para ilustrar o cálculo da taxa de exposição, podemos considerar uma gestação classificada no intervalo de 32 a 36 semanas.

Na abordagem inicial, estimamos a duração gestacional de 32 semanas; na abordagem intermediária, 34 semanas; e, na abordagem final, 36 semanas. Como a duração estimada das gestações varia entre as abordagens, os valores da taxa de exposição também apresentam pequenas diferenças. Por essa razão, realizaram-se testes estatísticos para verificar qual abordagem apresenta maior correlação com a variável quantitativa referente ao número exato de semanas de gestação, considerando os dados de 2013 a 2023.

Figura 1: Matriz de Correlação entre Taxas de Exposição



Fonte: Elaboração Própria

A Figura 1 exibe a matriz de correlação entre a taxa de exposição a OC calculada a partir do número exato de semanas de gestação, no período de 2013 a 2023, e as taxas de exposição estimadas a partir das três diferentes abordagens propostas. Podemos observar que todas as abordagens apresentaram correlação superior a 0,98 em relação à taxa de exposição obtida a partir da variável quantitativa exata. Os p-valores apresentados na Tabela 1 indicam que todas as correlações observadas foram estatisticamente significativas ao nível de 1% ($p - \text{valor} < 2 \times 10^{-16}$).

Tabela 1: P-valores da matriz de correlação entre as taxas de exposição

P-VALORES	Taxa Exata	Taxa – Início	Taxa – Meio	Taxa – Fim
Taxa Exata	2×10^{-16}	2×10^{-16}	2×10^{-16}	2×10^{-16}
Taxa – Início	2×10^{-16}	2×10^{-16}	2×10^{-16}	2×10^{-16}
Taxa – Meio	2×10^{-16}	2×10^{-16}	2×10^{-16}	2×10^{-16}
Taxa – Fim	2×10^{-16}	2×10^{-16}	2×10^{-16}	2×10^{-16}

Fonte: Elaboração Própria

Adicionalmente, foram realizados testes de equivalência TOST (*Two One-Sided Tests*), metodologia proposta por Schuirmann (1987) e posteriormente discutida por Lakens (2017), com o objetivo de avaliar a equivalência prática entre a taxa de exposição calculada a partir do número exato de semanas de gestação e as taxas obtidas pelas três abordagens baseadas nos intervalos gestacionais. Para os testes, adotaram-se limites de equivalência de $\pm 5\%$.

Tabela 2: Resultados dos testes de equivalência (TOST) com limites de $\pm 5\%$

Comparação TOST	Estatística T (Teste do Limite Inferior)	Estatística T (Teste do Limite Superior)	P-Valor (Inferior / Superior)	Decisão
Taxa Exata vs Início	26,28	-56,40	<0,01 / <0,01	Equivalentes
Taxa Exata vs Meio	55,96	-26,73	<0,01 / <0,01	Equivalentes
Taxa Exata vs Fim	82,24	-0,445	<0,01 / 0,328	Não Equivalentes

Fonte: Elaboração Própria

A hipótese nula da metodologia TOST corresponde à não equivalência entre as taxas comparadas, ou seja, assume que a diferença entre as taxas é inferior a um determinado limite

inferior de equivalência ou superior a um limite superior adotado. Em contrapartida, a hipótese alternativa assume equivalência prática entre as taxas, indicando que a diferença observada está dentro do intervalo de equivalência estabelecido. Quanto às estatísticas utilizadas nos testes, temos a estatística *t* de Student, por meio da qual podemos realizar os dois testes unilaterais associados aos limites inferior e superior de equivalência.

A Tabela 2 apresenta os resultados dos testes TOST realizados. Os resultados indicam que apenas as abordagens da semana inicial e da semana intermediária do intervalo gestacional apresentaram equivalência prática em relação à taxa de exposição calculada com base no número exato de semanas de gestação, enquanto a abordagem do fim não apresentou evidências suficientes de equivalência dentro dos limites adotados.

As análises exploratórias e resultados dos modelos de regressão aplicados às duas abordagens apresentaram bastante semelhanças entre si, com diferenças sutis ou, em certas análises, inexistentes. Além disso, a utilização do ponto médio do intervalo gestacional apresenta fundamentação estatística associada à representação de dados agrupados. Assumindo a hipótese de distribuição uniforme dentro de cada intervalo gestacional, o ponto médio corresponde ao valor esperado da categoria *e*, consequentemente, ao valor que minimiza o erro quadrático médio associado à representação do intervalo por um único valor.

Por este motivo, optamos por adotar a abordagem da semana intermediária do intervalo de gestação como padrão para as análises posteriores, com o objetivo de evitar duplicações desnecessárias desses procedimentos. Apesar da disponibilidade da duração exata da gestação no período mais recente, a abordagem da semana intermediária do intervalo gestacional foi adotada em ambos os períodos, com o objetivo de garantir a uniformidade das análises e evitar que eventuais diferenças nos resultados decorram de diferenças metodológicas.

2.3 Identificação e Classificação das Ondas de Calor

Nesta seção encontra-se descrita a metodologia utilizada para a análise dos dados meteorológicos do município do Rio de Janeiro, no período de 1970 a 2023, os quais, conforme explicado na Seção 2.1, são referentes à estação meteorológica SBRJ, localizada no Aeroporto Santos Dumont.

A principal variável meteorológica utilizada para a identificação das OC foi a temperatura média diária, obtida a partir da média aritmética entre a temperatura máxima e a temperatura mínima diárias fornecidas pelo DECEA.

Após a obtenção das temperaturas médias diárias para todo o período de 1970 a 2023, tornou-se essencial adotar um parâmetro que permitisse a classificação e identificação dos eventos de OC. Nesse contexto, o indicador utilizado é o *Excess Heat Factor* (EHF), metodologia proposta, primeiramente, por Nairn e Fawcett (2013) e adotada no estudo proposto por Santos et al. (2024), que estuda o impacto dos eventos de OC sobre as taxas de mortalidade registradas em 14 regiões metropolitanas brasileiras.

A partir do cálculo do EHF, cujos critérios de obtenção estão definidos nos parágrafos seguintes, foram determinados os dias de ocorrência de OC, bem como a intensidade das mesmas. Para a determinação do parâmetro EHF, é necessário, primeiramente, definir um período climatológico de referência que, neste estudo, foi o período de 1981 a 2010. A partir da série temporal climatológica de tal período de referência, calculou-se, através de algoritmo definido no software *R*, o percentil 95% das temperaturas médias diárias (T95), utilizada como parâmetro de referência para obtenção de um parâmetro denominado “Índice Diário de Significância” (EHIsig).

Tanto para o cálculo dos índices diários de significância (EHIsig), quanto para outro parâmetro denominado “Índice de Aclimação” (EHI_{accl}), utilizados no cálculo do EHF, é necessária a obtenção, para cada dia do período entre 1970 e 2023, dos valores das médias móveis de três dias consecutivos da temperatura média diária (DMT₃), considerando o dia em análise e os dois dias subsequentes. De forma complementar, para estimar o índice de aclimação, é necessária a obtenção do parâmetro DMT_{30ant}, calculado a partir da média das temperaturas médias diárias dos trinta dias anteriores ao dia em análise.

$$DMT_3 = \frac{T_i + T_{i+1} + T_{i+2}}{3}, \text{ onde } T_i \text{ é a temperatura média do } i - \text{ésimo dia em análise;}$$

$$DMT_{30ant} = \frac{1}{30} \times \sum_{j=i-30}^{i-1} T_j, \text{ em que } i \text{ representa o dia em análise.}$$

A partir das médias móveis (DMT₃), foram calculados os dois componentes que compõem o *Excess Heat Factor* (EHF). O primeiro indicador obtido foi o índice de significância (EHIsig), o qual representa o desvio térmico em relação ao limiar de calor extremo definido pelo percentil 95% das temperaturas médias diárias do período de referência (T95). Ou seja, o índice de significância (EHIsig) funciona como um índice de medição de

calor extremo, visto que o mesmo é obtido através da diferença entre a média móvel de três dias consecutivos da temperatura média diária (DMT_3) e o limiar de referência T95 (29,3).

$$EHIsig = DMT_3 - T95.$$

O outro índice obtido foi o índice de aclimação (EHI_{accl}), o qual representa o quanto a população do local, no caso o Rio de Janeiro, está aclimatada com o calor recente (últimos trinta dias). Esse índice é expresso pela diferença entre a média móvel de três dias consecutivos e a média das temperaturas médias diárias dos trinta dias anteriores ao dia em análise.

$$EHI_{accl} = DMT_3 - DMT_{30ant}.$$

Ou seja, caso um determinado dia apresente $EHI_{accl} > 0$, a população não está aclimatada e, portanto, o calor intenso não era esperado, enquanto que, se tal dia apresenta $EHI_{accl} \leq 0$, o calor já é esperado e a população já está aclimatada, reduzindo a severidade da OC.

Após a obtenção dos dois índices, calculou-se o EHF para cada dia do período entre 1970 e 2023, através do produto entre o EHI_{sig} e o máximo entre 1 e o índice EHI_{accl} , conforme definido por Nairn e Fawcett (2013). Devido à natureza do cálculo das médias móveis DMT_3 e DMT_{30ant} , não é possível obter os valores do EHI_{accl} e, consequentemente, do EHF, nem para os primeiros trinta dias da amostra, nem para os dois últimos.

$$EHF = EHI_{sig} \times \max(1, EHI_{accl}).$$

Assim, dias que apresentam $EHF > 0$ representam dias de OC no município do Rio de Janeiro. Quanto maior o EHI_{sig} , que indica calor extremo em relação ao parâmetro de referência T95, maior é a tendência de haver um EHF positivo e, portanto, ter um dia com OC incidente. Por sua vez, quanto maior o EHI_{accl} , o qual indica pouca aclimação da população frente ao evento de calor, maior é a magnitude do EHF.

Para este estudo, além de considerar os dias com valores positivos de EHF como dias de OC, foram incluídos os dois dias subsequentes a cada dia com EHF positivo como dias de incidência de OC, mesmo que estes não apresentassem EHF positivo. Assim, caso um determinado dia apresente EHF positivo e os dois dias seguintes não, os três dias ainda são

considerados dias com incidência de OC. Essa metodologia é justificada por haver a possibilidade de os efeitos do calor extremo se estenderem por alguns dias. Portanto, adotou-se esse critério um pouco mais conservador, em que todos os três dias são considerados pertencentes ao mesmo evento de OC.

Conforme estabelecido no Capítulo 1, além da determinação dos dias com incidência de OC, a severidade desses fenômenos climáticos também é objeto de estudo deste trabalho. Portanto, seguindo a metodologia adotada por Monteiro dos Santos et al. (2024), classifica-se a intensidade das OC através do parâmetro EHF_{85} , que representa o percentil 85% de todos os valores positivos de EHF obtidos anteriormente.

Nesse contexto, têm-se que cada dia que apresente EHF positivo e, portanto, seja classificado como dia de OC, apresenta a severidade da onda classificada como baixa, severa ou extrema. Lembrando que, caso o EHF de um determinado dia seja não positivo, o dia é classificado como “Sem Onda”.

As três categorias de severidade das ondas são classificadas da seguinte forma: um determinado dia é classificado como onda de baixa intensidade caso seu EHF seja maior que 0 e menor ou igual ao limiar EHF_{85} ; é classificada como severa caso o EHF seja maior que o limiar EHF_{85} , porém menor que o triplo do EHF_{85} ; enquanto que, para o caso de EHF maior que o triplo do limiar EHF_{85} , a onda é classificada como extrema.

Uma particularidade observada, no entanto, foi a ocorrência de dias com diferentes intensidades dentro de um mesmo evento de OC. Para padronizar a classificação, foi adotado o seguinte critério: para cada dia de ocorrência da OC, incluindo os dois dias subsequentes a um EHF positivo, prevalece a maior severidade registrada até aquele momento. Ou seja, se em um determinado evento de três dias de OC, o primeiro dia for classificado como severo e os dois dias seguintes forem classificados como dias de baixa intensidade, ao invés disso, todos os três dias passam a ser classificados como dias de onda severa. Esse procedimento garante uma representação mais coerente da continuidade e da intensidade dos efeitos do calor extremo.

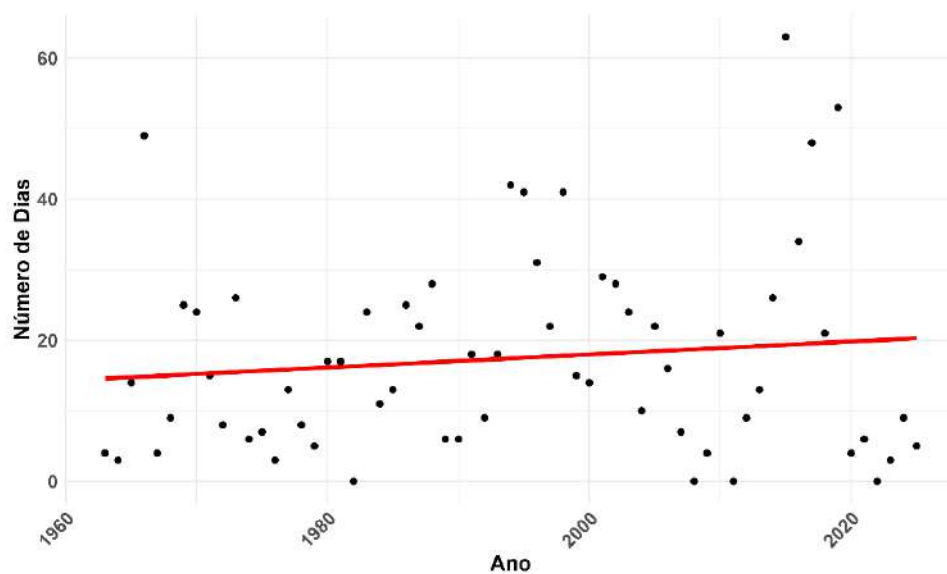
Após a classificação das OC, um modelo de regressão linear foi ajustado com o objetivo de avaliar a tendência temporal da ocorrência desses eventos ao longo dos anos. O modelo definido apresenta como variável resposta o número de dias com incidência de OC em cada ano, enquanto a variável explicativa corresponde ao ano de ocorrência.

O objetivo da definição de um modelo de regressão foi verificar a existência de tendência de crescimento no número de dias de incidência de OC ao longo dos anos, bem como identificar possíveis *outliers*, ou seja, anos em que a frequência de eventos climáticos

de calor extremo é muito mais significativa comparada aos demais anos da série temporal. Dessa forma, a existência de possíveis alterações nas frequências de nascimentos prematuros nos anos de maior frequência de OC pode ser um indicativo de associação entre as duas variáveis.

Após a determinação dos dias de OC e sua severidade, foram construídos gráficos que facilitam a visualização do comportamento dos eventos de OC durante o período de 1970 a 2023. A Figura 2 corresponde a um gráfico de dispersão que apresenta a série temporal do número de dias de OC por ano. Esse gráfico apresenta uma linha de tendência (em vermelho) construída a partir do modelo de regressão linear definido nesta seção, no qual a variável resposta é o número de dias de OC por ano.

Figura 2: Tendência do Número de Dias de Onda de Calor por Ano



Fonte: Elaboração Própria

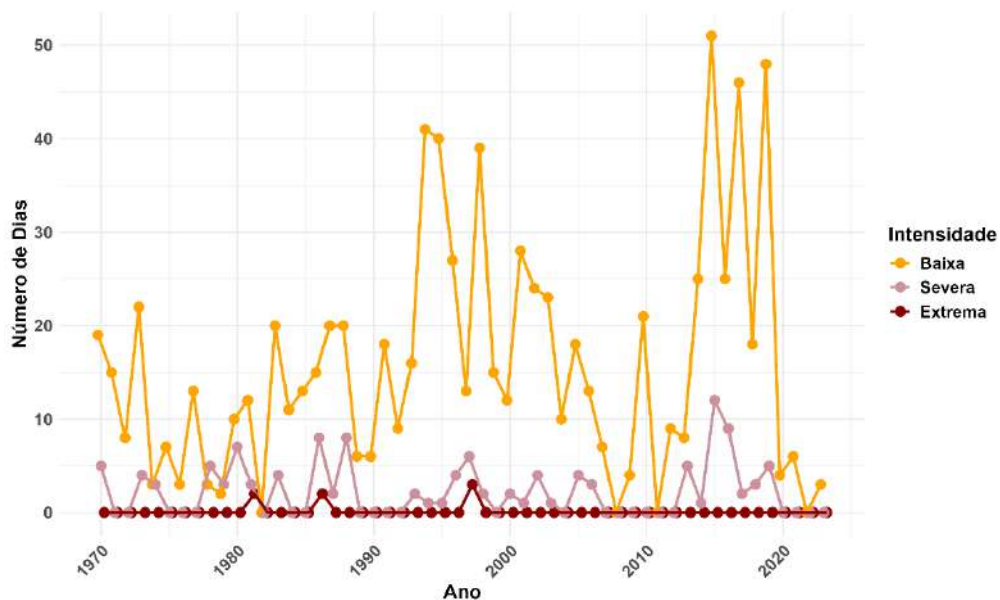
A análise do modelo de regressão linear definido na Figura 2 indica tendência de crescimento no número anual de dias de OC ao longo do período analisado, com aumento médio estimado de 0,092 dias por ano. A tendência observada, no entanto, não apresentou significância estatística ao nível de 5%, visto que o p-valor encontrado foi igual a 0,354. Adicionalmente, ao considerar separadamente os subperíodos de 2000 a 2012 e de 2013 a 2023, o período mais recente apresentou maior média anual de dias de OC (21,9 dias por ano) em comparação ao primeiro período (16,3 dias por ano).

Como o modelo de regressão linear não apresentou significância estatística, foi aplicada uma regressão quantílica para analisar o comportamento da série temporal em

diferentes partes da distribuição do número de dias de OC por ano. Esse método permite avaliar a tendência não apenas na média, mas também em quantis específicos da distribuição. Neste estudo, foi considerado o quantil 0,90, que representa os anos com maior número de dias de OC. Os resultados indicaram coeficiente positivo para o ano nesse quantil, com significância estatística ao nível de 10% (p -valor = 0,06), sugerindo aumento ao longo do tempo na ocorrência de anos com maior número de dias de OC no período de 1970 a 2023.

A Figura 3 mostra o número de dias de OC por ano, classificados de acordo com a severidade do evento. Embora a Figura 2 demonstre tendência de crescimento no número de dias de OC, a Figura 3 mostra que esse comportamento não é uniforme entre as distintas categorias de severidade. Enquanto as ondas de baixa intensidade parecem acompanhar a tendência de crescimento observada anteriormente, o comportamento anual das ondas caracterizadas como severas ou extremas não aparenta seguir um padrão de crescimento.

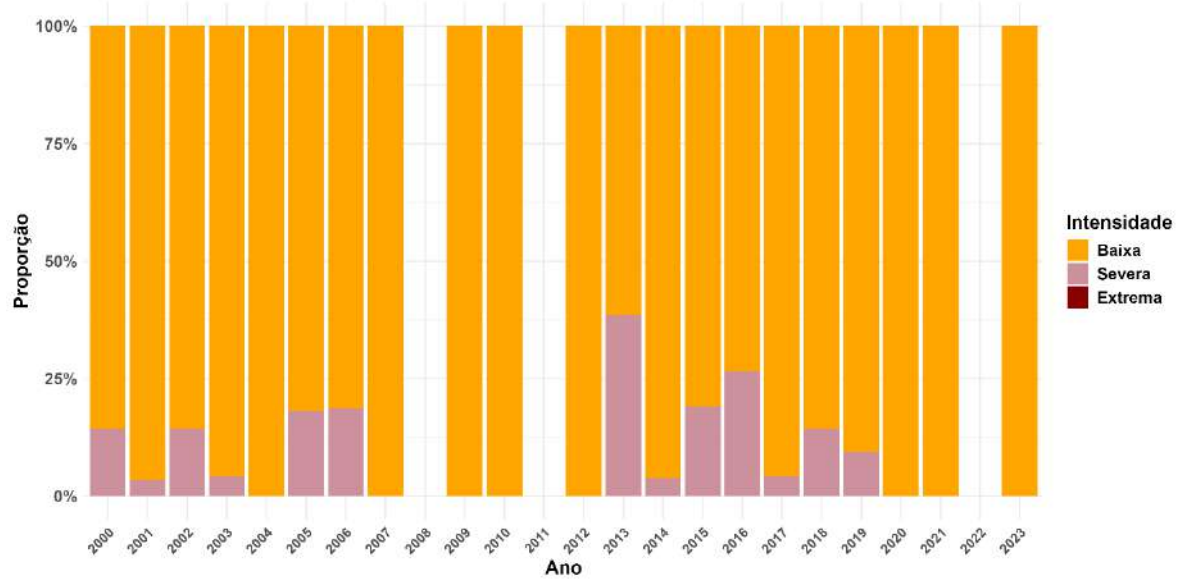
Figura 3: Número de Dias com Onda de Calor por Ano para cada Intensidade



Fonte: Elaboração Própria

A Figura 4, referente às proporções de dias anuais de OC por intensidade (barras empilhadas), evidencia que, ao longo do período de 2000 a 2023, todos os anos do período apresentam mais 70% dos eventos anuais com baixa intensidade, à exceção do ano de 2013, quando há maior participação relativa de dias de OC severa, representando 38% do total naquele ano, e dos anos de 2008, 2011 e 2022, quando não houve registros de OC. Dessa forma, as ondas de baixa intensidade predominam ao longo do período.

Figura 4: Proporção de Dias de Onda de Calor por Ano e Intensidade



Fonte: Elaboração Própria

A partir do contexto climático, definido nesta seção, e do contexto de prematuridade definido na Seção 2.2, foi possível prosseguir este estudo com a análise da associação entre a ocorrência de eventos de OC e o desfecho de prematuridade.

3 APLICAÇÃO DE MÉTODOS ESTATÍSTICOS E ANÁLISE DOS RESULTADOS

Neste capítulo foi analisada a associação entre a incidência de ondas de calor (OC) e o desfecho de prematuridade, incluindo a associação geral e sazonal. Assim, o capítulo encontra-se organizado em duas seções: “Relação entre Eventos de Onda de Calor e Prematuridade” e “Associação Sazonal entre Prematuridade e Ondas de Calor”. São apresentados os resultados dessas aplicações referentes à incidência de OC no município do Rio de Janeiro, entre 1970 e 2023, aos índices de saúde perinatal na cidade, no período de 2000 a 2023, bem como ao cruzamento entre essas variáveis, juntamente com as discussões associadas.

3.1 Relação entre Eventos de Onda de Calor e Prematuridade

A frequência de ocorrência de OC no município do Rio de Janeiro apresentou tendência de crescimento no período 1970 e 2023 (Figura 2), com taxa de aumento de 0,144 dias por ano (p -valor = 0,151). Tal tendência de crescimento observada, embora não seja estatisticamente significativa ao nível de 5%, coincidiu com resultados observados na literatura internacional, como aqueles apresentados em Fischer, Sippel e Knutti (2021), que demonstram um aumento na frequência de eventos climáticos de calor extremo no começo do século XXI.

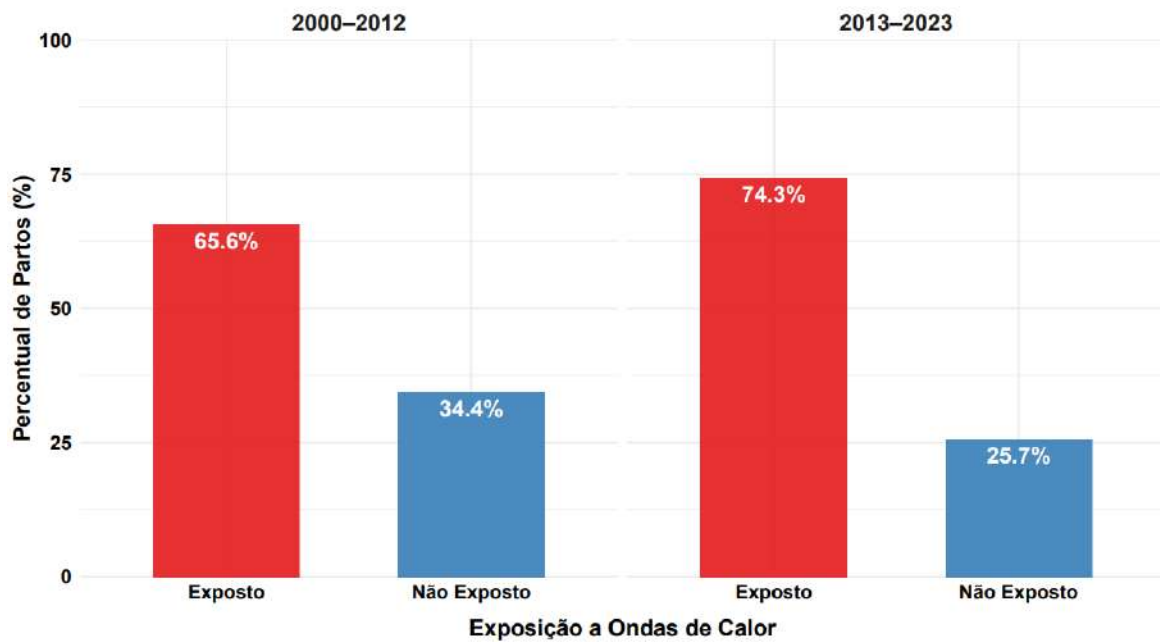
O estudo de Fischer, Sippel e Knutti (2021) aponta que o comportamento dos eventos de OC difere ao longo de diferentes momentos do período em estudo, com forte intensificação na última década. Similarmente, observamos um aumento considerável da frequência média de dias com OC no período de 2013 a 2023 (23,8 dias de OC por ano, em média) em relação àquela observada no período de 2000 a 2012 (17,3 dias por ano). Em virtude deste cenário, foi realizada a divisão do período em análise (2000 a 2023) nesses dois subperíodos, de forma a avaliar possíveis diferenças na ocorrência desses eventos entre os dois intervalos temporais.

Dessa forma, foi possível realizar uma análise exploratória específica para cada subperíodo, possibilitando uma análise gráfica menos generalista do comportamento dos nascimentos prematuros frente aos eventos de OC. Esta seção encontra-se subdividida em duas subseções: a primeira refere-se à análise exploratória dos dados, para a qual se considera a divisão do período em 2000 a 2012 e 2013 a 2023; e a segunda remete à especificação de um modelo de regressão logística.

3.1.1 Análise Exploratória

Com o intuito de realizar uma análise exploratória inicial da associação entre incidência de OC e prematuridade, foi criada uma variável binária com duas categorias: “Exposto”, quando a gestante vivencia um ou mais dias de exposição durante a gestação; e “Não-Exposto”, quando não houve qualquer ocorrência durante a gestação. Com isso, foi possível comparar a frequência relativa de partos cuja gestação apresenta exposição a pelo menos um dia de OC em cada um dos subperíodos definidos (Figura 5).

Figura 5: Frequência Relativa de Gestantes Expostas a Ondas de Calor



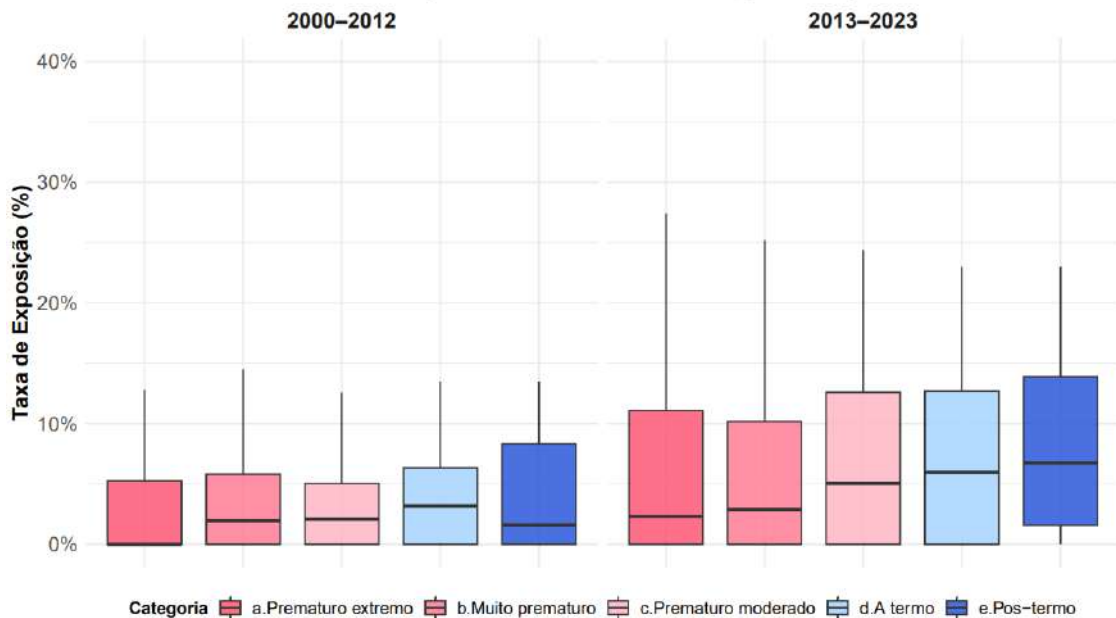
Fonte: Elaboração Própria

A diferença entre as frequências relativas de exposição entre os dois períodos reforça a tendência de crescimento de eventos de OC observada na Figura 2, visto que a frequência de gestantes expostas a pelo menos um dia de OC entre 2013 e 2023 (74,3%) apresentou um acréscimo de aproximadamente 8,7 pontos percentuais em relação ao período anterior (65,6%).

Em seguida, a fim de identificar a existência de associação entre prematuridade e exposição a OC, foram construídos dois gráficos do tipo boxplot, um para o período de 2000 a 2012 e outro para o período de 2013 a 2023, utilizando a taxa de exposição definida na Seção

2.2. O gráfico referente aos boxplots (Figura 6) explora a distribuição da taxa de exposição em cada categoria de prematuridade.

Figura 6: Boxplots da Taxa de Exposição por Categoria de Prematuridade

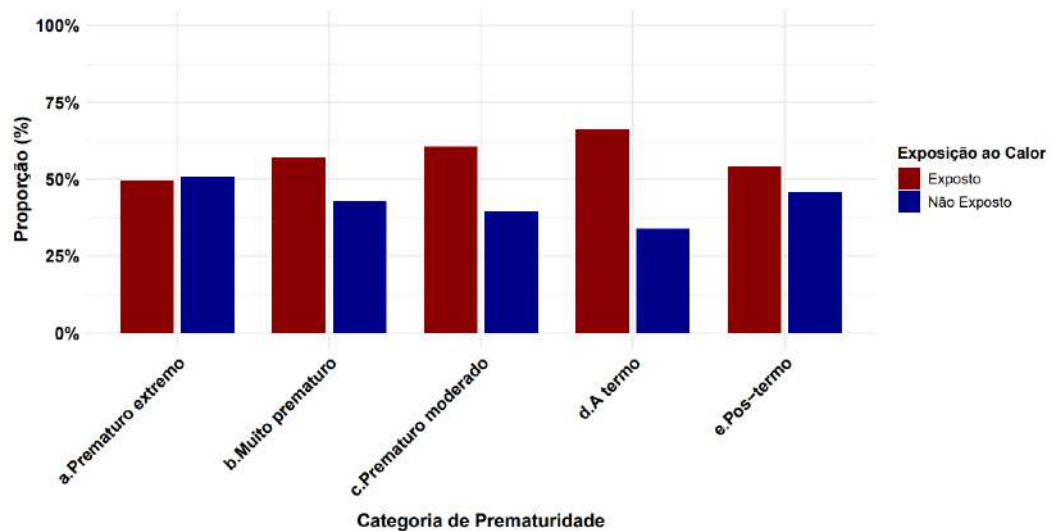


Fonte: Elaboração Própria

Conforme observado nos boxplots dos períodos de 2000 a 2012 e de 2013 a 2023 (Figura 6), as medianas e terceiros quartis das categorias “Prematuro Extremo” (até 27 semanas de gestação), “Muito Prematuro” (28 a 31 semanas) e “Prematuro Moderado” (32 a 36 semanas) apresentam valores semelhantes ou inferiores aos observados nas categorias “A termo” (37 a 41 semanas) e “Pós-termo” (42 semanas ou mais), que representam os partos não prematuros. Esse resultado sugere que, em um primeiro momento, não há indícios gráficos claros que corroborem a hipótese de que uma maior exposição a eventos de calor extremo esteja associada à prematuridade.

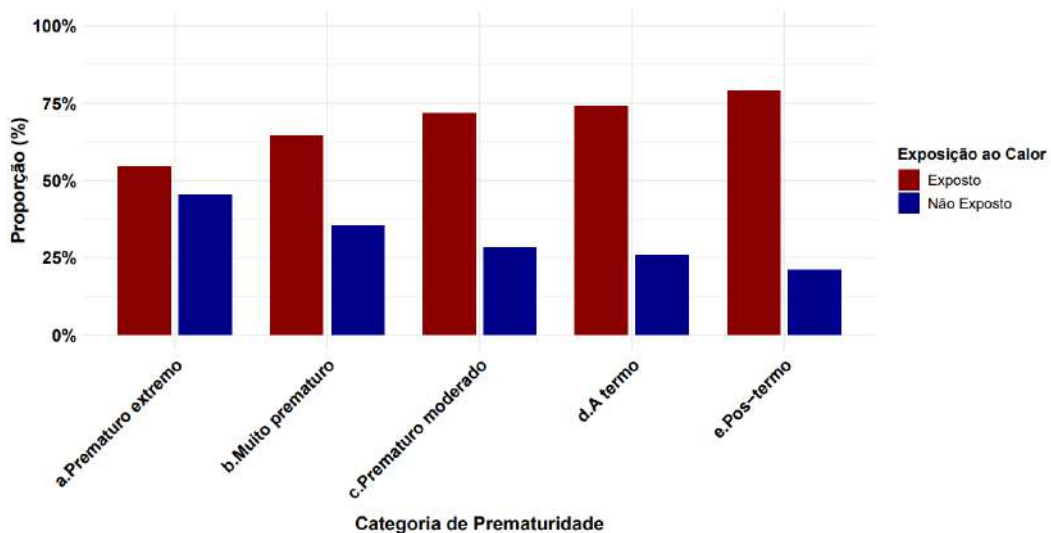
De forma a complementar a análise dos boxplots, foram construídos gráficos de barras, ilustrados nas Figuras 7 e 8, que representam a porcentagem de gestantes expostas e não expostas a pelo menos um dia de OC ao longo da gestação, separados por categoria de prematuridade. Os gráficos de barras apresentam padrão de comportamento semelhante ao observado na Figura 5, visto que exploram a variável binária formada pelas categorias “Exposto” e “Não-Exposto”, porém apresentam a separação desta análise de acordo com as categorias de prematuridade.

Figura 7: Frequência de Gestantes Expostas por Categoria de Prematuridade - 2000 e 2012



Fonte: Elaboração Própria

Figura 8: Frequência de Gestantes Expostas por Categoria de Prematuridade - 2013 e 2023



Fonte: Elaboração Própria

Além disso, assim como observado nos boxplots, os gráficos de barras (Figuras 7 e 8) não sugerem evidências de associação direta entre exposição a OC e a ocorrência de prematuridade. No entanto, a menor porcentagem de exposição a OC observada nas categorias de nascimentos prematuros pode ser justificada pelo menor tempo de gestação característico desses grupos. Em outras palavras, como nascimentos prematuros apresentam menos dias de gestação do que partos não prematuros, é natural que a porcentagem de

gestantes expostas siga esse mesmo padrão. Adicionalmente, as proporções de exposição observadas nas Figuras 7 e 8, são semelhantes às distribuições das taxas de exposição (Figura 6) dos grupos de “Prematuros Moderados” e “A termo”, não indicando, visualmente, um padrão que indique haver associação entre maior exposição ao calor e ocorrência de prematuridade.

3.1.2 Modelo de Regressão Logística

Apesar de a análise das Figuras 6, 7 e 8 oferecer um embasamento inicial, torna-se essencial a utilização de modelos e testes estatísticos que permitam verificar se a associação em estudo é estatisticamente significativa. Nesse contexto, considerando que variável resposta analisada nesse estudo corresponde à ocorrência ou não de prematuridade, caracterizando-se, portanto, como uma variável binária, a escolha de um modelo de regressão logística mostrou-se adequada. Conforme explicado por Hosmer, Lemeshow e Sturdivant (2013), o método de regressão logística é apropriado para modelar situações que apresentem uma variável resposta binária.

Dessa maneira, foi estabelecido como variável resposta do modelo o indicador binário referente à prematuridade, o qual assume valor igual a um quando um nascimento é prematuro (“Prematuro Extremo”, “Muito Prematuro” ou “Prematuro Moderado”) e zero caso contrário (“A termo” ou “Pós-termo”). Além disso, como a ocorrência da prematuridade pode ser influenciada por múltiplos indicadores, para além da variável “Taxa de Exposição”, optou-se pela definição de um modelo de regressão logística múltipla, no qual se considera apenas uma variável resposta, que consiste na ocorrência ou não de prematuridade, e quatro variáveis explicativas distintas, também denominadas covariáveis. Dessa forma, tornou-se possível avaliar o efeito de cada variável explicativa sobre o desfecho de prematuridade ao manter constantes as demais covariáveis incluídas no modelo.

Quanto às variáveis explicativas, o principal indicador utilizado foi a variável “Taxa de Exposição”, que representa a taxa de exposição da gestante a OC durante a gravidez, pois o alvo deste projeto é justamente avaliar essa associação. Além da taxa de exposição, outras covariáveis foram incluídas no modelo, como a idade materna, o sexo do recém-nascido e a raça/cor declarada do recém-nascido. Essas variáveis representam indicadores socioeconômicos e demográficos que podem apresentar alguma influência sobre a probabilidade de ocorrência de prematuridade, que é a variável resposta (CHERSICH et al., 2020).

Com relação à natureza das covariáveis utilizadas, temos que a idade materna foi incluída no modelo como variável quantitativa contínua, representando a idade da mãe, em anos, no momento do nascimento do bebê. Essa variável apresenta média de 26,94 anos, mediana de 27 anos e desvio padrão de 6,81 anos, indicando distribuição aproximadamente simétrica. O sexo do recém-nascido foi incorporado como variável categórica, com sutil predominância da categoria “Masculino” (51,0%), seguido de “Feminino” (48,9%), enquanto a categoria “Sem Informação” representou uma pequena porção dos registros (0,2%). Por fim, a variável raça/cor declarada do recém-nascido também foi incluída como variável categórica, sendo observada maior frequência das categorias “Branca” (43,8%) e “Parda” (43,0%), seguidas por “Preta” (8,5%), enquanto as categorias “Amarela” (0,3%), “Indígena” (0,1%) e “Sem informação” (4,3%) apresentaram menores proporções.

Outras variáveis disponíveis na base de dados, como o peso do recém-nascido e o tipo de parto (parto normal ou cesárea) não foram incluídas como variáveis explicativas pois, muitas vezes, são consequências da prematuridade. Essa escolha é justificada pois, em geral, um nascimento prematuro implica menor peso ao nascer e maior probabilidade de parto cesáreo devido a complicações associadas, de modo que incluí-las no modelo poderia comprometer a interpretação dos efeitos associados às demais variáveis de interesse.

Um potencial problema encontrado em modelos estatísticos é a colinearidade entre as variáveis explicativas, de forma que pode haver elevada correlação entre alguns desses indicadores. Por este motivo, verificamos os valores dos fatores de inflação da variância (VIF) das variáveis explicativas do modelo. Conforme descrito em Kutner et al. (2005), o VIF representa uma medida que permite o diagnóstico de multicolinearidade em modelos lineares generalizados, como o caso de regressão logística do modelo em questão. Para fins interpretativos, temos que valores de VIF próximos a 1 indicam baixa colinearidade entre as variáveis explicativas, enquanto valores superiores a 5 sugerem colinearidade moderada e valores acima de 10 indicam colinearidade elevada, podendo comprometer a análise das estimativas do modelo.

Conforme observado na Tabela 3, os VIFs referentes a cada uma das variáveis explicativas do modelo de regressão logística definido são próximos a 1. Estes valores sugerem baixa colinearidade entre as covariáveis do modelo, de modo que cada variável aparenta contribuir de forma independente, sem inflar a variância das estimativas dos demais coeficientes e, portanto, não apresentam elevada correlação entre si. Dessa forma, não se tornou necessária a remoção de nenhuma das variáveis explicativas selecionadas do modelo.

Tabela 3: Análise de Colinearidade – VIFs das Variáveis Explicativas do Modelo

Variáveis Explicativas	VIF	Graus de Liberdade	Interpretação
Taxa de Exposição	1,001	1	Baixa Colinearidade
Idade da Mãe	1,022	1	Baixa Colinearidade
Sexo do Bebê	1,001	2	Baixa Colinearidade
Raça/Cor do Bebê	1,024	5	Baixa Colinearidade

Fonte: Elaboração Própria

Quanto à estrutura do modelo em si, para modelar a relação entre a taxa de exposição a OC e a prematuridade, foi utilizada regressão logística com função logito (MONTGOMERY; RUNGER, 2018), em que o logito representa o logaritmo neperiano da razão de chances do evento de interesse ocorrer.

$$\text{logit}(p_i) = \log_e \left(\frac{p_i}{1 - p_i} \right),$$

em que p_i = Probabilidade de o nascimento "i" ser prematuro.

A escolha pela função de ligação logito foi motivada por sua capacidade de transformar a probabilidade, que varia entre zero e um, em uma escala contínua e ilimitada, definida no intervalo dos números reais, visto que $\frac{p_i}{1-p_i}$ varia de zero até infinito, uma vez que p_i está definido entre zero e um. Com isso, o logaritmo neperiano de $\frac{p_i}{1-p_i}$ varia no intervalo $(-\infty; \infty)$, que consiste no intervalo dos números reais, permitindo que a relação entre as variáveis explicativas e a chance do evento seja modelada de forma linear. Com isso, ajustamos o modelo de regressão logística com intercepto, quatro variáveis explicativas e o logito da probabilidade de prematuridade como variável resposta.

As variáveis categóricas do modelo, que neste caso são o sexo e a raça/cor do recém-nascido, foram decompostas em variáveis indicadoras (*dummies*). Essas variáveis assumem valor igual a 1 quando o indivíduo pertence a determinada categoria e valor 0 caso contrário, sendo uma das categorias de cada variável definida como referência. No caso da variável “Sexo do Bebê”, a categoria “Feminino” foi utilizada como referência, enquanto na variável “Raça/Cor do Bebê” a categoria “Amarela” foi adotada como referência no modelo. Por exemplo, a variável indicadora “Sexo do Bebê (Masculino)” assume valor 1 para recém-nascidos do sexo masculino e 0 caso contrário. O processo de decomposição das variáveis

categóricas em *dummies*, cuja explicação teórica pode ser encontrada em Montgomery e Runger (2018), resultou na estimação de coeficientes específicos para cada categoria em relação à sua respectiva categoria de referência, sendo que cada coeficiente representa o efeito específico de sua categoria sobre o logito da probabilidade de prematuridade.

$$\log_e \left(\frac{p_i}{1 - p_i} \right) = \beta_0 + \beta_1 \times (\text{Taxa de Exposição}) + \beta_2 \times (\text{Idade da Mãe}) + \sum_{j=3}^4 \beta_j \times Z_{\text{Sexo}, \text{Categoria } j} + \sum_{j=5}^9 \beta_j \times Z_{\text{Raça/Cor}, \text{Categoria } j}, \quad \text{em que:}$$

- p_i = Probabilidade de o nascimento “i” ser prematuro;
- β_0 = Intercepto, representando o logito de p_i quando todas as variáveis explicativas são zero;
- β_1 e β_2 = Efeitos da taxa de exposição a OC e da idade materna, respectivamente, sobre o logito da razão de chances de prematuridade;
- $Z_{\text{Sexo}, \text{Categoria } j}$ ($j = 3, 4$) = Variáveis indicadoras associadas às categorias da variável sexo do recém-nascido, sendo “Feminino” a categoria de referência;
- $Z_{\text{Raça/Cor}, \text{Categoria } j}$ ($j = 5, 6, 7, 8, 9$) = Variáveis indicadoras associadas às categorias da variável raça/cor do recém-nascido, sendo “Amarela” a categoria de referência;
- β_j ($j = 3, \dots, 9$) = Coeficientes associados às variáveis indicadoras, representando o efeito de cada categoria em relação às respectivas categorias de referência.

No modelo descrito, cada coeficiente β_i , em que $i = 1, 2, 3, \dots, 9$, representa a variação no logito da razão de chances de ocorrência de prematuridade associada a variações nas variáveis explicativas, mantendo as demais constantes. Para as variáveis quantitativas, como a taxa de exposição a OC (β_1) e a idade materna (β_2), os coeficientes indicam a variação no logito da razão de chances de prematuridade associada a um aumento unitário nessas variáveis. Já para as variáveis categóricas, representadas por *dummies*, os coeficientes $\beta_3, \dots, \beta_9$ indicam a diferença no logito da razão de chances em relação às respectivas categorias de referência. Neste caso, como a variável de interesse é a taxa de exposição a OC, seu coeficiente (β_1) indica como tal exposição impacta o logito da razão de chances de

ocorrência de nascimentos prematuros. Após o ajuste do modelo de regressão logística no software *R*, encontramos os coeficientes estimados ($\hat{\beta}_i$), seus respectivos erros padrão e p-valores, presentes na Tabela 4.

Tabela 4: Sumário do Modelo de Regressão Logística para Prematuridade

Variáveis Explicativas	Coeficiente Estimado ($\hat{\beta}_i$)	Erro Padrão	P-Valor
Intercepto	-2,5864	0,0505	$< 2 \times 10^{-16}$
Taxa de Exposição	0,4696	0,0434	$< 2 \times 10^{-16}$
Idade da Mãe	0,0143	0,0004	$< 2 \times 10^{-16}$
Sexo do Bebê (Masculino)	0,0525	0,0056	$< 2 \times 10^{-16}$
Sexo do Bebê (Sem Informação)	0,2067	0,0608	0,0007
Raça/Cor do Bebê (Branca)	-0,0945	0,0493	0,0551
Raça/Cor do Bebê (Indígena)	-0,0272	0,1064	0,7983
Raça/Cor do Bebê (Parda)	-0,0018	0,0493	0,9707
Raça/Cor do Bebê (Preta)	0,0575	0,0499	0,2496
Raça/Cor do Bebê (Sem Informação)	-0,0502	0,0509	0,3241

Fonte: Elaboração Própria

Os coeficientes estimados ($\hat{\beta}_i$) de cada variável explicativa do modelo, indicados na Tabela 4, representam os efeitos dessas covariáveis sobre o logito da razão de chances de ocorrência de prematuridade no modelo de regressão logística. De acordo com o modelo proposto, covariáveis que apresentam coeficientes positivos estão associadas a aumentos na chance de ocorrência de prematuridade, enquanto aquelas cujos coeficientes são negativos estão associadas à redução dessa chance. Quanto maior o valor absoluto do coeficiente, maior é a intensidade do efeito da variável sobre o desfecho de prematuridade. Outra estatística

importante representada na Tabela 4 são os p-valores das covariáveis, os quais avaliam a significância estatística dos seus efeitos correspondentes.

Observamos que a variável “Taxa de Exposição”, cujo comportamento representa o foco deste estudo, apresentou coeficiente estimado positivo ($\hat{\beta}_1 = 0,4696$) e p-valor próximo de zero. O sinal do coeficiente indica que aumentos na taxa de exposição a OC estão associados a aumentos na chance de ocorrência de parto prematuro, enquanto que seu p-valor próximo de zero ($p - \text{valor} \approx 0$) indica que esse efeito foi estatisticamente significativo, ao nível de 5%, para o período de 2000 a 2023.

Tabela 5: Razões de chances e Intervalos de Confiança de 95% das Variáveis Explicativas

Variáveis Explicativas	Razões de Chances (OR)	Intervalo de Confiança 95%
Intercepto	0,0753	[0,0682 ; 0,0831]
Taxa de Exposição	1,5993	[1,4690 ; 1,7412]
Idade da Mãe	1,0144	[1,0136 ; 1,0152]
Sexo do Bebê (Masculino)	1,0539	[1,0424 ; 1,0654]
Sexo do Bebê (Sem Informação)	1,2296	[1,0915 ; 1,3853]
Raça/Cor do Bebê (Branca)	0,9098	[0,8260 ; 1,0021]
Raça/Cor do Bebê (Indígena)	0,9732	[0,7900 ; 1,1989]
Raça/Cor do Bebê (Parda)	0,9982	[0,9063 ; 1,0994]
Raça/Cor do Bebê (Preta)	1,0592	[0,9604 ; 1,1682]
Raça/Cor do Bebê (Sem Informação)	0,9510	[0,8606 ; 1,0509]

Fonte: Elaboração Própria

Com o objetivo de obter uma interpretação mais clara do efeito relativo de cada covariável do modelo sobre a chance de ocorrência de prematuridade, buscamos obter, além dos coeficientes estimados e seus p-valores, as razões de chance (OR), associadas às variáveis explicativas, bem como seus respectivos intervalos de confiança a 95%. As razões de chances das covariáveis do modelo são apresentadas na Tabela 5.

Conforme observado na Tabela 5, a razão de chances unitária estimada para a variável “Taxa de Exposição” apresentou valor igual a 1,5993 ($OR_{\text{unitário}}$), com intervalo de confiança de 95% definido por [1,469;1,741]. Esse resultado indica que um aumento unitário na taxa de

exposição a OC está associado a um aumento de aproximadamente 59,9% nas chances de ocorrência de prematuridade.

Como a variável “Taxa de Exposição” assume valores restritos ao intervalo [0; 0,338] para os dados analisados no período de 2000 a 2023, variações reais dessa taxa correspondem, efetivamente, a frações de uma unidade, resultando em aumentos percentuais menores que 59,9% nas chances de prematuridade.

Em modelos de regressão logística, a razão de chances para um aumento de Δ unidades na variável explicativa, neste caso a “Taxa de Exposição”, é obtida a partir da razão entre as chances em dois pontos distintos da variável. Como as chances (odds) são representadas por $\frac{p_i}{1-p_i}$, temos que:

$$OR(\Delta) = \frac{\text{odds}(x + \Delta)}{\text{odds}(x)} = \frac{\frac{p_i(x + \Delta)}{1 - p_i(x + \Delta)}}{\frac{p_i(x)}{1 - p_i(x)}}, \text{ onde } x \text{ representa a Taxa de Exposição.}$$

Como, no modelo de regressão logística, o logito é dado por $\log_e \left(\frac{p_i(x)}{1-p_i(x)} \right) = \beta_0 + \beta_1 x + \dots$, temos que a razão de chances pode ser escrita na forma exponencial $\frac{p_i(x)}{1-p_i(x)} = \exp(\beta_0 + \beta_1 x + \dots)$. Analogamente, $\frac{p_i(x+\Delta)}{1-p_i(x+\Delta)} = \exp(\beta_0 + \beta_1(x + \Delta) + \dots)$. Substituindo essas expressões na razão de chances entre os pontos x e $x + \Delta$, obtemos a expressão de $OR(\Delta)$:

$$OR(\Delta) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1(x + \Delta) + \dots)}{\exp(\hat{\beta}_0 + \hat{\beta}_1 x + \dots)} = \exp(\hat{\beta}_1 \times \Delta) = (OR_{\text{unitário}})^\Delta.$$

A expressão em questão, apresentada por Hosmer, Lemeshow e Sturdivant (2013), apresenta a formulação mencionada pois os coeficientes do modelo de regressão logística encontram-se estimados na escala do logito (*log-odds*). Ou seja, temos que as chances de probabilidade, chamadas de *odds*, são representadas por $\frac{p_i}{1-p_i}$ e, conseqüentemente, os logitos, chamados de *log-odds*, são representados por $\log_e \left(\frac{p_i}{1-p_i} \right)$.

Devido a essa mudança de escala do modelo proporcionada pela função logito, temos que variações lineares nas variáveis explicativas, em particular na variável “Taxa de Exposição”, implicam variações multiplicativas nas chances de ocorrência de prematuridade.

Retomando os resultados do modelo exibidos na Tabela 4, o coeficiente estimado para a variável “Taxa de Exposição”, corresponde a $\hat{\beta}_1 = 0,4696$. Considerando que o valor máximo observado para tal variável na base de dados é de 0,338, o efeito máximo esperado da exposição sobre as chances de prematuridade pôde ser estimado por meio da razão de chances (OR) relativa a essa variação, definida como a razão entre a probabilidade de ocorrência do evento (prematuridade) e a probabilidade de não ocorrência, isto é, $\frac{p_i}{1-p_i}$, permitindo comparar como essas chances se alteram com a variação na exposição.

$$OR(0,338) = \exp(\hat{\beta}_1 \times \Delta) = \exp(0,4696 \times 0,338) = 1,1720;$$

$$\text{Aumento da Chance } (0,338) = (1,1720 - 1) \times 100\% \approx 17,20\% .$$

Esse resultado indicou que, no maior nível de exposição observado na base, há um aumento de cerca de 17,2% nas chances de ocorrência de prematuridade.

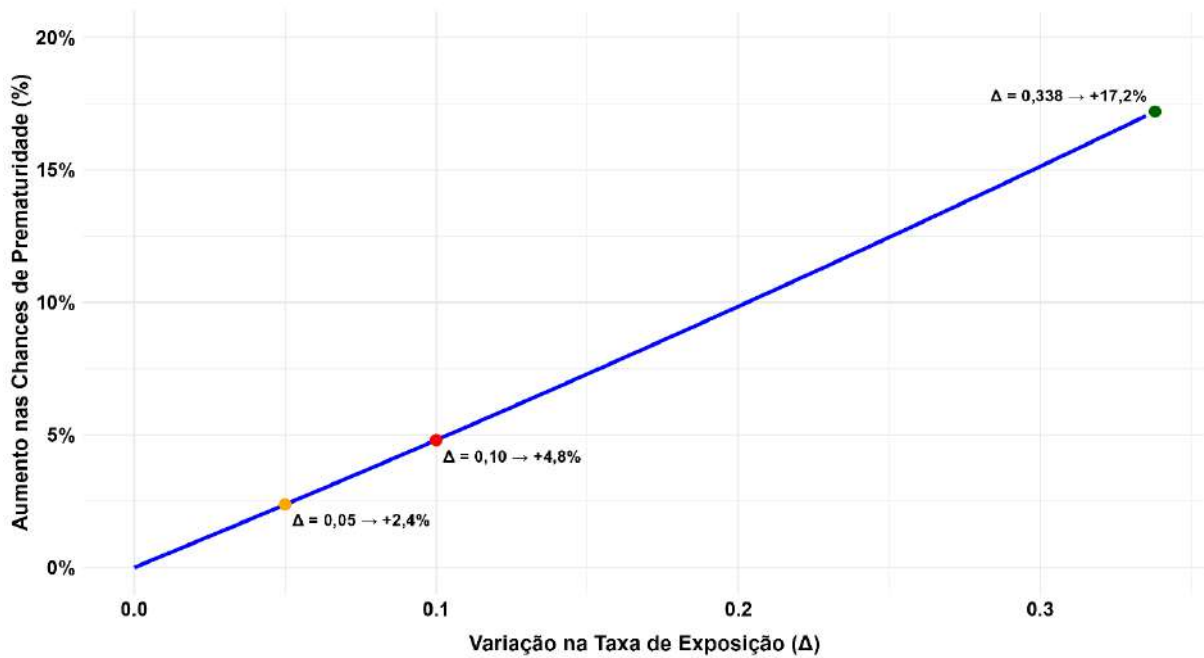
Analogamente, estimamos a razão de chances associada a um aumento de 0,10 unidade na taxa de exposição a OC ($\Delta = 0,10$). Nesse caso, ao elevar a razão de chances unitária ($OR_{\text{unitário}} = 1,5993$) ao expoente 0,1, obtivemos $OR(0,1) = 1,0481$, o que implica um aumento aproximado de 4,8% nas chances de ocorrência de parto prematuro.

$$OR(0,1) = 1,5993^{0,1} = 1,0481;$$

$$\text{Aumento da Chance } (0,1) = (1,0481 - 1) \times 100\% \approx 4,81\%.$$

Portanto, embora a análise visual dos boxplots e gráficos de barras (Figuras 6, 7 e 8) não evidencie aumento na ocorrência de prematuridade associado à exposição a OC, os resultados obtidos através do modelo de regressão logística, como o coeficiente e a razão de chances estimados, bem como o p-valor associado, indicaram que incrementos na taxa de exposição estão associados a aumentos estatisticamente significativos nas chances de ocorrência de prematuridade.

Figura 9: Aumento nas Chances de Prematuridade em Função da Variação da Exposição



Fonte: Elaboração Própria

Após o ajuste do modelo, analisamos o efeito da variação na taxa de exposição sobre as chances de ocorrência de nascimentos prematuros. A partir do coeficiente estimado ($\beta = 0,4696$), construímos a Figura 9, que apresenta o aumento percentual nas chances de prematuridade em função de diferentes variações na exposição (Δ). A relação observada é exponencial, característica do modelo logístico especificado, indicando que aumentos na exposição estão associados a elevações nas chances de prematuridade. Conforme indicado na Figura 9, um aumento de 0,05 na taxa de exposição está associado a um acréscimo aproximado de 2,4% nas chances de prematuridade. Para uma variação de 0,10, esse aumento é de cerca de 4,8%, enquanto que, para a variação máxima observada nos dados ($\Delta = 0,338$), o aumento nas chances atinge aproximadamente 17,2%.

Esses resultados evidenciam que a variável de exposição a OC possui um efeito consistente sobre a ocorrência de prematuridade. Em particular, mesmo variações relativamente pequenas na exposição já estão associadas a aumentos no risco, o que reforça a relevância do fator analisado no contexto deste estudo.

Ao longo da Seção 3.2 deste estudo, dissertaremos sobre a relação entre eventos de OC e prematuridade em determinados períodos do ano, com ênfase em partos cuja gestação transcorre todo o período do verão.

3.2 Associação Sazonal entre Prematuridade e Ondas de Calor

O objetivo desta seção é investigar a sazonalidade da associação entre a ocorrência de OC e a prematuridade, com foco específico em gestações que abrangem todo o período do verão. Ou seja, os dados tratados ao longo desta seção são referentes às mulheres cujas gestações compreendem o período de 21 de dezembro de um dado ano até 20 de março do ano seguinte, cobrindo, portanto, toda a estação do verão.

Esta seção encontra-se subdividida em quatro subseções: na primeira apresenta-se o embasamento teórico para a escolha do verão como estação-alvo desta análise; a segunda remete à análise exploratória das gestações que transcorrem integralmente durante o verão; a terceira aborda os testes não paramétricos utilizados na comparação entre as taxas de exposição das diferentes categorias de prematuridade; e, por fim a quarta subseção descreve o modelo de regressão logística, análogo ao descrito na Subseção 3.1.2.

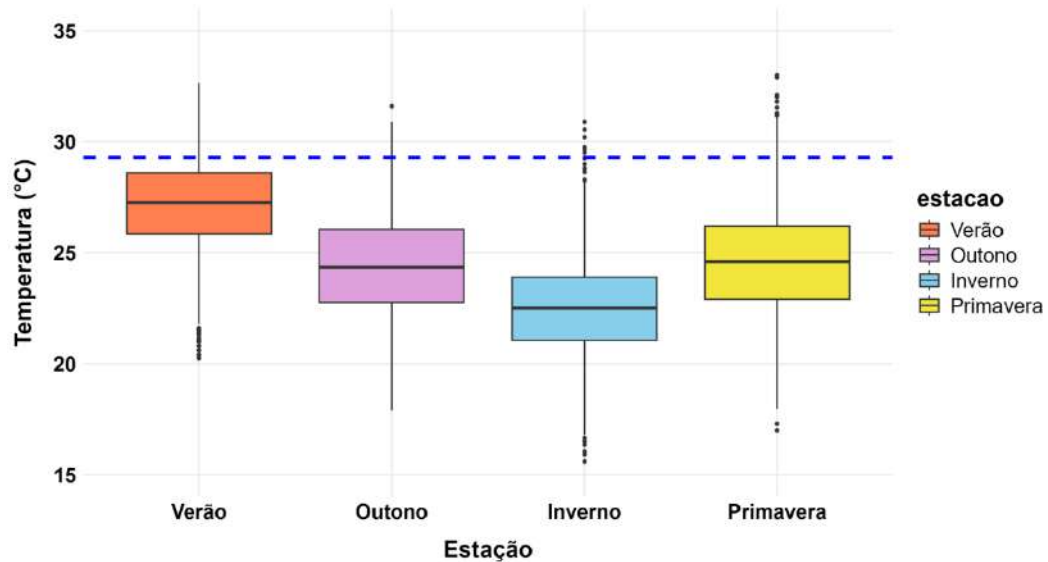
3.2.1 Análise das Estações do Ano

Antes da construção dos gráficos referentes à associação entre a incidência de OC e prematuridade sobre as gestações que passam por todo o verão, é essencial justificar a escolha do verão como foco da análise desta subseção. A Figura 10 ilustra a distribuição das temperaturas médias diárias na cidade do Rio de Janeiro entre 2000 a 2023, categorizada por estação do ano, permitindo a obtenção de uma análise visual clara sobre o comportamento climático de cada estação.

De forma complementar, a Figura 10 apresenta uma linha azul tracejada que representa o limiar T95, o qual corresponde ao percentil 95 da temperatura média diária definida no período de referência de 1981 a 2010, conforme descrito na Seção 2.3. Esse parâmetro (T95) é utilizado para o cálculo do EHF (MONTEIRO DOS SANTOS et al., 2024) e, conseqüentemente, para determinar se um determinado dia é caracterizado pela incidência de OC.

Complementarmente, a Figura 11 indica o número de dias por ano em que há ocorrência de OC, categorizada por estação do ano, avaliada entre o período de 2000 e 2023. A análise dessa série temporal permite verificar quais estações do ano apresentam maior concentração anual de dias de OC.

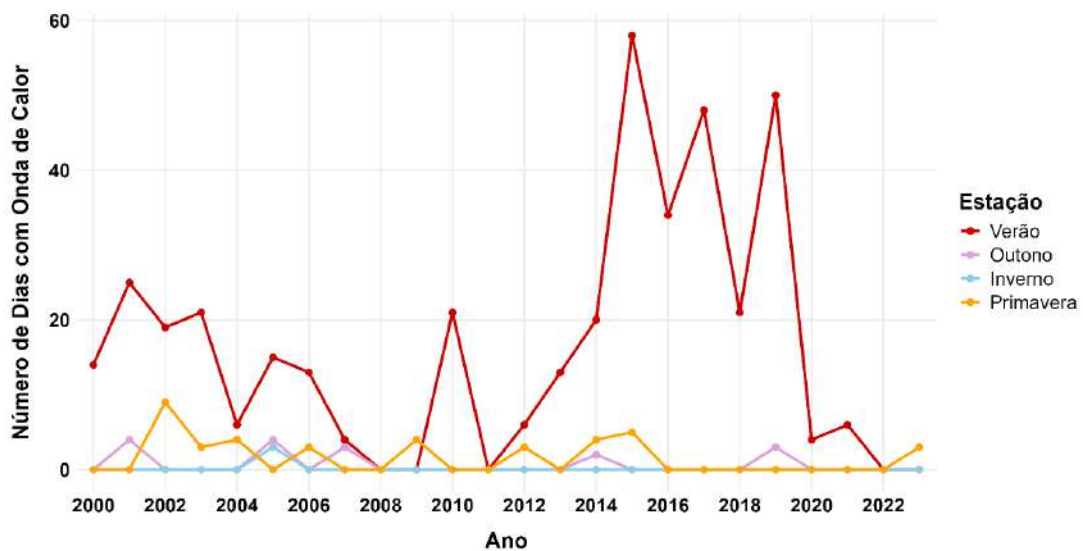
Figura 10: Boxplot das Temperaturas Diárias separadas por Estação do Ano (2000 -2023)



Fonte: Elaboração Própria

Conforme observado na Figura 10, o primeiro quartil, a mediana e o terceiro da distribuição das temperaturas diárias no verão são mais elevados do que os observados para o outono, inverno e primavera. Com relação à série temporal (Figura 11), observa-se que, em praticamente todos os anos do período de 2000 a 2023, o verão é a estação que apresenta o maior número de dias com incidência de OC.

Figura 11: Série Temporal do Número de Dias com Ocorrência de Ondas de Calor por Estação do Ano (2000–2023)



Fonte: Elaboração Própria

Dessa forma, conforme retratam as Figuras 10 e 11, o verão foi a estação cujas temperaturas diárias e cujo número de dias com incidência de OC mostraram-se mais elevados em comparação às demais estações ao longo do período de 2000 a 2023. As gestações que compreendem todo o verão podem, portanto, indicar exposição das gestantes a longos períodos de incidência de OC. Tendo em vista esses resultados, a escolha da análise das gestações que transcorrem toda essa estação como foco desta seção é justificada. Com isso, foi possível realizar uma comparação dos resultados obtidos com os resultados gerais discutidos na Seção 3.1. Cabe ressaltar que a abordagem utilizada para estimar o número de semanas gestacionais foi análoga à utilizada na Subseção 3.1.2.

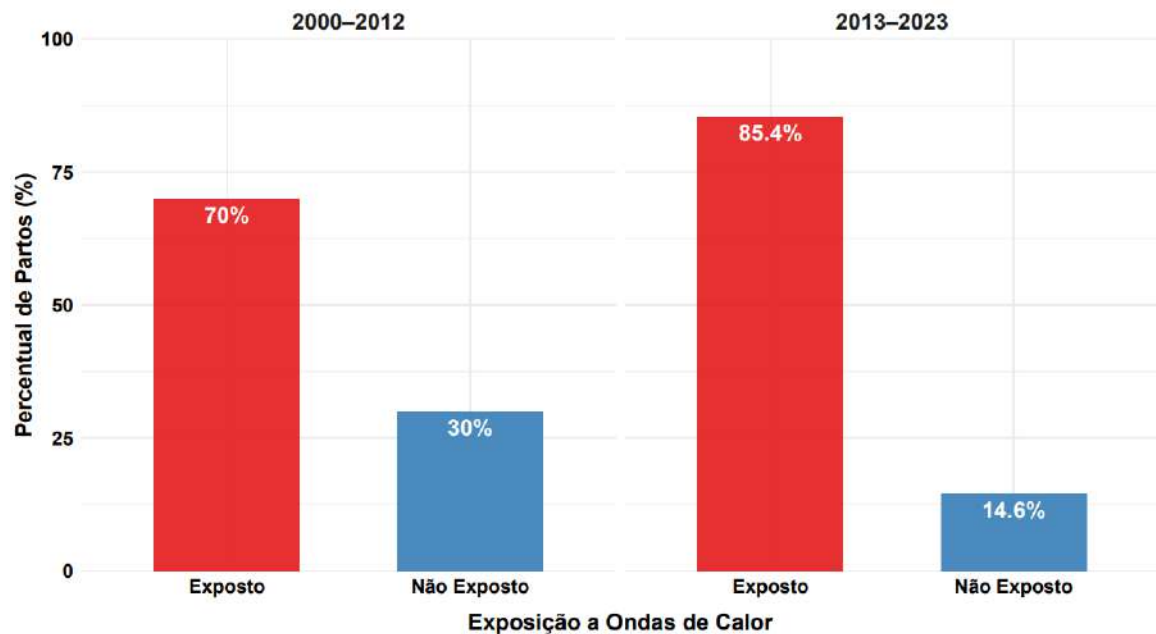
3.2.2 Análise Exploratória das Gestações no Verão

Nesta subseção encontram-se dispostos os gráficos referentes às gestações que transcorrem todo o período do verão, considerando os intervalos de 2000 a 2012 e de 2013 a 2023.

Antes da construção de testes estatísticos propriamente ditos, foram construídos gráficos que possibilitaram uma análise visual da distribuição das taxas de exposição a OC, separadas por categoria de prematuridade, nesse novo subgrupo sazonal de gestações. A primeira etapa consistiu na construção do gráfico representado na Figura 12, o qual indica a frequência relativa de gestantes expostas a pelo menos um dia de OC ao longo de toda a gestação. Neste caso, apenas as gestações que transcorrem integralmente pelo verão foram alvos desta análise.

A partir da análise dos gráficos da Figura 12, verificamos que a frequência relativa de gestantes expostas a pelo menos um dia de OC no período de 2013 a 2023 (85,4%) foi superior à frequência de aproximadamente 70% observada no período de 2000 a 2012. A tendência de crescimento é coerente com o comportamento identificado na análise geral apresentada na Figura 2. Analogamente ao processo realizado na Subseção 3.1.1, construímos boxplots (Figura 13) que retratam a distribuição da taxa de exposição a OC, estratificados por categoria de prematuridade.

Figura 12: Frequência Relativa de Gestantes Expostas a Ondas de Calor - Verão

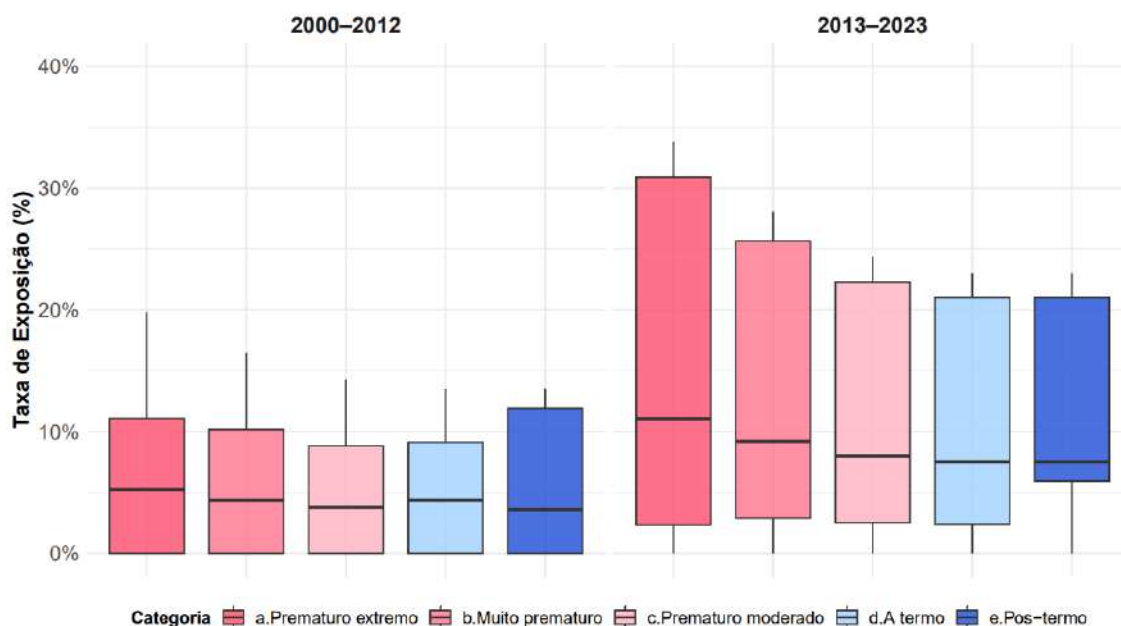


Fonte: Elaboração Própria

O gráfico de boxplots, ilustrado na Figura 13, retrata, tanto para o período de 2000 a 2012 quanto para o período de 2013 e 2023, uma diferença nas distribuições das taxas de exposição entre as categorias de prematuridade. Conforme observado na Figura 13, tais diferenças foram mais acentuadas no período mais recente. Diferentemente do padrão observado na Subseção 3.1.1, a Figura 13, referente aos partos que passam por todo o verão, evidencia que, no período de 2013 a 2023, a distribuição da taxa de exposição das categorias “Prematuro Extremo” e “Muito Prematuro” apresenta medidas como mediana e terceiro quartil superiores àsquelas observadas para a categoria de partos não prematuros (“A termo” e “Pós-termo”).

Embora esse padrão de comportamento evidenciado na Figura 13 seja mais sutil no período de 2000 a 2012, os resultados obtidos indicam uma possível associação entre a exposição a OC durante o verão e a ocorrência de nascimentos prematuros. O gráfico apresentado na Figura 14 complementa essa análise, visto que representa as médias da taxa de exposição a OC, separadas por categoria de prematuridade. Além disso, as barras de erro presentes na Figura 14 representam os intervalos de confiança de 95% das médias estimadas, indicando a precisão das estimativas obtidas para cada categoria.

Figura 13: Boxplots da Taxa de Exposição por Categoria de Prematuridade- Verão

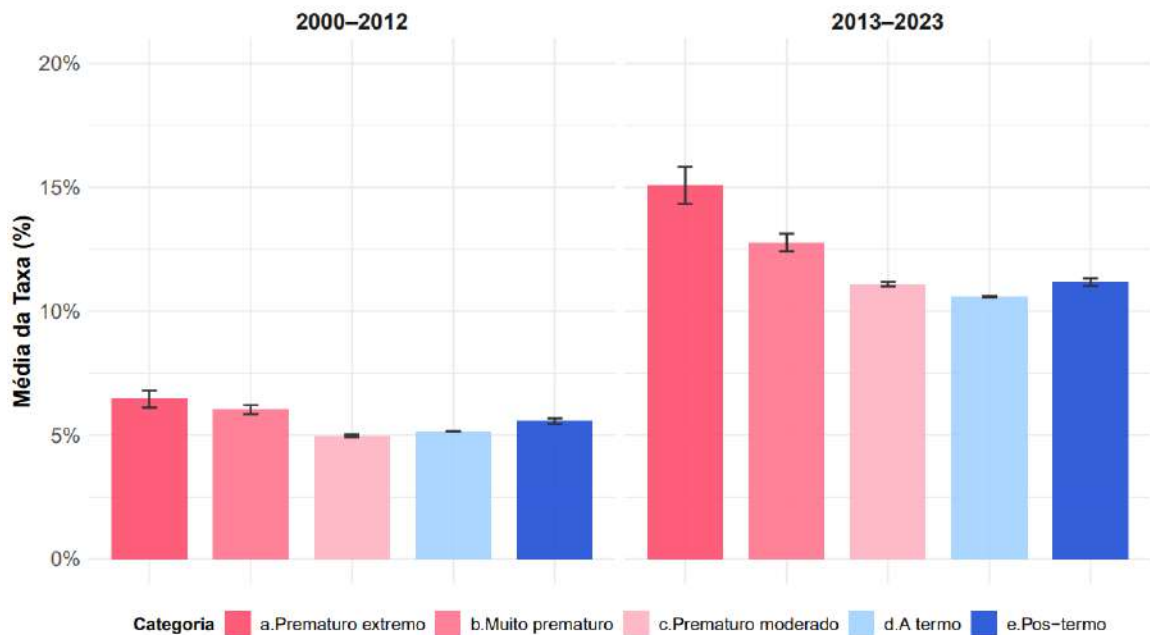


Fonte: Elaboração Própria

Os resultados do gráfico das médias foram coerentes com aqueles observados nos boxplots. No período de 2000 a 2012, observado na Figura 14, as maiores médias observadas correspondem aos nascimentos prematuros extremos (6,31%) e muito prematuros (6,18%), enquanto os partos a termo apresentam média de 5,14%. Já a categoria de prematuros moderados (4,99%) apresenta média inferior à observada entre os partos a termo, comportamento este que coincide com o padrão observado no boxplot (Figura 13) para o período em questão.

Para o período de 2013 a 2023, conforme observado na Figura 14, todas as cinco categorias apresentam médias de exposição superiores às encontradas no período anterior. As médias das taxas de exposição das três categorias que representam nascimentos prematuros, “Prematuro Extremo” (14,7%), “Muito Prematuro” (12,7%) e “Prematuro Moderado” (11,1%), são superiores à média da taxa de exposição dos partos a termo (10,6%). A discrepância entre os grupos de nascimentos prematuros e os partos não prematuros no período mais recente reforça, portanto, o padrão observado nos boxplots representados na Figura 13.

Figura 14: Média da Taxa de Exposição por Categoria de Prematuridade – Verão



Fonte: Elaboração Própria

Com o objetivo de embasar os resultados descritivos obtidos, verificamos se as diferenças encontradas são estatisticamente significativas, permitindo avaliar a associação entre elevadas taxas de exposição a OC e desfechos de prematuridade em gestações que transcorrem ao longo de todo o verão.

3.2.3 Análise Inferencial das Gestações no Verão – Testes Não Paramétricos

Com o objetivo de comparar as distribuições das taxas de exposição entre as diferentes categorias por meio de testes estatísticos inferenciais, foi utilizado um teste não paramétrico. A escolha desse tipo de teste ocorreu pois, como as taxas de exposição nas categorias de prematuridade apresentaram comportamento assimétrico, em especial no período de 2013 a 2023 (Figura 13), os pressupostos de normalidade exigidos para a aplicação de testes paramétricos não foram atendidos.

Nesse contexto, como o primeiro objetivo é verificar a existência de alguma categoria de prematuridade cuja distribuição da taxa de exposição apresente diferença estatisticamente significativa em relação às demais, optamos pela utilização do teste não paramétrico de Kruskal-Wallis (Kruskal & Wallis, 1952), o qual serve como possível evidência inicial de associação entre exposição ao calor extremo e prematuridade. O teste de Kruskal-Wallis permite a comparação entre múltiplas categorias de prematuridade e, como neste caso temos

cinco categorias, adequa-se melhor a esse contexto do que o teste de Mann-Whitney, por exemplo, o qual permite a comparação entre apenas dois grupos.

A hipótese nula do teste de Kruskal-Wallis assume a ausência de diferenças estatisticamente significativas nas distribuições das taxas de exposição entre todas as categorias, ou seja, indica equivalência entre suas distribuições. Dessa forma, a rejeição da hipótese nula indica a existência de evidências estatísticas, a um certo nível de significância, de diferenças entre as distribuições das taxas de exposição de pelo menos um par de categorias de prematuridade. A estatística do teste (H) é baseada na comparação de postos entre os grupos e, sob a hipótese nula de igualdade entre as distribuições das categorias, possui distribuição assintótica qui-quadrado com “k-1” graus de liberdade, em que “k” representa o número de categorias analisadas. Os resultados do teste de Kruskal-Wallis encontram-se na Tabela 6.

Os resultados da Tabela 6 indicaram rejeição da hipótese nula, ao nível de significância de 1%, tanto para o período de 2000 a 2012 quanto para o período de 2013 a 2023, visto que os p-valores encontrados dos dois testes foram menores que 0,01. Ou seja, há evidências estatisticamente significativas de que pelo menos uma das categorias de prematuridade apresenta diferenças na distribuição das taxas de exposição em relação a alguma outra categoria. Esse resultado corrobora o padrão observado na Figura 13, através da qual foi possível identificar níveis de exposição mais elevados para as categorias “Prematuro Extremo” e “Muito Prematuro” em comparação aos partos a termo.

Tabela 6: Resultados do teste de Kruskal–Wallis por categoria de prematuridade - Verão

Período	Graus de Liberdade	Estatística H	P-Valor	Decisão
2000-2012	4	127	$< 2 \times 10^{-16}$	Rejeita Hipótese Nula
2013-2023	4	2148	$< 2 \times 10^{-16}$	Rejeita Hipótese Nula

Fonte: Elaboração Própria

Tendo em vista a rejeição das hipóteses nulas dos testes de Kruskal-Wallis, foi realizada uma análise relacionada à categoria “A termo”, que representa a ampla maioria dos partos não prematuros (98,2%), visto que o objetivo deste estudo é investigar o impacto das OC sobre nascimentos prematuros em comparação aos não prematuros. Portanto, a etapa

seguinte ao teste de Kruskal-Wallis, que consiste na comparação entre múltiplas categorias, foi a utilização de um teste que possibilitasse a comparação individual de cada categoria de prematuridade em relação ao grupo de partos A termo, o qual corresponde à categoria de referência. Nesse contexto, um teste não-paramétrico que atende a essa demanda é o teste de Mann-Whitney, visto que este teste permite a comparação entre dois grupos distintos e, por ser não-paramétrico, não exige normalidade dos dados (MANN; WHITNEY, 1947).

Dessa forma, as distribuições das taxas de exposição referentes a cada uma das categorias de prematuridade, representadas pelos grupos “Prematuro Extremo”, “Muito Prematuro” e “Prematuro Moderado”, foram comparadas individualmente com o grupo “A termo”, totalizando três testes de Mann-Whitney realizados para cada período analisado (2000 a 2012 e 2013 a 2023). A hipótese nula desse teste assume que a distribuição da taxa de exposição da categoria em questão é estatisticamente equivalente àquela relativa à categoria “A termo”, enquanto a hipótese alternativa indica que há diferenças estatisticamente significativas entre elas. O teste de Mann-Whitney utilizado consiste na ordenação conjunta das observações dos dois grupos comparados, sendo a estatística de teste (W) construída a partir da soma dos postos atribuídos às observações de cada categoria. Sob a hipótese nula de igualdade entre as distribuições, essa estatística apresenta distribuição assintótica aproximadamente normal para amostras grandes, como é o caso deste estudo.

A Tabela 7 exhibe os resultados dos três testes de Mann-Whitney para os períodos de 2000 a 2012 e 2013 a 2023. Alguns indicadores presentes na Tabela 7 incluem as médias das taxas percentuais de exposição para cada categoria de prematuridade, bem como a média do grupo “A termo”, a taxa de aumento relativo da categoria de prematuridade testada em relação aos partos a termo e os p-valores dos testes de Mann-Whitney. O cálculo da estatística “Taxa de Aumento”, indicada na Tabela 7, foi obtido por meio da diferença entre a média da categoria em questão e a média da categoria “A termo”, seguida da razão desse resultado pela média do grupo “A termo”.

Os resultados dos testes relativos ao período de 2013 a 2023 (Tabela 7) indicam que as categorias “Prematuro Extremo” e “Muito Prematuro” apresentam taxas médias de exposição mais elevadas que aquelas do grupo “A termo”, com diferenças estatisticamente significativas em relação a esse grupo de referência. Quanto ao período de 2000 a 2012, apesar de a taxa de exposição da categoria “Prematuro Moderado” ser inferior àquela observada nas gestações não prematuras, as duas categorias de prematuridade mais acentuada também apresentam, assim como no período subsequente, médias de exposição superiores às do grupo “A termo”. No período de 2013 a 2023, as taxas de aumento das médias de exposição em relação ao

grupo “A termo” apresentaram valores percentuais maiores do que aqueles observados no período de 2000 a 2012, especialmente para a categoria “Prematuro Extremo”.

Tabela 7: Teste de Mann–Whitney: Comparação da taxa média de exposição a ondas de calor entre categorias de prematuridade e a categoria A termo – Verão

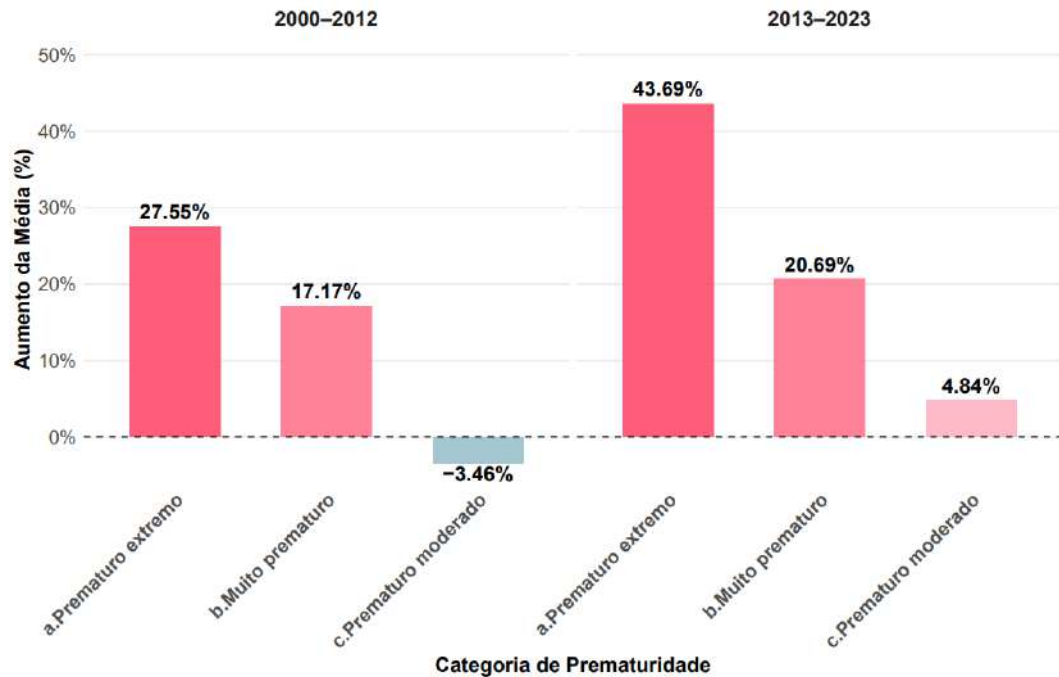
Período	Grupo	Média da Categoria	Taxa de Aumento (%)	P-Valor
2000-2012	A Termo	5,14	—	—
	Prematuro Extremo	6,56	27,55	< 0,01
	Muito Prematuro	6,02	17,17	< 0,01
	Prematuro Moderado	4,96	-3,46	< 0,01
2013-2023	A Termo	10,58	—	—
	Prematuro Extremo	15,2	43,69	< 0,01
	Muito Prematuro	12,77	20,69	< 0,01
	Prematuro Moderado	11,09	4,84	< 0,01

Fonte: Elaboração Própria

Como os p-valores dos testes de Mann-Whitney das três categorias foram, em ambos os períodos, menores que 0,01, rejeitamos as hipóteses nulas dos testes a um nível de significância de 1%. Ou seja, há evidências estatísticas significativas que indicam diferença entre as distribuições das taxas de exposição de cada uma das três categorias de prematuridade em relação à taxa de exposição das gestações do grupo “A termo”. Esse resultado reforçou a hipótese de associação entre maior exposição a OC e prematuridade.

Complementarmente aos testes, a Figura 15, que aborda o período de 2000 a 2012 e 2013 a 2023, permite a visualização gráfica do aumento percentual da média da taxa de exposição de cada categoria de prematuridade em relação ao grupo “A termo”.

Figura 15: Alteração na Média da Exposição em Relação ao Grupo A Termo



Fonte: Elaboração Própria

No período de 2013 a 2023, conforme observado na Figura 15 e na Tabela 7, as médias das taxas de exposição das três categorias que caracterizam nascimentos prematuros foram superiores à média do grupo “A termo”, que representa os partos não prematuros. Em contrapartida, no período de 2000 a 2012, a categoria “Prematuro Moderado” apresentou média de exposição inferior à observada no grupo “A termo”, enquanto as categorias “Prematuro Extremo” e “Muito Prematuro” registraram médias superiores. Essa diferença indica que os impactos da exposição a OC podem ser mais intensos para os grupos de maior grau de prematuridade.

Assim, a combinação dos resultados dos testes não paramétricos de Kruskal-Wallis e Mann-Whitney forneceu evidências significativas que sugerem associação entre elevadas taxas de exposição a OC e aumento da frequência de nascimentos prematuros que transcorrem toda a estação do verão, especialmente para as categorias “Prematuro Extremo” e “Muito Prematuro”, as quais representam maior risco.

3.2.4 Modelo de Regressão Logística para Gestações no Verão

Nesta subseção apresenta-se, assim como realizado para o modelo geral definido na Subseção 3.1.2, a aplicação do modelo de regressão logística, bem como as tabelas e cálculos referentes aos resultados do ajuste do modelo.

Nesse contexto, utilizamos, novamente, um modelo de regressão logística com função logito, previamente apresentado na Subseção 3.1.2. Neste modelo, a variável resposta é a indicadora binária de prematuridade, assumindo valor “um” quando o parto é prematuro (“Prematuro Extremo”, “Muito Prematuro” ou “Prematuro Moderado”) e “zero” quando não é (“A termo” ou “Pós-termo”), e as variáveis explicativas são as mesmas do modelo anterior, em que o principal parâmetro corresponde à taxa de exposição a OC ao longo da gestação.

Com isso, assim como realizado no modelo anterior, verificamos se há presença de colinearidade entre as variáveis explicativas incluídas neste modelo, pois correlações elevadas entre as covariáveis podem causar interpretações inadequadas dos efeitos dos coeficientes. A metodologia utilizada para verificar a colinearidade foi, novamente, a análise do fator de inflação da variância (VIF) (KUTNER et al., 2005) das covariáveis.

Tabela 8: Análise de Colinearidade – VIFs das Variáveis Explicativas do Modelo - Verão

Variáveis Explicativas	VIF	Graus de Liberdade	Interpretação
Taxa de Exposição	1,004	1	Baixa Colinearidade
Idade da Mãe	1,021	1	Baixa Colinearidade
Sexo do Bebê	1,001	2	Baixa Colinearidade
Raça/Cor do Bebê	1,024	5	Baixa Colinearidade

Fonte: Elaboração Própria

Conforme observado na Tabela 8, todos os valores dos fatores de inflação da variância encontrados para este modelo, assim como para o modelo geral, foram próximos a 1. Este resultado indica baixa multicolinearidade entre as variáveis explicativas e, portanto, não se justifica a remoção de nenhuma covariável do modelo.

Os coeficientes estimados ($\hat{\beta}_i$), seus respectivos erros-padrão e p-valores referentes ao modelo de regressão logística encontram-se na Tabela 9. Os resultados do modelo ajustado indicam que a variável “Taxa de Exposição”, foco deste estudo, apresenta coeficiente estimado $\hat{\beta}_1$ igual a 1,4457 e p-valor aproximadamente igual a zero (p – valor $< 0,0001$). O

sinal positivo do coeficiente estimado, juntamente com seu p-valor inferior a 0,05, indica que aumentos na taxa de exposição a OC estão associados, de forma estatisticamente significativa ao nível 5%, a maiores chances de ocorrência de parto prematuro. Observamos, ainda, que o coeficiente estimado da “Taxa de Exposição” apresentou valor mais elevado que aquele do modelo definido na Subseção 3.1.2 ($\hat{\beta}_1 = 0,4696$), apresentado na Tabela 4, o qual considera o conjunto geral de todos os nascimentos entre 2000 e 2023.

Tabela 9: Sumário do Modelo de Regressão Logística para Prematuridade - Verão

Variáveis Explicativas	Coeficiente Estimado ($\hat{\beta}_1$)	Erro Padrão	P-Valor
Intercepto	-3,0069	0,0816	$< 2 \times 10^{-16}$
Taxa de Exposição	1,4457	0,0594	$< 2 \times 10^{-16}$
Idade da Mãe	0,0149	0,0006	$< 2 \times 10^{-16}$
Sexo do Bebê (Masculino)	0,0539	0,0088	$8,38 \times 10^{-10}$
Sexo do Bebê (Sem Informação)	0,2005	0,0973	0,0393
Raça/Cor do Bebê (Branca)	-0,0356	0,0796	0,6541
Raça/Cor do Bebê (Indígena)	-0,0839	0,1771	0,6359
Raça/Cor do Bebê (Parda)	0,0532	0,0796	0,5038
Raça/Cor do Bebê (Preta)	0,1172	0,0806	0,1461
Raça/Cor do Bebê (Sem Informação)	0,0323	0,0821	0,6943

Fonte: Elaboração Própria

A maior magnitude do efeito observada no modelo ajustado para os partos cujas gestações transcorrem todo o verão sugere uma possível associação mais intensa entre a exposição a OC e a ocorrência de nascimentos prematuros restritos a esse período.

Tabela 10: Razões de Chances e Intervalos de Confiança das Variáveis Explicativas - Verão

Variáveis Explicativas	Razão de Chances (OR)	Intervalo de Confiança 95%
Intercepto	0,0494	[0,0421 ; 0,0580]
Taxa de Exposição	4,2448	[3,7782 ; 4,7690]
Idade da Mãe	1,0150	[1,0137 ; 1,0163]
Sexo do Bebê (Masculino)	1,0554	[1,0374 ; 1,0737]
Sexo do Bebê (Sem Informação)	1,2220	[1,0098 ; 1,4789]
Raça/Cor do Bebê (Branca)	0,9650	[0,8256 ; 1,1278]
Raça/Cor do Bebê (Indígena)	0,9196	[0,6498 ; 1,3013]
Raça/Cor do Bebê (Parda)	1,0546	[0,9024 ; 1,2326]
Raça/Cor do Bebê (Preta)	1,1243	[0,9600 ; 1,3168]
Raça/Cor do Bebê (Sem Informação)	1,0328	[0,8793 ; 1,2130]

Fonte: Elaboração Própria

As razões de chances (OR) das variáveis explicativas deste modelo, juntamente com seus respectivos intervalos de confiança a 95%, estão apresentadas na Tabela 10. Quanto à variável de interesse “Taxa de Exposição”, a razão de chances estimada (OR) foi de 4,2448, com intervalo de confiança de 95% definido por [3,778;4,769]. Esse coeficiente estimado indica que um aumento de uma unidade na taxa de exposição a OC está associado a um aumento de aproximadamente 324,5% nas chances de ocorrência de parto prematuro.

A variável “Taxa de Exposição”, assim como observado no modelo geral descrito na Subseção 3.1.2, assume valores apenas no intervalo [0; 0,338] para gestações que transcorrem ao longo de todo o verão. Dessa forma, os aumentos nas chances de ocorrência de prematuridade correspondem a frações do efeito estimado para a variação unitária da taxa (324,5%).

No modelo descrito nesta subseção, foi utilizada novamente a expressão que permite o cálculo das variações nas chances de ocorrência de prematuridade a partir do aumento de uma unidade da variável “Taxa de Exposição” (HOSMER; LEMESHOW; STURDIVANT, 2013). Para estimar a razão de chances de prematuridade associada a um aumento de 0,10 unidade na “Taxa de Exposição” ($\Delta=0,10$), utilizamos o coeficiente estimado para essa variável ($\hat{\beta}_1 =$

1,4457), indicada na Tabela 9, ou, de forma equivalente, a razão de chances unitária dessa variável, presente na Tabela 10, a qual corresponde a 4,2248 ($OR_{\text{unitário}}$).

$$OR(\Delta) = \exp(\hat{\beta}_1 * \Delta) = (OR_{\text{unitário}})^\Delta = (4,2448)^{0,1} ;$$

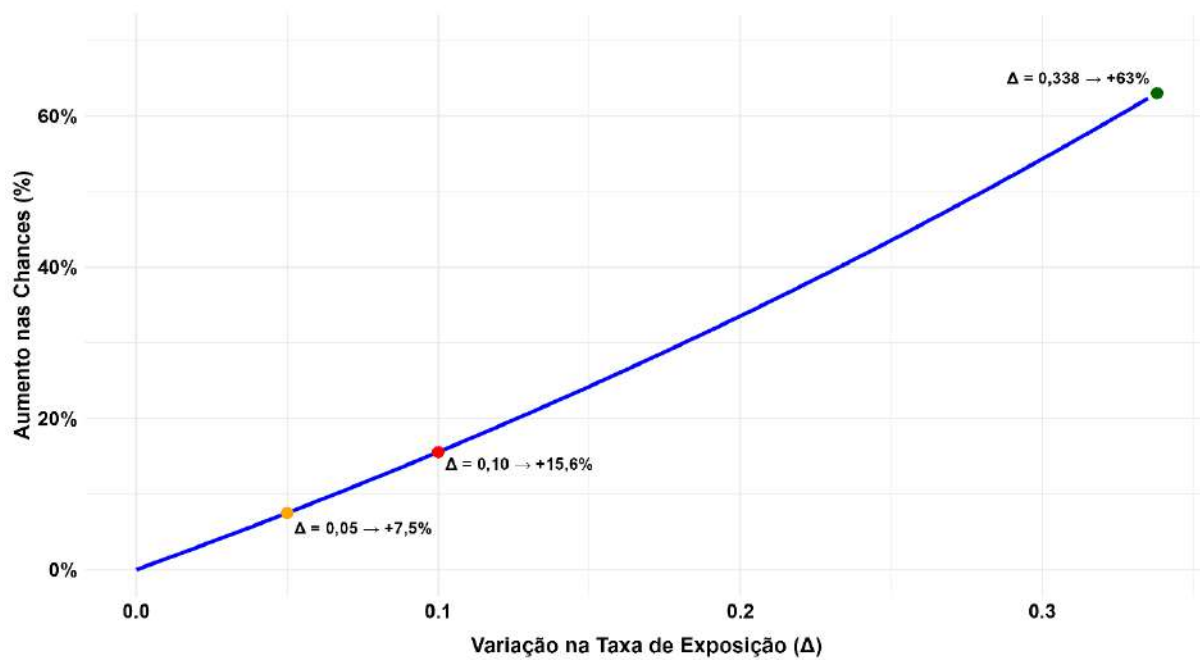
$$\text{Aumento da Chance (0,1)} = (1,1555 - 1) * 100\% \approx 15,55\%.$$

Dessa forma, um incremento de 0,10 unidade na taxa de exposição a OC está associado a um aumento estimado de, aproximadamente, 15,6% nas chances de ocorrência de prematuridade. Essa estimativa foi obtida ao elevar a razão de chances unitária da “Taxa de Exposição” ao expoente 0,10. Os resultados do modelo de regressão logística corroboram, portanto, a hipótese levantada após a análise da Figura 13 de que existe associação positiva entre a exposição a OC e a ocorrência de prematuridade em gestações que transcorrem por todo o verão.

Enquanto o modelo geral, apresentado na Subseção 3.1.2, retornou um aumento de aproximadamente 4,8% nas chances de prematuridade para um aumento de 0,1 na taxa de exposição, o modelo restrito às gestações que transcorrem todo o verão apresentou incremento de 15,6% nas chances de prematuridade, valor este que supera o triplo do valor do modelo geral. Este resultado sugere que exposições concentradas no período do verão possivelmente apresentam um impacto mais intenso sobre o desfecho de prematuridade.

A Figura 16 ilustra esse comportamento ao longo de diferentes variações na taxa de exposição. Observamos uma relação exponencial entre a variação da exposição e o aumento percentual nas chances de prematuridade, com incrementos mais acentuados à medida que a exposição aumenta, atingindo aproximadamente 63% para a maior variação considerada ($\Delta = 0,338$). Quando comparado ao modelo geral (Figura 9), observamos que o aumento estimado nas chances de prematuridade, para uma mesma variação da taxa de exposição, é mais elevado para as gestações que transcorrem integralmente durante o verão. Para uma variação de 0,10 na taxa de exposição, o modelo geral indica um aumento aproximado de 4,8% nas chances de prematuridade, enquanto no modelo restrito ao período de verão esse incremento atinge cerca de 15,6%.

Figura 16: Efeito da Variação da Exposição sobre as Chances de Prematuridade– Verão



Fonte: Elaboração Própria

4 CONSIDERAÇÕES FINAIS

Este estudo investiga a relação estatística entre a exposição a eventos extremos de calor durante o período gestacional e a ocorrência de nascimentos prematuros na cidade do Rio de Janeiro, no período de 2000 a 2023.

Os resultados evidenciaram uma tendência crescente na frequência de dias com ondas de calor (OC), bem como a existência de associação estatisticamente significativa entre prematuridade e exposição a esses eventos, obtida por modelos de regressão logística. Em particular, foi determinado pelo modelo que um aumento de 0,10 na taxa de exposição gestacional implica um incremento percentual aproximado de 4,8% nas chances de ocorrência de um parto prematuro, considerando o ajuste por demais fatores demográficos e socioeconômicos. Ressaltamos que a razão de chances (OR) igual a 1,599 representa o efeito global da variável de exposição no modelo, enquanto o aumento de 4,8% nas chances de prematuridade corresponde ao efeito marginal associado a um incremento absoluto de 0,10 na taxa de exposição, definida como a proporção de dias gestacionais com ocorrência de OC, o que equivale a um aumento de 10 pontos percentuais nessa proporção. No entanto, a análise visual das distribuições não revelou padrões claros de diferença na exposição entre os diferentes grupos de prematuridade.

Os resultados obtidos neste estudo foram consistentes com as evidências da literatura internacional que apontam associação positiva entre exposição a OC e aumento do risco de prematuridade. O estudo de Darrow et al. (2024), realizado nas 50 maiores regiões metropolitanas dos Estados Unidos, estimou que o efeito da exposição a OC sobre as taxas de prematuridade é estatisticamente significativo, porém de baixa magnitude, com razões de taxas (*rate ratios*) próximas de 1,02. A razão de taxas, utilizada no estudo de Darrow et al., é uma medida diferente da razão de chances (*odds ratio*), porém permite, nesse contexto, comparar a taxa de ocorrência de prematuridade entre gestantes expostas e não expostas a OC. Tal medida é calculada pela razão entre a taxa de ocorrência de prematuridade no grupo exposto a OC e a taxa correspondente observada no grupo não exposto.

Em outro artigo, publicado por Fitch et al. (2025b), os autores relataram, a partir de um estudo realizado em oito estados estadunidenses, uma razão de chances de 1,015 para a ocorrência de nascimentos prematuros após exposição a OC. Ao compararmos os resultados deste presente estudo com aqueles obtidos por Fitch et al., podemos observar que a intensidade do efeito registrada na cidade do Rio de Janeiro (OR = 1,599) foi mais elevada do que aquela obtida no estudo dos estados americanos (OR=1,015). Essa diferença na

intensidade do efeito pode estar associada a fatores socioeconômicos e estruturais, como o acesso aos serviços de saúde e as condições de vulnerabilidade social, que podem contribuir para uma maior suscetibilidade aos efeitos do calor, uma vez que populações mais vulneráveis tendem a ser desproporcionalmente afetadas por eventos climáticos extremos (IPCC, 2022).

Em contrapartida, ao restringirmos a análise às gestações que atravessam integralmente o verão, período caracterizado pela maior frequência e intensidade das OC, os resultados sugerem uma associação mais expressiva. No período mais recente (2013-2023), em especial, as categorias “Prematuro Extremo” e “Muito Prematuro”, que correspondem aos casos mais acentuados de prematuridade, apresentaram distribuição da exposição e taxas médias de exposição superiores às observadas nos partos “A termo”.

A análise descritiva do número de dias de OC por ano, estratificada também por estação do ano, embasa a escolha das gestações que transcorrem integralmente durante o verão como alvo da análise sazonal. Isso se justifica pois, especialmente entre o período de 2012 e 2020, foi observado que o número de dias de OC no verão foi significativamente superior ao das demais estações.

Esse padrão observado na capital fluminense durante o verão também foi descrito por Monteiro dos Santos et al. (2024), que apontaram que, principalmente a partir de meados da década de 2010, os eventos de OC na região Sudeste do Brasil passam a se concentrar com maior frequência no verão. De acordo com os autores, o aumento da frequência desses eventos foi acompanhado por uma maior intensidade e duração dos mesmos nesse período, reforçando a escolha do verão como estação central para a análise neste trabalho.

Nesse contexto, o modelo de regressão logística para o subgrupo de gestações que abrangem todo o período do verão revela que um aumento de 10% na taxa de exposição está associado a um acréscimo de cerca de 15,6% nas chances de prematuridade, efeito esse que corresponde, aproximadamente, ao triplo do valor observado na análise do modelo geral. Esse resultado destaca o verão como um período de risco potencialmente mais elevado em comparação às demais estações, sugerindo que a exposição concentrada ao calor extremo, característica de tal estação, pode exercer um impacto mais evidente sobre o desfecho de prematuridade.

A interpretação das razões de chances obtidas nos modelos complementa essa particularidade das gestações que transcorrem ao longo de todo o verão. Enquanto o modelo geral apresentou razão de chances de 1,599 (1,47 - 1,74) para a variável “Taxa de Exposição”, indicando um aumento no risco, o modelo restrito ao verão gerou uma razão de chances de

4,24 (3,78 - 4,77). Essa diferença na magnitude das razões de chances sugere um aumento do risco associado à exposição concentrada no período do verão.

O efeito de sazonalidade encontrado foi consistente com o resultado obtido por Fitch et al. (2025a), em que se analisa o efeito das OC, restritas ao período de maior temperatura no hemisfério Norte (entre maio e setembro), sobre o desfecho de prematuridade em oito estados estadunidenses. Os resultados do estudo indicaram que essa associação é positiva, porém com magnitude inferior à encontrada no modelo do presente estudo, visto que, entre todos os subgrupos analisados por Fitch et al., o aumento máximo nas chances de ocorrência de prematuridade observada foi de aproximadamente 7,9%.

Após as análises dos modelos, concluímos que os resultados do presente trabalho contribuem para a literatura ao fornecer evidências da associação entre eventos de OC e prematuridade em centros urbanos com clima tropical, como a cidade do Rio de Janeiro. A conclusão sobre essa associação foi obtida por meio de testes estatísticos não paramétricos, como os testes de Kruskal-Wallis e Mann-Whitney, bem como modelos de regressão logística do tipo logito.

No entanto, a natureza dos dados de saúde apresentou desafios, como a imprecisão na data exata de concepção para as gestações ocorridas entre 2000 e 2012. Tais limitações foram contornadas, neste estudo, através da análise de correlação e de testes de equivalência (TOST), conforme descrito na Seção 2.2.

Nesse contexto, a associação entre a prematuridade e a exposição a OC mostrou-se complexa e de difícil quantificação. Ainda assim, a aplicação de modelos estatísticos mais robustos permitiu identificar e quantificar esse impacto, reforçando a relevância das abordagens metodológicas adotadas, sobretudo ao evidenciar que os efeitos da exposição a ondas de calor sobre o desfecho de prematuridade são mais pronunciados em períodos de maior exposição térmica (verão), especialmente no período mais recente, entre 2013 e 2023.

Diante das evidências estatísticas apresentadas, este estudo reforça a necessidade de considerar a exposição ao calor excessivo, incluindo a sazonalidade, no planejamento da saúde perinatal, de forma que a abordagem estatística adotada fornece um suporte fundamental para esse planejamento. Além disso, a identificação do verão como um período de risco potencialmente mais elevado para nascimentos prematuros, justifica a implementação de estratégias específicas de adaptação para essa estação. Nesse contexto, é importante destacar que a qualidade da atenção pré-natal no Brasil apresenta desigualdades associadas a fatores socioeconômicos, regionais e demográficos, como renda, idade e local de residência (TOMASI et al., 2017), reforçando a necessidade de elaboração de estratégias adaptadas às

diferentes realidades territoriais e níveis de vulnerabilidade característicos do contexto brasileiro. Dessa forma, algumas possíveis estratégias a serem adotadas podem estar relacionadas a um maior cuidado no acompanhamento de gestantes durante o verão, bem como a disponibilização de cursos profissionalizantes que aumentem a capacitação dos profissionais de saúde diante desse cenário.

Nesse contexto, este estudo oferece um arcabouço para a ampliação de pesquisas acerca da relação entre eventos extremos de calor e desfechos perinatais adversos. Uma possível continuidade deste projeto consiste na construção de modelos preditivos capazes de avaliar a probabilidade de ocorrência de prematuridade a partir dos indicadores demográficos e socioeconômicos utilizados neste estudo, bem como de outras variáveis relevantes. Outra possível análise a ser realizada consiste na investigação do impacto da severidade das ondas de calor sobre o desfecho de prematuridade, utilizando a categorização das ondas em “Baixa”, “Severa” e “Extrema”, descrita na Seção 2.3 deste trabalho, como base para a avaliação desse efeito. Dessa forma, os resultados apresentados neste trabalho podem contribuir para o aprofundamento das discussões acerca dos impactos dos eventos extremos de calor sobre a prematuridade, bem como para o desenvolvimento de estratégias específicas de acompanhamento de gestantes no período pré-natal.

REFERÊNCIAS

SANTOS, Djacinto Monteiro dos; LIBONATI, Renata; GARCIA, Beatriz N. et al. Twenty-first-century demographic and social inequalities of heat-related deaths in Brazilian urban areas. *PLOS ONE*, v. 19, n. 1, e0295766, 2024. Disponível em: <https://doi.org/10.1371/journal.pone.0295766>. Acesso em: 22 jul. 2025.

FISCHER, E. M.; SIPPEL, S.; KNUTTI, R. Increasing probability of record-shattering climate extremes. *Nature Climate Change*, Londres, v. 11, p. 689–695, 2021. Disponível em: <https://doi.org/10.1038/s41558-021-01092-9>. Acesso em: 29 dez. 2025.

CHERSICH, Matthew Francis et al. Associations between high temperatures in pregnancy and risk of preterm birth, low birth weight, and stillbirths: systematic review and meta-analysis. *BMJ*, Londres, v. 371, p. m3811, 2020. Disponível em: <https://doi.org/10.1136/bmj.m3811>. Acesso em: 29 dez. 2025.

ORGANIZAÇÃO MUNDIAL DE SAÚDE et al. Born too soon: decade of action on preterm birth. Geneva: *World Health Organization*, 2023. Disponível em: <https://borntoosoonaction.org/>. Acesso em: 3 fev. 2026.

R CORE TEAM. R: A language and environment for statistical computing. *Vienna: R Foundation for Statistical Computing*, 2025. Disponível em: <https://www.r-project.org/>. Acesso em: 27 maio 2026.

WICKHAM, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016. Disponível em: <https://ggplot2.tidyverse.org/>. Acesso em: 27 maio 2026.

WICKHAM, Hadley et al. *Welcome to the tidyverse*. *Journal of Open Source Software*, v. 4, n. 43, p. 1686, 2019. Disponível em: <https://joss.theoj.org/papers/10.21105/joss.01686?>. Acesso em: 27 maio 2026.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde e Ambiente. Perfil epidemiológico dos nascimentos prematuros no Brasil, 2012 a 2022. *Boletim Epidemiológico*, Brasília, DF, v. 55, n. 13, 30 set. 2024. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2024/boletim-epidemiologico-volume-55-no-13.pdf>. Acesso em: 3 fev. 2026.

BRASIL. Ministério da Saúde. Departamento de Informática do Sistema Único de Saúde (DATASUS). *Transferência de arquivos*. Brasília, DF: Ministério da Saúde, 2025. Disponível em: <https://datasus.saude.gov.br/transferencia-de-arquivos/>. Acesso em: 29 dez 2025.

NAIRN, J.; FAWCETT, R. The Excess Heat Factor: A metric for heatwave intensity and its use in classifying heatwave severity. *International Journal of Environmental Research and Public Health*, Basel, v. 12, n. 1, p. 227–253, 2014. Disponível em: <https://doi.org/10.3390/ijerph120100227>. Acesso em: 22 jul. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. Preterm birth. Geneva: OMS, 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Acesso em: 29 dez. 2025.

BRASIL. Empresa Brasileira de Serviços Hospitalares (EBSERH). Prematuridade – uma questão de saúde pública: como prevenir e cuidar. Natal, 30 nov. 2023. Disponível em: <https://www.gov.br/ebserh/pt-br/hospitais-universitarios/regiao-nordeste/huab-ufrn/comunicacao/noticias/prematuridade-2013-uma-questao-de-saude-publica-como-prevenir-e-cuidar>. Acesso em: 30 dez. 2025.

LAKENS, D. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, v. 8, n. 4, p. 355–362, 2017. Disponível em: <https://doi.org/10.1177/1948550617697177>. Acesso em: 16 out. 2025.

SCHUIRMANN, D. J. A comparison of the Two One-Sided Tests Procedure and the Power Approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, Rockville, v. 15, n. 6, p. 657–680, dez. 1987. Disponível em: <https://doi.org/10.1007/BF01068419>. Acesso em: 16 out. 2025.

HOSMER, David W.; LEMESHOW, Stanley; STURDIVANT, Rodney X. *Applied logistic regression*. 3. ed. Hoboken: John Wiley & Sons, 2013. Disponível em: <https://doi.org/10.1002/9781118548387>. Acesso em: 11 dez. 2025.

KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, v. 47, n. 260, p. 583–621, 1952. Disponível em: <https://doi.org/10.1080/01621459.1952.10483441>. Acesso em: 10 fev. 2026.

MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, v. 18, n. 1, p. 50–60, 1947. DOI: 10.1214/aoms/1177730491. Disponível em: <https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-18/issue-1/On-a-Test-of-Whether-one-of-Two-Random-Variables/10.1214/aoms/1177730491.full>. Acesso em: 26 fev. 2026.

KUTNER, Michael H.; NACHTSHEIM, Christopher J.; NETER, John; LI, William. *Applied Linear Statistical Models*. 5. ed. New York: McGraw-Hill/Irwin, 2005.

MONTGOMERY, Douglas C.; RUNGER, George C. *Applied Statistics and Probability for Engineers*. 7th ed. Hoboken: Wiley, 2018.

DARROW, L. A. et al. Preterm and Early-Term Delivery After Heat Waves in 50 US Metropolitan Areas. *JAMA Network Open*, [S. l.], v. 7, n. 5, p. e2412055, 2024. Disponível em: <https://doi.org/10.1001/jamanetworkopen.2024.12055>. Acesso em: 4 fev. 2026.

FITCH, A. et al. Heat Waves and Early Birth: Exploring Vulnerability by Individual- and Area-Level Factors. *GeoHealth*, [S. l.], v. 9, n. 4, p. e2025GH001348, 2025a. Disponível em: <https://doi.org/10.1029/2025GH001348>. Acesso em: 4 fev. 2026.

FITCH, A. et al. Preterm and early-term delivery after heat waves in eight US states: A case-crossover study using the High-resolution Urban Meteorology for Impacts Dataset (HUMID). *medRxiv*, [S. l.], 2025b. Pré-publicação (preprint). Disponível em: <https://doi.org/10.1101/2025.03.04.25323373>. Acesso em: 4 fev. 2026.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Cambridge: Cambridge University Press, 2022.

TOMASI, Elaine et al. Qualidade da atenção pré-natal na rede básica de saúde do Brasil: indicadores e desigualdades sociais. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 33, n. 3, e00195815, 2017. Disponível em: <https://doi.org/10.1590/0102-311X00195815>. Acesso em: 29. abr 2026.