



UNIVERSIDADE FEDERAL DE PELOTAS
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA
MESTRADO EM EPIDEMIOLOGIA



IDENTIFICAÇÃO E DESCRIÇÃO DE GRUPOS DE RISCO DE CRIANÇAS NUNCA
VACINADAS EM PAÍSES DE BAIXA E MÉDIA RENDA POR MEIO DE ÁRVORES
DE DECISÃO

Dissertação de mestrado

Mestrando: Thiago Melo Santos
Orientador: Prof. Dr. Alúcio J. D. Barros

Pelotas, RS
Fevereiro de 2021

S237i Santos, Thiago Melo

Identificação e descrição de grupos de risco de crianças nunca vacinadas em países de baixa e média renda por meio de árvores de decisão / Thiago Melo Santos ; Aluisio J. D. Barros, orientador. – Pelotas : Universidade Federal de Pelotas, 2020.

230 f. : il.

Dissertação (Mestrado) – Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal de Pelotas, 2020.

1. Epidemiologia. 2. Cobertura vacinal 3. Árvores de decisão I. Barros, Aluisio J. D., orient. II. Título.

CDD 614.4

Ficha catalográfica: M. Fátima S. Maia CRB 10/1347

THIAGO MELO SANTOS

IDENTIFICAÇÃO E DESCRIÇÃO DE GRUPOS DE RISCO DE CRIANÇAS NUNCA
VACINADAS EM PAÍSES DE BAIXA E MÉDIA RENDA POR MEIO DE ÁRVORES
DE DECISÃO

Banca examinadora:

Prof. Dr. Aluísio J. D. de Barros
Universidade Federal de Pelotas
Presidente da banca

Prof^a. Dr^a. Rejane Sobrino Pinheiro
Universidade Federal do Rio de Janeiro
Examinadora externa

Prof. Dr. Fernando Pires Hartwig
Universidade Federal de Pelotas
Examinador interno

Pelotas, 26 de fevereiro de 2021

Agradecimentos

Em primeiro lugar, agradeço a meus pais. Foi o investimento contínuo de vocês na minha educação, acima das suas próprias necessidades, que tornou possível o meu ingresso na academia. Que eu possa fazer pelas próximas gerações tanto quando vocês fizeram pela minha.

Ao Aluísio, pelo apadrinhamento e sabedoria. Eu não poderia ter escolhido um orientador que melhor conseguisse balancear liberdade criativa com proteção e direcionamento. Agradeço enormemente por me mostrar onde eu deveria estar e pelo voto de confiança em um desconhecido. Que eu possa um dia orientar alguém como você me orientou e continua a orientar.

A todos os meus professores e professoras, da infância à vida adulta. Vocês me apresentaram a prática da ciência e a paixão pelo conhecimento. E de quebra me convenceram do prazer de ensinar. Que eu possa reproduzir o que vocês me ensinaram e ajudar a formar indivíduos como vocês me formaram.

À Bianca, pela interação diária, amizade e troca. Você me apoiou em todas as bases que eu poderia pedir ao longo do mestrado e contribuiu constantemente para esta dissertação. Que eu possa te dar o mesmo nível de suporte e contribuição.

À Cíntia, pelo trabalho e disposição constantes. Você me ajudou na produção dos materiais que compõe o artigo desta dissertação, emprestando seu talento, habilidade e carinho. Que eu possa manter seu profissionalismo e leveza em todo lugar que eu trabalhe.

A todos os colegas de mestrado, da numera e do Centro Internacional de Equidade em Saúde, especialmente os membros da sala 331. Vocês me receberam de braços abertos, me aceitaram como um de vocês e tornaram meus últimos anos alguns dos mais prazerosos e produtivos da minha vida. Que eu possa retribuir o carinho e dedicação de vocês.

A todos os membros do Programa de Pós-Graduação em Epidemiologia da UFPel, pelo nível de excelência com que vocês conduzem cada atividade. Tenho orgulho no peito de fazer parte de um programa nacional que alcança esse nível de

qualidade e reconhecimento. Que eu possa contribuir para a ciência brasileira como vocês contribuem diariamente.

Finalmente, a todos os funcionários e funcionárias, gestores e gestoras, ativistas e membros de organizações sociais, governo, universidades, serviços de saúde e órgãos financiadores que lutaram e lutam pela saúde e educação pública no Brasil. Esta dissertação é também produto do trabalho de vocês. As condições exponencialmente melhores de vida da minha geração é também produto do trabalho de vocês. Que eu possa, assim como vocês, ajudar a manter o momentum de avanço do Brasil na direção de uma nação mais igualitária, consciente e que não deixa ninguém para trás.

Resumo

Santos, Thiago Melo. **Identificação e descrição de grupos de risco de crianças nunca vacinadas em países de baixa e média renda por meio de árvores de decisão** [dissertação de mestrado]. Pelotas: Programa de Pós-Graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal de Pelotas; 2021.

A fim de “não deixar ninguém para trás”, a Agenda de Imunização 2030 tem como meta a redução de desigualdades em vacinação. Árvores de decisão são métodos estatísticos robustos que podem ser usadas para identificar grupos de risco de não vacinação entre crianças – denominadas crianças dose zero – em países de baixa e média renda (PBMR). Árvores de classificação e regressão (CART, em inglês) foram criadas para 92 PBMRs e para todos os países combinados usando dados de inquéritos Demographic and Health Surveys (DHS) e Multiple Indicator Cluster Surveys (MICS) a fim de identificar grupos de risco de dose zero entre crianças de 12 a 23 meses. Os grupos de risco identificados foram caracterizados em termos de riqueza, educação da mãe, local de residência e sexo. Três grupos de risco foram identificados e aquele com a maior prevalência de dose zero (42%) representava 4% de todas as crianças, mas correspondia a uma dentre cada quatro crianças dose zero. Este grupo foi composto por crianças cujas mães não receberam nenhuma dose da vacina antitetânica, não fizeram pré-natal e cujo parto não foi realizado em uma unidade de saúde. Quanto maior a prevalência de dose zero em um grupo, mais pobres, rurais e com mães menos instruídas eram as crianças desse grupo. As árvores nacionais apresentaram um resultado semelhante. Existe uma oportunidade de melhor integração entre estes serviços para que os Objetivos do Desenvolvimento Sustentável possam ser alcançados. As árvores de decisão se mostraram uma técnica promissora para identificação de grupos de risco na pesquisa epidemiológica.

Palavras-chave

Vacinação, imunização, cuidado pré-natal, árvore de decisão, CART, dose zero

Abstract

Santos, Thiago Melo. **Identification and description of never vaccinated children risk groups in low- and middle-income countries using decision trees** [thesis]. Pelotas (BR): Postgraduate Program in Epidemiology, Medical School, Federal University of Pelotas; 2021.

In order to “leave no one behind”, the Immunization Agenda 2030 targets reduction in vaccination inequalities as a goal. Decision trees are robust statistical tools that can be used for identifying risk groups of non-vaccination in children, or zero dose children, from low- and middle-income countries (LMIC). Classification and Regression Trees were created for 92 LMICs and for all countries combined using Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) in order to identify zero dose risk groups among children aged 12-23 months. Risk groups were characterized in terms of wealth, mother’s education, place of residence and sex. Three risk groups were identified and the one with the highest zero dose prevalence (42%) represented 4% of all children, but they accounted for one in every four zero-dose children. This group was composed of children whose mother did not received any doses of the tetanus vaccine, who did not have any antenatal care visits and whose delivery was not in a health facility. The higher the zero-dose prevalence in a group, the poorer, more rural and with a less educated mother the children in that group. National trees had a similar result. There is an opportunity for better integration between those services in order to achieve the Sustainable Development Goals. Decision trees have shown to be a promising technique for identifying risk groups in epidemiological research.

Keywords

Vaccination, immunisation, unvaccinated, antenatal care, decision tree, CART, zero dose

Apresentação

Esta dissertação foi elaborada por Thiago Melo Santos, sob orientação do Professor Aluísio J. D. Barros, como parte do mestrado em Epidemiologia do Programa de Pós-Graduação em Epidemiologia da Universidade Federal de Pelotas. O projeto de pesquisa foi desenvolvido no Centro Internacional de Equidade em Saúde | Pelotas.

Esta dissertação é composta pelos seguintes itens:

1. **Projeto de pesquisa:** projeto defendido, avaliado e aprovado no dia 19 de setembro de 2019, com a banca composta por Fernando P. Hartwig e Aluísio J. D. Barros;
2. **Modificações no projeto;**
3. **Relatório de trabalho de campo:** descreve as atividades do consórcio de pesquisa que compõe a maioria das dissertações de mestrado do Programa de Pós-Graduação em Epidemiologia. Apesar de a presente dissertação não ter sido realizada no consórcio de pesquisa, todos os alunos do mestrado participaram das atividades de campo;
4. **Relatório das atividades no Centro Internacional de Equidade em Saúde | Pelotas (ICEH):** descreve as atividades de trabalho no ICEH realizadas durante o mestrado, incluindo aquelas diretamente relacionadas a esta dissertação;
5. **Artigo original:** artigo original baseado nos resultados da dissertação, formatado de acordo com as normas da revista Vaccine;
6. **Comunicação para a imprensa.**

Sumário

Agradecimentos	4
Resumo	6
Abstract	7
Apresentação	8
1. PROJETO DE PESQUISA	11
1.1. Introdução	15
1.1.1. Dose zero de vacinação	15
1.1.2. Árvores de decisão	17
1.1.3. Árvores de decisão e não vacinação	19
1.2. Justificativa	19
1.3. Revisão da literatura	20
1.3.1. Dose zero	21
1.3.1.1. Estratégia de busca	21
1.3.1.2. Prevalência de dose zero	30
1.3.1.3. Fatores associados à dose zero	32
1.3.2. Árvores de decisão	35
1.3.2.1. Conceitos básicos	35
1.3.2.2. Implementações	38
1.3.2.3. Sobreajuste e poda	40
1.3.2.4. Revisão da literatura	42
1.4. Marco teórico e modelo conceitual	50
1.4.1. Desfecho	50
1.4.2. Acesso e intenção	51
1.4.3. Características	52
1.4.4. Fatores regionais/nacionais	52
1.4.5. Ambientes	56
1.5. Objetivos	62
1.5.1. Objetivo geral	62
1.5.2. Objetivos específicos	62
1.6. Hipóteses	62
1.7. Métodos	63
1.7.1. Fonte dos dados	63
1.7.2. População-alvo	63
1.7.2.1. População-alvo do estudo	63
1.7.2.2. População-alvo do DHS	64
1.7.2.3. População-alvo do MICS	65
1.7.2.4. Países estudados	65
1.7.3. Definição operacional do desfecho	66
1.7.4. Definição operacional das variáveis independentes	67
1.7.5. Processamento e análise de dados	69
1.8. Aspectos éticos	69
1.9. Divulgação dos resultados	70
1.10. Financiamento	70
1.11. Cronograma	71
1.12. Referências	72
1.13. Anexos	78
1.13.1. Anexo 1	78
1.13.2. Anexo 2	83
2. MODIFICAÇÕES NO PROJETO	84

3. RELATÓRIO DE TRABALHO DE CAMPO	87
3.1. Introdução.....	88
3.2. Comissões de trabalho de campo	90
3.2.1. Elaboração do projeto de pesquisa	90
3.2.2. Elaboração do questionário e manual de instruções	90
3.2.3. Gestão do banco de dados	91
3.2.4. Checagem das inconsistências e controle de qualidade	91
3.2.5. Comunicação de divulgação	92
3.2.6. Seleção e treinamento de pessoal	92
3.2.7. Logística	93
3.2.8. Financeiro.....	93
4. RELATÓRIO DE ATIVIDADES NO CENTRO INTERNACIONAL DE EQUIDADE EM SAÚDE PELOTAS	105
4.1. Projeto GAVI.....	106
4.2. Rotina de indicadores	107
4.3. Árvores de decisão	108
4.4. Outras atividades.....	108
5. ARTIGO ORIGINAL	110
Abstract	112
Introduction.....	113
Methods	115
Data source and children sample	115
Outcome	115
Potential indicators of vulnerability.....	116
Analysis methods.....	117
Ethical clearance	119
Results	120
Zero-dose prevalence.....	120
Pooled decision tree	120
Characterization of the risk groups	122
Country-specific trees	123
Discussion	124
Conclusion.....	129
Declaration of Competing Interest.....	130
Acknowledgements	130
Contributors.....	130
References	130
Supplementary materials.....	135
6. COMUNICAÇÃO PARA A IMPRENSA	230

1. PROJETO DE PESQUISA



UNIVERSIDADE FEDERAL DE PELOTAS
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA
MESTRADO EM EPIDEMIOLOGIA



IDENTIFICAÇÃO E DESCRIÇÃO DE GRUPOS DE RISCO DE CRIANÇAS NUNCA
VACINADAS EM PAÍSES DE BAIXA E MÉDIA RENDA POR MEIO DE ÁRVORES DE
DECISÃO

Projeto de Pesquisa

Mestrando: Thiago Melo Santos
Orientador: Prof. Dr. Aluísio J. D. Barros

Pelotas, RS
Setembro de 2019

LISTA DE FIGURAS

Figura 1 – Exemplo ilustrativo de uma árvore de decisão	18
Figura 2 – Estratégia de busca – dose zero.....	21
Figura 3 – Fluxograma do processo de busca na literatura sobre dose zero de vacinação	22
Figura 4 – Prevalência de dose zero de vacinação por BCG, OPV, DTP e MCV para países de baixa e média renda	31
Figura 5 – Exemplo de partições realizadas por árvore de decisão	36
Figura 6 – Estratégia de busca – árvores de decisão	43
Figura 7 – Fluxograma do processo de busca na literatura epidemiológica sobre árvores de decisão	43
Figura 8 – Principais objetivos para utilização das árvores de decisão na pesquisa epidemiológica de acordo com a revisão da literatura	47
Figura 9 – Distribuição dos métodos utilizados nos artigos	48
Figura 10 – Número de artigos publicados ao longo dos anos	48
Figura 11 – Modelo conceitual da dose zero de vacinação	61
Figura 12 – Cronograma proposto do projeto de dissertação	71
Figura 13 – Reprodução da seção de imunização do inquérito DHS – Angola 2015-2016	78
Figura 14 – Reprodução da seção de imunização do inquérito MICS – Guiné-Bissau 2014	81
Figura 15 – Reprodução do formulário de consentimento do inquérito DHS – Angola 2015-2016	83
Figura 16 – Reprodução do formulário de consentimento do inquérito MICS – Guiné-Bissau 2014.....	83

LISTA DE ABREVIATURAS E SIGLAS

DHS: *Demographic and Health Survey*

CART: *Classification and Regression Tree* ou Árvore de Classificação e Regressão

CTree: *Conditional Inference Tree* ou Árvore de Inferência Condicional

GAVI: *Global Alliance for Vaccines and Immunization* ou Aliança Global para Vacinas e Imunização

IC95%: Intervalo de confiança com nível de confiança de 95%

IIQ: Intervalo interquartil

MICS: *Multiple Indicator Cluster Survey*

OMS: Organização Mundial da Saúde

PAHO: *Pan American Health Organization* ou Organização Pan-Americana de Saúde

p.p.: Pontos percentuais

RO: Razão de *odds*

UNICEF: *United Nations Children's Fund* ou Fundo das Nações Unidas para a Infância

USAID: *United States Agency for International Development* ou Agência dos Estados Unidos para o Desenvolvimento Internacional

1.1. Introdução

1.1.1. Dose zero de vacinação

Vacinação em massa é uma das intervenções de saúde pública mais custo efetivas existentes (Rogers, 2019), sendo responsável pela prevenção de 2 a 3 milhões de mortes por ano (Mcgavin *et al.*, 2018). Apesar da atuação na promoção da vacinação de instituições públicas e privadas internacionais, tais como a Organização Mundial da Saúde (OMS), o Fundo das Nações Unidas para a Infância (UNICEF) e a Aliança Global para Vacinas e Imunização (GAVI), em 2017, 19.9 milhões de crianças ainda não haviam sido adequadamente vacinadas em todo o mundo. Quatro das seis regiões da OMS apresentaram surtos de sarampo: Américas, Mediterrâneo Oriental, Europa e Sudeste Asiático. Mesmo países com sistemas de saúde consolidados, como a Alemanha, voltaram a ser classificados como regiões endêmicas de sarampo (WHO, 2018).

O Plano de Ação Global de Vacinas 2011–2020, endossado na Assembleia Mundial da Saúde de 2012, define como meta para todos os países uma cobertura mínima de 90% até 2020 para todas as vacinas presentes nos cronogramas nacionais de vacinação. No monitoramento da vacinação, uma das métricas mais utilizadas é a cobertura de DTP3 (percentual de crianças com até um ano de idade que receberam três doses da vacina contra difteria, tétano e coqueluche) (Vanderende *et al.*, 2018). Esta cobertura se encontra estagnada em torno de 85% e, assim como a maioria das metas estabelecidas no plano, muito provavelmente não alcançará 90% até 2020 (WHO, 2018). Em 2017, os países com maior número absoluto de crianças não cobertas pela DTP3 eram a Nigéria (4,0 milhões de crianças), Índia (2,9 milhões), Paquistão (1,3 milhões), Indonésia (1,3 milhões) e Etiópia (0,9 milhões) (Vanderende *et al.*, 2018).

A pesquisa e monitoramento de vacinação são amplas e tendem a focar em dados de cobertura, particularmente de múltiplas doses de vacinas como a DTP3, pois expressam a capacidade dos sistemas de saúde de atingirem as crianças em mais de uma ocasião. Já a pesquisa sobre crianças que nunca receberam uma dose de qualquer vacina, ou crianças dose zero, é marcadamente reduzida (Bosch-Capblanch *et al.*, 2012). O foco em cobertura de vacinas específicas pode representar um panorama incompleto dos programas globais de vacinação, principalmente no que

tange a dose zero, já que o aumento da cobertura específica para uma vacina não implica necessariamente na redução do número de crianças nunca vacinadas.

Uma vez que muitas destas crianças nunca tiveram contato com o sistema de saúde de seus respectivos países, elas só conseguem ser acessadas por meio de pesquisas domiciliares e não por sistemas de monitoramento rotineiros, o que por sua vez limita a disponibilidade dos dados (Bosch-Capblanch *et al.*, 2012). Portanto, estudar estas crianças requer inquéritos domiciliares globais com coleta de dados sobre vacinação, tais como os *Demographic and Health Surveys* (DHS) e os *Multiple Indicator Cluster Surveys* (MICS).

Uma análise realizada em 2012 de mais de 241 destes inquéritos em 96 países de baixa e média renda encontrou uma prevalência de dose zero de vacinação de 9,9% entre crianças de 12 a 59 meses (considerando todos os inquéritos e anos juntos, de 1986 a 2007), com alguns países atingindo uma prevalência nacional superior a 20%. A dose zero de vacinação foi definida como nunca ter recebido nenhuma dose das seguintes vacinas (Bosch-Capblanch *et al.*, 2012):

- BCG (bacilo Calmette-Guérin): contra tuberculose;
- OPV (oral polio vaccine): vacina oral contra a poliomielite;
- DTP: qualquer vacina contra difteria, tétano e coqueluche;
- MCV: qualquer vacina contendo antígeno do sarampo;

A dose zero esteve significativamente associada com o status de vacinação por tétano da mãe da criança (razão de odds ajustada de 2,53 para aquelas com menos de 2 doses da vacina), a renda familiar (2,30 para o grupo mais pobre em relação ao mais rico), autonomia de decisão da mãe em relação ao companheiro em caso de doença da criança (2,19 para as mães que tomam decisões sozinhas), educação da mãe (1,87 para o grupo com menor nível educacional) e educação do parceiro (1,61 para o grupo com menor nível educacional). Apesar disso, nenhum preditor investigado esteve associado com a dose zero em todos os inquéritos na análise não ajustada (Bosch-Capblanch *et al.*, 2012). Tal resultado indica que a associação entre possíveis preditores e dose zero pode variar entre diferentes regiões investigadas. Desse modo, medidas de associação agregadas podem não ser a forma mais adequada de investigar o problema.

Com o prazo do *Plano de Ação Global de Vacinas 2011–2020* chegando ao fim, é importante que as estimativas de prevalência de dose zero de vacinação sejam atualizadas e a evolução individual dos países seja avaliada com base nos inquéritos mais recentes. Além disso, ainda há bastante espaço para investigação dos fatores associados à dose zero, tanto para evidenciar possíveis barreiras a programas de imunização quanto para identificar as crianças mais vulneráveis a não serem atingidas por tais programas, permitindo uma melhor alocação de recursos e direcionamento de intervenções.

1.1.2. Árvores de decisão

Modelos de regressão estão entre as ferramentas mais utilizadas na pesquisa epidemiológica, inclusive na investigação da cobertura de vacinação. Apesar disso, eles tendem a ser pouco flexíveis em termos de interações complexas – tais como modificação de efeito – entre as variáveis de exposição, pois requerem que termos de interação sejam explicitamente especificados. Quando diferentes subgrupos de uma amostra apresentam mecanismos marcadamente distintos para a ocorrência do desfecho, a análise de regressão exige que diferentes modelos sejam especificados para cada subgrupo separadamente, cuja separação fica a critério do investigador. A criação de modelos *ad hoc* pode aumentar a probabilidade de erro tipo I e o risco de sobreajuste dos dados (Wolfson e Venkatasubramaniam, 2018).

Exemplo de múltiplos mecanismos para as mesmas variáveis são os fatores associados à dose zero de vacinação investigados por meio de regressão logística (Bosch-Capblanch *et al.*, 2012): em diferentes inquéritos encontramos relações distintas entre as exposições e os desfechos, muito provavelmente porque os mecanismos para a dose zero não são os mesmos em todos os países.

Se realizada a análise separadamente por inquérito/país, há o risco do não reconhecimento de padrões regionais importantes. Se a análise agregada globalmente é realizada, associações que são válidas apenas regionalmente (ou para determinado subgrupo) podem ser diluídas. Já a criação de subgrupos amostrais pode ser de algum modo arbitrária e pouco prática quando considera-se a natureza multivariável do fenômeno a ser investigado.

Uma possível solução é a utilização de árvores de decisão construídas por meio de algoritmos de aprendizagem de máquina. Estas são uma família de métodos estatísticos particularmente úteis para investigar relações não lineares entre variáveis e identificar subgrupos homogêneos em relação ao desfecho dentro de uma amostra. Tratam-se de métodos não-paramétricos que não requerem a especificação prévia nem de termos de interação entre as exposições, nem do tipo de relação esperada entre exposições e desfecho (Henrard *et al.*, 2015; Wolfson e Venkatasubramaniam, 2018). Isso permite maior flexibilidade na investigação de subgrupos amostrais, uma vez que não é necessário escolher previamente estratificadores ou combinações de estratificadores – como quintis de renda, moradia rural/urbana ou sexo – tipicamente utilizados na pesquisa de saúde global.

As árvores de decisão se baseiam no particionamento recursivo desta amostra em subgrupos binários que são progressivamente criados de modo hierárquico. Tal processo de divisão pode ser representado como uma árvore de decisões – de onde recebe seu nome – com cada decisão representando uma partição binária (Venkatasubramaniam, Ashwini *et al.*, 2017).

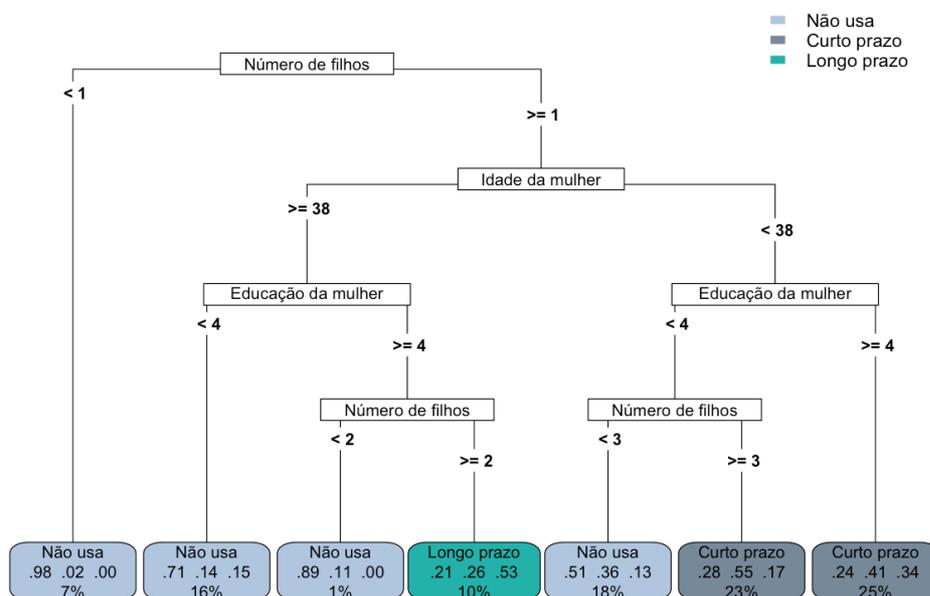


Figura 1 – Exemplo ilustrativo de uma árvore de decisão utilizada para investigar o uso de contraceptivos entre mulheres, baseada em uma amostra da National Indonesia Contraceptive Prevalence Survey de 1987. Dentro de cada nó terminal (caixas na parte inferior da árvore), são apresentados: 1) a previsão do algoritmo da forma de uso de contraceptivo para as mulheres pertencentes àquele nó (no primeiro nó, o algoritmo prevê que as mulheres pertenceriam à categoria “Não usa”; 2) a fração de mulheres do nó que pertencem, na realidade, às três categorias de uso de contraceptivos na ordem da legenda (no primeiro nó, 98% das mulheres não usam contraceptivos, 2,0% usam contraceptivos de curto prazo e 0,0% de longo prazo); 3) o percentual de mulheres da amostra que pertencem àquele nó (7% das mulheres na amostra pertencem ao primeiro nó, ou seja, não tem filhos)

Árvores de decisão são ferramentas bem estabelecidas para modelagem estatística que não pertencem ao grupo de algoritmos “caixa preta” dentro do qual outros algoritmos de aprendizagem de máquina são alocados – como redes neurais ou máquinas de vetores de suporte – pela complexidade com que relacionam as variáveis estudadas. Apesar disso, as árvores de decisão ainda têm aplicação restrita na pesquisa epidemiológica, mas interesse recente tem sido dado a elas, considerando as três maneiras distintas como podem ser aplicadas: modelagem explicativa, predição de desfecho e identificação de subgrupos (Wolfson e Venkatasubramaniam, 2018).

1.1.3. Árvores de decisão e não vacinação

Considerando a possibilidade de relações não lineares entre os fatores de risco a serem investigados e a natureza do método das árvores de decisão, a investigação da dose zero de vacinação em crianças por meio destes algoritmos utilizando os dados dos inquéritos DHS e MICS mais recentes mostra-se uma possibilidade promissora. No contexto da investigação de não imunização, árvores de decisão são úteis para:

- identificar sistematicamente subgrupos de crianças da amostra global que apresentam características similares entre si;
- investigar as variáveis que melhor explicam a dose zero dentro dos subgrupos criados;
- realizar predições sobre o perfil das crianças mais vulneráveis a serem dose zero;

1.2. Justificativa

A estagnação na taxa de cobertura de vacinação em valores abaixo das metas estabelecidas no Plano de Ação Global de Vacinas 2011–2020, somada aos recentes surtos de doenças como sarampo em várias regiões do mundo, são indicativos de que a vacinação, apesar de custo-efetiva, ainda não atinge um nível de cobertura adequado para todas as populações. Um grupo particularmente vulnerável e pouco estudado são as crianças com dose zero de vacinas – e que chegam a representar

28,5% das crianças entre 12 e 59 meses em países como a Etiópia. Considerando que muitas das estimativas envolvem dados das décadas de 1990 e 2000, é necessário que sejam atualizadas as prevalências nacionais de dose zero utilizando inquéritos mais recentes, particularmente aqueles posteriores a 2010.

É vantajoso que os fatores associados à não imunização sejam investigados por metodologias adequadas para lidar com a complexidade dos múltiplos mecanismos regionais que dão origem à dose zero. As árvores de decisão são possíveis candidatas para tal avaliação, uma vez que são desenhadas para identificar subgrupos vulneráveis e, por meio disso, investigar diferentes fatores associados em distintas populações. Enquanto métodos como regressão logística exigem uma especificação prévia de estratificadores e suas combinações, essas especificações são criadas pelas próprias árvores como consequência da busca por grupos de risco.

As estimativas atualizadas e os grupos de risco para dose zero são informações importantes na tomada de decisão de instituições internacionais como a GAVI, que organizam, fomentam e financiam não só a imunização de crianças, mas também a pesquisa técnico-científica na área. Este projeto é inspirado em um conjunto de análises sobre vacinação nos países de baixa e média renda – incluindo a investigação de desigualdades na distribuição de dose zero – em desenvolvimento pelo Centro Internacional de Equidade em Saúde | Pelotas, em parceria com a GAVI.

1.3. Revisão da literatura

Uma vez que este projeto trata da aplicação de modelos de árvore de decisão para identificação de grupos de risco para dose zero de vacinação, a revisão da literatura foi realizada separadamente para os dois tópicos e encontra-se apresentada nas seguintes seções:

- **Seção 1.3.1 – Dose zero:** trata da prevalência de dose zero e fatores associados em países de baixa e média renda;
- **Seção 1.3.2 – Árvores de decisão:** trata do método estatístico a ser utilizado, de suas implementações e de como ele tem sido aplicado à pesquisa epidemiológica;

1.3.1. Dose zero

1.3.1.1. Estratégia de busca

A revisão da literatura foi realizada por meio da busca indexada nas bases PubMed e Web of Science. Para o PubMed, foram selecionados todos os artigos com título ou resumo que contenham pelo menos um dos seguintes termos: *unvaccinated*, *non vaccination*, *non-vaccination*, *zero dose*, *non immunized*, *non-immunized*, *non immunization* ou *non-immunization*, aplicando o filtro da própria plataforma para artigos que tratem de crianças (do nascimento até os 18 anos). Para o Web of Science, os mesmos termos foram utilizados, mas além de conter pelo menos um destes, o artigo também deveria conter no título ou resumo os termos *child* ou *children*, de modo a simular o filtro para crianças disponível no PubMed. A estratégia encontra-se apresentada na Figura 2.

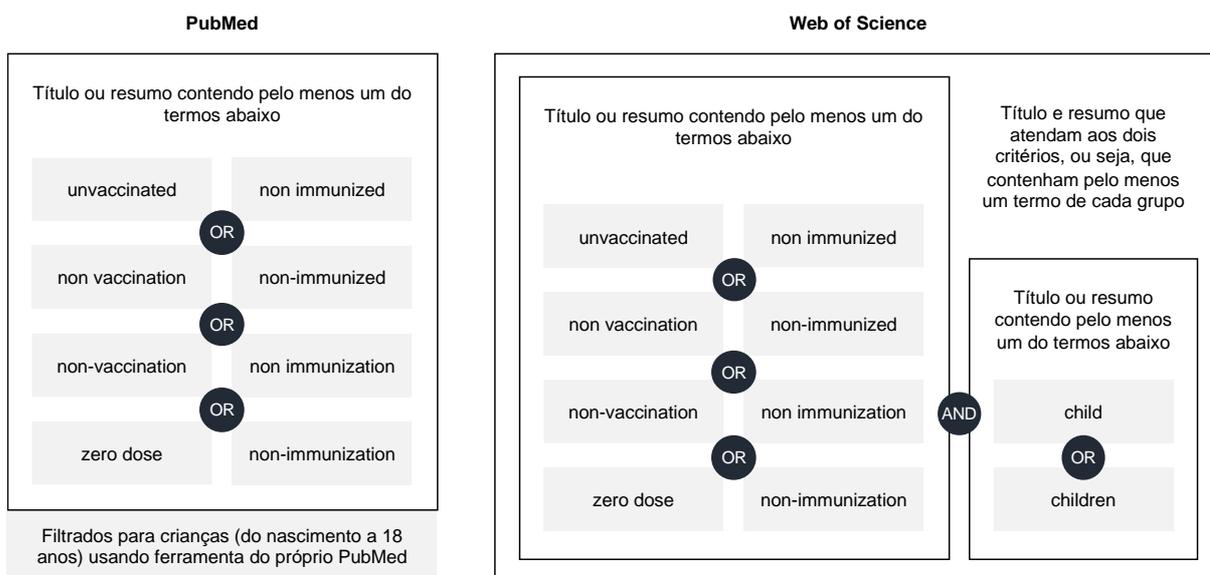


Figura 2 – Estratégia de busca utilizada para as bases do PubMed e Web of Science.

Foram selecionados os artigos que apresentassem medidas de prevalência de dose zero de vacinação e/ou fatores associados em países de baixa ou média renda para as vacinas BCG, OPV, DTP e MCV individuais ou em conjunto. A Figura 3 apresenta o fluxograma da busca.

De modo geral, o foco da pesquisa é maior na cobertura de vacinação (principalmente de doses completas). Poucos artigos foram encontrados que tratassem da dose zero de vacinas em países de baixa e média renda.

Particularmente para esta revisão, os únicos países encontrados com pesquisas individuais sobre a não vacinação por várias vacinas foram o Afeganistão, El Salvador, Índia, Nigéria e Quênia. De modo conjunto entre países, apenas uma pesquisa foi encontrada que observasse a dose zero e ela tratava de 96 países (Bosch-Capblanch et al., 2012). O resultado da busca encontra-se sumarizado no Quadro 1.

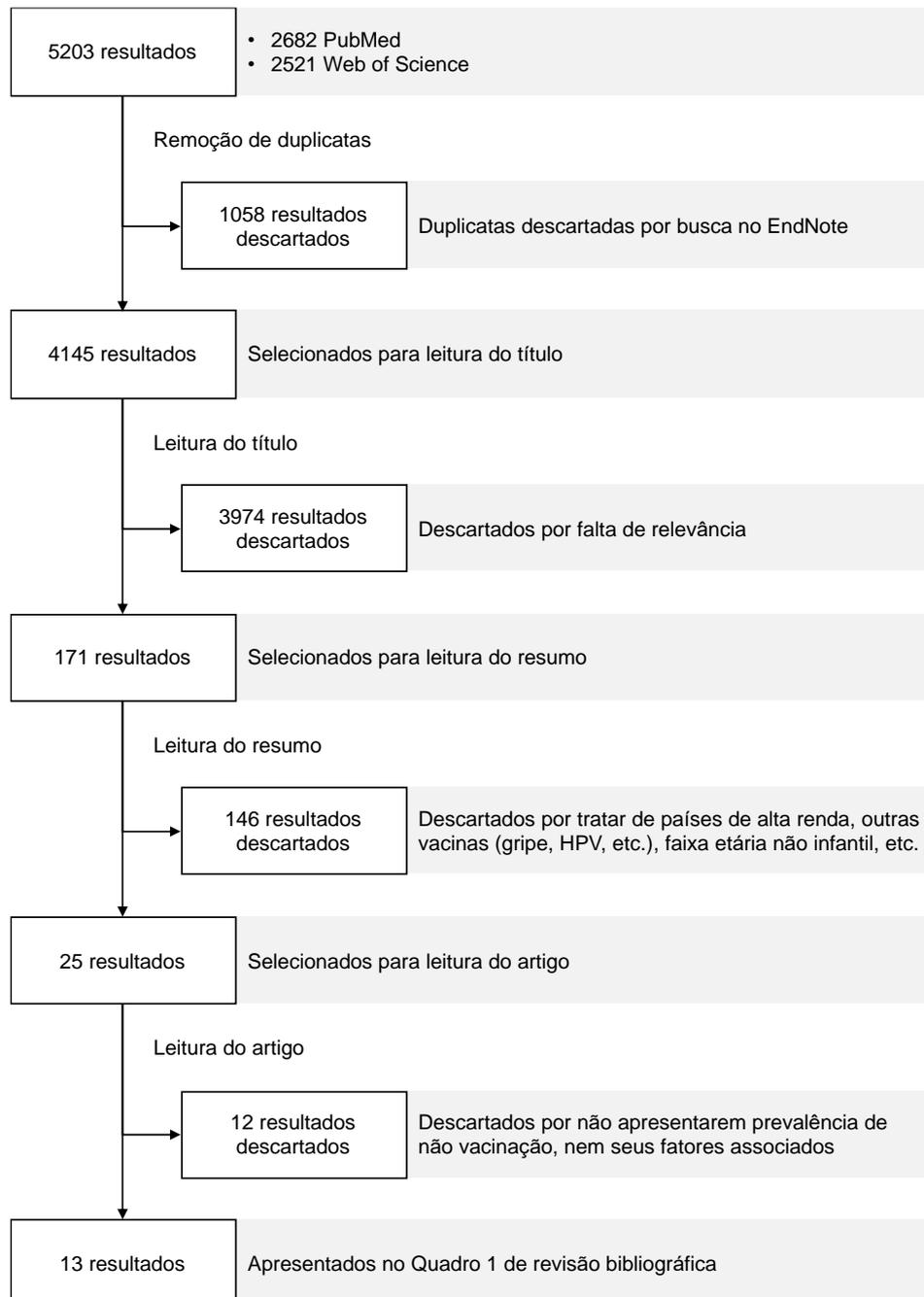


Figura 3 – Fluxograma do processo de busca na literatura sobre dose zero de vacinação. Busca realizada em 22 de junho de 2019.

Quadro 1 – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Masters <i>et al.</i> , 2019)	Quênia, 2014	2014 Kenyan Demographic and Health Survey	BCG, MCV, OPV e pentavalente (DTP, Haemophilus influenzae tipo B (Hib) e hepatite B)	Para crianças de 12 a 23 meses: Dose zero: 1,6% (1%; 2,3%) Pentavalente: 2,5% (1,8%; 3,2%) OPV: 1,8% (1,2%; 2,5%) BCG: 3,3% (2,6%; 4,1%) MCV: 12,9% (11,6%; 14,2%)	Após modelo de regressão logística multinomial: As crianças cujos respondentes eram do grupo étnico Somali tiveram uma razão de odds de dose zero de 59,4 (3,2; > 999) em relação ao grupo de referência (Kikuyu) e os muçulmanos de 0,05 (0,00; 0,62) em relação aos protestantes e outros cristãos.
(Subaiya <i>et al.</i> , 2018)	Quênia, 2016	Pesquisa nacional de cobertura de vacinação por MR	MR (sarampo-rubéola)	Para crianças de 9 a 59 meses: Dose zero: 0,8% (0,6%; 1,2%)	Foram apresentados fatores associados à cobertura de duas doses da vacina e não à dose zero.
(Shenton, Wagner, Carlson, <i>et al.</i> , 2018)	Afeganistão, 2015	2015 Afghanistan Demographic and Health Survey	BCG, MCV, OPV e pentavalente (DTP, Haemophilus influenzae tipo B (Hib) e hepatite B)	Para crianças de 1 a 4 anos: Dose zero: 17,0% (15,0%; 19,0%)	Razão de odds de dose zero versus vacinação completa ajustada para covariáveis sociodemográficas: 0,66 (0,47; 0,92) para moradia urbana versus rural; 0,58 (0,47; 0,71) para visita a uma instituição de saúde nos últimos 12 meses; 1,5 (1,1; 2,2) para parto instituição privada e 2,80 (2,3; 3,5) para parto não institucional desacompanhado (comparado a parto em instituição governamental); 0,55 (0,33; 0,91) para primário incompleto, 0,38 (0,26; 0,59) para secundário completo e 0,24 (0,06; 0,98) para superior comparados a nenhuma educação; 1,5 (1,2; 1,8) para 2 crianças abaixo de 5 anos na casa, 1,6 (1,2; 2,0) para 3 e 1,5 (1,2; 2,0) para 4 ou mais, comparados a zero ou 1 criança; 0,72 (0,51; 0,99) para os ricos e 0,57 (0,37; 0,87) para os mais ricos no índice de riqueza comparados aos mais pobres;

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Shenton, Wagner, Bettampadi, <i>et al.</i> , 2018)	Índia, 2012-2013	District-Level Household and Facility Survey and Annual Health Survey	BCG, MCV, OPV e DTP	Para crianças de 12 a 48 meses: Dose zero: 7,5% (7,4%; 7,7%)	Das crianças com apenas uma dose de vacinação, 90% haviam recebido BCG; Razão de odds de dose zero versus vacinação completa ajustada para covariáveis sociodemográficas: 1,2 (1,1; 1,3) para moradia urbana versus rural; 1,5 (1,4; 1,6) para muçulmanos e 1,8 (1,6; 2,1) para cristãos comparados aos hindus. 1,3 (1,3; 1,4) para filhos de mães sem escolaridade comparados aos filhos de mães com 1 a 5 anos de educação; 1,6 (1,5; 1,7) para parto não institucional acompanhado e 2,4 (2,2; 2,5) para parto não institucional desacompanhado (comparado a parto em instituição governamental); 0,14 (0,13; 0,14) para 1 a 2 visitas pré-natal, 0,07 (0,07; 0,07) para 3 a 6 e 0,04 (0,04; 0,05) para 7 ou mais, comparado a nenhuma visita;
(Mcgavin <i>et al.</i> , 2018)	Nigéria, 2013	2013 Nigerian Demographic and Health Survey	BCG, MCV, OPV e DTP	Para crianças de 12 a 24 meses: Dose zero: 26,6% (intervalo de confiança não apresentada)	Razão de odds de vacinação completa versus dose zero ajustada por modelo de regressão logística multinomial: 0,34 (0,22; 0,53) para grupo mais baixo de índice de riqueza e 3,7 (2,3; 5,9) para o grupo mais alto comparados ao grupo intermediário; 0,42 (0,28; 0,63) para religião islâmica comparada ao grupo de outros cristãos; 0,33 (0,14; 0,78) para o grupo étnico Fulani e 0,25 (0,11; 0,56) para o grupo Hausa comparados ao grupo de referência Yoruba; 0,44 (0,32; 0,62) para parto em casa comparado ao parto em hospital público;

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Mcgavin <i>et al.</i> , 2018)	Nigéria, 2013	2013 Nigerian Demographic and Health Survey	BCG, MCV, OPV e DTP	Para crianças de 12 a 24 meses: Dose zero: 26,6% (intervalo de confiança não apresentado)	(continuação) 0,13 (0,06; 0,29) para crianças de mães sem escolaridade, 0,22 (0,11; 0,45) para educação primária e 0,33 (0,17; 0,66) para educação secundária comparadas às crianças de mães com educação superior.
(Francis <i>et al.</i> , 2018)	Índia, 1998-2008	India's District Level Household and facility Survey's	BCG, MCV, OPV e DTP	Para crianças de 12 a 23 meses: Dose zero: 1998-99: 18,3% (17,9%; 18,8%) 2002-04: 20,0% (19,4%; 20,6%) 2007-08: 9,4% (9,0%; 9,7%) Combinado: 14,5% (14,3%; 14,9%)	Razão de odds de dose zero versus vacinação completa ajustada em regressão multivariável: 1,4 (1,3; 1,5) para residência urbana comparada a residência rural; 2,0 (1,9; 2,2) para muçulmanos comparados aos hindus; 2,6 (2,3; 2,9) para filhos de mães sem escolaridade, 1,5 (1,3; 1,7) para 1 a 5 anos e 1,2 (1,0; 1,3) para 6 a 8 anos, comparados aos filhos de mães com 9 ou mais anos de educação; 1,7 (1,5; 2,1) para zero visitas pré-natal, comparadas a 7 ou mais; 2,8 (2,6; 3,0) para crianças de mães com vacina contra tétano; 1,5 (1,4; 1,7) para parto não institucional comparado ao parto institucional governamental;
(Chido-Amajuoyi <i>et al.</i> , 2018)	Nigéria, 2013	2013 Nigeria Demographic Health Survey	Qualquer vacina	Para crianças de 12 a 59 meses: Dose zero: 20,5% (intervalo de confiança não apresentado)	Razão de odds de dose zero versus pelo menos uma dose ajustada por fatores sociodemográficos: 0,79 (0,67; 0,94) para crianças de mães acima de 35 anos e 0,85 (0,74; 0,98) para 26 a 25 anos, comparadas às de 16 a 25 anos; 1,5 (1,1; 2,1) para muçulmanos comparados aos cristãos;

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Chido-Amajuoyi <i>et al.</i> , 2018)	Nigéria, 2013	2013 Nigeria Demographic Health Survey	Qualquer vacina	Para crianças de 12 a 59 meses: Dose zero: 20,5% (intervalo de confiança não apresentado)	(continuação) 1,7 (1,4; 1,9) para crianças de mães desempregadas comparadas às empregadas; 2,2 (1,6; 3,0) para crianças de mães com educação secundária incompleta, 2,4 (1,8; 3,3) para alguma ou completa educação primária e 4,3 (3,1; 6,0) para nenhuma educação comparadas às mães com educação secundária completa ou superior; 1,8 (1,2; 2,6) para os ricos na escala de índice de riqueza, 2,41 (1,7; 3,5) para os intermediários, 2,7 (1,7; 4,1) para os pobres e 3,2 (2,0; 5,1) para os mais pobres quando comparados aos mais ricos; 1,3 (1,1; 1,6) para as famílias que não possuem televisão ou rádio;
(Uthman <i>et al.</i> , 2017)	Nigéria, 2013	2013 Nigeria Demographic Health Survey	Poliomielite	Para crianças de 12 a 23 meses: Poliomielite: 21,3% (19,5%; 23,2%)	Razão de odds de dose zero para poliomielite ajustada por fatores individuais e da vizinhança: 1,9 (1,3; 2,7) para os mais pobres no índice de riqueza e 1,7 (1,3; 2,2) para os intermediários em relação aos mais ricos; 2,8 (2,1; 3,7) para crianças de mães sem escolaridade e 1,9 (1,5; 2,5) para educação primária comparadas àquelas com educação secundária ou superior; 1,5 (1,2; 1,8) para crianças de mães que não estavam trabalhando;
(Shrivastwa <i>et al.</i> , 2015)	Índia, 2008	District Level Household and Facility Survey 2008	BCG, DTP e MCV	Para crianças de 12 a 36 meses: Dose zero: 12,4% (12,0%; 12,8%)	Razão de odds de dose zero versus vacinação completa ajustada em regressão logística multinomial: 1,8 (1,7; 1,9) para residência urbana comparada a residência rural; 2,2 (2,0; 2,4) para muçulmanos

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Shrivastwa <i>et al.</i> , 2015)	Índia, 2008	District Level Household and Facility Survey 2008	BCG, DTP e MCV	Para crianças de 12 a 36 meses: Dose zero: 12,4% (12,0%; 12,8%)	(continuação) comparados aos hindus; 0,76 (0,70; 0,82) para os classificados como pobres no índice de riqueza, 0,60 (0,55; 0,66) para os intermediários, 0,46 (0,40; 0,52) para os ricos e 0,32 (0,27; 0,37) para os mais ricos, comparados aos mais pobres; 1,2 (1,2; 1,3) para meninas em relação aos meninos; 0,64 (0,60; 0,69) para filhos de mães com 1 a 5 anos de educação, 0,41 (0,38; 0,45) para 6 a 12 anos, 0,27 (0,19; 0,37) para 13 ou mais anos em relação às mães sem escolaridade; 1,4 (1,2; 1,7) para parto em instituição privada e 1,8 (1,6; 2,0) para parto não institucional comparados ao parto em instituição governamental; 0,25 (0,21; 0,29) para filhos de mães vacinadas por tétano;
(Suarez-Castaneda <i>et al.</i> , 2014)	El Salvador, 2011	Inquérito domiciliar realizado pela própria pesquisa	BCG, rotavírus, pentavalente, MMR (sarampo, caxumba e rubéola) e OPV	Para crianças de 23 a 59 meses: Dose zero: nenhuma criança foi encontrada (de uma amostra de 2550 crianças)	Não foram apresentados fatores associados à dose zero;
(Bosch-Capblanch <i>et al.</i> , 2012)	96 países de baixa ou média renda, 1986-2007	Demographic and Health Surveys (DHS) e Multiple Indicator Cluster Survey (MICS)	BCG, MCV, OPV e DTP	Para crianças de 12 a 59 meses: Dose zero: 9,9% (intervalo de confiança não apresentado) para todas as pesquisas e todos os países, com uma mediana de 5,3%, atingindo um máximo de 28,5%. Os países com maior prevalência de dose zero foram:	Razão de odds de dose zero mediana (e intervalo interquartil) por regressão logística multivariável: 1,9 (1,3; 2,9) para responsáveis com menor nível de educação; 1,6 (1,2; 2,5) para parceiros(as) com menor nível de educação; 2,5 (1,6; 3,9) para não vacinação do responsável por tétano antes do parto;

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Bosch-Capblanch <i>et al.</i> , 2012)	96 países de baixa ou média renda, 1986-2007	Demographic and Health Surveys (DHS) e Multiple Indicator Cluster Survey (MICS)	BCG, MCV, OPV e DTP	(continuação) 28,5% na Etiópia, 2005; 28,2% em Comores, 2000; 27,2% no Zimbábue, 2005; 26,6% na República Democrática do Laos, 2000; 26,3% no Sudão do Sul, 2000; 22,6% na Nigéria, 2003; 19,9% no Níger, 2006; 19,9% em Madagáscar, 2004; 17,9% na República Centro-Africana, 2000; 16,7%, no Chade, 2004;	(continuação) 2,2 (1,5; 3,1) para membro da família tomando decisão sozinho em caso de doença da criança; 1,3 (0,98; 1,8) para os mais pobres comparados aos menos pobres no índice de riqueza; 1,8 (1,0; 2,7) para os mais pobres comparados aos moderadamente pobres; 1,8 (1,0; 3,1) para os mais pobres comparados aos ricos; 2,3 (1,0; 5,3) para os mais pobres comparados aos mais ricos;
(Antai, D., 2009)	Nigéria, 2003	2003 Nigeria Demographic and Health Survey	BCG, MCV, OPV e DTP	Estudo realizado com crianças de 12 meses a 5 anos. Prevalência não apresentada.	Razão de odds de dose zero ajustada para variáveis socioeconômicas e indicadores de empoderamento por meio de regressão logística: 1,6 (1,2; 2,2) para islâmicos em relação aos cristãos; 2,2 (1,4; 3,6) para filhos de mães sem escolaridade versus filhos de mães com ensino secundário ou superior; 0,19 (0,06; 0,66) para filhos de pais que não trabalham em relação àqueles com ocupação classifica como: profissional/técnico/gerencial; 1,9 (1,2; 2,9) para os classificados como pobres no índice de riqueza e 1,7 (1,2; 2,4) para os intermediários em relação aos ricos; 0,60 (0,40; 0,91) para os que apresentam frequência de leitura de jornais; 0,74 (0,61; 0,89) para os que apresentam frequência de ouvir ao rádio; 0,78 (0,60; 1,00) para os que apresentam frequência de assistir TV;

Quadro 1 (continuação) – Revisão: estudos sobre dose zero por BCG, MCV, OPV ou DTP em países de baixa ou média renda

Autores e ano de publicação	País e ano investigado	Fonte dos dados	Vacinas investigadas	Prevalência de dose zero e intervalo de confiança (95%)	Fatores associados selecionados significativos ($\alpha = 5\%$) e outros pontos
(Tshikuka <i>et al.</i> , 2007)	Chade, 2001-2002	National Immunization Days Rounds	OPV	Para crianças de 0 a 59 meses: 5,7% em 2001 e 5,4% em 2002 (intervalos de confiança não apresentado)	Medidas de associação não apresentadas

1.3.1.2. Prevalência de dose zero

As prevalências de dose zero de vacinação para os países de baixa e média renda investigados por Bosch-Capblanch *et al.* estão apresentados na Figura 4. A prevalência mediana de dose zero entre todos os países e anos investigados foi de 5,3% (intervalo interquartil: 1,9% a 12,4%) e a maioria dos países (56) teve uma prevalência de dose zero abaixo de 5%. Aqueles com a maior fração de crianças nunca vacinadas foram a Etiópia (1997, 28,5%), Comores (2000, 28,2%) e Zimbábue (2005, 27,2%) (Bosch-Capblanch *et al.*, 2012).

A variação anual foi estimada por Bosch-Capblanch *et al.* para os países com mais de um inquérito disponível, comparando os relatórios mais recentes com os mais antigos, por país. 48 países apresentaram mudanças estatisticamente significativas, dos quais 38 reduziram a proporção de dose zero com redução anual mediana de 0,9 pontos percentuais (intervalo interquartil: 0,4 p.p. a 1,4 p.p.) e 10 observaram aumento na proporção de dose zero, com crescimento anual mediano de 0,4 p.p. (IIQ: 0,2 p.p. a 1,4 p.p.). Dentre os 10, 4 pioraram não só em termos de dose zero, mas também de vacinação completa: Colômbia de 1986 a 2005, Quênia de 1989 a 2003, Uganda de 1988 a 2006 e Zimbábue de 1988 a 2005 (Bosch-Capblanch *et al.*, 2012).

Utilizando os dados dos outros artigos identificados na revisão, a variação temporal das prevalências da Nigéria, Índia e Quênia pôde ser avaliada. Os valores atualizados estão apresentados na Tabela 1. Enquanto o Quênia apresentou marcada redução na prevalência de dose zero, tanto a Nigéria em 2013 quanto a Índia em 2012/2013 apresentaram valores um pouco superiores aos encontrados originalmente em 2003 e 2006, respectivamente.

Tabela 1 – Comparação das prevalências de dose zero de vacinação encontradas por Bosch-Capblanch *et al.* em 2012 com valores mais recentes de outras pesquisas para países disponíveis.

País	De acordo com Bosch-Capblanch <i>et al.</i> , 2012		De acordo com artigo mais recente		
	Ano	Prevalência	Ano	Prevalência	Fonte
Nigéria	2003	22,6%	2013	26,6%	(McGavin <i>et al.</i> , 2018)
Índia	2006	6,7%	2012 / 2013	7,53%	(Shenton, Wagner, Bettampadi, <i>et al.</i> , 2018)
Quênia	2003	6,1%	2014	1,6%	(Masters <i>et al.</i> , 2019)

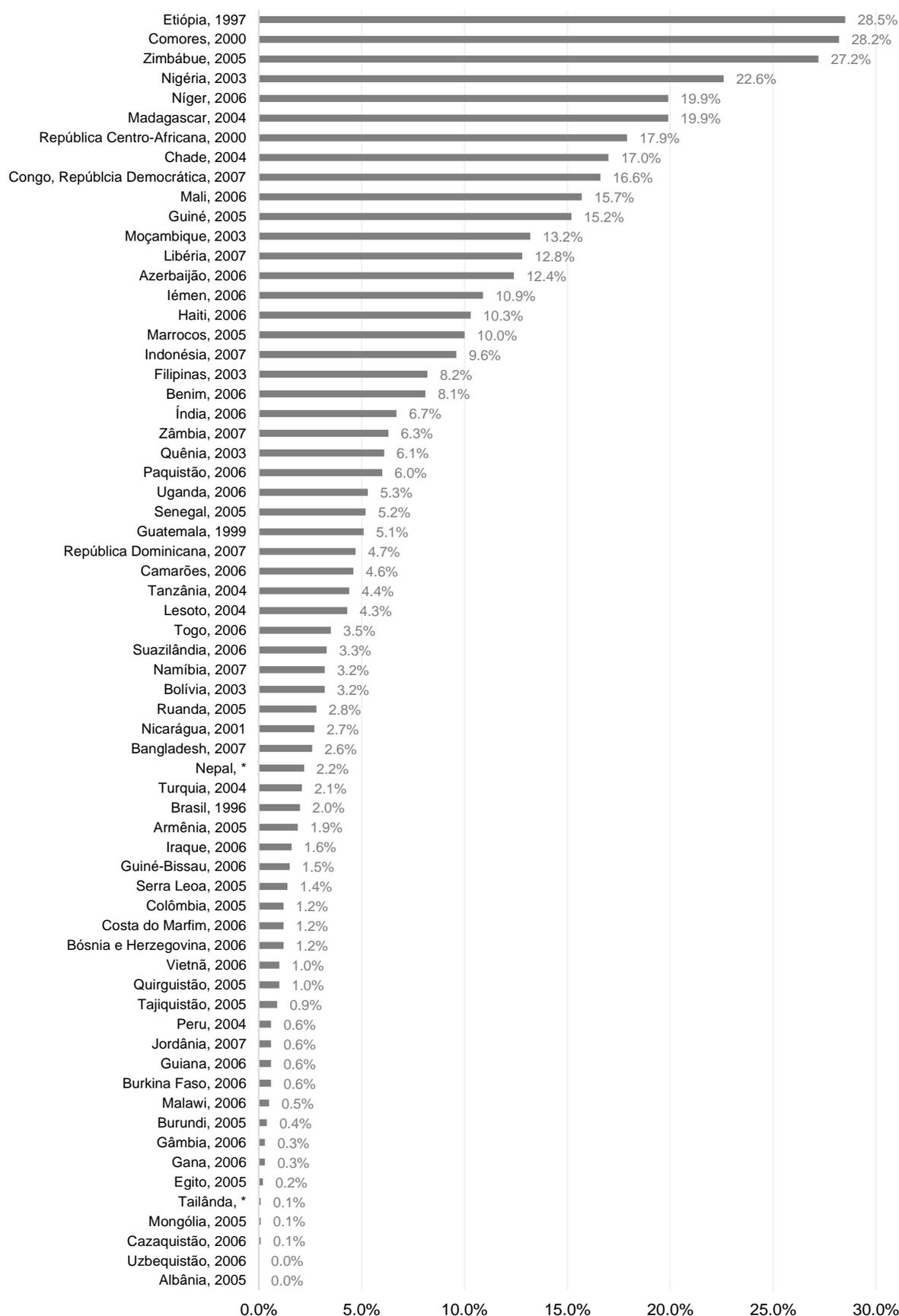


Figura 4 – Prevalência de dose zero de vacinação por BCG, OPV, DTP e MCV para países de baixa e média renda de acordo com dados do último relatório DHS ou MICS disponível (Bosch-Capblanch *et al.*, 2012). *anos inconsistentes no artigo original.

1.3.1.3. Fatores associados à dose zero

Com o foco das pesquisas em cobertura vacinal, boa parte da investigação sobre fatores associados a falhas na vacinação é direcionada para doses incompletas. Portanto, são limitados os artigos que fazem a distinção entre os fatores associados à dose zero versus a vacinação incompleta (Favin *et al.*, 2012).

Em uma revisão da literatura de 1999 a 2009 sobre fatores associados tanto para a dose zero quanto para a subvacinação, 19 fatores foram encontrados em 11 artigos envolvendo a dose zero. Destes, 8 estavam relacionados à atitude e ao conhecimento parentais, 6 aos sistemas de imunização, 4 às características familiares e 1 às limitações ou deficiências na comunicação sobre a vacinação. Dos 8 fatores relacionadas à atitude dos pais, 3 estavam diretamente relacionados às crenças culturais e religiosas contra a vacinação. Dos 6 relacionados ao sistema de imunização, 4 envolviam dificuldade ao acesso a serviços de imunização por distância, moradia em área remota ou conflito civil (Rainey *et al.*, 2011).

Em termos de fatores associados, alguns apresentam maior consistência na direção da associação e significância estatística entre diferentes pesquisas. Dos inquéritos de DHS e MICS investigados por Bosch-Capblanch *et al.*, os fatores com maior frequência de razão de odds de dose zero significativa ($RO \neq 1$, nível de significância de 5%) foram:

- se a mãe decide sozinha em caso de doença da criança (87% dos inquéritos com dados disponíveis sobre esse quesito apresentaram razão de odds positiva);
- menor nível de educação do parceiro (84%);
- vacinação da mãe por tétano antes do nascimento (77%);
- menor nível de educação da mãe (77%);
- não possuir rádio (69%);

Já moradia urbana ou rural apresenta resultados variáveis entre os estudos. 50% dos inquéritos apresentaram razão de odds menor do que 1 (ou seja, o limite superior do intervalo de confiança de 95% não inclui 1) para residência urbana em relação a rural, 43% igual a 1 (intervalo de confiança continha o 1) e 7% maior do que 1. Tal resultado é corroborado por outros estudos: a Índia em 2012/13 apresentou

uma razão de odds de 1,23 (IC95%: 1,14; 1,33) para moradia urbana versus rural (Shenton, Wagner, Bettampadi, *et al.*, 2018), enquanto o Afeganistão em 2015 apresentou uma razão de odds de 0,66 (0,47; 0,92) (Shenton, Wagner, Carlson, *et al.*, 2018).

As minorias étnicas também apresentam associações marcadamente variadas entre os diferentes inquéritos: em 21% destes as minorias apresentaram razão de odds de dose zero menor do que 1, em 42% igual a 1 e em 38% maior do que 1 (Bosch-Capblanch *et al.*, 2012).

Esses resultados não indicam necessariamente que a associação entre desfecho (dose zero) e exposição (residência ou pertencer a uma minoria étnica) é espúria ou não causal. Os diferentes resultados observados podem ser indicativos de que diferentes mecanismos são responsáveis pela conexão entre a exposição e a imunização das crianças em diferentes localidades/países/contextos.

Tomemos os casos dos grupos étnicos Somali no Quênia e Yoruba na Nigéria como exemplo.

Com base no inquérito DHS do Quênia de 2014, foi encontrada uma razão de odds de dose zero para os Somalis de 59,4 (IC95%: 3,2; >999) quando comparados aos Kikuyu (o maior grupo étnico do Quênia). Os Somalis representavam 3,8% (316 indivíduos de 4051) da amostra, enquanto os Kikuyu 17,5%. Das crianças dose zero investigadas no inquérito, 47,6% eram Somalis (Masters *et al.*, 2019). Os Somalis quenianos moram primordialmente na região nordeste do País, fronteira com a Somália e onde fica localizado o campo de refugiados Dadaab, o maior do mundo. O campo recebe imigrantes resultantes de guerras civis, fome e secas que atingiram a Somália durante décadas, levando ao deslocamento de uma população com baixa cobertura de vacinação prévia e que representa um maior desafio no melhoramento da cobertura (pelo difícil acesso, falta de familiaridade com os costumes e serviços governamentais, discriminação social, etc.) (Masters *et al.*, 2019).

Com base no inquérito DHS da Nigéria de 2013, foi encontrada uma razão de odds de vacinação completa (versus dose zero) para os membros do grupo étnico Hausa de 0,25 (0,11; 0,56) em relação ao grupo Yoruba. Enquanto o grupo Hausa, o mais prevalente, representava 34,0% da amostra e possuía uma cobertura completa

de vacinação de apenas 7,9%, o grupo Yoruba representava 11,1% da amostra e possuía uma cobertura completa de vacinação de 43,1% (McGavin *et al.*, 2018). Tal disparidade está associada à situação socioeconômica drasticamente diferente entre os grupos étnicos da Nigéria, e que reflete a capacidade e propensão deste grupo à migração para regiões com maiores oportunidades econômicas e mais fácil acesso à vacinação (Antai, Diddy, 2009). Enquanto os grupos Hausa/Fulani/Kanuri apresentam marcada concentração nos estratos mais pobres do índice de riqueza, os grupos Igbo e Yoruba, por sua vez, apresentam maior concentração nos mais ricos, apesar de serem minoritários em termos absolutos (Adedini *et al.*, 2015).

Os Somalis no Quênia e os Yoruba na Nigéria são exemplos de grupos minoritários que apresentam medidas de associação à dose zero em direções opostas, justamente por apresentarem contextos e mecanismos na sua relação com a vacinação excepcionalmente diferentes. O mesmo pode acontecer com residência urbana/rural (dependendo de onde a maioria da população reside) e com religião (que varia em termos de crença de um país para outro, e mesmo de entendimento sobre o que é aceitável ou não, dentro de uma mesma denominação religiosa).

1.3.2. Árvores de decisão

1.3.2.1. Conceitos básicos

As árvores de decisão pertencem a uma categoria mais ampla de métodos estatísticos denominados **métodos de aprendizagem supervisionada**. O termo supervisionado refere-se ao fato de lidarem com a construção de um modelo de previsão de um desfecho **conhecido**, baseando-se em uma série de variáveis independentes. Em contraste, há problemas de aprendizagem não supervisionada, nos quais não existe um desfecho conhecido e busca-se apenas descrever os padrões e associações entre um conjunto de variáveis independentes (Hastie *et al.*, 2009).

Quando o desfecho a ser predito é uma variável quantitativa, trata-se de uma tarefa de **regressão**. Quando a variável é qualitativa, a tarefa é dita de **classificação**. As árvores de decisão podem ser usadas tanto para tarefas de regressão quanto de classificação, mas – uma vez que o foco da pesquisa é o status de dose zero, uma variável dicotômica – enfoque será dado às árvores usadas para classificação (Hastie *et al.*, 2009).

Por fundamento, métodos baseados em árvores dividem geometricamente em regiões distintas o espaço formado pelas variáveis independentes e aplicam sobre elas um modelo simples – como cálculo de uma constante – para previsão do desfecho. Se existe apenas uma variável independente, a reta que contém os valores desta variável é dividida formando regiões representadas por segmentos de reta; se existem duas variáveis, o plano que contém seus valores é dividido formando regiões representadas por retângulos e assim por diante (Hastie *et al.*, 2009). Matematicamente, a relação entre a estimativa do desfecho \hat{y} e as variáveis independentes x estabelecida pela árvore de decisão pode ser expressa da forma:

$$\hat{y} = \hat{f}(x) = \sum_{m=1}^M c_m I\{x \in R_m\}$$

em que $I\{x \in R_m\}$ representa a função identidade que retorna 1 se x pertencer a uma região R_m e 0 se não pertencer. Como as regiões são mutualmente exclusivas, teremos que se $x \in R_m$, então $\hat{y} = c_m$, uma constante que tipicamente é a média do desfecho na região R_m (Hastie *et al.*, 2009). Se o desfecho é uma variável dicotômica, como ser ou não dose zero, c_m equivale à proporção de casos positivos na região R_m .

Uma representação visual desse processo para duas variáveis independentes encontra-se apresentada na Figura 5.

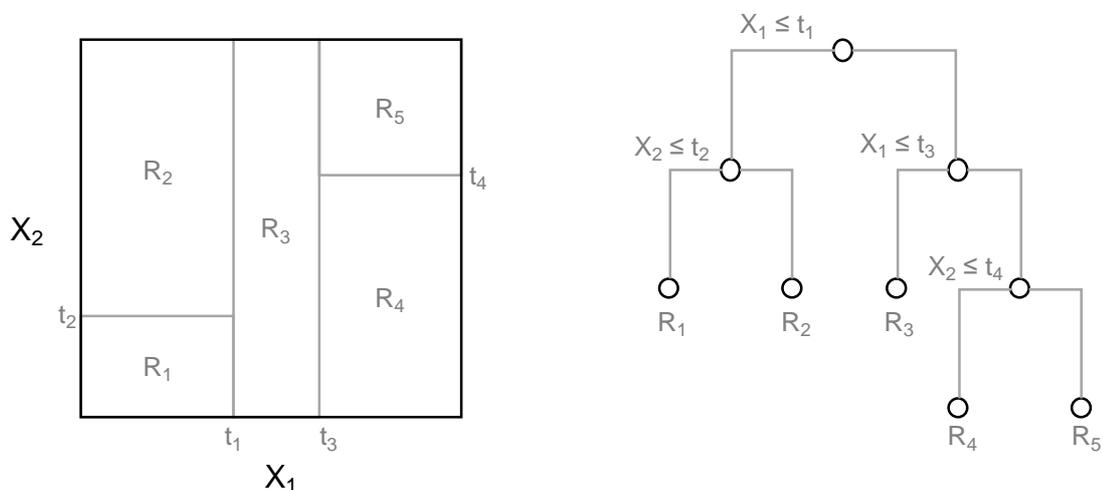


Figura 5 – Exemplo de partições realizadas no espaço formado por duas variáveis independentes X_1 e X_2 e a árvore de decisão correspondente. t_m representa os fictícios pontos de corte e R_m os subgrupos da amostra criados a partir da partição. Para cada subgrupo um modelo seria aplicado individualmente, de modo a estabelecer uma previsão do desfecho de interesse. Adaptado (Hastie *et al.*, 2009).

Para a situação específica da dose zero de vacinação, X_1 poderia ser, por exemplo, a idade da criança e X_2 a idade da mãe no parto. Para cada retângulo formado, seria calculada a proporção de crianças dose zero. Por exemplo, para as crianças tais que $X_1 < t_1$ e $X_2 > t_2$, a estimativa do desfecho gerada pela árvore de decisão seria simplesmente a proporção de crianças dose zero na região R_2 . A partição seria realizada de modo a encontrar subgrupos homogêneos em relação ao desfecho, ou seja, com a proporção de crianças dose zero o mais próximo possível ou de zero ou de um. Desse modo, as crianças tenderiam a ser classificadas em grupos de alto ou baixo risco para dose zero. A funcionalidade principal da árvore de decisão reside então em encontrar os pontos de corte que permitam a criação de subgrupos internamente homogêneos, mas heterogêneos entre si (Henrard *et al.*, 2015).

A forma de divisão apresentada na Figura 5 é denominada **partição binária recursiva**, em que cada separação é realizada de modo a formar exatamente dois subgrupos, também chamados de **nós**, e que podem ou não ser novamente divididos até que uma **regra de parada** seja atingida e um nó terminal seja criado. O objetivo de cada partição é minimizar uma medida de **impureza dos nós** previamente estabelecida (Flach, 2012; Venkatasubramaniam, Ashwini *et al.*, 2017).

As medidas de impureza mais comumente empregadas em árvores de decisão incluem taxas de erros de classificação, índice de Gini e medidas de entropia. Especificamente para a classificação de um desfecho dicotômico, a proporção de uma das categorias – como a de crianças dose zero em um nó – pode ser representada por p e da categoria complementar por $1 - p$. Para cada nó criado na investigação de uma possível partição j , as respectivas medidas de impureza são calculadas a partir de p , da forma:

- Erro de classificação: $Imp(D_{j,k}) = 1 - \max(p_{j,k}, 1 - p_{j,k})$
- Índice Gini: $Imp(D_{j,k}) = 2p_{j,k}(1 - p_{j,k})$
- Entropia¹: $Imp(D_{j,k}) = -p_{j,k} \log_2(p_{j,k}) - (1 - p_{j,k}) \log_2(1 - p_{j,k})$

sendo $Imp(D_{j,k})$ a impureza do nó $D_{j,k}$, com $k \in \{1,2\}$ para uma partição binária (Hastie *et al.*, 2009; Flach, 2012).

Como cada partição j dá origem a dois nós, uma média ponderada de impureza M_j^{imp} destes nós é calculada de modo a mensurar a qualidade da possível partição j :

$$M_j^{imp} (\{D_{j,1}, D_{j,2}\}) = \frac{n_{D_{j,1}}}{n_{D_j}} Imp(D_{j,1}) + \frac{n_{D_{j,2}}}{n_{D_j}} Imp(D_{j,2})$$

com n representando o número de observações em cada nó e $D_j = D_{j,1} \cup D_{j,2}$ representando o nó que deu origem a $D_{j,1}$ e $D_{j,2}$ quando particionado (Flach, 2012).

O índice de Gini e medidas de entropia são preferenciais em relação a medidas de erros de classificação, uma vez que são mais sensíveis a mudanças nas probabilidades dos nós. Isso pode ser exemplificado imaginando a partição de um nó com 800 indivíduos, dos quais metade possui o desfecho de interesse. Tal nó pode ser denotado como (400; 400), representando 400 indivíduos com o desfecho e 400 sem o desfecho, respectivamente. As medidas de impureza de duas possibilidades de partição estão apresentadas na Tabela 2 (Hastie *et al.*, 2009):

¹ A medida de entropia também pode ser calculada dividida por 2, de modo alcançar o mesmo valor máximo (0.5) que as outras medidas de impureza alcançam quando $p = 0.5$

Tabela 2 – Medidas de impureza de duas possíveis partições de um nó (400, 400)

Nós resultantes da partição	Erro de classificação	Índice Gini	Medida de entropia
(300; 100) e (100; 300)	0,25	0,38	0,81
(200; 400) e (200; 0)	0,25	0,33	0,69

A segunda partição é considerada preferencial, uma vez que produz um nó puro, ou seja, com todos os indivíduos com o mesmo status do desfecho. De acordo com a Tabela 2, tanto o índice de Gini quanto a medida de entropia realizariam a partição favorecendo a segunda possibilidade, uma vez que resulta em menor impureza. Já a medida do erro de classificação não consegue diferenciar entre as duas possibilidades (Hastie *et al.*, 2009).

Uma importante limitação das árvores de decisão com partição binária recursiva é o fato de utilizarem algoritmos ditos *gulosos*, ou seja, que realizam escolhas localmente otimizadas (como a escolha de uma primeira partição) e que nunca são reconsideradas. É possível que uma partição realizada nos estágios iniciais de construção da árvore, embora localmente otimizada, dê origem a uma árvore que, como um todo, é menos otimizada. Embora uma alternativa seja a utilização de algoritmos de *backtracking*, que reconsideram partições já realizadas, estes algoritmos podem aumentar significativamente a complexidade computacional do problema a ser estudado (Flach, 2012).

1.3.2.2. Implementações

O método mais popular para construção de árvores de decisão é denominado **CART** (*Classification and Regression Tree*). No método CART, a comparação entre os pontos de divisão acontece simultaneamente entre todas as variáveis independentes e seus respectivos possíveis pontos de corte. No caso de uma variável independente contínua, o algoritmo avalia as possibilidades de corte ao longo dos valores disponíveis desta variável. No caso de uma variável categórica, o algoritmo cria combinações de categorias para testar as possíveis partições (Henrard *et al.*, 2015; Wolfson e Venkatasubramaniam, 2018).

Dadas as possibilidades de partições investigadas, a impureza média dos nós resultantes de cada partição é então calculada e o ponto de corte escolhido é aquele

que apresenta maior redução desta impureza. Esse processo é realizado continuamente até que uma regra de parada seja atingida. As **regras de parada** podem incluir testes de hipótese, um número mínimo de observações a ser mantido em cada nó ou ainda um limite mínimo de redução nos erros de classificação (Hastie *et al.*, 2009; Flach, 2012; Wolfson e Venkatasubramaniam, 2018).

Pela natureza do processo de busca simultânea e exaustiva em cada divisão do método CART, há o risco de uma seleção enviesada das variáveis utilizadas na construção do modelo em favor daquelas que apresentam um maior número de possíveis divisões. Além disso, o algoritmo não consegue distinguir uma mudança na capacidade preditiva que seja estatisticamente significativa de uma insignificante, uma vez que não leva em consideração as propriedades das *distribuições* das medidas (Hothorn *et al.*, 2006).

Considerando essas limitações, uma alternativa mais recente proposta para o CART é o método **CTree** (*Conditional Inference Tree*), que segue dois estágios para realizar a divisão. Primeiramente, a variável de divisão é selecionada por meio de uma medida de correlação calculada entre cada variável independente e o desfecho. Aquela com a associação mais forte é escolhida. Em seguida, o melhor ponto de divisão dentre os possíveis valores da variável selecionada é calculado e a divisão é realizada (Wolfson e Venkatasubramaniam, 2018).

No método CTree, a significância estatística é contemplada, uma vez que antes de cada divisão é realizado um teste estatístico apropriado baseado em uma *hipótese nula global* de que nenhuma variável independente possui uma associação univariada com o desfecho. Caso esta hipótese seja rejeitada, aquela variável com a associação mais forte é selecionada como candidata para divisão. Esta divisão só ocorrerá se o valor p da associação individual entre esta variável e o desfecho (ajustado para múltiplos testes) for inferior a um limite previamente estabelecido. Caso isso não aconteça, a divisão não ocorre e o nó é considerado um nó terminal. Portanto, a expectativa de erro tipo I está contemplada tanto globalmente quanto em cada nó (Venkatasubramaniam, Ashwini *et al.*, 2017; Wolfson e Venkatasubramaniam, 2018).

A desvantagem do algoritmo CTree reside na natureza inferencial do próprio método: com uma amostra grande o suficiente, mesmo as menores diferenças são consideradas estatisticamente significativas. Exemplo disso é a comparação dos

métodos CART e CTree realizada por meio de dados coletados do *Box Lunch Study* (BLS), um ensaio clínico randomizado realizado entre 2010 e 2013 para investigar o efeito da disponibilidade de diferentes tamanhos de porção na ingestão calórica e ganho de peso. Ao investigar o efeito do tamanho da amostra ($n = 30, 250, 500, 1000, 3000, 5000$) na árvore de decisão criada, o método CART manteve o número de nós terminais estável e sempre abaixo de 25, enquanto o método CTree apresentou crescimento exponencial, atingindo número de nós terminais superior a 200, tornando a interpretação humana do modelo impraticável (French *et al.*, 2014; Venkatasubramaniam, Ashwini *et al.*, 2017).

1.3.2.3. Sobreajuste e poda

As árvores de decisão, assim como outros algoritmos de aprendizagem de máquina, podem sofrer de sobreajuste. Sobreajuste é o produto de uma modelagem excessiva dos dados analisados, de modo a explicar não só as tendências gerais do fenômeno a ser estudado, mas também anomalias presentes nos dados (como erros de medida ou fenômenos temporais muito específicos). O problema associado ao sobreajuste é que o modelo criado, apesar de descrever em detalhe os dados a partir dos quais foi criado, pode não ser generalizável para novos dados, mesmo que tratem do mesmo fenômeno (Flach, 2012; Murphy, 2012).

Para controlar sobreajuste na construção de uma árvore de decisão, pode-se utilizar uma estratégia denominada **poda**, que consiste na exclusão (ou prevenção da criação) de nós com pouco poder preditivo. A poda pode ser realizada por meio de regras de parada, como as citadas na seção 2.2.2, e por **validação cruzada**. Esta consiste em (Flach, 2012; Patel e Upadhyay, 2012):

- Dividir aleatoriamente o banco de dados a partir do qual a árvore será criada em k partes de mesmo tamanho (com k tipicamente igual a 10);
- Usar $k - 1$ partes em conjunto para construir uma árvore e 1 parte (que foi reservada) para avaliar a capacidade preditiva do modelo, comparando os resultados previstos pela árvore de decisão com valores reais do desfecho;
- O processo é repetido k vezes, cada vez com uma parte diferente do banco reservada para teste. Deste modo, todo o banco é usado tanto para construção da árvore, quanto para sua validação;

- Com base nos resultados observados (e de acordo com o algoritmo específico que será utilizado), constrói-se uma nova árvore utilizando todo o banco. Esta árvore é a árvore podada;

Uma das vantagens de se utilizar validação cruzada é que os diferentes resultados obtidos das k árvores criadas fornecem uma estimativa da variabilidade na performance dos algoritmos e, portanto, sua dependência de variações nos bancos usados para construção das árvores. As medidas não são de modo algum independentes, uma vez que são criadas a partir de bancos com grande sobreposição, mas são úteis para identificação e tratamento de sobreajuste (Flach, 2012).

Neste projeto, a método CART será implementado utilizando o pacote *rpart* do software R. O algoritmo do pacote *rpart* pode utilizar a validação cruzada para determinar o **parâmetro de complexidade** (α) que será usado para construir a árvore podada. Este parâmetro é uma medida do aumento mínimo na capacidade preditiva do modelo necessário para que a divisão de um determinado nó seja realizada, essencialmente operando como uma regra de parada para a construção da árvore podada. Caso o usuário não especifique o parâmetro de complexidade – calculado por meio de validação cruzada ou determinado arbitrariamente – um valor padronizado de 0,01 é utilizado (Therneau *et al.*, 2019).

O parâmetro de complexidade α é utilizado para balancear a capacidade preditiva da árvore e o número de nós terminais que a compõe, por meio uma **função de risco ajustada** R_α , análoga ao R^2 ajustado utilizado na avaliação de modelos de regressão linear multivariável. A função R_α de uma árvore T é expressa da forma:

$$R_\alpha(T) = R(T) + \alpha|T|$$

em que $R(T)$ é a função de risco não ajustada da árvore (que representa uma medida ponderada das impurezas dos nós terminais da árvore T) e $|T|$ é o seu número de nós terminais. α varia de 0 a $+\infty$, de modo que (Therneau e Atkinson, 2019):

- quando $\alpha = 0$, a árvore podada equivale à árvore sem poda, ou seja, à árvore com o número máximo de nós permitidos pelas outras regras terminais utilizadas;

- quando $\alpha = +\infty$ a árvore podada equivale a uma árvore com apenas um nó terminal, ou seja, sem nenhuma partição;

A função de risco ajustada tende a ser maior nos extremos de α e atingir um valor mínimo em algum ponto entre 0 e $+\infty$ onde existe um balanço entre uma árvore muito complexa com baixa impureza (mas possivelmente sobreajustada) e uma árvore muito simples com alta impureza (Therneau e Atkinson, 2019).

Durante a validação cruzada:

- são criadas árvores a partir da combinação de diferentes valores de α e os diferentes sub-bancos compostos por $k - 1$ partes do banco de dados original;
- usando as partes reservadas para testes, uma função de risco é calculada para cada combinação;
- soma-se a função de risco de todas as combinações criadas a partir de um mesmo valor de α , determinando uma função de risco ajustada para cada α ;
- o resultado é apresentado em uma tabela para o usuário;

Tipicamente, o parâmetro de complexidade selecionado é aquele com a menor função de risco ajustada. Uma descrição mais detalhada de quais valores de α são avaliados durante a validação cruzada pode ser encontrada em outro lugar (Therneau e Atkinson, 2019).

1.3.2.4. Revisão da literatura

A revisão da literatura foi realizada por meio da busca indexada nas bases PubMed e Web of Science. Para o PubMed, foram selecionados todos os artigos com título ou resumo que contenham qualquer um dos termos *epidemiology*, *epidemiological*, *cohort*, *case control study*, *prospective study*, *retrospective study*, *cross sectional study*, *ecologic study*, *transverse study*, *prevalence study* ou *observational study* e com qualquer campo que contenham *classification and regression tree* ou *conditional inference tree*, aplicando o filtro da própria plataforma para artigos que tratem de humanos. Para o Web of Science, os mesmos termos foram utilizados, mas nenhum filtro foi aplicado, uma vez que a plataforma não apresentava essa funcionalidade. A estratégia encontra-se apresentada na Figura 6.

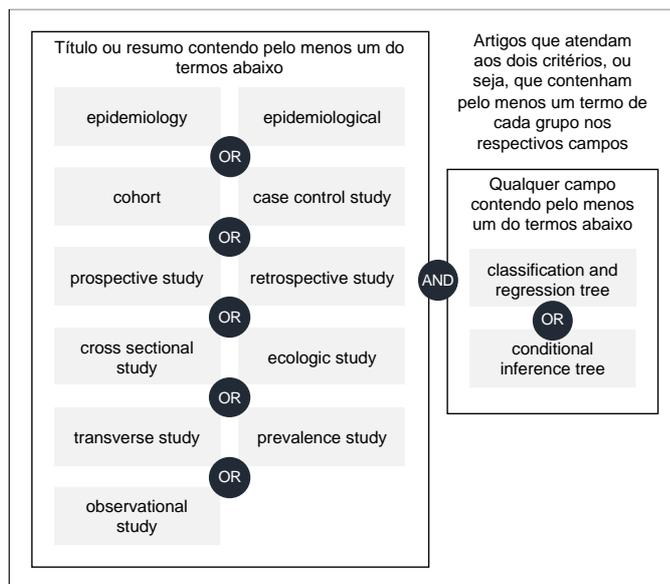


Figura 6 – Estratégia de busca utilizada para as bases do PubMed e Web of Science. No Pubmed, após realizada a busca, também foi aplicado um filtro para artigos que tratem de humanos.

A Figura 7 apresenta o fluxograma das etapas da busca e seus respectivos resultados. Diferentemente da busca realizada para a dose zero de vacinação, o objetivo desta revisão era descrever os principais modos de utilização das árvores de decisão na pesquisa epidemiológica. Portanto, nenhum artigo foi descartado pelo título e todos tiveram os resumos lidos.

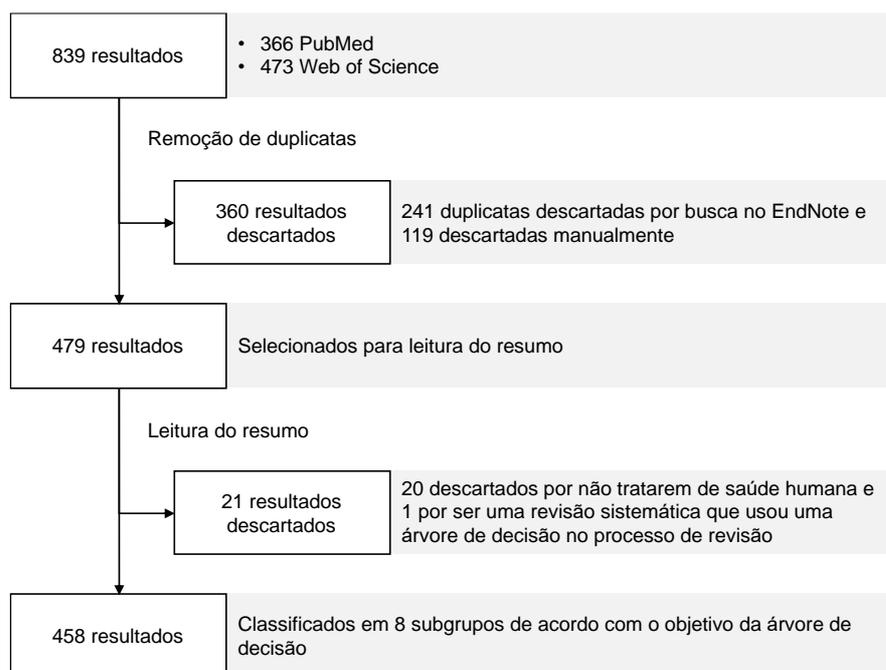


Figura 7 – Fluxograma do processo de busca na literatura epidemiológica sobre árvores de decisão. Busca realizada em 01 de agosto de 2019.

Um resumo dos principais pontos discutidos pelos artigos encontrados com foco nos métodos CART e CTree aplicado à pesquisa epidemiológica estão apresentados no Quadro 2.

Quadro 2 – Revisão: estudos com foco nos métodos CART e CTree aplicado à pesquisa epidemiológica

Autores e ano de publicação	País	Método discutido	Descrição
(Venkatasubramaniam, A. <i>et al.</i> , 2017)	Reino Unido	CART e CTree	Comparação de CART e CTree na análise de um estudo sobre tamanho das porções alimentares. Discute o fato de CTree utilizar testes estatísticos de hipótese formais para construção de uma árvore de decisão, o que facilita sua interpretação no contexto epidemiológico. Apresenta uma ferramenta gráfica desenvolvida para melhor apresentação dos subgrupos criados pela árvore de decisão. Discute o impacto do tamanho da amostra na acurácia e número de nós terminais da árvore criada.
(Bertolet <i>et al.</i> , 2016)	EUA	CART	Propõe um método para identificar subgrupos de fatores de risco que interagem entre si e variam com o tempo ajustando para fatores de confundimento utilizando CART em combinação com regressão de Cox. O desfecho é a sobrevivência de participantes de um estudo que possuíam tanto diabetes quanto 2 quanto doença coronariana.
(Henrard <i>et al.</i> , 2015)	Bélgica	CART	Comparação de CART e <i>random forest</i> (outro algoritmo de machine learning baseado em árvores de decisão, mas que tipicamente é classificado como “caixa preta”) com regressão linear e logística multivariável para estudo de hemofilia. Considera que a maior vantagem do CART é a facilidade na implementação, entendimento e interpretação. A criação de subgrupos cria uma ferramenta de fácil utilização na prática clínica. Outra vantagem destacada é o fato de não fazer suposições sobre a distribuição das variáveis e lidar de maneira flexível com multicolinearidade. Apresenta um quadro comparativo entre os métodos.
(Gaudart <i>et al.</i> , 2015)	França e Mali	CART	Apresenta um novo método para investigar dados espaciais com árvores que realizam cortes oblíquos, denominado SpODT (<i>spatial oblique decision tree</i>). O algoritmo é baseado no CART, que, quando aplicado a dados espaciais dá origem a zonas retangulares resultantes de partições perpendiculares de latitude e longitude. Como SpODT cria zonas oblíquas, tende a ter maior acurácia na epidemiologia espacial.
(Willke <i>et al.</i> , 2012)	EUA	CART	Compara uma série de modelos (como meta-análise, regressão quantílica e métodos não paramétricos selecionados), incluindo CART, para estudo da heterogeneidade de efeito de tratamento, ou HTE. Considera que a simplicidade do método CART o torna particularmente útil para os estágios iniciais de análise exploratória. Como não faz suposições sobre a distribuição das variáveis a serem investigadas, pode ser usada em situações em que outras técnicas de modelagem falham, principalmente quando exigem uma distribuição normal. Apresenta um quadro comparativo entre as técnicas, mas específico para HTE.
(Schwender <i>et al.</i> , 2004)	Alemanha e Inglaterra	CART	Compara uma série de modelos (como <i>random forest</i> , <i>LogitBoost</i> e <i>k</i> vizinhos mais próximos), incluindo CART, para estudo de interações genéticas em um estudo caso-controle de câncer de mama. Todos os métodos foram considerados adequados para investigação das interações genéticas. Um possível problema associado ao CART é sua instabilidade, uma vez que as partições se mostraram bem dependentes de pequenas alterações nos valores das variáveis independentes.

Todos os 458 artigos foram classificados em 8 subgrupos definidos a partir da leitura dos 100 resumos mais recentes. Os subgrupos foram criados de modo a tentar representar os objetivos para utilização das árvores de decisão mais prevalentes na pesquisa epidemiológica e as áreas da saúde que se destacaram por sua utilização. As definições utilizadas estão apresentadas a seguir:

- **Prognóstico:** dados indivíduos doentes/com um agravo de saúde, as árvores de decisão foram utilizadas para realizar previsões sobre o desenvolvimento da doença/agravo, com base em uma série de fatores de risco;
- **Diagnóstico:** dada uma série de fatores de exposição (como histórico familiar, resultados de exames e características biológicas), as árvores de decisão foram usadas para estabelecer critérios de diagnóstico de uma doença/agravo (ou diferentes expressões da mesma doença/agravo);
- **Grupos de risco – geral:** dado um desfecho conhecido, as árvores de decisão foram usadas para encontrar grupos de risco do desfecho baseado em uma série de exposições. A busca de grupos de risco para dose zero de vacinação de enquadra neste grupo;
- **Grupos de risco – pesquisa genética:** um subgrupo com a mesma definição de “Grupos de risco – geral”, porém trata exclusivamente de artigos que envolvam pesquisa genética. As árvores de decisão foram usadas para investigar a interação entre diferentes genes entre si e com outros fatores biológicos, socioeconômicos e comportamentais. Este grupo foi criado de modo a destacar a recorrência de utilização das árvores de decisão na pesquisa genética;
- **Ponto de corte – geral:** as árvores de decisão foram usadas para auxiliar no processo de categorização de variáveis contínuas. Para cada variável a ser categorizada, uma árvore era criada com base apenas nesta variável. A associação entre os estratos da variável categorizada e o desfecho de interesse era então investigado por meio de outro modelo estatístico;
- **Ponto de corte – medicamentos:** um subgrupo com a mesma definição de “Ponto de corte – geral”, porém trata exclusivamente da pesquisa em fármacos. As árvores de decisão foram usadas para estabelecer doses, limites de toxicidade ou qualquer ponto de corte associado ao uso de um medicamento.

Este grupo foi criado de modo a destacar a recorrência de utilização das árvores de decisão na pesquisa farmacêutica;

- **Diretrizes e orientações:** com foco administrativo dos serviços de saúde e da pesquisa clínica e epidemiológica, as árvores de decisão foram utilizadas para estabelecer diretrizes ou orientações a serem seguidas, criando critérios para elegibilidade de pessoas em pesquisa, definições de melhores práticas para tratamentos, triagem de pacientes e outras decisões clínicas;
- **Método:** o artigo tratava especificamente das árvores de decisão, mesmo que comparada a outros métodos, e da sua aplicação na pesquisa sobre saúde humana;

A distribuição dos subgrupos está apresentada na Figura 8. A distribuição dos métodos utilizados nos artigos – apenas CART, apenas CTree ou uma combinação de ambos – está apresentada na Figura 9. A distribuição do número de artigos publicados ao longo dos anos está apresentada na Figura 10.

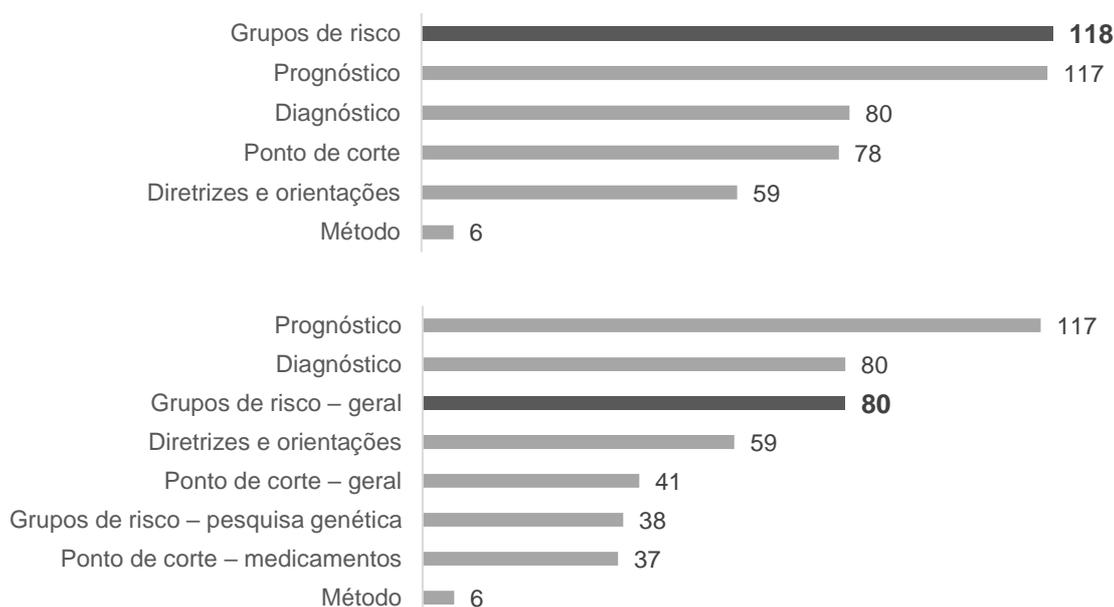


Figura 8 – Principais objetivos para utilização das árvores de decisão na pesquisa epidemiológica de acordo com a revisão da literatura. Na parte superior, as seis grandes categorias foram apresentadas. Na parte inferior, o subgrupo “Grupos de risco” foi apresentado dividido em “Grupos de risco – geral” e “Grupos de risco – pesquisa genética”, enquanto o subgrupo “Ponto de corte” foi apresentado dividido em “Ponto de corte – geral” e “Ponto de corte – medicamentos”. Em destaque os subgrupos “Grupos de risco” e “Grupos de risco – geral” aos quais a árvore de decisão aplicada a identificação de crianças dose zero pertenceria.



Figura 9 – Distribuição dos métodos utilizados nos artigos. A grande maioria utiliza CART (*Classification and regression tree*) e apenas 3 o utilizam em combinação com CTree (*Conditional inference tree*).

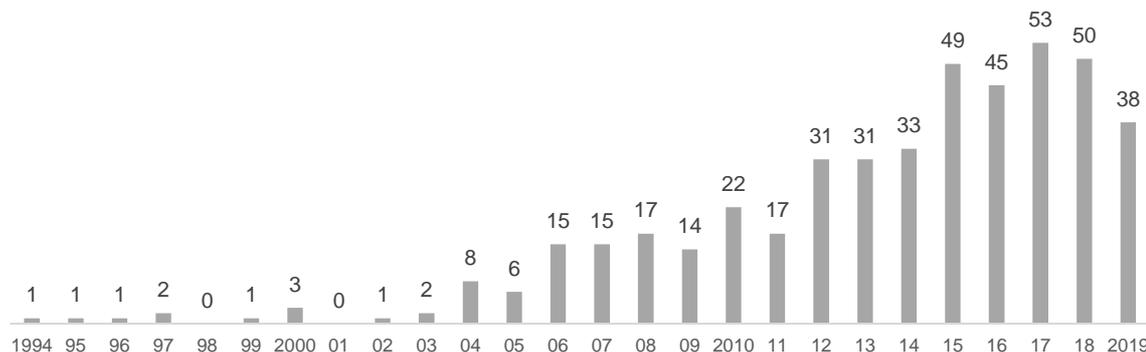


Figura 10 – Número de artigos publicados ao longo dos anos. O primeiro artigo que pertencia à revisão da literatura realizada foi publicado em 1994. O primeiro artigo que usava o método CTree foi publicado em 2006.

A revisão da literatura realizada indica algumas características da utilização de árvores de decisão (especificamente as implementações CART e CTree) na pesquisa epidemiológica:

- O método ainda é pouco utilizado, uma vez que apenas 458 resultados válidos foram encontrados, em sua maioria na pesquisa clínica;
- A implementação CART é predominante na literatura, sendo utilizada em 99% dos artigos investigados, enquanto CTree era utilizada em 4%;
- Artigos específicos de discussão metodológica são raros: apenas 6 foram encontrados na busca;
- A utilização de árvores de decisão na pesquisa epidemiológica é recente. O primeiro artigo encontrado foi de 1994 e mais de 10 artigos por ano só passaram a ser publicados a partir de 2006. O primeiro artigo que utilizava CTree foi publicado em 2006;

Em relação ao objetivo de sua utilização, as árvores de decisão foram usadas tipicamente para estabelecer diagnósticos ou prognósticos, encontrar grupos de risco para um determinado desfecho ou ajudar no processo de categorização de variáveis contínuas. Duas áreas de aplicação se destacaram: na busca de interação entre

genes e entre genes e outras exposições e na definição de limites de toxicidade ou qualquer ponto de corte associado ao uso de um medicamento.

As árvores de decisão têm sido particularmente aplicadas nas situações em que são esperadas múltiplas interações entre os fatores de exposição de interesse. Além disso, sua característica visual, relativa facilidade na interpretação e semelhança com ferramentas heurísticas de tomada de decisão na prática clínica, tornam as árvores de decisão adequadas para estabelecimento de ferramentas de diagnóstico e prognóstico (Smadja *et al.*, 2013; Shi *et al.*, 2017).

1.4. Marco teórico e modelo conceitual

O modelo conceitual foi estruturado em 5 níveis: **ambientes**, **fatores regionais/nacionais**, **características**, **acesso/intenção** e **desfecho**; e encontra-se apresentado na Figura 11, ao final da seção.

O primeiro nível, **ambientes**, trata dos elementos mais distais na cadeia da dose zero de vacinas, representando características nacionais/regionais que compõem o ambiente macro nos quais as crianças que receberão ou não alguma dose de vacina se inserem e que afetam a vacinação de maneira mais distal.

O segundo nível, **fatores regionais/nacionais**, trata de estruturas mais tangíveis – como a presença de um sistema de saúde público – e mais proximais da dose zero, influenciadas pelo ambiente e que diretamente afetam as características de vida e vacinação das crianças. O primeiro e o segundo níveis juntos compõem os **fatores contextuais**.

O terceiro nível, **características**, trata diretamente das condições de vida e saúde da criança e de sua família (e que são afetadas pelos fatores contextuais). Neste nível, não estamos tratando das características gerais de desenvolvimento econômico ou nível de analfabetismo de uma país/região e sim do nível de riqueza e educação dos membros da família da criança em estudo.

No quarto nível, **acesso/intenção**, temos a culminação dos fatores contextuais e das características individuais da criança e de sua família nos fatores mais proximais da dose zero apresentados: acesso e intenção de vacinação. Por fim, no quinto e último nível, o **desfecho**: a dose zero de vacinação.

1.4.1. Desfecho

Uma diferença crucial entre um modelo para dose zero de vacinação e outro que envolvesse a vacinação de forma geral, é a ausência de fatores associados a múltiplas doses. Para que uma criança esteja apropriadamente coberta pelas vacinas do calendário de vacinação do seu país/região, múltiplos contatos com o sistema de saúde ou com campanhas de vacinação são necessários (Bosch-Capblanch *et al.*, 2012). Isso ocorre porque o período da vida no qual a criança é vacinada é importante para o sucesso da imunização e cada vacina/dose possui um período recomendado específico. Uma dose administrada muito tarde deixa a criança exposta

desnecessariamente, enquanto uma dose administrada muito cedo pode não ser efetiva (Suarez-Castaneda *et al.*, 2014).

1.4.2. Acesso e intenção

Para que a criança receba pelo menos uma dose de qualquer vacina, a interação dos seguintes fatores proximais é determinante:

- **Acesso a vacinação:** que inclui todos os fatores necessários para que a vacina esteja disponível para ser recebida por uma criança, caso os responsáveis tenham a intenção de vaciná-la. Dentre eles, inclui-se não só o suprimento da vacina em si e o seu monitoramento, mas também a gestão da cadeia de suprimentos, comunicação dos programas de vacinação, parcerias com a comunidade, supervisão cooperativa da imunização e o uso local de dados de desempenho e gestão financeira (Steinglass, 2013).
- **Intenção de vacinação:** que inclui os fatores necessários para que os responsáveis tomem a decisão de vacinar ou não a criança, independentemente da disponibilidade da vacina. Dentre os fatores associados a esta decisão, incluem-se as crenças religiosas, culturais e étnicas, o nível e a qualidade do conhecimento sobre a segurança e efetividade das vacinas, a relação pessoal com o sistema de saúde (e as experiências negativas ou positivas passadas, chegando a extremos de humilhação pública e abuso verbal por parte dos profissionais de saúde), a suspeita dos interesses comerciais da indústria farmacêutica e a influência de modismos e da naturopatia extrema (Favin *et al.*, 2012; Justich, 2015).

É importante destacar que a *intenção de vacinação* não implica necessariamente em *desejo ou vontade própria de vacinação*, uma vez que políticas públicas de vacinação compulsória podem afetar a intenção de vacinação, mesmo que a contragosto. Na Europa, por exemplo, 11 países dispõem de políticas de vacinação compulsória, tais como a Itália, que possui multas para os pais de crianças não vacinadas que frequentem o ensino primário (Draeger *et al.*, 2019);

Tanto o acesso quanto a intenção de vacinação estão relacionados às **características individuais e familiares** da criança e também aos **fatores contextuais** aos quais está exposta. Porém, estas características individuais e

familiares são também afetadas pelos fatores contextuais, de modo que estes podem afetar o acesso e a intenção de vacinação diretamente ou por meio destas características individuais e familiares. Tomemos o desenvolvimento econômico do país como exemplo. Ele pode impactar negativa e diretamente o acesso à vacinação no caso da ausência de um sistema de transporte e logística eficientes (que dificulta a cadeia de transporte a frio das vacinas) ou indiretamente por meio de desigualdades socioeconômicas que tornam inviáveis para algumas famílias arcar com os custos do cartão de vacinação e das vacinas (em clínicas particulares) ou do transporte até os locais de vacinação e do tempo de afastamento do trabalho (Rainey *et al.*, 2011; Favin *et al.*, 2012).

1.4.3. Características

As **características individuais e familiares** da criança tem efeitos variados em diferentes regiões do mundo, mesmo dentro dos países de baixa e média renda, devido a mecanismos que se expressam de maneira diferente nas complexas interações sociais entre os indivíduos e o ambiente. Ao mesmo tempo em que maior nível educacional dos pais, maior renda e posse de bens de comunicação (como rádio e TV) estão associados consistentemente com menor chance de dose zero (Bosch-Capblanch *et al.*, 2012), outros fatores como a religião familiar apresentam associações significativas díspares para diferentes grupos religiosos. Muçulmanos e cristãos apresentaram maior frequência de dose zero quando comparados aos hindus na Índia (Shenton, Wagner, Bettampadi, *et al.*, 2018) e muçulmanos possuíam maior frequência quando comparados aos cristãos na Nigéria (Chido-Amajuoyi *et al.*, 2018). Já no Quênia, no entanto, os muçulmanos apresentaram frequência de dose zero significativamente menor quando comparados aos protestantes e outros cristãos (Masters *et al.*, 2019).

1.4.4. Fatores regionais/nacionais

Em termos dos fatores contextuais mais proximais, há quatro categorias de destaque: **fatores de saúde populacional, fatores do sistema de saúde público e privado, fatores socioeconômicos e fatores socioculturais.**

Dentre os **fatores de saúde populacional**, se destacam a carga populacional de doença e a ocorrência de surtos epidêmicos devido ao seu impacto na vacinação, tanto no sentido de ampliar quanto de reduzir a vacinação. De um lado, maiores

cargas de doenças e surtos epidêmicos é acompanhada de maiores investimentos nacionais e internacionais em atividades suplementares de imunização, tais como campanhas para controlar surtos de sarampo, febre amarela, difteria e meningite meningocócica (WHO, 2009). Do outro lado, pode ocorrer também a queda na cobertura de vacinação (e o conseqüente aumento à susceptibilidade a doenças como sarampo) que se segue a epidemias graves ao ponto de disrupção dos sistemas de saúde regionais, como na epidemia de ebola na África Ocidental de 2013 (Takahashi *et al.*, 2015), ou por meio de sobrecarga destes sistemas devido ao aumento da carga de doenças como HIV/AIDS não acompanhado de investimentos correspondentes (Streefland *et al.*, 1999).

Em termos de **fatores do sistema de saúde público e privado**, tratamos não só da aplicação das vacinas na população, mas também de todo o processo de pesquisa, desenvolvimento, produção, armazenamento, logística e comercialização das vacinas por meio de parcerias público-privadas que envolvem governos, a indústria farmacêutica, empresas de biotecnologia, órgãos como a OMS e a GAVI, além de instituições financiadoras como a Bill & Melinda Gates Foundation e o Wellcome Trust (Rappuoli *et al.*, 2019).

Historicamente, o desenvolvimento de uma nova vacina (que requer entre 15 e 20 anos e recursos financeiros da ordem de 1 bilhão de dólares americanos) é voltado para dois mercados distintos: i) países de alta renda – de onde as indústrias farmacêuticas obtêm a maior parte do lucro que justifica economicamente o desenvolvimento; e ii) países de baixa e média renda – onde os preços de venda são significativamente reduzidos. Entretanto, esse modelo não está mais se mostrando sustentável considerando os novos desafios de doenças como ebola e zika, uma vez que são encontrados predominantemente em países de baixa e média renda (Rappuoli *et al.*, 2019).

Já a vacinação em si ocorre tanto via programas de imunização de rotina quanto via campanhas mais intermitentes. A imunização de rotina sofre de barreiras críticas relacionadas à fragilidade dos sistemas de saúde, logística e infraestrutura dos países de baixa e média renda, devido à falta de investimento político e financeiro, à gestão e ao monitoramento falhos, à falta de profissionais de saúde qualificados e à complexidade da cadeia fria de transporte e armazenamento das vacinas. Já as

campanhas de imunização podem ser utilizadas justamente para atingir as regiões de mais difícil acesso ou em um único movimento alcançar um maior número de pessoas do que os sistemas de rotina geralmente conseguem (WHO, 2009).

Do modelo de campanhas mais focadas, tem-se como exemplo a estratégia *Reaching Every District*, de uma parceria da OMS, UNICEF e outras instituições, com foco nas populações mais difíceis de serem alcançadas, principalmente na África e Ásia, e que envolve microplanejamento geográfico, social e cultural com objetivo de identificar e alcançar estas populações. Do modelo de campanhas mais massivas, têm-se as “semanas de imunização” em muitos países americanos e europeus e os “dias da saúde infantil” de muitos países africanos, que utilizam estratégias de mobilização em massa para tentar atingir cobertura universal de vacinação (Who, 2009).

Do ponto de vista dos **fatores socioeconômicos** contextuais, é especialmente importante analisar a decisão e acesso à vacinação não apenas como um fenômeno individualizado ou restrito à esfera familiar, mas sim como pertencendo a um contexto social mais abrangente. A campanha de vacinação contra poliomielite realizada em Israel após um surto epidêmico da doença em 2013 é um exemplo do impacto dos fatores socioeconômicos comunitários na vacinação individual. Após a campanha, a maior taxa de aceitação da vacina foi observada nas populações de menor nível socioeconômico e menor índice de desigualdade. Uma possível explicação proposta é a relação inversa entre desigualdade e solidariedade – que já foi explorada em outros contextos, como imigração – uma vez que uma característica particular da campanha realizada em Israel foi o foco em solidariedade e segurança da família como motivos para vacinação (Tur-Sinai *et al.*, 2019).

Por outro lado, existem registros da desigualdade social como fator associado à uma maior vacinação, como na Indonésia. No nível das vilas, a administração dos serviços de saúde promovia vacinação por meio das elites locais que, devido às relações de poder e dependência, estava associada a uma maior conformidade da maioria mais pobre, que se sentia obrigada a participar das atividades de saúde promovidas (Streefland *et al.*, 1999).

Quando avalia-se a escolha informada das famílias de vacinar ou não suas crianças, o papel da educação é crucial, considerando não só o fato de ser um dos

fatores mais consistentemente associados com a dose zero no nível individual (Bosch-Capblanch *et al.*, 2012), mas também o impacto de mudanças sociais como a recente redução da confiança pública em instituições como o governo e a indústria farmacêutica (Iriart, 2017). Tal fenômeno não impacta apenas a população geral, mas também os profissionais de saúde, resultando até mesmo em médicos e médicas que realizam recomendações contrárias a vacinação, particularmente dentro da corrente homeopática (Justich, 2015).

As crenças contrárias à vacinação têm diversas origens: religiosas, culturais, étnicas e até mesmo conspiratórias, que estão associadas ao acesso e proliferação de informação por meio de fontes pouco controladas, como a internet (Justich, 2015). Nesse contexto, os **fatores socioculturais** assumem papel de destaque. As relações comunitárias são impactantes na reputação desenvolvida pelas vacinas, uma vez que vizinhos ou parentes, por exemplo, trocam informações sobre as suas experiências, como ser maltratado por um profissional de saúde ou possíveis efeitos colaterais que ocorrem após a vacinação e que os pais associam com a vacina, criando culturas locais de vacinação (Streefland *et al.*, 1999).

Embora o impacto da religião na vacinação seja conflitante entre diferentes regiões e estudos, isso de modo algum reduz a importância do tópico. Particularmente no continente africano – que abriga boa parte dos países com as maiores taxas de dose zero – a religião representa uma fonte significativa de identidade e conexão social, culminando em líderes religiosos com influência direta sobre as decisões de saúde dos membros de seus respectivos grupos religiosos (Antai, D., 2009; Bosch-Capblanch *et al.*, 2012).

A Nigéria, um dos países com maior prevalência de dose zero, possui um sistema de saúde complexo que mistura a medicina ortodoxa com práticas de cura tradicional, como herbalistas, curandeiros e divinadores islâmicos. O impacto dos líderes religiosos na vacinação no país é de tal modo marcado que, em 2003, uma campanha de imunização infantil teve que ser suspensa nos estados do norte do país devido à campanha contrária realizada por esses líderes conclamando que os pais não permitissem que seus filhos fossem imunizados. Alegações foram realizadas de que as vacinas contra a poliomielite estavam contaminadas com agentes

antifertilidade, cancerígenos e com HIV, como parte de um plano dos governos ocidentais para reduzir a população muçulmana global (Antai, D., 2009).

Por outro lado, o engajamento social pode ter impacto positivo na aceitação da imunização, seja por meio do desenvolvimento de movimentos sociais organizados pró-vacinação que utilizam de redes sociais e de engajamento comunitário para gerar pressão regulatória (e que surgem como reação contrária ao crescimento dos movimentos críticos à vacinação em países como os Estados Unidos e a Austrália), seja por meio de interação comunitária diária, que permite a comunicação e transmissão de informações básicas sobre as campanhas de vacinação, como local, horário e público-alvo (Rainey *et al.*, 2011; Vanderslott, 2019).

1.4.5. Ambientes

Por fim, na mais distal das camadas do modelo proposto, são apresentados os **ambientes ecológico, político, econômico, social e cultural**.

Provavelmente o exemplo mais marcante do impacto que o **ambiente ecológico**, aliado ao **ambiente econômico**, têm sobre a cadeia de vacinação vem das doenças tropicais negligenciadas, como dengue, Chicungunha, doença de Chagas, lepra, cisticercose e esquistossomose. Apesar de afetarem mais de um bilhão de pessoas, estas doenças foram sistematicamente negligenciadas em maior ou menor grau por pesquisadores, financiadores e legisladores. Isso está associado ao fato de estarem geograficamente presentes primordialmente em países de baixa e média renda (tipicamente localizados em regiões tropicais e subtropicais, onde os patógenos e vetores são endêmicos), serem quase que exclusivamente restritas às pessoas vivendo abaixo da linha da pobreza e, mais recentemente, terem que concorrer com o foco global no combate a HIV/AIDS, tuberculose e malária. Sem incentivo financeiro, o investimento em pesquisa e desenvolvimento de vacinas para as doenças tropicais negligenciadas por parte da indústria farmacêutica têm sido insuficiente (Feasey *et al.*, 2009; Bethony *et al.*, 2011; Peeling *et al.*, 2017).

Em termos do **ambiente político**, o impacto na cadeia de vacinação é multifacetado. De um lado, há regimes que alcançam maior conformidade com as políticas de saúde pública – como vacinação – por meio de ameaças, intimidações e supressão de liberdades individuais, como o regime Derg na Etiópia nos anos 70 e 80. De outro lado, a instabilidade política de uma região pode interromper a

imunização de rotina, devido à deterioração dos sistemas de saúde, sem necessariamente afetar as campanhas de vacinação mais esporádicas (Streefland *et al.*, 1999). Porém, em termos gerais, para o sucesso de novos programas de imunização, um processo sólido de tomada de decisão e forte vontade política são cruciais para que sejam realizadas as necessárias medidas de mobilização social e integração de políticas públicas, que incluem inclusão de novas vacinas, atualizações dos calendários de vacinação e a criação e melhoramento de infraestrutura do sistema de saúde (Hardt *et al.*, 2016).

Em termos legais, a discussão internacional a respeito do papel do Estado ao legislar sobre o processo de imunização – até mesmo coerciva – ganhou destaque na última década, considerando o crescimento dos movimentos contrários à vacinação e o ressurgimento de doenças como o sarampo em países de alta renda. Exemplo emblemático é o surto de sarampo no Disneyland Park, nos Estados Unidos, que se iniciou em dezembro de 2014 e já havia atingido Canadá e México até fevereiro de 2015 e que acentuou o debate nacional sobre o direito dos governos estaduais de impedirem crianças não vacinadas de frequentarem aulas em caso de surtos (Yang e Silverman, 2015).

No extremo do impacto do ambiente político, estão conflitos tais como invasões internacionais e guerras civis. O colapso do sistema de saúde de um território resultante do ataque às instituições de saúde e da redução do número de profissionais da saúde (por êxito, morbidade ou mortalidade), unido aos deslocamentos populacionais forçados, são responsáveis por quebra na continuidade da imunização e redução das taxas de cobertura. Exemplo disso é a região norte da Síria, como na cidade de Kobanê, que, devido à Guerra Civil Síria ainda em andamento, apresenta baixas taxas de cobertura vacinais e alto risco de doenças como sarampo, cólera e pneumonia (De Lima Pereira *et al.*, 2018).

O **ambiente econômico** (nacional e internacional) possui impacto direto no custeamento dos programas de imunização. Nos países de baixa renda, o financiamento insuficiente é uma das principais barreiras para o acesso mais universal à vacinação, tornando esses países dependentes de fontes externas de recursos financeiros e econômicos para instalação e manutenção destes programas (Hardt *et al.*, 2016).

Embora ajuda internacional por meio de organizações como a Pan American Health Organization (PAHO), GAVI e UNICEF tenham um papel fundamental no financiamento dos programas de vacinação, a busca por sustentabilidade financeira a longo prazo tem levado ao desenvolvimento de estratégias públicas e privadas em diversos países para levantamento de recursos (WHO, 2009; Hardt *et al.*, 2016). No Vietnã, Tajiquistão e Haiti, impostos sobre artigos de luxo, álcool e tabaco são usados para esse fim. Na Costa Rica, parte do orçamento do Fundo Nacional de Imunização vem da loteria nacional. No Butão, o Fundo Fiduciário da Saúde, criado para estabilizar o financiamento do sistema de saúde que inclui programas de vacinação, é mantido por doações públicas e privadas (Hardt *et al.*, 2016).

Em termos do **ambiente social**, as diversas formas de organização da sociedade podem impactar a vacinação infantil de maneira desigual entre estratos, castas e outras formas de divisão social.

Na Índia, Castas e Tribos Registradas são grupos sociais oficialmente reconhecidos pelo governo indiano como compostos por pessoas desfavorecidas e para os quais existem ações afirmativas específicas, como cotas em escolas e em vagas de trabalho no governo (Coffey *et al.*, 2019). Crianças pertencentes a esses grupos apresentam maior risco de dose zero, o que pode estar relacionado ao histórico de discriminação que, por sua vez, reduz o acesso aos serviços de saúde, incluindo serviços de imunização (Shenton, Wagner, Bettampadi, *et al.*, 2018).

Já em Myanmar, às pessoas pertencentes ao povo Rohingya é negada identidade legal, certidões de nascimento e até mesmo acesso a serviços essenciais como vacinação infantil. A expulsão sistemática dos Rohingya – descrita como uma forma de limpeza étnica, acompanhada de prisões, violência e tortura – já levou cerca de 700 mil pessoas para Bangladesh, metade da qual se instalou no distrito de Cox's Bazar. Uma pesquisa realizada em 2018 nesse distrito encontrou que 61,7% das crianças refugiadas até dois anos (e 42,9% das até 4 anos) nunca havia recebido uma dose de vacina injetável enquanto em Myanmar e 57,5% nunca havia recebido uma dose de vacina oral (Bhatia *et al.*, 2018).

Além das formas de divisão social, a relação social com o espaço físico também pode ser determinante. O estilo de vida nômade, por exemplo, é um fator de risco significativo para baixa cobertura vacinal. Em uma pesquisa realizada em 2016 entre

as crianças de um ano pertencentes a grupos pastores nômades no Quênia, 42,2% destas haviam recebido zero doses de vacinas. 63,6% das mães investigadas na mesma pesquisa afirmaram ter levado os filhos a um estabelecimento de saúde no último mês, das quais apenas 55% foram questionadas sobre a vacinação dos filhos nestes estabelecimentos (Pertet *et al.*, 2018).

Em termos do **ambiente cultural**, um inquérito de 2015 realizado em 67 países sobre a confiança populacional na imunização indica uma atitude global positiva em relação às vacinas, mas com significativa variabilidade nacional e regional. A Europa concentra sete dos dez países com menor confiança na imunização, com a França se destacando com 41% dos seus respondentes discordando que vacinas são seguras (contra uma média global de 13%) (Larson *et al.*, 2016).

Na pesquisa, foram investigadas as opiniões sobre quatro critérios relacionados à vacinação: importância para as crianças, segurança, efetividade e compatibilidade religiosa. Uma tendência consistente entre quase todos os países foi observada: uma maior proporção de pessoas considera as vacinas efetivas do que as considera seguras. Isso sugere uma disposição a assumir riscos relacionados à vacinação dado seu efeito protetor contra doenças (Larson *et al.*, 2016).

Em alguns países, como Tailândia e Mongólia, a incompatibilidade religiosa é um forte fator de opiniões negativas sobre as vacinas. Apesar disso, o grupo religioso não é suficiente para explicar a incompatibilidade com a imunização: a Arábia Saudita apresentou 100% de respondentes muçulmanos e baixa rejeição à vacinação, enquanto os muçulmanos na Nigéria e Paquistão apresentaram moderado nível de objeção religiosa e estes países possuem um histórico de rejeição à vacina da poliomielite ligado ao fundamentalismo religioso. Os contextos políticos, culturais e históricos modificam o contexto religioso e afetam a formação de crenças pessoais em relação à vacinação, que são compartilhadas entre membros da mesma comunidade, sem que necessariamente exista uma rejeição teológica explícita das vacinas (Larson *et al.*, 2016).

Em conclusão, o modelo proposto não pretende ser exaustivo em relação a descrever as relações sociais em suas diversas facetas e como elas podem afetar a vacinação de uma criança. Primordialmente, seu objetivo é expressar a complexidade

do desafio global de imunização e criar uma estrutura básica para ser usada como guia na realização e interpretação das análises a serem realizadas.

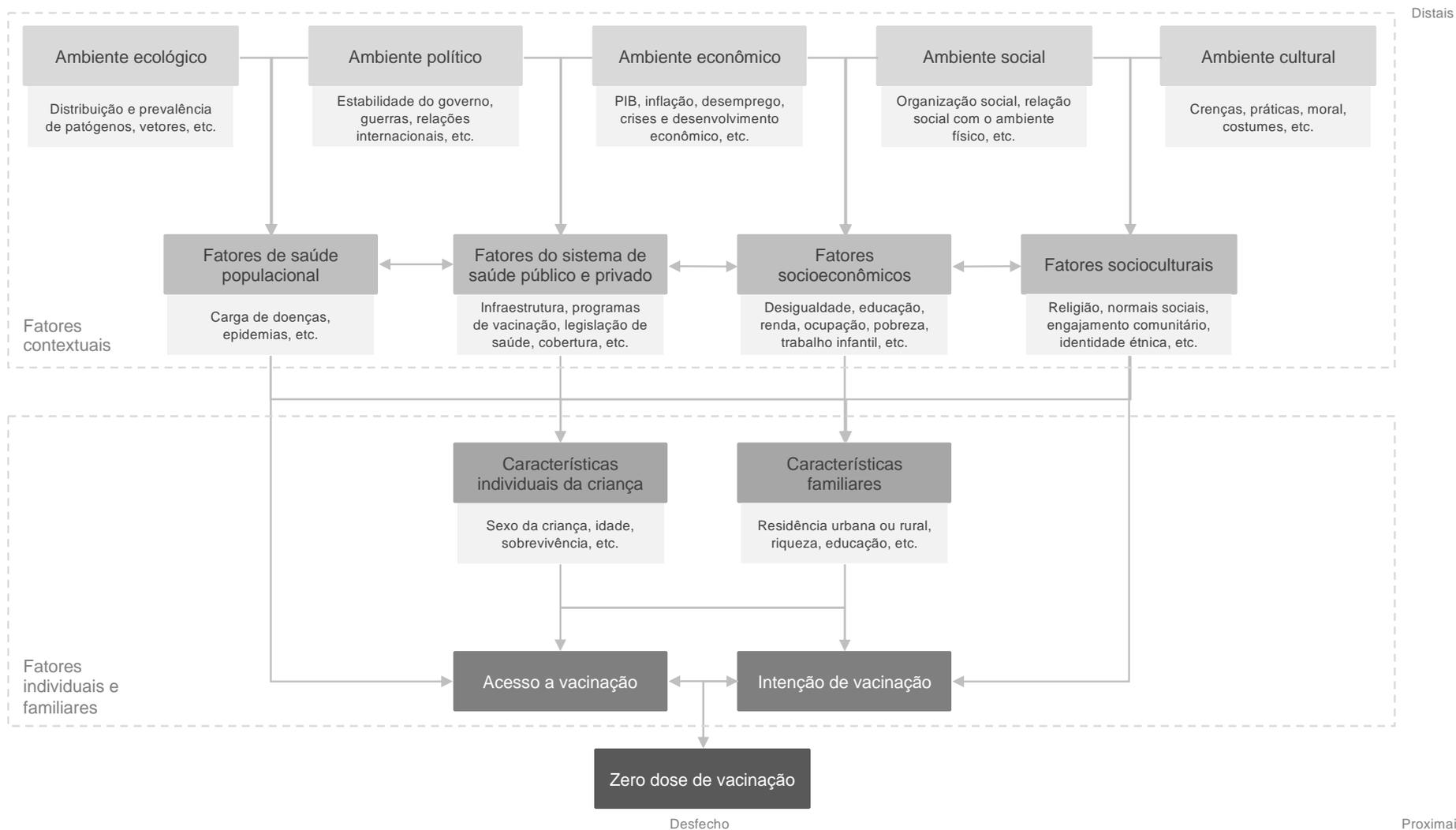


Figura 11 – Modelo conceitual da dose zero de vacinação. Os fatores associados à dose zero são apresentados dos mais distais ao mais proximais de cima para baixo. Abaixo de cada agrupamento de fatores (como Ambiente político), seguem alguns exemplos de fatores deste agrupamento (como estabilidade do governo, guerras e relações internacionais).

1.5. Objetivos

1.5.1. Objetivo geral

- Investigar grupos de risco de dose zero de vacinação por meio de algoritmos de árvore de decisão e determinar a prevalência de dose zero nos inquéritos DHS e MICS de países de baixa e média renda;

1.5.2. Objetivos específicos

- Explorar o uso das árvores de decisão no contexto de epidemiologia e saúde global;
- Estimar a prevalência de dose zero de vacinação nos países de baixa e média renda em seus relatórios DHS e MICS;
- Investigar a associação da dose zero de vacinação com os principais estratificadores de desigualdade em saúde (moradia rural/urbana, educação da mãe, índice de riqueza, etc.) por meio de regressão logística multivariável;
- Construir árvores de decisão global e por país para dose zero de vacinação e descrever os grupos de risco identificados;
- Investigar a aplicabilidade das árvores de decisão como método de análise epidemiológica para dose zero, comparando qualitativamente os resultados obtidos por regressão logística multivariável com os das árvores de decisão construídas;

1.6. Hipóteses

- A prevalência de dose zero de vacinação entre todas as crianças de 12 a 59 meses de todos os inquéritos e todos os países em conjunto será inferior a 10%;
- Serão observadas grandes disparidades entre países em termos de prevalência de dose zero de vacinação, com as maiores prevalências atingindo valores superiores a 25%;
- Os fatores que estarão consistentemente associados com menores prevalências de dose zero serão:
 - Vacinação da mãe por tétano antes da gravidez;
 - Maior nível educacional do responsável pela criança;

- Maior nível educacional do companheiro(a) do responsável pela criança;
- Pertencimento aos níveis superiores do índice de riqueza;
- Os outros fatores associados terão associações variantes entre os distintos países/regiões;

Não serão levantadas hipóteses mais específicas sobre o resultado das árvores de decisão a serem construídas, uma vez que a novidade da análise no contexto epidemiológico e o nível de complexidade dos subgrupos criados torna o exercício demasiadamente especulativo.

1.7. Métodos

1.7.1. Fonte dos dados

Este projeto será realizado utilizando dados coletados em estudos transversais de base populacional nacionalmente representativos promovidos individualmente para países de baixa e média renda pela *United States Agency for International Development (USAID)* e *United Nations Children's Fund (UNICEF)* em parceria com os governos dos países investigados nos seus respectivos inquéritos, DHS e MICS.

O delineamento transversal é particularmente útil na determinação de prevalência de desfechos de saúde (tais como a dose zero de vacinação) e identificação de fatores associados, principal objetivo deste estudo. Além disso, os inquéritos selecionados têm sido extensivamente utilizados como base para decisões de saúde pública em países em desenvolvimento e apresentam alta comparabilidade entre si, uma vez que os times responsáveis trabalham em proximidade para harmonizar as metodologias de coleta de dados (Unicef, 2014).

1.7.2. População-alvo

1.7.2.1. População-alvo do estudo

A população-alvo do estudo serão todas crianças de 12 a 59 meses com dados de vacinação disponíveis via DHS e MICS de todos os países para os quais as pesquisas foram realizadas a partir de 2010 (93 países). Os critérios de definição da população-alvo dos dois inquéritos são apresentados a seguir.

1.7.2.2. População-alvo do DHS

A população-alvo, para cada país, são todas mulheres em idade reprodutiva (entre 15 e 49 anos) e as crianças com menos de 5 anos com habitação residencial. A maioria dos inquéritos inclui também todos os homens entre 15 e 59 anos (International, 2012).

- São excluídas da amostragem regiões de extrema inacessibilidade, violência ou instabilidade;
- Em alguns países, apenas mulheres já casadas são incluídas no processo amostral;

Tomemos os relatórios da Índia de 2015-16 e do Paquistão de 2017-18 como exemplo:

Na Índia, todas as mulheres entre 15 e 49 anos que moravam em residência ou que passaram a noite anterior em um dos domicílios selecionados foram consideradas elegíveis para responder ao questionário, assim como todos os homens entre 15 e 54 anos que habitavam ou passaram a noite anterior no domicílio (International Institute for Population Sciences e Icf, 2017).

No Paquistão, todas as mulheres que já foram casadas e com idade entre 15 e 49 anos que moravam em residência ou que passaram a noite anterior em um dos domicílios selecionados foram consideradas elegíveis para responder ao questionário, assim como todos os homens que já foram casados e com idade entre 15 e 49 anos que habitavam ou passaram a noite anterior no domicílio (National Institute of Population Studies e Icf, 2019).

Em ambos os casos, medições e testes de biomarcadores tiveram critérios de inclusão e exclusão específicos e que dependiam da natureza técnica e ética das próprias medições e testes. Na Índia, a altura e o peso foram medidos para crianças com idade entre 0 a 59 meses, mulheres de 15 a 49 anos e homens de 15 a 54 anos, enquanto o teste de anemia via amostra de sangue foi realizado para crianças com idade entre 6 a 59 meses, mulheres de 15 a 49 anos e homens de 15 a 54 anos. Já no Paquistão, a altura e o peso foram medidos para crianças com idade de 0 a 59 meses e mulheres de 15 a 49 anos, mas amostras de sangue não foram coletadas

(International Institute for Population Sciences e Icf, 2017; National Institute of Population Studies e Icf, 2019).

1.7.2.3. População-alvo do MICS

A população-alvo, para cada país, são todas as mulheres de 15 a 49 anos, homens de 15 a 49 anos e crianças com menos de 5 anos, membros de uma residência (geralmente, mas não exclusivamente, moradores). Tomemos o relatório da Nigéria de 2016-17 como exemplo: ele utiliza a população-alvo padrão do MICS e não acrescenta nenhum novo critério de inclusão ou exclusão ((Unicef), 2017).

1.7.2.4. Países estudados

Serão incluídos todos os 93 países que tenham inquéritos MICS ou DHS realizados a partir de 2010. A lista completa de países, inquéritos e anos correspondentes está apresentada na Tabela 3.

Tabela 3 – Países a serem investigados no estudo e os respectivos anos e inquéritos disponíveis a partir de 2010

País	Ano	Inquérito	País	Ano	Inquérito
Afganistão	2010	MICS	Mauritânia	2011	MICS
Afganistão	2015	DHS	Mauritânia	2015	MICS
África do Sul	2016	DHS	México	2015	MICS
Albânia	2017	DHS	Moçambique	2011	DHS
Algéria	2012	MICS	Moçambique	2015	DHS
Angola	2015	DHS	Moldávia	2012	MICS
Armênia	2010	DHS	Mongólia	2010	MICS
Armênia	2015	DHS	Mongólia	2013	MICS
Bangladesh	2011	DHS	Montenegro	2013	MICS
Bangladesh	2014	DHS	Myanmar	2015	DHS
Belize	2011	MICS	Namíbia	2013	DHS
Belize	2015	MICS	Nepal	2010	MICS
Benin	2011	DHS	Nepal	2011	DHS
Benin	2014	MICS	Nepal	2014	MICS
Bósnia e Herzegovina	2011	MICS	Nepal	2016	DHS
Burkina Faso	2010	DHS	Níger	2012	DHS
Burundi	2010	DHS	Nigéria	2011	MICS
Burundi	2016	DHS	Nigéria	2013	DHS
Camarões	2011	DHS	Nigéria	2016	MICS
Camarões	2014	MICS	Panamá	2013	MICS
Camboja	2010	DHS	Paquistão	2012	DHS
Camboja	2014	DHS	Paquistão	2017	DHS
Cazaquistão	2010	MICS	Paraguai	2016	MICS
Cazaquistão	2015	MICS	Peru	2010	DHS
Chade	2010	MICS	Peru	2011	DHS
Chade	2014	DHS	Peru	2012	DHS
Colômbia	2010	DHS	Peru	2013	DHS
Comores	2012	DHS	Peru	2014	DHS
Congo Brazzaville	2011	DHS	Peru	2015	DHS
Congo Brazzaville	2014	MICS	Peru	2016	DHS
Congo, República Democrática	2010	MICS	Qatar	2012	MICS
Congo, República Democrática	2013	DHS	Quênia	2014	DHS
Costa do Marfim	2011	DHS	Quirguistão	2012	DHS

País	Ano	Inquérito	País	Ano	Inquérito
Costa do Marfim	2016	MICS	Quirguistão	2014	MICS
Costa Rica	2011	MICS	República Centro Africana	2010	MICS
Cuba	2010	MICS	República Dominicana	2013	DHS
Cuba	2014	MICS	República Dominicana	2014	MICS
Egito	2014	DHS	Ruanda	2010	DHS
El Salvador	2014	MICS	Ruanda	2014	DHS
Estado da Palestina	2010	MICS	São Tomé e Príncipe	2014	MICS
Estado da Palestina	2014	MICS	Senegal	2010	DHS
Etiópia	2011	DHS	Senegal	2012	DHS
Etiópia	2016	DHS	Senegal	2014	DHS
Filipinas	2013	DHS	Senegal	2015	DHS
Filipinas	2017	DHS	Senegal	2016	DHS
Gabão	2012	DHS	Senegal	2017	DHS
Gâmbia	2010	MICS	Serra Leoa	2010	MICS
Gâmbia	2013	DHS	Serra Leoa	2013	DHS
Gana	2011	MICS	Serra Leoa	2017	MICS
Gana	2014	DHS	Sérvia	2014	MICS
Guatemala	2014	DHS	Suazilândia	2010	MICS
Guiana	2014	MICS	Suazilândia	2014	MICS
Guiné	2012	DHS	Sudão	2010	MICS
Guiné	2016	MICS	Sudão	2014	MICS
Guiné-bissau	2014	MICS	Sudão do Sul	2010	MICS
Haiti	2012	DHS	Suriname	2010	MICS
Haiti	2016	DHS	Tailândia	2012	MICS
Honduras	2011	DHS	Tailândia	2015	MICS
Íemen	2013	DHS	Tajiquistão	2012	DHS
Índia	2015	DHS	Tajiquistão	2017	DHS
Indonésia	2012	DHS	Tanzânia	2010	DHS
Iraque	2011	MICS	Tanzânia	2015	DHS
Iraque	2018	MICS	Timor-Leste	2016	DHS
Jamaica	2011	MICS	Togo	2010	MICS
Jordânia	2012	DHS	Togo	2013	DHS
Jordânia	2017	DHS	Trinidad e Tobago	2011	MICS
Kosovo	2013	MICS	Tunísia	2011	MICS
Laos	2011	MICS	Turcomenistão	2015	MICS
Laos	2017	MICS	Ucrânia	2012	MICS
Lesoto	2014	DHS	Uganda	2011	DHS
Libéria	2013	DHS	Uganda	2016	DHS
Macedônia	2011	MICS	Vietnam	2010	MICS
Malawi	2010	DHS	Vietnam	2013	MICS
Malawi	2013	MICS	Zâmbia	2013	DHS
Malawi	2015	DHS	Zimbábue	2010	DHS
Maldivas	2016	DHS	Zimbábue	2014	MICS
Mali	2012	DHS	Zimbábue	2015	DHS
Mali	2015	MICS			

1.7.3. Definição operacional do desfecho

O desfecho de interesse é a **dose zero de vacinação** em crianças (12 a 59 meses) de países de baixa e média renda.

Nos inquéritos MICS e DHS, as informações sobre vacinação são coletadas por meio de cartões de vacinação ou, em sua ausência ou caso não seja apresentado ao entrevistador, as mães são questionadas sobre as vacinas da criança, incluindo

aquelas recebidas em campanhas e programas de imunização. Uma vacina é considerada dada se tiver sido registrada no cartão ou informada pela mãe (Croft, 2018). Reprodução das perguntas utilizadas nos inquéritos para coleta dessas informações estão apresentadas no Anexo I.

A criança dose zero será definida como aquela que nunca recebeu sequer uma dose das seguintes vacinas:

- DTP, ou tríplice bacteriana: vacinas contra difteria, tétano e coqueluche;
- MCV, ou vacinas contendo o antígeno do sarampo;
- BCG, ou Bacillus Calmette-Guérin: vacinas contra tuberculose;
- OPV, ou vacinas por via oral contra a poliomielite;

A criança que apresentar registros ausentes – ou seja, nem positivo, nem negativo – para todos os quatro grupos de vacinas será considerada como “informação ausente” e será removida da análise. A criança que apresentar registro de não vacinação para os quatro grupos será considerada “dose zero”. As crianças restantes serão classificadas como “pelo menos uma dose” (Bosch-Capblanch *et al.*, 2012).

1.7.4. Definição operacional das variáveis independentes

As seguintes variáveis independentes serão avaliadas no estudo:

- **País onde o inquérito foi realizado;**
- **Regiões subnacionais:** baseadas nos domínios de amostragem usados em cada inquérito;
- **Número de membros da residência:** membros são aqueles que o entrevistado apontou como moradores usuais da residência;
- **Número de filhos vivos na residência;**
- **Número de filhos já falecidos da residência;**
- **Área de residência:** baseada em critérios estabelecidos por cada país e definida como rural ou urbana;
- **Quintis de riqueza:** a classificação socioeconômica se dá por índice de posse de bens (como televisores e geladeiras), de características da edificação da casa (como materiais das paredes e abastecimento de água) e de outras variáveis relacionadas à situação econômica (como propriedade da casa). Um

único índice é preparado por meio de análise de componentes principais e, com base nele, os quintis são calculados;

- **Posse de rádio;**
- **Posse de TV;**
- **Posse de geladeira;** ²
- **Sexo do chefe da residência;**
- **Ocupação dos pais:** status de emprego da mãe e do pai (quando disponível) reportada pela mãe da criança durante a entrevista;
- **Educação materna:** nível educacional codificado em seis categorias: nenhum, primário incompleto, primário completo, secundário incompleto, secundário completo e ensino superior;
- **Estado civil da mãe:** no momento da entrevista, classificado em i) nunca esteve casada ou em união³; ii) atualmente casada ou em união ou iii) divorciada, separada ou viúva;
- **Idade materna:** anos completos de idade da mãe no momento da entrevista;
- **Grupo étnico da mãe:** baseado em etnia e cor da pele autorreferidas ou no idioma falado em casa, dependendo do inquirido;
- **Religião da mãe:** religião informada na entrevista a partir de pergunta com opções que variam por país de acordo com as religiões mais prevalentes nacionalmente;
- **Vacinação materna por vacina antitetânica:** número de doses de toxoide tetânico recebidas pela mãe da criança durante a última gravidez nos últimos 3 ou 5 anos antes da entrevista (período recordatório varia por país);
- **Número de visitas pré-natal:** informada pela mãe de acordo com a gravidez da criança em questão;
- **Instituição do parto:** classificada em três categorias principais – institucional governamental, institucional privada e não institucional – de acordo com local do parto informado pela mãe da criança;
- **Sexo da criança;**

² A posse de rádio e TV serão investigadas pelo seu possível impacto como meios de comunicação, particularmente para campanhas de vacinação. Entretanto, rádio e TV podem também ser indicadores do nível de riqueza da família e, apesar de estar planejado controle para quintil de riqueza, há risco de confundimento residual. Portanto, a posse de geladeira será usada como variável de comparação.

³ União: declarar viver junto com um homem como se estivessem casados.

- **Ordem de nascimento da criança;**

1.7.5. Processamento e análise de dados

A associação entre a dose zero de vacinação e as variáveis independentes de interesse será investigada por três métodos no nível global (incluindo todos os países e relatórios disponíveis na mesma análise):

- Regressão logística multivariável⁴;
- Árvore de decisão via método CART
- Árvore de decisão via método CTree

A escolha da regressão logística multivariável se dá pela comparação desejada com os resultados já obtidos por meio de regressão logística na investigação de dose zero em 96 países de baixa e média renda (Bosch-Capblanch *et al.*, 2012). A escolha dos métodos CART e CTree é justificada pelos motivos já apresentados em mais detalhes na seção 2.5.

A investigação também será realizada individualmente para cada país, por meio de regressão logística e do método de construção de árvores de decisão que for considerado mais adequado ao problema, considerando o resultado da análise global e da análise de uma subamostra aleatória de 3 inquéritos. Serão usados como critério de decisão: a acurácia de predição e a interpretabilidade do modelo.

As análises serão realizadas por meio dos pacotes `stats` (regressão logística), `rpart` (método CART, com medida de impureza de nós via índice de Gini) e `partykit` (método CTree) do programa estatístico R, versão 3.6.0.

1.8. Aspectos éticos

A pesquisa envolverá a utilização apenas de dados secundários já coletados nos inquéritos DHS e MICS. Em ambos os casos, as aprovações por comitês de ética já foram realizadas no nível nacional de acordo com as especificações e exigências

⁴ A regressão logística será especificada de acordo com modelo hierárquico de análise a ser construído com base no modelo conceitual apresentado na seção 4, considerando possíveis interações entre as variáveis e associações não lineares a serem determinadas na análise global e na análise de uma subamostra aleatória de 3 inquéritos.

de cada país, além dos critérios globais estabelecidos pelas instituições responsáveis pelos inquéritos. As informações (incluindo bancos de dados de posicionamento geográfico) são divulgadas de modo a manter a confidencialidade e não rastreabilidade dos respondentes.

Mais informações sobre os aspectos éticos de cada pesquisa podem ser encontradas nas respectivas páginas dos programas responsáveis pelos inquéritos: DHS (<https://www.dhsprogram.com/>) e MICS (<http://mics.unicef.org/>). Um exemplo de termo de consentimento utilizado no inquérito de 2015/2016 de Angola está apresentado no Anexo 1.

1.9. Divulgação dos resultados

A divulgação dos resultados do projeto será realizada por meio do volume final da dissertação de mestrado do qual este projeto faz parte, de publicação de artigo científico em revista apropriada da área e de um aplicativo web do tipo *Shiny app* a ser desenvolvido em linguagem R e publicado na plataforma shinyapps.io.

1.10. Financiamento

Este projeto é financiado pela Bill & Melinda Gates Foundation (número do grant: OPP1148933) e pelo Wellcome Trust (número do grant: 101815/Z/13/Z) por meio do Centro Internacional de Equidade em Saúde | Pelotas.

1.11. Cronograma

O cronograma proposto para execução do projeto de dissertação está apresentado na Figura 12.

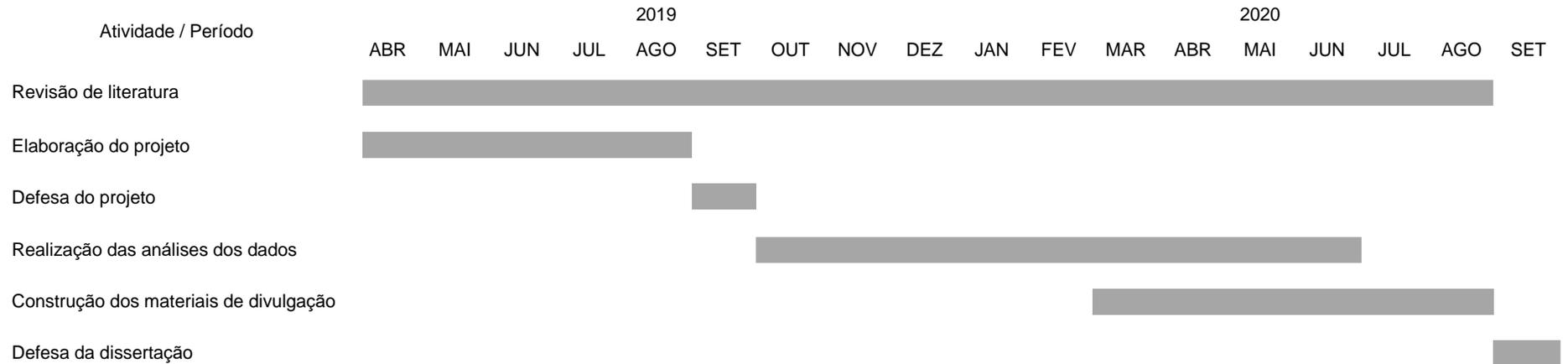


Figura 12 – Cronograma proposto do projeto de dissertação. A análise dos dados está descrita na seção 7.5. Os materiais de divulgação a serem criados estão apresentados na seção 9.

1.12. Referências

(UNICEF), N. B. O. S. N. A. U. N. C. S. F. **2017 Multiple Indicator Cluster Survey 2016-17, Survey Findings Report**. National Bureau of Statistics and United Nations Children's Fund. Abuja, Nigeria. 2017

ADEDINI, S. A. et al. Ethnic differentials in under-five mortality in Nigeria. **Ethnicity & health**, v. 20, n. 2, p. 145-162, 2015. ISSN 1465-3419 1355-7858. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24593689> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4337727/> >.

ANTAI, D. Faith and child survival: the role of religion in childhood immunization in Nigeria. **J Biosoc Sci**, v. 41, n. 1, p. 57-76, Jan 2009. ISSN 0021-9320.

ANTAI, D. Inequitable childhood immunization uptake in Nigeria: a multilevel analysis of individual and contextual determinants. **BMC infectious diseases**, v. 9, p. 181-181, 2009. ISSN 1471-2334. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/19930573> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2787508/> >.

BERTOLET, M.; BROOKS, M. M.; BITTNER, V. Tree-based identification of subgroups for time-varying covariate survival data. **Statistical Methods in Medical Research**, v. 25, n. 1, p. 488-501, Feb 2016. ISSN 0962-2802. Disponível em: < <Go to ISI>://WOS:000370685000029 >.

BETHONY, J. M. et al. Vaccines to combat the neglected tropical diseases. **Immunological reviews**, v. 239, n. 1, p. 237-270, 2011. ISSN 1600-065X 0105-2896. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21198676> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3438653/> >.

BHATIA, A. et al. **The Rohingya in Cox's Bazar: When the Stateless Seek Refuge**. 2018. 105-122.

BISSAU, G.-B. M. D. E. E. F. E. D. G. D. P.; (INE)., I. N. D. E. **Inquérito aos Indicadores Múltiplos (MICS5) 2014, Relatório Final**. 2014

BOSCH-CAPBLANCH, X.; BANERJEE, K.; BURTON, A. Unvaccinated children in years of increasing coverage: how many and who are they? Evidence from 96 low- and middle-income countries. **Trop Med Int Health**, v. 17, n. 6, p. 697-710, Jun 2012. ISSN 1360-2276.

CHIDO-AMAJUOYI, O. G. et al. Prevalence and correlates of never vaccinated Nigerian children, aged 1-5years. **Vaccine**, v. 36, n. 46, p. 6953-6960, Nov 12 2018. ISSN 0264-410x.

COFFEY, D. et al. Local Social Inequality, Economic Inequality, and Disparities in Child Height in India. **Demography**, 2019/07/15 2019. ISSN 1533-7790. Disponível em: < <https://doi.org/10.1007/s13524-019-00794-2> >.

CROFT, T. N., AILEEN M. J. MARSHALL, COURTNEY K. ALLEN ET AL. **Guide to DHS Statistics**. Rockville, Maryland, USA. 2018

DE LIMA PEREIRA, A. et al. Infectious Disease Risk and Vaccination in Northern Syria after 5 Years of Civil War: The MSF Experience. **PLoS currents**, v. 10, p. ecurrents.dis.bb5f22928e631dff9a80377309381feb, 2018. ISSN 2157-3999. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/29511602>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5815631/> >.

DRAEGER, E.; BEDFORD, H. E.; ELLIMAN, D. A. C. Should measles vaccination be compulsory? **BMJ**, v. 365, p. l2359, 2019. Disponível em: < <http://www.bmj.com/content/365/bmj.l2359.abstract> >.

FAVIN, M. et al. Why children are not vaccinated: a review of the grey literature. **International Health**, v. 4, n. 4, p. 229-238, 2012/12/01/ 2012. ISSN 1876-3413. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1876341312000496> >.

FEASEY, N. et al. Neglected tropical diseases. **British Medical Bulletin**, v. 93, n. 1, p. 179-200, 2009. ISSN 0007-1420. Disponível em: < <https://doi.org/10.1093/bmb/ldp046> >. Acesso em: 7/19/2019.

FLACH, P. **Machine Learning: The Art and Science of Algorithms that Make Sense of Data**. Cambridge: Cambridge University Press, 2012. ISBN 9781107096394. Disponível em: < <https://www.cambridge.org/core/books/machine-learning/621D3E616DF879E494B094CC93ED36A4> >.

FRANCIS, M. R. et al. Factors associated with routine childhood vaccine uptake and reasons for non-vaccination in India: 1998-2008. **Vaccine**, v. 36, n. 44, p. 6559-6566, Oct 22 2018. ISSN 0264-410x.

FRENCH, S. A. et al. Portion size effects on weight gain in a free living setting. **Obesity**, v. 22, n. 6, p. 1400-1405, 2014/06/01 2014. ISSN 1930-7381. Disponível em: < <https://doi.org/10.1002/oby.20720> >. Acesso em: 2019/08/01.

GAUDART, J. et al. SPODT: An R Package to Perform Spatial Partitioning. **Journal of Statistical Software**, v. 63, n. 16, p. 1-23, Jan 2015. ISSN 1548-7660. Disponível em: < <Go to ISI>://WOS:000349847100001 >.

HARDT, K. et al. Vaccine strategies: Optimising outcomes. **Vaccine**, v. 34, n. 52, p. 6691-6699, 2016/12/20/ 2016. ISSN 0264-410X. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0264410X16310301> >.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Springer, 2009. ISBN 9780387848846. Disponível em: < <https://books.google.com.br/books?id=eBSgoAEACAAJ> >.

HENRARD, S.; SPEYBROECK, N.; HERMANS, C. Classification and regression tree analysis vs. multivariable linear and logistic regression methods as statistical tools for

studying haemophilia. **Haemophilia**, v. 21, n. 6, p. 715-722, Nov 2015. ISSN 1351-8216. Disponível em: <<Go to ISI>://WOS:000363508900144 >.

HOTHORN, T.; HORNIK, K.; ZEILEIS, A. Unbiased Recursive Partitioning: A Conditional Inference Framework. **Journal of Computational and Graphical Statistics**, v. 15, n. 3, p. 651-674, 2006. ISSN 10618600. Disponível em: <<http://www.jstor.org/stable/27594202>>.

INSTITUTO NACIONAL DE ESTATÍSTICA, I. N. E. A.; MINISTÉRIO DA SAÚDE, M. A.; ICF. **Angola Inquérito de Indicadores Múltiplos e de Saúde (IIMS) 2015-2016**. INE, MINSA, and ICF. Luanda, Angola. 2017

INTERNATIONAL, I. **Demographic and Health Survey Sampling and Household Listing Manual**. Calverton, Maryland USA: ICF International 2012.

INTERNATIONAL INSTITUTE FOR POPULATION SCIENCES, I. I.; ICF. **India National Family Health Survey NFHS-4 2015-16**. IIPS and ICF. Mumbai, India. 2017

IRIART, J. A. Individual autonomy vs. collective protection: non-vaccination of children among higher income/educational level segments as a challenge to public health. **Cad Saude Publica**, v. 33, n. 2, p. e00012717, Mar 9 2017. ISSN 0102-311x.

JUSTICH, P. R. Refusal to have children vaccinated: A challenge to face. **Arch Argent Pediatr**, v. 113, n. 5, p. 443-8, Oct 2015. ISSN 0325-0075.

LARSON, H. et al. **The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey**. 2016.

MASTERS, N. B. et al. Childhood vaccination in Kenya: socioeconomic determinants and disparities among the Somali ethnic community. **Int J Public Health**, v. 64, n. 3, p. 313-322, Apr 2019. ISSN 1661-8556.

MCGAVIN, Z. A. et al. Childhood full and under-vaccination in Nigeria, 2013. **Vaccine**, v. 36, n. 48, p. 7294-7299, Nov 19 2018. ISSN 0264-410x.

MURPHY, K. P. **Machine Learning: A Probabilistic Perspective**. The MIT Press, 2012. 1096 ISBN 0262018020, 9780262018029.

NATIONAL INSTITUTE OF POPULATION STUDIES, N. P.; ICF. **Pakistan Demographic and Health Survey 2017-18**. NIPS/Pakistan and ICF. Islamabad, Pakistan. 2019

PATEL, N.; UPADHYAY, S. Study of Various Decision Tree Pruning Methods with their Empirical Comparison in WEKA. **Int. J. Comput. Appl.**, v. 60, p. 20-25, 12/18 2012.

PEELING, R. W.; BOERAS, D. I.; NKENGASONG, J. Re-imagining the future of diagnosis of Neglected Tropical Diseases. **Computational and structural biotechnology journal**, v. 15, p. 271-274, 2017. ISSN 2001-0370. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/28352456>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5358522/>>.

PERTET, A. M. et al. Under vaccination of children among Maasai nomadic pastoralists in Kenya: is the issue geographic mobility, social demographics or missed opportunities? **BMC Public Health**, v. 18, n. 1, p. 1389, 2018/12/19 2018. ISSN 1471-2458. Disponível em: < <https://doi.org/10.1186/s12889-018-6309-5> >.

RAINEY, J. J. et al. Reasons related to non-vaccination and under-vaccination of children in low and middle income countries: findings from a systematic review of the published literature, 1999-2009. **Vaccine**, v. 29, n. 46, p. 8215-21, Oct 26 2011. ISSN 0264-410x.

RAPPUOLI, R.; BLACK, S.; BLOOM, D. E. Vaccines and global health: In search of a sustainable model for vaccine development and delivery. **Science Translational Medicine**, v. 11, n. 497, p. eaaw2888, 2019. Disponível em: < <http://stm.sciencemag.org/content/11/497/eaaw2888.abstract> >.

ROGERS, K. M. Immunization Noncompliance: Gauging the Cause, Effect, and Management in the School Setting. **NASN Sch Nurse**, v. 34, n. 3, p. 140-144, May 2019. ISSN 1942-602x.

SCHWENDER, H. et al. A pilot study on the application of statistical classification procedures to molecular epidemiological data. **Toxicology Letters**, v. 151, n. 1, p. 291-299, Jun 2004. ISSN 0378-4274. Disponível em: < <Go to ISI>://WOS:000222226100033 >.

SHENTON, L. M. et al. Factors Associated with Vaccination Status of Children Aged 12-48 Months in India, 2012-2013. **Matern Child Health J**, v. 22, n. 3, p. 419-428, Mar 2018. ISSN 1092-7875.

SHENTON, L. M. et al. Vaccination status of children aged 1-4years in Afghanistan and associated factors, 2015. **Vaccine**, v. 36, n. 34, p. 5141-5149, Aug 16 2018. ISSN 0264-410x.

SHI, K. Q. et al. Classification and regression tree analysis of acute-on-chronic hepatitis B liver failure: Seeing the forest for the trees. **Journal of Viral Hepatitis**, v. 24, n. 2, p. 132-140, Feb 2017. ISSN 1352-0504. Disponível em: < <Go to ISI>://WOS:000393867500006 >.

SHRIVASTWA, N. et al. Predictors of vaccination in India for children aged 12-36 months. **Vaccine**, v. 33 Suppl 4, p. D99-105, Nov 27 2015. ISSN 0264-410x.

SMADJA, D. et al. Detection of Subclinical Keratoconus Using an Automated Decision Tree Classification. **American Journal of Ophthalmology**, v. 156, n. 2, p. 237-246, Aug 2013. ISSN 0002-9394. Disponível em: < <Go to ISI>://WOS:000322748200005 >.

STEINGLASS, R. **Routine immunization: An essential but wobbly platform.** 2013.

STREEFLAND, P.; CHOWDHURY, A. M.; RAMOS-JIMENEZ, P. Patterns of vaccination acceptance. **Soc Sci Med**, v. 49, n. 12, p. 1705-16, Dec 1999. ISSN 0277-9536 (Print)

0277-9536.

SUAREZ-CASTANEDA, E. et al. Routine childhood vaccination programme coverage, El Salvador, 2011-In search of timeliness. **Vaccine**, v. 32, n. 4, p. 437-44, Jan 16 2014. ISSN 0264-410x.

SUBAIYA, S. et al. Use of the revised World Health Organization cluster survey methodology to classify measles-rubella vaccination campaign coverage in 47 counties in Kenya, 2016. **PLoS One**, v. 13, n. 7, p. e0199786, 2018. ISSN 1932-6203.

TAKAHASHI, S. et al. Reduced vaccination and the risk of measles and other childhood infections post-Ebola. **Science**, v. 347, n. 6227, p. 1240-2, Mar 13 2015. ISSN 0036-8075.

THERNEAU, T., ATKINSON, B. An Introduction to Recursive Partitioning Using the RPART routines, Mayo Foundation. 2019 Disponível em < <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf> >

THERNEAU, T., ATKINSON, B., RIPLEY, B. Package 'rpart'. 2019 Disponível em < <https://cran.pau.edu.tr/web/packages/rpart/rpart.pdf> >

TSHIKUKA, J. G.; GUEIM, M.; DIALLO, P. M. Households that harbour zero-dose children after door-to-door national immunization campaigns. **Trop Doct**, v. 37, n. 4, p. 216-7, Oct 2007. ISSN 0049-4755 (Print) 0049-4755.

TUR-SINAI, A. et al. Vaccination uptake and income inequalities within a mass vaccination campaign. **Israel Journal of Health Policy Research**, v. 8, n. 1, p. 63, 2019/07/15 2019. ISSN 2045-4015. Disponível em: < <https://doi.org/10.1186/s13584-019-0324-6> >.

UNICEF. Statistics and Monitoring – Multiple Indicator Cluster Survey (MICS). 23 de outubro 2014. Disponível em: < https://www.unicef.org/statistics/index_24302.html >. Acesso em: 10 de julho de 2019.

UTHMAN, O. A. et al. Children who have received no routine polio vaccines in Nigeria: Who are they and where do they live? **Hum Vaccin Immunother**, v. 13, n. 9, p. 2111-2122, Sep 2 2017. ISSN 2164-5515.

VANDERENDE, K. et al. Global Routine Vaccination Coverage - 2017. **MMWR Morb Mortal Wkly Rep**, v. 67, n. 45, p. 1261-1264, Nov 16 2018. ISSN 0149-2195.

VANDERSLOTT, S. Exploring the meaning of pro-vaccine activism across two countries. **Social Science & Medicine**, v. 222, p. 59-66, 2019/02/01/ 2019. ISSN 0277-9536. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0277953618307160> >.

VENKATASUBRAMANIAM, A. et al. Decision trees in epidemiological research. **Emerging Themes in Epidemiology**, v. 14, p. 12, Sep 2017. ISSN 1742-7622. Disponível em: < <Go to ISI>://WOS:000411437000001 >.

VENKATASUBRAMANIAM, A. et al. Decision trees in epidemiological research. **Emerging themes in epidemiology**, v. 14, p. 11-11, 2017. ISSN 1742-7622. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28943885> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5607590/> >.

WHO, S. A. G. O. E. O. I. **2018 Assessment report of the Global Vaccine Action Plan**. World Health Organization. Geneva, Switzerland. 2018

WHO, U., WORLD BANK. **State of the world's vaccines and immunization**. 3rd edition. Geneva: World Health Organization, 2009. ISBN 978 92 4 156386 4

WILLKE, R. J. et al. From concepts, theory, and evidence of heterogeneity of treatment effects to methodological approaches: a primer. **Bmc Medical Research Methodology**, v. 12, p. 12, Dec 2012. ISSN 1471-2288. Disponível em: < <Go to ISI>://WOS:000314130000001 >.

WOLFSON, J.; VENKATASUBRAMANIAM, A. Branching Out: Use of Decision Trees in Epidemiology. **Current Epidemiology Reports**, v. 5, n. 3, p. 221-229, 2018/09/01 2018. ISSN 2196-2995. Disponível em: < <https://doi.org/10.1007/s40471-018-0163-y> >.

YANG, Y. T.; SILVERMAN, R. D. Social distancing and the unvaccinated. **N Engl J Med**, v. 372, n. 16, p. 1481-3, Apr 16 2015. ISSN 0028-4793.

1.13. Anexos

1.13.1. Anexo 1

Reprodução da seção de imunização dos questionários DHS e MICS que contém as perguntas utilizadas na construção da variável de desfecho (dose zero) do projeto. Os exemplos das Figuras 13 e 14 pertencem, respectivamente, ao inquérito DHS de Angola 2015-2016 (Instituto Nacional De Estatística *et al.*, 2017) e ao inquérito MICS de Guiné-Bissau 2014 (Bissau e (Ine). 2014).

SECÇÃO 5A. IMUNIZAÇÃO (ÚLTIMA CRIANÇA)			
NO.	PERGUNTAS E FILTROS	CODIFICAÇÃO DAS CATEGORIAS	PASSE A
501A	VERIFIQUE 215 NO HISTORIAL DE NASCIMENTOS: ALGUM NASCIMENTO DESDE JANEIRO DE 2012? UM NASCIMENTO OU MAIS <input type="checkbox"/> DESDE JANEIRO DE 2012 ↓	NENHUM NASCIMENTO <input type="checkbox"/> DESDE JANEIRO DE 2012	→ 601
502A	REGISTE O NOME E NÚMERO DE ORDEM DE NASCIMENTO DA ÚLTIMA CRIANÇA NASCIDA DESDE JANEIRO DE 2012 NA PERGUNTA 212. NOME DA ÚLTIMA CRIANÇA _____	Nº DE ORDEM DE NASCIMENTO <input type="text"/> <input type="text"/>	
503A	VERIFIQUE 216 PARA A CRIANÇA: VIVA <input type="checkbox"/> ↓	FALECIDA <input type="checkbox"/>	→ 501B
504A	O(A) (NOME DA CRIANÇA) tem o cartão de vacina ou algum outro documento que regista as vacinas que apanhou?	SIM, SÓ TEM CARTÃO 1 SIM, SÓ TEM OUTRO DOCUMENTO 2 SIM, TEM CARTÃO E OUTRO DOCUMENTO ... 3 NÃO, NEM CARTÃO NEM DOCUMENTO 4	→ 507A → 507A
505A	Em algum momento o(a) (NOME DA CRIANÇA) teve o cartão de vacina?	SIM 1 NÃO 2	
506A	VERIFIQUE 504A: CÓDIGO '2' MARCADO <input type="checkbox"/> ↓	CÓDIGO '4' MARCADO <input type="checkbox"/>	→ 511A
507A	Por favor, posso ver o cartão ou documento que regista as vacinas que o (a) (NOME DA CRIANÇA) apanhou?	SIM, SÓ VIU O CARTÃO 1 SIM, SÓ VIU OUTRO DOCUMENTO 2 SIM, VIU O CARTÃO E OUTRO DOCUMENTO ... 3 NÃO, NEM CARTÃO NEM DOCUMENTO 4	→ 511A

Figura 13.A – Reprodução da seção de imunização do inquérito DHS – Angola 2015-2016.

SECÇÃO 5A. IMUNIZAÇÃO (ÚLTIMA CRIANÇA)

NO.	PERGUNTAS E FILTROS	CODIFICAÇÃO DE CATEGORIAS	PASSE A																																																																														
	NOME DA ÚLTIMA CRIANÇA _____	Nº DE ORDEM DE NASCIMENTO <input type="text"/> <input type="text"/>																																																																															
508A	<p>COPIE AS DATAS DE VACINAÇÃO DO CARTÃO. REGISTE '44' NA COLUNA DE 'DIA' SE O CARTÃO INDICA QUE A CRIANÇA RECEBEU UMA DOSE, MAS A DATA NÃO FOI REGISTRADA.</p> <table border="1"> <thead> <tr> <th></th> <th>DIA</th> <th>MÊS</th> <th>ANO</th> <th>DIA</th> <th>MÊS</th> <th>ANO</th> <th>DIA</th> <th>MÊS</th> <th>ANO</th> <th>DIA</th> <th>MÊS</th> <th>ANO</th> </tr> </thead> <tbody> <tr> <td>AO NASCER</td> <td colspan="3">PÓLIO</td> <td colspan="3">BCG</td> <td colspan="3">HEPATITE B</td> <td colspan="3" rowspan="2"></td> </tr> <tr> <td>AOS 2 MESES</td> <td colspan="3">PÓLIO 1ª DOSE</td> <td colspan="3">PENTAVALENTE 1ª DOSE</td> <td colspan="3">PNEUMO 1ª DOSE</td> <td colspan="3">ROTAVÍRUS 1ª DOSE</td> </tr> <tr> <td>AOS 4 MESES</td> <td colspan="3">PÓLIO 2ª DOSE</td> <td colspan="3">PENTAVALENTE 2ª DOSE</td> <td colspan="3">PNEUMO 2ª DOSE</td> <td colspan="3">ROTAVÍRUS 2ª DOSE</td> </tr> <tr> <td>AOS 6 MESES</td> <td colspan="3">PÓLIO 3ª DOSE</td> <td colspan="3">PENTAVALENTE 3ª DOSE</td> <td colspan="3">PNEUMO 3ª DOSE</td> <td colspan="3">VITAMINA A 1ª DOSE</td> </tr> <tr> <td>AOS 9 MESES (15 MESES PARA SARAMPO)</td> <td colspan="3">SARAMPO 1ª DOSE</td> <td colspan="3">FEBRE AMARELA DOSE ÚNICA</td> <td colspan="3">VITAMINA A 2ª DOSE</td> <td colspan="3">SARAMPO 2ª DOSE</td> </tr> </tbody> </table>		DIA	MÊS	ANO	DIA	MÊS	ANO	DIA	MÊS	ANO	DIA	MÊS	ANO	AO NASCER	PÓLIO			BCG			HEPATITE B						AOS 2 MESES	PÓLIO 1ª DOSE			PENTAVALENTE 1ª DOSE			PNEUMO 1ª DOSE			ROTAVÍRUS 1ª DOSE			AOS 4 MESES	PÓLIO 2ª DOSE			PENTAVALENTE 2ª DOSE			PNEUMO 2ª DOSE			ROTAVÍRUS 2ª DOSE			AOS 6 MESES	PÓLIO 3ª DOSE			PENTAVALENTE 3ª DOSE			PNEUMO 3ª DOSE			VITAMINA A 1ª DOSE			AOS 9 MESES (15 MESES PARA SARAMPO)	SARAMPO 1ª DOSE			FEBRE AMARELA DOSE ÚNICA			VITAMINA A 2ª DOSE			SARAMPO 2ª DOSE				
	DIA	MÊS	ANO	DIA	MÊS	ANO	DIA	MÊS	ANO	DIA	MÊS	ANO																																																																					
AO NASCER	PÓLIO			BCG			HEPATITE B																																																																										
AOS 2 MESES	PÓLIO 1ª DOSE			PENTAVALENTE 1ª DOSE			PNEUMO 1ª DOSE						ROTAVÍRUS 1ª DOSE																																																																				
AOS 4 MESES	PÓLIO 2ª DOSE			PENTAVALENTE 2ª DOSE			PNEUMO 2ª DOSE			ROTAVÍRUS 2ª DOSE																																																																							
AOS 6 MESES	PÓLIO 3ª DOSE			PENTAVALENTE 3ª DOSE			PNEUMO 3ª DOSE			VITAMINA A 1ª DOSE																																																																							
AOS 9 MESES (15 MESES PARA SARAMPO)	SARAMPO 1ª DOSE			FEBRE AMARELA DOSE ÚNICA			VITAMINA A 2ª DOSE			SARAMPO 2ª DOSE																																																																							
509A	<p>VERIFIQUE 508A: TODAS AS VACINAS, DE 'POLIO' ATÉ 'SARAMPO 2ª DOSE' FORAM REGISTRADAS?</p> <p>NÃO <input type="checkbox"/></p> <p>SIM <input type="checkbox"/></p>		→ 526A																																																																														
510A	<p>O(A) (NOME DA CRIANÇA) apanhou alguma vacina que não consta (cartão/documento), incluindo vacinas recebidas em campanhas de vacinação?</p> <p>REGISTE 'SIM' SÓ SE A INQUIRIDA MENCIONAR PELO MENOS UMA DAS VACINAS EM 508A QUE NÃO FOI REGISTRADA.</p>	<p>SIM 1 (INDAGUE PARA IDENTIFICAR AS VACINAS E REGISTE '66' NO COLUNA DE DIA QUE CORRESPONDE ÀS VACINAS EM 508A)</p> <p>(DEPOIS, PASSE A 526A)</p> <p>NÃO 2 NÃO SABE 8</p>	→ 526A																																																																														
511A	<p>O(A) (NOME DA CRIANÇA) apanhou alguma vacina para prevenir doenças, incluindo vacinas recebidas em campanhas de vacinação ou nos dias de saúde infantil?</p>	<p>SIM 1 NÃO 2 NÃO SABE 8</p>	→ 526A																																																																														

Figura 13.B – Reprodução da seção de imunização do inquérito DHS – Angola 2015-2016.

SECÇÃO 5A. IMUNIZAÇÃO (ÚLTIMA CRIANÇA)

NO.	PERGUNTAS E FILTROS	CODIFICAÇÃO DE CATEGORIAS	PASSE A
	NOME DA ÚLTIMA CRIANÇA _____	Nº DE ORDEM DE NASCIMENTO <input type="text"/> <input type="text"/>	
512A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina BCG contra a tuberculose, isto é, uma injeção no braço ou ombro que geralmente deixa uma cicatriz?	SIM 1 NÃO 2 NÃO SABE 8	
513A	Nas primeiras 24 horas depois do nascimento, o(a) (NOME DA CRIANÇA) apanhou a vacina contra a Hepatite B, isto é, uma injeção na coxa para prevenir Hepatite B?	SIM 1 NÃO 2 NÃO SABE 8	
514A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina oral contra a pólio, isto é, duas gotas na boca para prevenir pólio?	SIM 1 NÃO 2 NÃO SABE 8	→ 517A
515A	O(A) (NOME DA CRIANÇA) apanhou a primeira vacina oral contra a pólio nas primeiras duas semanas depois do parto ou mais tarde?	NAS PRIMEIRAS DUAS SEMANAS 1 MAIS TARDE 2	
516A	Quantas vezes o(a) (NOME DA CRIANÇA) apanhou a vacina oral contra a pólio?	NÚMERO DE VEZES <input type="text"/>	
517A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina pentavalente, isto é, uma injeção que se toma na coxa ao mesmo tempo que as gotas de pólio?	SIM 1 NÃO 2 NÃO SABE 8	→ 519A
518A	Quantas vezes o(a) (NOME DA CRIANÇA) apanhou a vacina pentavalente?	NÚMERO DE VEZES <input type="text"/>	
519A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina pneumocócica, isto é, uma injeção que se toma na coxa para prevenir a pneumonia?	SIM 1 NÃO 2 NÃO SABE 8	→ 521A
520A	Quantas vezes o(a) (NOME DA CRIANÇA) apanhou a vacina pneumocócica?	NÚMERO DE VEZES <input type="text"/>	
521A	O (A) (NOME DA CRIANÇA) alguma vez apanhou a vacina contra rotavírus, isto é, um líquido na boca para prevenir a diarreia?	SIM 1 NÃO 2 NÃO SABE 8	→ 523A
522A	Quantas vezes o(a) (NOME DA CRIANÇA) apanhou a vacina rotavírus?	NÚMERO DE VEZES <input type="text"/>	
523A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina contra o sarampo, isto é, uma injeção no braço para prevenir o sarampo?	SIM 1 NÃO 2 NÃO SABE 8	→ 526A
524A	Quantas vezes o(a) (NOME DA CRIANÇA) apanhou a vacina contra o sarampo?	NÚMERO DE VEZES <input type="text"/>	
525A	O(A) (NOME DA CRIANÇA) alguma vez apanhou a vacina contra a febre amarela?	SIM 1 NÃO 2 NÃO SABE 8	
526A	Nos últimos 7 dias, o(a) (NOME DA CRIANÇA) tomou: a) Sulfato ferroso como este(s) aqui (MOSTRAR IMAGENS DE AMPOLAS OU COMPRIMIDOS)? b) Comprimidos para desparasitação (MOSTRAR IMAGENS DOS COMPRIMIDOS)? c) Algum suplemento nutricional com fórmula integral?	SIM NÃO NS a) SULFATO FERROSO 1 2 8 b) DESPARASITAÇÃO 1 2 8 c) SUPL. NUTRICIONAL 1 2 8	
527A	CONTINUE COM 501B.		

Figura 13.C – Reprodução da seção de imunização do inquérito DHS – Angola 2015-2016.

VACINAÇÃO		IM		
Se existe um cartão de vacina disponível, copie as datas no IM3 para cada tipo de vacina e vitamina A marcada no cartão. IM6-IM17 são feitas somente quando não existe cartão disponível.				
IM1. EXISTE UM CARTÃO DE VACINAS PARA (NOME) ONDE ESTÃO REGISTRADAS TODAS AS VACINAS? SE SIM: POSSO VER O CARTÃO, POR FAVOR?		Sim, vi	1	1⇒IM3
		Sim, não vi	2	2⇒IM6
		Não tem cartão	3	
IM2. TEVE CARTÃO DE VACINA PARA (NOME)?		Sim	1	1⇒IM6
		Não	2	2⇒IM6
IM3. a) Copiar as datas das vacinas para cada vacina a partir do cartão ou ficha. b) Escreva '44' na coluna dia se o cartão indicar que a vacina foi feita mas que a data não foi marcada.		Data de Vacinação		
		Dia	Mês	Ano
BCG	BCG			
POLIO AO NASCER	OPV0			
POLIO 1	OPV1			
POLIO 2	OPV2			
POLIO 3	OPV3			
PENTA 1	PENTA 1			
PENTA 2	PENTA 2			
PENTA 3	PENTA 3			
SARAMPO	SARAMPO			
FEBRE AMARELA	YF			
VITAMINA A (PRIMEIRA DOSE)	VIT A1			
VITAMINA A (SEGUNDA DOSE)	VIT A2			
IM4. VERIFIQUE IM3. SERÁ QUE TODAS AS VACINAS (DE BCG À FEBRE AMARELA) E AS VITAMINAS FORAM REGISTRADAS ?				
<input type="checkbox"/> Sim ⇒ Siga para IM19				
<input type="checkbox"/> Não ⇒ Continue com IM5				
IM5. ALÉM DESTAS VACINAS REGISTRADAS NO CARTÃO, O/A (NOME) RECEBEU OUTRAS VACINAS, INCLUINDO AS VACINAS RECIBIDAS DURANTE AS CAMPANHAS DE VACINAÇÕES?				
<input type="checkbox"/> Sim ⇒ Volte para IM3 e insista sobre o tipo de vacina e escreva '66' na coluna de dia correspondente para cada vacina mencionada. Quando terminar siga para IM19				
<input type="checkbox"/> Não/NS ⇒ siga para IM19				

Figura 14.A – Reprodução da seção de imunização do inquérito MICS – Guiné-Bissau 2014.

<p>IM6. SERÁ QUE O/A (NOME) RECEBEU VACINAS QUE EVITAM A CONTAMINAÇÃO DE DOENÇAS, INCLUINDO VACINAS RECEBIDAS NA CAMPANHA NACIONAL DE VACINAÇÃO?</p>	<p>Sim.....1 Não.....2 NS8</p>	<p>2⇒IM19 8⇒IM19</p>
<p>IM7. SERÁ QUE O/A (NOME) RECEBEU VACINA BCG CONTRA A TUBERCULOSE, QUER DIZER UMA INJEÇÃO NO BRAÇO QUE DEIXA, GERALMENTE UMA CICATRIZ?</p>	<p>Sim.....1 Não.....2 NS8</p>	
<p>IM8. SERÁ QUE O/A (NOME) RECEBEU VACINA SOBRE A FORMA DE GOTAS NA BOCA PARA O/A PROTEGER CONTRA DOENÇAS – COMO A PÓLIO?</p>	<p>Sim.....1 Não.....2 NS8</p>	<p>2⇒IM11 8⇒IM11</p>
<p>IM9. SERÁ QUE A PRIMEIRA VACINA CONTRA A PÓLIO FOI RECEBIDA NAS DUAS PRIMEIRAS SEMANAS APÓS O NASCIMENTO?</p>	<p>Sim.....1 Não.....2</p>	
<p>IM10. QUANTAS VEZES RECEBEU A VACINA CONTRA A PÓLIO?</p>	<p>Número de vezes</p>	
<p>IM11A. SERÁ QUE O/A (NOME) RECEBEU VACINA DE PENTAVALENTE, QUER DIZER UMA INJEÇÃO DADA NO BRAÇO ESQUERDO PARA EVITAR DE CONTRAIR O TETANO, A TOSSE CONVULSA, A DIFTERIA, A HEPATITE B E A HAEMOPHILUS INFLUENZA DE TIPO B? REFORCE QUE A VACINA DE PENTA É ADMINISTRADA MUITAS VEZES AO MESMO TEMPO QUE A PÓLIO</p>	<p>Sim.....1 Não.....2 NS8</p>	<p>2⇒IM16 8⇒IM16</p>
<p>IM12A. QUANTAS VEZES RECEBERAM A VACINA DE PENTA?</p>	<p>Numero de vezes</p>	
<p>IM16. SERÁ QUE O/A (NOME) RECEBEU VACINA CONTRA SARAMPO – INJEÇÃO DADA NAS COSTAS NA IDADE DE 9 MESES OU MAIS PARA EVITAR APANHAR SARAMPO?</p>	<p>Sim.....1 Não.....2 NS8</p>	
<p>IM17. SERÁ QUE O/A (NOME) RECEBEU VACINA CONTRA FEBRE-AMARELA: QUER DIZER UMA INJEÇÃO QUE É DADA NO BRAÇO NA IDADE DE 9 MESES OU MAIS PARA PREVENIR CONTRA FEBRE-AMARELA. FEITA ALGUMAS VEZES AO MESMO TEMPO QUE A VACINA DO SARAMPO.</p>	<p>Sim.....1 Não.....2 NS8</p>	
<p>IM19. DIZ-ME POR FAVOR, SE (NOME) PARTICIPOU NUMA JORNADA NACIONAL DAS SEGUINTE CAMPANHAS DE VACINAÇÃO E/OU DA VITAMINA A OU JORNADA DE SAÚDE DA CRIANÇA;</p> <p>[A] DEZEMBRO 2012, CAMPANHA VACINAÇÃO CONTRA O SARAMPO E DE VITAMINA A E MEBENDAZOL</p> <p>[B] MAIO 2013, CAMPANHA DE VACINAÇÃO ANTIPOLIO E DE VITAMINA A E MEBENDAZOL</p> <p>[C] NOVEMBRO 2013, CAMPANHA DE VACINAÇÃO ANTIPOLIO E DE VITAMINA A E MEBENDAZOL</p>	<p>S N NS</p> <p>Campanha Dez. 2012.....1 2 8</p> <p>Campanha Maio 20131 2 8</p> <p>Campanha Nov. 2013.....1 2 8</p>	

Figura 14.B – Reprodução da seção de imunização do inquérito MICS – Guiné-Bissau 2014.

1.13.2. Anexo 2

Modelos de formulário de consentimento utilizado na aplicação do inquérito DHS em Angola 2015-2016 (Instituto Nacional De Estatística *et al.*, 2017) e do inquérito MICS de Guiné-Bissau 2014 (Bissau e (Ine). 2014).

APRESENTAÇÃO E CONSENTIMENTO

Bom dia/boa tarde. O meu nome é _____. Sou Inquiridor(a) do Instituto Nacional de Estatística e a minha identificação é esta (MOSTRAR CARTÃO). Estamos a realizar um inquérito nacional sobre vários aspectos de saúde. A informação recolhida através deste inquérito vai apoiar o governo na planificação e na melhoria dos serviços de saúde. O seu agregado familiar foi seleccionado para o inquérito. Todas as respostas serão confidenciais e não serão partilhadas com mais ninguém, além dos membros da equipa do inquérito.

A sua participação neste inquérito é voluntária e se tiver qualquer pergunta que não queira responder pode nos dizer e passaremos para a pergunta seguinte. Pode interromper a entrevista a qualquer momento. Contudo, esperamos que participe no inquérito já que suas respostas são muito importantes. Em caso de precisar mais informação sobre o inquérito, pode contactar ao INE ou os Serviços Provinciais do INE.

Tem alguma pergunta?
Posso iniciar a entrevista?

ASSINATURA DO INQUIRIDOR _____ DATA _____

O INQUIRIDO ACEITA
SER ENTREVISTADO . . 1
↓

O INQUIRIDO NÃO ACEITA
SER ENTREVISTADO . . 2 → FIM

Figura 15 – Reprodução do formulário de consentimento do inquérito DHS – Angola 2015-2016.

NÓS FAZEMOS PARTE DO INSTITUTO NACIONAL DE ESTATÍSTICAS. ESTAMOS A REALIZAR UM INQUÉRITO SOBRE A SITUAÇÃO DAS CRIANÇAS, DAS FAMÍLIAS E DOS AGREGADOS. GOSTARÍAMOS DE FALAR CONSIGO SOBRE ESTES ASPECTOS. A NOSSA CONVERSA TOMARÁ **75** MINUTOS. TODAS INFORMAÇÕES QUE SERÃO PRESTADAS SÃO ESTRITAMENTE CONFIDENCIAIS E ANÓNIMAS. PODEMOS COMEÇAR AGORA?

SIM, PERMISSÃO CONCEDIDA ⇨ VA À **HH18** PARA REGISTAR A HORA E COMEÇAR A ENTREVISTA.

NÃO, PERMISSÃO NÃO CONCEDIDA ⇨ **CIRCULE '04'** EM HH9. DISCUZA ESTE RESULTADO COM SEU CHEFE DE EQUIPA.

HH9. Resultado da entrevista do agregado familiar:

Completa	01
Não havia membros no alojamento com competência para responder ao questionário	02
Membros do agregado totalmente ausentes por longa duração	03
Recusa	04
Alojamento desocupado/Morada não é alojamento	05
Alojamento destruído	06
Alojamento não encontrado	07
Outro (especificar)	96

Figura 16 – Reprodução do formulário de consentimento do inquérito MICS – Guiné-Bissau 2014.

2. MODIFICAÇÕES NO PROJETO

Foram realizadas as seguintes modificações em relação ao Projeto de Pesquisa original:

- A faixa etária das crianças estudadas foi restrita de 12 a 59 meses para 12 a 23 meses. Esta modificação foi realizada de modo a reduzir a quantidade de informações faltantes nos bancos, uma vez que as perguntas relacionadas ao pré-natal eram exclusivas para a última criança nascida viva. Em países com alta paridade e pequenos intervalos entre os nascimentos, as últimas crianças nascidas vivas tendem a ser mais novas, e a faixa etária entre 23 e 59 meses tende a apresentar maior quantidade de informação faltante;
- A lista de inquéritos investigados foi atualizada em julho de 2020. Foram selecionados os inquéritos MICS e DHS mais recentes, nacionalmente representativos, publicados a partir de 2010 e referentes a países de baixa e média renda de acordo com a classificação do Banco Mundial no ano mediano dos inquéritos (2015), resultando em 92 países;
- Foram selecionados apenas os inquéritos mais recentes de modo a 1) ter apenas um inquérito por país na amostra que seria utilizada na análise agregada e 2) garantir a viabilidade da condução da análise em tempo hábil, uma vez que a harmonização dos bancos de dados mostrou-se mais demorada do que inicialmente previsto;
- A definição operacional do desfecho (zero dose) foi atualizada para estar alinhada com uma série de análises desenvolvidas pelo Centro Internacional de Equidade em Saúde | Pelotas para a GAVI, seguindo recomendações da OMS no tratamento de informações faltantes de vacinação. Crianças sem informação para uma determinada vacina foram classificadas como não tendo recebido esta vacina;
- Os indicadores de região subnacional e grupo étnico foram removidos da análise. Os dois indicadores apresentavam um número elevado de categorias em diversos inquéritos. Somado ao restrito tamanho amostral na nova faixa etária selecionada, muitas destas categorias apresentavam poucas crianças, tornando a análise propensa a sobreajuste;
- Os indicadores de ocupação dos pais, educação do parceiro e ordem de nascimento foram removidos da análise, uma vez que não estavam disponíveis nos inquéritos MICS;
- As regressões logísticas planejadas não foram realizadas. À medida em que evoluiu o conhecimento e entendimento dos autores sobre as árvores de decisão, ficou clara

uma diferença fundamental entre os dois métodos: enquanto as árvores trabalham na identificação de grupos de risco, as regressões são marcadamente voltadas para a identificação de fatores de risco. As árvores lidam naturalmente com complexas interações entre os indicadores investigados, criando subgrupos amostrais que não seriam avaliados por uma técnica de regressão logística. Entretanto, as árvores não controlam para efeitos dos outros indicadores, de modo que os grupos criados podem ser muito distintos em relação a indicadores não apresentados na árvore. Desse modo, a comparação com o resultado de uma regressão logística não seria particularmente informativa e não foi realizada;

- Apenas o método CART foi apresentado no artigo original. Ambos os métodos CART e Ctree foram investigados e o método CART mostrou-se muito superior ao Ctree no contexto de inquéritos como DHS e MICS, que permitem a análise de centenas ou milhares de crianças por país. Na presença de tamanhos amostrais dessa ordem de grandeza, o método Ctree produziu árvores com um número de nós que as tornavam ininteligíveis. Este resultado era esperado, uma vez que o método Ctree se baseia no conceito de significância estatística para determinar se uma divisão deve ou não ocorrer. Com um tamanho amostral grande o suficiente, o número de divisões (e nós) cresce rapidamente, sem que isso represente uma real característica do problema investigado. Portanto, apenas as análises do método CART foram apresentadas;
- Os resultados nacionais foram apresentados na forma de perfis individuais incluídos nos materiais suplementares do artigo original, ao invés de serem apresentados em aplicativo Shiny, como inicialmente planejado. As árvores atualmente produzidas pelo pacote rpart (e que seriam apresentadas no aplicativo Shiny) são de difícil interpretação e versões manualmente editadas dos perfis foram criadas;
- O texto da seção 1.3.2.3 “Sobreajuste e Poda” do Projeto de Pesquisa foi alterado para melhor esclarecimento de como o pacote rpart determina o valor do parâmetro de complexidade utilizado na construção de uma árvore. Particularmente, foram alterados o terceiro parágrafo da página 40 e a página 41;

3. RELATÓRIO DE TRABALHO DE CAMPO

3.1. Introdução

O Programa de Pós-graduação em Epidemiologia (PPGEpi) da UFPel foi criado no ano de 1991, a partir de um trabalho em conjunto de docentes do Departamento de Medicina Social (DMS) da universidade. Desde 1999, o PPGEpi realiza, bianualmente, uma estratégia pioneira de pesquisa chamada de “Consórcio de Pesquisa”, no qual um estudo transversal de base populacional é realizado em Pelotas, no sul do Rio Grande do Sul (BARROS *et al.*, 2008), com participação dos alunos de mestrado do PPGEpi.

A utilização da metodologia de consórcio de pesquisa apresenta algumas vantagens, tais como: a possibilidade de ampliação do tamanho amostral, dando origem a um estudo de base populacional; menor tempo e gasto financeiro individual na execução do trabalho de campo; oportunidade de desenvolvimento de um ambiente de criação e reflexão coletiva entre os mestrandos, com trocas de experiências e conhecimentos, possibilitando que projetos de diferentes áreas da saúde possam estar interligados.

As turmas de mestrado que seguiram o formato de consórcio de pesquisa estudaram diferentes temas e populações. A população-alvo dos consórcios iniciais era composta por residentes na zona urbana do município de Pelotas. Porém, desde 2013, as turmas de mestrado vêm diferindo em relação a este padrão e estudando outras populações, como moradores da zona rural, idosos e estudantes universitários. A turma atual de mestrandos, que ingressou em março de 2019 no PPGEpi propôs o retorno da realização do consórcio de pesquisa na zona urbana do município, visando possibilitar comparações com os achados de outros estudos com a mesma população e também a realização de novas avaliações temáticas na área da saúde.

O consórcio na zona urbana do município de Pelotas, RS, foi realizado por dez mestrandos que integram a turma 2019/2020 do PPGEpi, com participação e orientação do corpo docente do programa e teve como população-alvo indivíduos residentes na zona urbana deste município, com 18 anos ou mais.

Ao longo dos quatro primeiros bimestres do curso de mestrado, nas disciplinas de Prática de Pesquisa I a IV, ocorreu o planejamento de todo o trabalho de campo do estudo pelos mestrandos e docentes, desde a escolha dos temas até o planejamento

logístico da coletas de dados. Nessa pesquisa foram investigados os temas específicos de cada mestrando, conforme descrito na Tabela 1.

Tabela 1 – Mestrandos, orientadores e temas do Consórcio de Pesquisa do PPGEpi. Pelotas 2019/2020

Mestrando(a)	Orientador(a)	Tema
Anna Muller Pereira	Mariângela Freitas da Silveira	Consumo de alimentos ultraprocessados e fatores associados na população adulta da zona urbana da cidade de Pelotas/RS
Barbara Berrutti	Luciana Tovo Rodrigues	Prevalência de insônia e fatores associados em adultos da zona urbana da cidade de Pelotas, RS
Bruna Venturin	Luiz Augusto Facchini	Percepção de apoio social em adultos residentes na zona urbana de Pelotas, RS
Eloisa Porciúncula da Silva	Denise Petrucci Gigante	Reavaliação da insegurança alimentar nos domicílios de Pelotas-RS em intervalo de 13 anos
Eveline Bordignon	Joseph Murray	Vitimização por violência urbana entre adultos: tendências temporais e fatores associados
Gabriela Ávila Marques	Fernando César Wehrmeister	Uso de inaladores dosimetrados em adultos com doença respiratória crônica: evolução temporal em residentes da zona urbana de Pelotas 2012 a 2020
Lucas Gonçalves de Oliveira	Helen Gonçalves	Crença em um mundo justo e fatores associados em adultos de uma cidade de médio porte ao sul do Brasil
Marina de Borba Oliveira	Ana Maria Baptista Menezes	Prevalência e utilização de benzodiazepínicos em adultos da zona urbana de Pelotas, RS
Paulo Victor Cesar de Albuquerque	Elaine Tomasi	Autopercepção de discriminação em serviços de saúde entre adultos da zona urbana de Pelotas, RS
Rafaela do Carmo Borges	Flávio Fernando Demarco	Utilização de serviços odontológicos no último ano na área urbana de Pelotas

Por meio dos projetos individuais de cada mestrando, foi elaborado um projeto geral intitulado “Avaliação da saúde de adultos residentes na zona urbana do município de Pelotas, RS”. Este projeto mais amplo contemplou o delineamento do estudo, os objetivos e as justificativas de todos os temas de pesquisa dos mestrandos, além da metodologia, processo de amostragem e outras características da execução do estudo.

O projeto geral foi encaminhado para avaliação do Comitê de Ética em Pesquisa (CEP) da Faculdade de Medicina (FAMED) da UFPel. Em novembro de 2019, recebeu a aprovação com o número de protocolo 3.676.549. O parecer contendo a aprovação para o estudo encontra-se no Anexo I.

O presente relatório descreve o processo de construção desse estudo.

3.2. Comissões de trabalho de campo

O Consórcio de Pesquisa busca, além dos resultados específicos de cada projeto, capacitar os mestrandos para o trabalho de campo em equipe. Para que isso fosse possível, foram estabelecidas comissões a fim de garantir agilidade, melhor distribuição de tarefas e bom andamento do trabalho de campo.

Todos os mestrandos participaram de comissões, podendo um mesmo aluno atuar em mais de uma comissão. Ainda, este consórcio contou com a colaboração de alunos vinculados ao Centro Internacional de Equidade em Saúde (ICEH) do Centro de Pesquisas Epidemiológicas (Daniel Leventhal e eu), participando das comissões e do trabalho de campo durante os quatro primeiros meses do estudo. Os projetos de dissertação dos alunos do ICEH não previam a utilização dos dados coletados pelo consórcio.

As atividades relacionadas a cada comissão e seus responsáveis estão descritas a seguir.

3.2.1. Elaboração do projeto de pesquisa

Os responsáveis pela elaboração do projeto geral foram as mestrandas Bruna Venturin, Eloisa Porciúncula e Marina de Borba Oliveira. A equipe reuniu justificativas, objetivos gerais e específicos e hipóteses dos projetos individuais dos dez mestrandos na composição de um único documento sobre o estudo, denominado "projetão".

O "projetão" também contemplou aspectos comuns a todos os mestrandos, como: descrição do PPGEpi e da forma de pesquisa adotada pelo programa, delineamento do estudo, população-alvo, amostra e processo de amostragem, instrumentos utilizados, logística, estudo pré-piloto e piloto, processamento e análise de dados, aspectos éticos, orçamento, cronograma e referências bibliográficas.

3.2.2. Elaboração do questionário e manual de instruções

Os responsáveis por esta comissão foram os mestrandos Gabriela Marques, Lucas Oliveira e eu. A equipe elaborou um instrumento único contendo as perguntas de cada mestrando de maneira padronizada e um manual de instruções com todas as

informações sobre o instrumento geral, bem como procedimentos a serem tomados em cada pergunta.

A versão digital do questionário foi inserida no Research Eletronic Data Capture (RedCap) pelos mestrandos responsáveis pelo banco de dados (Lucas Oliveira e eu).

3.2.3. Gestão do banco de dados

Os responsáveis por essa comissão foram os mestrandos Lucas Oliveira e eu. A comissão ficou responsável pelo desenvolvimento da versão eletrônica do questionário, pela sua inserção na plataforma RedCap, pela instalação do aplicativo em todos os equipamentos e pela atualização de todos os tablets.

Essa comissão ficou encarregada ainda de todos os processos da gestão do banco de dados, incluindo o descarregamento semanal dos questionários já preenchidos, reparo de possíveis erros ao longo do trabalho de campo, limpeza do banco e correções que porventura fossem necessárias após a realização de checagem de inconsistências, além de envio periódico a todos os demais mestrandos do banco de dados atualizado.

3.2.4. Checagem das inconsistências e controle de qualidade

Os responsáveis por essa comissão foram os mestrandos Marina de Borba Oliveira e Lucas Oliveira. A comissão ficou responsável pela realização da checagem semanal das inconsistências do bloco geral, assim como por reunir as modificações necessárias nos blocos individuais dos mestrandos e professores, repassando as alterações a serem realizadas para a comissão do banco de dados por meio de planilha compartilhada do Google Sheets. Para isso, a comissão contou com um aplicativo na plataforma shinyapps desenvolvido por mim utilizando a linguagem de programação R versão 3.6.2. Este instrumento teve como princípio a checagem de uma série de regras lógicas que os dados coletados deveriam seguir e, que, casos as regras não fossem respeitadas, indicavam possíveis correções necessárias. Assim, a base do processo de checagem tornou-se automatizada, de modo que a comissão realizava a inserção semanal do banco de dados no aplicativo, o que gerava uma segunda planilha onde estavam contidos os dados das entrevistas cujas variáveis apresentavam possíveis erros indicados pelo aplicativo. A partir desta segunda planilha, a comissão avaliava a

necessidade de modificação das variáveis em questão por meio da interpretação do banco de dados, discussão com as entrevistadoras e, se necessário, novo contato telefônico com o entrevistado.

A comissão ficou ainda responsável pelo sorteio aleatório de 10% da amostra, realizado por meio do mesmo aplicativo, e posterior contato telefônico para aplicação de questionário reduzido para os indivíduos sorteados. O processo do controle de qualidade será discutido em maiores detalhes no item oito do presente relatório.

3.2.5. Comunicação de divulgação

Os responsáveis por essa comissão foram os mestrandos Rafaela Borges e Lucas Oliveira. A comissão foi responsável pela divulgação do estudo antes, durante e depois do trabalho de campo. Ferramentas como o site do programa, rádios, jornais, Facebook e Instagram foram utilizadas para divulgação da pesquisa. Auxiliaram essa comissão os mestrandos Daniel Leventhal e Eveline Bordignon.

Antes e durante o trabalho de campo a equipe também ficou responsável por ligações telefônicas e envio de e-mails aos meios de comunicação para ampliar a divulgação entre os residentes da cidade.

Até a elaboração deste relatório, o trabalho de divulgação não havia sido concluído. Após a conclusão dos trabalhos individuais de cada mestrando, será elaborado e discutido um material para divulgação dos resultados para a comunidade.

3.2.6. Seleção e treinamento de pessoal

Esta comissão esteve sob responsabilidade dos mestrandos Daniel Leventhal, Gabriela Marques e Rafaela Borges. Eles foram responsáveis pela realização de entrevistas com as auxiliares de pesquisa inscritas para seleção, pelo planejamento da logística do treinamento – envolvendo a elaboração do cronograma e do material utilizado no decorrer da semana de treinamentos – e pela execução e correção da prova teórica. A nota final foi composta pela média desta com uma atividade prática, realizada durante o estudo piloto, a qual foi acompanhada pelos mestrandos.

3.2.7. Logística

As responsáveis por essa comissão foram as mestrandas Eloisa Porciúncula e Gabriela Marques. A equipe foi responsável pela gestão do trabalho de campo propriamente dito. A comissão organizou os setores censitários sorteados para participar do estudo, dividiu-os entre todos os mestrandos, de modo que fosse de responsabilidade de cada um o processo de reversão de recusas, bem como o contato com os domicílios onde as auxiliares de pesquisa não obtiveram êxito em suas tentativas de entrevistas. Além disso, essa comissão foi a responsável por elaborar escalas de trabalho, de modo que os alunos envolvidos cooperassem de maneira semelhante em todas as funções: listagem e reconhecimento de domicílios e supervisão do trabalho de campo. Tal comissão também tinha como atribuição a elaboração semanal de planilhas para registro interno e apresentação de resultados sobre o andamento da pesquisa utilizando como ferramenta principal a plataforma Monday.

3.2.8. Financeiro

As responsáveis por essa comissão foram as mestrandas Anna Muller e Rafaela Borges. A comissão ficou encarregada de todas as questões relacionadas ao controle financeiro, orçamento e previsão de compras durante todo o Consórcio de Pesquisa. Também ficaram responsáveis pelo controle de pagamento dos auxiliares de pesquisa e do caixa feito pela turma para possíveis gastos não incluídos no financiamento pela CAPES.

3.3. Elaboração de relatórios

Esta comissão foi formada pelas mestrandas Bruna Venturin, Eloisa Porciúncula e Marina de Borba Oliveira. A equipe foi responsável pela elaboração do presente documento, contando com a colaboração dos integrantes das demais comissões que compuseram este consórcio.

3.4. Questionário

Nos quadros 1 e 2 apresenta-se os temas abordados no questionário principal e no subestudo sobre inaladores dosimetrados, respectivamente.

Quadro 1. Bloco, número de questões abordadas no questionário do consórcio de pesquisa 2019/2020

Bloco	Questões	Assuntos
A	222	Gerais (idade, escolaridade, trabalho, tabagismo e consumo de álcool); Alimentação, Atividades Físicas, Sono, Consumo de Medicamentos (benzodiazepínicos e inaladores dosimetrados), Asma e DPOC, Doenças Autorreferidas, Serviços de Saúde (utilização e discriminação), Saúde Bucal, Percepções (apoio social e justiça) e Violência Urbana.
B	38	Renda, bens e insegurança alimentar.

Quadro 2. Bloco, número de questões abordadas no questionário do subestudo do consórcio de pesquisa 2019/2020

Bloco	Questões	Assuntos
1	8	Medicamento e características de obtenção e utilização
<i>Checklist</i>	<i>8 - 11</i>	Observação da Técnica

3.5. Manual de instruções

O manual de instruções foi construído para auxiliar os mestrandos e facilitar o fluxo de treinamento das auxiliares de pesquisa e o trabalho de campo. A versão impressa do manual de instruções fez parte do kit disponibilizado para as auxiliares e que foi levado a campo.

O manual possuía informações necessárias para a compreensão do questionário, incluindo orientações sobre os dados que se pretendia coletar, explicações sobre cada pergunta, opções de respostas e instruções para perguntas em que as opções deveriam ou não ser lidas. Também possuía as definições de termos utilizados no questionário e o telefone de todos os mestrando supervisores. Cada um dos mestrando foi responsável pela elaboração do manual referente ao seu bloco de questões.

3.6. Amostra e processo de amostragem

Para definição do tamanho de amostra final, cada mestrando calculou o tamanho de amostra necessário para seu tema de interesse. Considerou-se acréscimo de 10%

para perdas e recusas e 15% para controle de fatores de confusão e possível efeito do delineamento.

Durante a oficina de amostragem, coordenada pelos professores Aluísio Barros e Fernando Hartwig, foi definido o maior tamanho de amostra necessário para contemplar os temas de interesse de todos os mestrandos, levando em consideração questões logísticas e financeiras. A amostra mínima necessária era de 3.400 indivíduos, residentes em 1700 domicílios da cidade, considerando-se uma média de dois adultos por domicílio. Para alcançar a amostra desejada, definiu-se que seriam sorteados 100 setores censitários e visitados, em média, 17 residências por setor. O número de residências visitadas por setor seria proporcional ao aumento do tamanho de cada setor censitário entre os anos de 2010 e o tamanho verificado durante a listagem de domicílios realizada pelos mestrandos em 2019/20 (o número final de domicílios sorteados por setor variou entre 10 e 25).

O processo de amostragem foi realizado em múltiplos estágios. De acordo com o Censo Demográfico de 2010, do Instituto Brasileiro de Geografia e Estatística (IBGE), foram listados todos os 488 setores censitários do município de Pelotas, ordenados pela informação de renda média do domicílio.

O número total de domicílios da zona urbana de Pelotas (107.152) foi dividido por 100 (número de setores a serem visitados), para obter-se o pulo sistemático (1072), respeitando-se a probabilidade proporcional ao número de domicílios por setor. Em seguida, sorteou-se de modo aleatório o número 955 (entre 1 e 1072) por meio do programa Microsoft Excel, correspondendo a um domicílio pertencente ao primeiro setor definido. A seleção dos demais setores ocorreu por meio da soma do pulo sistemático ao número do domicílio do setor inicial (955) e assim sucessivamente até o término da listagem.

3.7. Estudos pré-piloto e piloto

Os estudos pré-piloto e piloto tiveram a colaboração dos mestrandos Bárbara Berruti, Daniel Leventhal e Eveline Bordignon para sua organização. Com o objetivo de detectar possíveis falhas de compreensão das questões ou do modo de preenchimento, tanto por parte de entrevistador(a) quanto de entrevistado(a), do dia 04 ao dia 15 de novembro de 2019, foi realizado o estudo pré-piloto, em que cada mestrando ficaria responsável pela aplicação de no mínimo dez questionários. No total, foram aplicados aproximadamente 140 questionários e discutidas as principais questões que tiveram intercorrências durante a aplicação.

Em seguida, os mestrandos se reuniram e avaliaram todas as dúvidas, inconsistências e dificuldades encontradas, organizando uma nova versão do questionário para aplicação do estudo piloto, sendo uma parte do processo de seleção e treinamento das auxiliares em pesquisa.

O estudo piloto foi realizado durante a seleção e treinamento das auxiliares de pesquisa, sendo parte avaliativa do processo. Foi realizado nos dias 21 e 22 de novembro de 2019. No total foram aplicados 36 questionários em um setor censitário não sorteado para participar da pesquisa.

Após o procedimento exposto acima, os mestrandos se reuniram, avaliaram e corrigiram os questionários e as incompatibilidades encontradas naquele momento, redigindo uma versão final do questionário.

3.8. Trabalho de campo

O trabalho de campo do presente estudo consistiu em três etapas principais: listagem de domicílios, reconhecimento de domicílios sorteados e entrevistas. Tais etapas ocorreram entre os meses de novembro de 2019 e março de 2020, sendo que a listagem de domicílios teve início algumas semanas antes das demais, possibilitando um adequado fluxo de entrevistas. O processo teve início no subdistrito Fragata, passando por Três Vendas, Areal, Laranjal, Barragem, São Gonçalo e sendo finalizado no subdistrito Centro.

Cabe ressaltar que, em todas as fases, tanto os mestrandos quanto as auxiliares de pesquisa apresentavam-se com adequada identificação do estudo para comprovar o propósito e a idoneidade da pesquisa (camiseta, crachá e documentos afins).

Na primeira etapa – listagem de domicílios – os alunos, geralmente em duplas, faziam a visita inicial ao setor censitário, no qual se coletava o número de todas as residências pertencentes ao setor em questão; quando não havia numeração no domicílio, coletava-se o máximo possível de informações para identificação posterior do mesmo, caso fosse sorteado para entrevista. A listagem dos domicílios ocorria de forma sistemática, prezando-se pela representatividade dos sorteados em relação às características gerais do setor como um todo. Desta forma, primeiramente eram listadas as casas pertencentes ao perímetro do setor, iniciando-se pelo ponto mais ao Norte e percorrendo o setor em sentido horário até o ponto inicial. Em seguida, eram listadas as residências contidas nas ruas internas do setor, também partindo do ponto mais ao Norte em direção ao Sul e obedecendo simultaneamente à orientação de Oeste para Leste do mapa do setor. Eram registrados os nomes das ruas em questão e numerados os segmentos a que cada uma das ruas pertencia, de acordo com a orientação geográfica de Norte para Sul, em sentido-horário, e de Oeste para Leste, com base nos mapas de cada setor, organizados pelas mestrandas Bruna Venturin, Bárbara Berrutti, Eveline Bordignon e Paulo Victor de Albuquerque, obtidos previamente à visita ao setor, com auxílio de imagens via Google Earth.

Essa listagem sistematizada, com o número dos domicílios de cada setor censitário, era digitada em planilhas da Microsoft Excel. As planilhas continham informações descritivas dos domicílios, como o número do segmento, o nome da rua, o número da casa, nome do edifício, sua cor, entre outros, que fossem relevantes para a futura identificação e localização dos mesmos pelas auxiliares de pesquisa. Em posse dessa planilha, os mestrandos Eloísa Porciúncula, Marina de Borba Oliveira e eu eramos responsáveis pelo sorteio dos domicílios onde seriam realizadas as entrevistas.

Utilizando a lista de domicílios válidos por setor, o sorteio dos domicílios a serem entrevistados foi realizado em plataforma shinyapps, desenvolvida para automatizar o processo de amostragem, de acordo com a seguinte sequência:

- Uma correção do número n de domicílios a serem sorteados por setor foi realizada de modo a considerar a expansão ou redução no número de domicílios em cada setor encontrados na etapa da listagem dos domicílios em relação ao valor fornecido pelo Censo Demográfico de 2010, da forma:
 - Calculava-se: $n = 17 \times (\text{n}^\circ \text{ de domicílios no setor encontrados na listagem dos domicílios}) / (\text{n}^\circ \text{ de domicílios no setor de acordo com o Censo})$;
 - n era arredondado para o valor inteiro imediatamente superior a n ;
 - Se n era maior do que 25, n era considerado igual a 25, de modo a estabelecer um limite superior para a expansão;
 - Se n era menor do que 10, n era considerado igual a 10, de modo a estabelecer um limite inferior para a redução;
- De posse de n , o pulo sistemático era determinado como sendo o quociente da divisão do número de domicílios em cada setor encontrados na etapa da listagem dos domicílios por n ;
- Um domicílio era sorteado aleatoriamente dentre a lista ordenada geograficamente dos domicílios do setor e o pulo era aplicado sistematicamente até o final da lista, retornando então para o início da lista até que n domicílios tivessem sido sorteados;
- O modelo de sorteio sistemático foi realizado de modo a garantir uma distribuição homogênea de domicílios ao longo do setor sendo sorteado;

Na etapa seguinte – reconhecimento de domicílios sorteados – preferencialmente os mesmos mestrandos que fizeram a listagem de domicílios do setor censitário em questão procediam com uma primeira visita às residências sorteadas, com o intuito de apresentar e explicar a pesquisa a algum morador e coletar o máximo possível de informações (nomes e número de adultos residentes; telefone para contato; melhores dias e/ou horários para realização de entrevistas). Quando nenhum morador se fazia presente na residência durante essa fase da pesquisa, tentava-se conseguir alguma dessas informações com vizinhos, além de entrega de correspondência contendo carta de apresentação e cópia de matéria de jornal local impresso sobre o estudo.

Para realização da etapa final de campo – as entrevistas – as auxiliares de pesquisa recebiam uma escala com as informações das residências nas quais havia possibilidade de entrevistas no turno em questão (manhã ou tarde), bem como o mapa

do setor censitário onde as entrevistas daquele turno seriam realizadas. De posse do material completo (mochila com tablet, manual de instruções, questionário impresso, cartões necessários para a aplicação de determinados questionários, bloco para anotações, estojo contendo canetas, lápis e borracha, termos de consentimento livre e esclarecido) e corretamente identificadas como colaboradoras do estudo, as auxiliares de pesquisas saíam do Centro de Pesquisas Epidemiológicas (CPE) em horário previamente combinado e eram levadas ao ponto de encontro definido pelos mestrandos como seguro e adequado dentro do setor onde as entrevistas do turno seriam realizadas. Após a chegada ao setor, com o auxílio do mapa e da escala, as auxiliares se dirigiam aos domicílios, sempre com supervisão de algum dos mestrandos (presencial no setor ou via remota). Após o término do turno, as auxiliares repassavam o status das entrevistas (entrevista realizada, não finalizada, recusa, inelegível, agendar) para o mestrando responsável pelo turno de coleta de dados e o mesmo atualizava os dados coletados na plataforma Monday diariamente.

Semanalmente, as auxiliares de pesquisa levavam os tablets até o CPE para download dos dados, realizado pela comissão de gestão de banco de dados. Em período semelhante ou um pouco mais estendido, a depender da necessidade, realizavam-se reuniões com os mestrandos e as auxiliares de pesquisa para discutir falhas observadas pelo controle de qualidade ou reforçar questões relevantes ao bom andamento da pesquisa.

As etapas acima citadas foram realizadas até a segunda semana de março de 2020, quando a pandemia de Covid-19 levou a Universidade Federal de Pelotas a cancelar as atividades presenciais. Seguindo determinação do colegiado do Programa de Pós-graduação em Epidemiologia, definiu-se o cancelamento deste consórcio. Este ocorreu em função da impossibilidade de continuação da pesquisa após uma possível normalização das atividades, considerando a imprevisibilidade desta normalização e a discrepância de cenário social, econômico e de saúde com o qual a continuidade desta pesquisa se depararia.

3.9. Controle de qualidade

A qualidade dos dados coletados foi avaliada por meio da replicação de um questionário reduzido para 10% dos entrevistados, mediante contato telefônico. As

ligações foram realizadas pela Comissão de Checagem das Inconsistências e Controle de Qualidade, compostas pelos mestrandos Marina de Borba Oliveira e Lucas Oliveira. O sorteio era realizado semanalmente por meio de aplicativo na plataforma shinyapps. Caso não se obtivesse contato com o entrevistado após 3 tentativas em horários diferentes, novo sorteio era realizado, até atingir os 10% previstos das entrevistas realizadas no período. As entrevistas que apresentassem possível inconsistência (conforme discutido no item 2.3.1 do presente relatório) foram excluídas do controle de qualidade. Os dados foram inseridos em formulário específico na plataforma RedCap.

Para avaliar a concordância entre as questões utilizou-se o coeficiente de Kappa. O coeficiente de Kappa variou entre 0,78 (questão sobre consumo de bebidas alcoólicas) e 1,0 (questão sobre o entrevistado saber ler/escrever), indicando que a concordância entre o questionário reduzido, para o controle de qualidade, e o questionário utilizado na pesquisa foi entre concordância substantiva e quase perfeita (Landis; Koch, 1977). Os valores do coeficiente de Kappa podem ser conferidos na tabela 2.

Tabela 2 – Estatística Kappa realizada entre o questionário de pesquisa e o questionário reduzido do controle de qualidade

Questão	Kappa	p-valor
O(a) Sr.(a) sabe ler ou escrever?	1,0	<0,001
O(a) Sr.(a) tem filhos?	1,0	<0,001
O(a) Sr.(a) consome alguma bebida de álcool?	0,78	<0,001
Algum médico ou profissional de saúde disse que o(a) Sr.(a) tem asma?	0,88	<0,001
Alguma vez na vida, o(a) Sr.(a) já consultou com o dentista?	1,0	<0,001
Algum médico ou profissional de saúde disse que o(a) Sr.(a) tem problema de coração atual ou antigo?	0,8	<0,001

3.10. Resultados gerais

A coleta de dados foi interrompida no dia 18 de março de 2020, devido ao contexto atual da epidemia do novo coronavírus, seguindo as recomendações da Organização Mundial de Saúde e da Universidade Federal de Pelotas. A comissão de organização trabalhou nas semanas posteriores fazendo a contagem de TCLEs e conferência da planilha disponível no Monday. A comissão de gestão e conferência do banco de dados detectou e corrigiu as inconsistências.

No total foram realizadas 827 entrevistas, descritas conforme a Tabela 3. A maioria dos participantes do estudo foram do sexo feminino; idade igual ou superior a 60 anos; cor da pele autodeclarada branca; casados e com ensino fundamental incompleto.

Tabela 3 – Frequência absoluta (n) e relativa (%) das variáveis sociodemográficas dos participantes da pesquisa (n=827)

Variáveis	n	%
Sexo (n=827)		
Masculino	277	33.5
Feminino	550	66.5
Idade (n=827)		
18-28	137	16.6
29-39	103	12.5
40-49	112	13.5
50-59	153	18.5
60 ou mais	322	38.9
Cor da pele/etnia (n=821)		
Branca	615	74.9
Preta	103	12.6
Parda	94	11.5
Amarela	4	0.5
Indígena	5	0.6
Estado civil (n=827)		
Solteiro(a)	294	35.6
Casado(a)	346	41.8
Separado(a) ou divorciado(a)	71	8.6
Viúvo(a)	116	14
Escolaridade (n=822)		
Analfabeto(a)	54	6.5
Fundamental incompleto	346	42.1
Ensino fundamental completo ou médio incompleto	141	17.2
Ensino médio completo ou superior incompleto	194	23.6
Ensino superior completo ou pós-graduação incompleta	64	7.8
Pós-graduação completa	23	2.8
Nível Socioeconômico - ABEP (650)		
A	25	3.9
B1	42	6.5
B2	125	19.2
C1	179	27.5
C2	181	27.9
D-E	98	15.0

3.11. Orçamento

A Comissão de Aperfeiçoamento de Pessoal no Nível Superior (CAPES/PROEX) foi responsável pelo financiamento parcial do consórcio de pesquisa, disponibilizando o valor de R\$ 90,000.00 (noventa mil reais) destinados ao pagamento das entrevistas, transporte e impressões. Ainda, o colegiado do Programa de Pós-Graduação em Epidemiologia disponibilizou R\$5,000.00 (cinco mil reais), que seriam acrescidos ao orçamento, caso fosse necessário. Os mestrandos da turma 2019/2020, colaboraram com R\$11,027.00 (onze mil e vinte e sete reais), totalizando, portanto, R\$106,027.00 (cento e seis mil e vinte e sete reais) disponíveis para o trabalho de campo. Por fim, o PPGE disponibilizou espaço físico e linha telefônica para o andamento do trabalho. Os gastos totais estão detalhados na tabela 4.

Tabela 4 – Gastos Totais do Consórcio de Pesquisa 2019/2020

Item	Quantidade	Custo total (R\$)
Crachás	28	R\$ 8.00
Camisetas	56	R\$ 1,176.00
Cópias/Impressões	20855	R\$ 4,171.00
Entrevistas	827	R\$ 16,540.00
Transporte	-	R\$ 4,366.00
Seguros de vida	25	R\$ 656.25
Material de campo	-	R\$ 1,899.05
Software Monday	-	R\$ 664.00
Total		R\$ 29,513.40

3.12. Cronograma

O cronograma do Consórcio encontra-se apresentado na Figura 1.

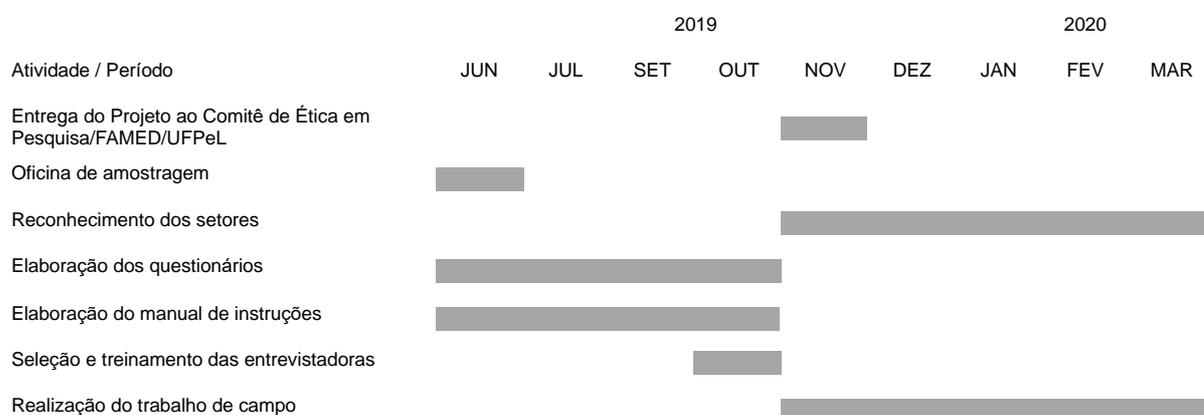


Figura 1 – Cronograma do Consórcio de Pesquisa

3.13. Referências

BARROS, A. J. D. et al. O Mestrado do Programa de Pós-graduação em Epidemiologia da UFPel baseado em consórcio de pesquisa: uma experiência inovadora. **Revista Brasileira de Epidemiologia**. 2008; 11(supl1): 133-144.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. **Biometrics**. 1977 Mar;33(1):159-74. PMID: 843571.

3.14. Anexos

UFPEL - FACULDADE DE MEDICINA DA UNIVERSIDADE FEDERAL DE PELOTAS	
---	---

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AVALIAÇÃO DA SAÚDE DE ADULTOS RESIDENTES NA ZONA URBANA DO MUNICÍPIO DE PELOTAS, RS: Consórcio de Pesquisa do Mestrado em Epidemiologia 2019/2020

Pesquisador: Luciana Tovo Rodrigues

Área Temática:

Versão: 1

CAAE: 24342919.0.0000.5317

Instituição Proponente: Faculdade de Medicina

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.676.549

Apresentação do Projeto:

Após alguns anos realizando estudos com populações específicas, como residentes na zona urbana, idosos e universitários, a turma de mestrandos 2019/2020, do Programa de Pós-Graduação em Epidemiologia, propõe o retorno da realização do conhecimento sobre a situação de saúde dos adultos residentes na zona urbana do município, de delineamento transversal e realizado sob a forma de trabalho conjunto dos mestrandos – denominado como consórcio de pesquisa. A presente proposta prevê a realização de entrevistas realizadas em domicílios com adultos residentes na zona urbana de Pelotas, com 18 anos ou mais. Serão amostrados 1.700 domicílios, correspondendo a estimados 3.400 indivíduos, de novembro de 2019 a julho de 2020. O instrumento que avaliará os aspectos mencionados será aplicado por intermédio de entrevistadoras e respondido através do questionário construído pelos mestrandos. Como fruto deste trabalho conjunto, espera-se que, por sua relevância em saúde, possam ser construídas e divulgadas informações a favor da melhoria da saúde e qualidade de vida dessa população, a partir do conhecimento da situação de saúde dos residentes na zona urbana do município motive a busca por melhores condições de saúde e mudança de hábitos de vida, assim como construção de políticas públicas.

Objetivo da Pesquisa:

Avaliar condições sociodemográficas, hábitos de vida e alimentação, comportamentais, de saúde e

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

Página 01 de 04

Figura A1 – Aprovação do projeto no CEP

4. RELATÓRIO DE ATIVIDADES NO CENTRO INTERNACIONAL DE EQUIDADE EM SAÚDE | PELOTAS

Esta dissertação foi realizada junto ao Centro Internacional de Equidade em Saúde | Pelotas, ou ICEH em inglês, centro pesquisa no qual atuei durante o programa de mestrado.

O ICEH foi desenvolvido por um time de pesquisadores da Universidade Federal de Pelotas e é liderado por Alúcio Barros e Cesar Victora, possuindo como missão o monitoramento de equidade em saúde no mundo, especialmente em países de baixa e média renda. O ICEH é atualmente Centro Colaborador da Organização Mundial da Saúde e Centro Regional para a América Latina e Caribe do Countdown to 2030.

Dentre as suas principais atividades está a análise e harmonização de inquéritos MICS e DHS, essenciais para o desenvolvimento desta dissertação. Por meio destes e outros inquéritos nacionais, o ICEH recalcula indicadores de saúde reprodutiva, materna, neonatal e infantil, incluindo indicadores de vacinação.

4.1. Projeto GAVI

Enquanto no ICEH, atuei em projeto de análise sobre vacinação em países de baixa e média renda encomendado pela GAVI e que continua em andamento. O projeto é constituído de uma série de entregas na forma de relatórios, apresentações e artigos científicos acordadas pelas duas instituições.

Em 2019, três grandes entregas foram realizadas:

- **Relatório GAVI – *cascade***: um relatório de 425 páginas detalhando a cascata de imunização em 91 países separadamente e combinados em regiões da UNICEF, países atendidos e não atendidos pela GAVI e todos juntos. A cascata de imunização é um indicador desenvolvido para o projeto que reflete o número de vacinas (e não doses) que uma criança recebeu;
- **Relatório GAVI – *equity***: um relatório de 322 páginas detalhando a cascata de imunização de acordo com indicadores chave de equidade: quintis de riqueza, sexo da criança, área de residência e índice de empoderamento da mulher. A análise foi realizada para 87 países, combinados de maneira semelhante ao relatório anterior;
- **Relatório GAVI – *health services***: um relatório de 320 páginas detalhando a integração entre indicadores de vacinação e outros serviços de saúde primários

essenciais: ao menos 4 visitas pré-natal, parto institucional, procura de cuidados para diarreia, febre e pneumonia e fontes de água e saneamento melhorados;

Em 2020, seis grandes entregas foram realizadas:

- **Artigo *Zero-dose children and the immunisation cascade: understanding immunisation pathways in low and middle-income countries***: artigo submetido à revista *Vaccine* e aguarda revisão. Trata da cascata de imunização em 92 países de baixa e média renda e apresenta importantes aspectos de desigualdade encontrados;
- **Artigo *Assessing the overlap between immunization and essential health interventions: opportunities for expanding immunization and primary health care***: artigo em processo de finalização. Trata da cobertura de ao menos 4 visitas pré-natal, parto institucional, procura de cuidados para doenças e sintomas infantis comuns e local para higiene em casa de acordo com o estado de vacinação na criança;
- **Apresentação GAVI – *religion***: análises sobre zero dose de acordo com a religião da família e com pertencer ou não a uma maioria religiosa;
- **Apresentação GAVI – *ethnicity***: análises sobre zero dose de acordo com o grupo étnico da família e com pertencer ou não a uma maioria étnica;
- **Apresentação GAVI – *multiple deprivation***: análises sobre a coocorrência de não vacinação e baixo peso e altura para idade, ausência de água e saneamento no domicílio e mãe sem educação;
- **Apresentação GAVI – *decision trees***: árvores de decisão para identificação de grupos de risco de zero dose em 10 países;

4.2. Rotina de indicadores

Uma das atividades de rotina dentro do ICEH das quais participei foi a compilação, análise e harmonização de indicadores de saúde reprodutiva, saneamento e malária dos novos inquéritos MICS e DHS publicados ao longo de 2019 e 2020, além de reanálises de inquéritos anteriores. Além de contribuir para a expansão do banco de indicadores do ICEH, esta atividade também deu origem a um artigo científico ainda em desenvolvimento sob a liderança de Carolina Coll no tópico de mutilação genital feminina, investigando a relação entre a opinião da mãe sobre a

continuação da prática de mutilação e seu nível de empoderamento com a prevalência de mutilação de genital das respectivas filhas.

4.3. Árvores de decisão

Considerando os resultados preliminares promissores desta dissertação, o método das árvores de decisão está sendo investigado para outros dois desfechos além da zero dose. O primeiro – violência contra mulher por parceiro íntimo – constitui um artigo científico liderado por Carolina Coll e está em processo de finalização. O segundo – uso de método contraceptivo moderno – liderado por Aluísio Barros ainda está nos estágios iniciais de desenvolvimento.

De posse dos resultados da aplicação das árvores de decisão em três desfechos diferentes (zero dose, violência por parceiro íntimo e uso de contraceptivo moderno), existe a intenção de publicação de um artigo metodológico voltado para a aplicação das árvores no contexto de saúde global/epidemiologia, com foco particular na interpretação dos resultados.

4.4. Outras atividades

Uma das maiores oportunidades de desenvolvimento oferecidas pelo ICEH é o contato e trabalho em equipe com pesquisadores e pesquisadoras das mais diversas áreas da saúde. Isso possibilitou a contribuição na publicação de quatro artigos na área de nutrição e um na área de vacinação:

- *Ultra-processed foods and the nutrition transition: Global, regional and national trends, food systems transformations and political economy drivers* na revista *Obesity Reviews*;
- *First-food systems transformations and the ultra-processing of infant and young child diets: The determinants, dynamics and consequences of the global rise in commercial milk formula consumption*, na revista *Maternal and Child Nutrition*;
- *Consumption of breast milk, formula and other non-human milk by children aged under two years: analysis of 86 low and middle income countries* na revista *Public Health Nutrition*;
- *Monitoring breastfeeding indicators in high-income countries: Levels, trends and challenges* na revista *Maternal and Child Nutrition*;

- Patterns in Wealth-related Inequalities in 86 Low- and Middle-Income Countries: Global Evidence on the Emergence of Vaccine Hesitancy na revista *American Journal of Preventive Medicine*;

Além de publicações acadêmicas, outras importantes entregas nas quais estive envolvido no ICEH foram:

- Contribuição na produção de diversos materiais gráficos, particularmente por meio da linguagem de programação R;
- Desenvolvimento de uma ferramenta online gratuita para criação de equiplots (gráficos utilizados para representar desigualdades absolutas). A ferramenta pode ser acessada por meio do link: equidade.org/equiplot_creator
- Participação na tradução para o Português do manual *Inequality Monitoring in Immunization* da Organização Mundial da Saúde;
- Participação na automatização da produção dos perfis de equidade apresentados pelo *Countdown to 2030*: countdown2030.org/equity-profiles

5. ARTIGO ORIGINAL

1 **Finding children with high risk of non-vaccination in 92 low- and middle-**
2 **income countries: a decision tree approach**

3

4 Thiago Melo Santos^a

5 1160 Marechal Deodoro, Pelotas, RS, 96020-220, Brazil

6 Bianca de Oliveira Cata-Preta^a

7 1160 Marechal Deodoro, Pelotas, RS, 96020-220, Brazil

8 Cesar Gomes Victora^a

9 1160 Marechal Deodoro, Pelotas, RS, 96020-220, Brazil

10 Aluisio JD Barros^a

11 1160 Marechal Deodoro, Pelotas, RS, 96020-220, Brazil

12

13 ^a Federal University of Pelotas; International Center for Equity in Health

14

15 Corresponding author information: Thiago Melo Santos

16 Email address: thiagomelosts@gmail.com

17 Institutional address: Marechal Deodoro, 1160 - Pelotas, RS. Zip code: 96020-220.

18 Phone/fax +55 (53) 3284 – 1300.

19 **Abstract**

20 **Introduction** In order to “leave no one behind” and the concentration of unvaccinated
21 children in low- and middle-income countries (LMICs), the Immunization Agenda 2030
22 has reduction in vaccination inequalities as a goal. Decision trees are robust statistical
23 tools that can be used for identifying high-risk groups of children who receive no
24 vaccines at all (zero dose).

25 **Methods** The Classification and Regression Trees approach was used to create
26 decision trees for 92 LMICs separately and for all countries combined using data from
27 Demographic and Health Surveys and Multiple Indicator Cluster Surveys to identify
28 groups with the highest risk of children 12-23 months being zero dose. The risk groups
29 identified were characterized in terms of wealth, maternal education, place of
30 residence and sex.

31 **Results** Three high-risk groups were identified in the analysis combining all countries.
32 The one with the highest zero-dose prevalence (42%) included 4% of all children, but
33 24% of zero-dose children. This group was composed of children whose mothers
34 received no tetanus vaccine, had no antenatal care visits and whose delivery was not
35 in a health facility. The groups with higher zero-dose prevalence were poorer, more
36 rural and the mothers were less educated. National trees presented similar results,
37 generally.

38 **Interpretation** Children who have been left out by vaccination services have also been
39 left out by other primary health care services, especially those related to antenatal
40 care. There is an opportunity for better integration between those services in order to
41 achieve the Sustainable Development Goals.

42 **Keywords**

43 Vaccination, immunisation, unvaccinated, antenatal care, decision tree, CART

44 **Introduction**

45 With the end of the decade, the Global Vaccine Action Plan 2011-2020 oversaw
46 important advances in vaccination coverage, especially in low- and middle-income
47 countries (LMICs), but many of its targets have not been met [1]. Between 2010 and
48 2018, it was estimated that the measles vaccine prevented about 23 million deaths,
49 but large measles outbreaks occurred in all World Health Organization (WHO) regions
50 in 2018 [1,2]. Since 2010, 116 countries have introduced vaccines not previously used,
51 but coverage of essential vaccines have stagnated and polio has not yet been
52 eradicated worldwide [1,3]. The COVID-19 pandemic, regional outbreaks of diseases
53 such as the Ebola Virus Disease and the threat of new pandemics represent an
54 important challenge for health systems on the continuation and improvement of routine
55 vaccination services in all countries [3].

56 The Immunization Agenda 2030 (IA2030) recognized the contribution of
57 vaccination in 14 of the 17 Sustainable Development Goals (SDGs), especially SDG
58 3: “Ensure healthy lives and promote well-being for all at all ages”. With over 13 million
59 children receiving no vaccines each year, the IA2030 selected “leave no one behind”
60 as a motto and adopted a strategy that should both target reduction in inequalities and
61 be tailored to the national context of each country [3]. Identifying risk groups for non-
62 vaccination in LMICs could be useful for reducing inequalities and tailoring
63 interventions that could reach and successfully vaccinate the children that are being
64 left behind. Furthermore, zero-dose children – those who failed to receive any vaccine
65 – represent a particularly vulnerable group being left behind by multiple vaccination
66 services and interventions.

67 Vaccination estimates that use administrative data tend to be less reliable – due
68 to inaccuracies in the population denominator and the reported number of doses – and

69 survey data are often considered a more reliable alternative [4]. Demographic and
70 Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) are nationally
71 representative household surveys that collect information on both vaccination and
72 sociodemographic indicators that can be used for identifying risk groups of zero-dose
73 children and their characteristics. According to 241 DHS and MICS surveys from 1986
74 to 2007 from 96 LMICs, the percentage of zero-dose children was 9.9% among those
75 aged 12-59 months. Zero-dose status was generally more common in children from
76 the poorest households and from mothers who were less educated, not vaccinated
77 with the tetanus toxoid vaccine and who decided alone about child's care when the
78 child was sick [5].

79 In order to identify high-risk groups of zero-dose children, the analysis of sample
80 subgroups (according to indicators of wealth or education, for example) is a common
81 option, but it tends to be restricted in terms of the number of subgroups tested and the
82 different combinations of indicators that are used for stratification. As the number of
83 combinations and indicators grow, the analysis gets progressively more complex and
84 the sample size smaller, hindering the identification of high-risk groups. A possible
85 alternative is the use of decision trees. They are robust statistical tools that provide a
86 framework for subgroup identification that simplify that process and allow for a greater
87 number of indicators and their combinations to be tested, without the need to make
88 estimates for all possible combinations [6]. In the context of non-vaccination, decision
89 trees can be used for identifying subgroups of children at a higher risk of being zero
90 dose, without previous specification of what subgroups should be tested.

91 Our objective is to identify high risk groups of zero-dose children in LMICs, using
92 a decision tree approach, in order to determine defining features that can be used for

93 both targeting and tailoring vaccination policies and interventions at national level and
94 globally.

95

96 **Methods**

97 *Data source and children sample*

98 We analyzed 210,509 children from 92 LMICs with DHS or MICS surveys
99 conducted since 2010 and with data released until July 2020. For each country, we
100 selected the most recent survey with available information for the Bacillus Calmette-
101 Guérin (BCG), poliomyelitis (polio), diphtheria, tetanus, pertussis (DTP), and measles
102 vaccines. DHS and MICS are household surveys whose responsible agencies
103 – respectively the United States Agency for International Development (USAID) and
104 the United Nations Children's Fund (UNICEF) – collaborate in order to ensure that both
105 are harmonized and comparable [7]. We classified each country according to their
106 World Bank income level using the median survey year, 2015 [8].

107 The proportion of LMICs represented by the surveys in our study in each
108 UNICEF region was 92% in West & Central Africa, 79% of Eastern & Southern Africa,
109 58% of Middle East & North Africa, 55% of Eastern Europe & Central Asia, 75% of
110 South Asia, 48% of East Asia & the Pacific and 58% of Latin America & Caribbean.
111 Regarding the World Bank income level in 2015, the proportions were 90%, 73% and
112 46% of the low-, lower-middle and upper-middle income countries, respectively.

113 *Outcome*

114 We assessed vaccination status for each child using information from the
115 vaccination cards and, when the card was not available or was not shown, using the
116 caregiver's report. We analyzed children aged 12-23 months living with their mothers
117 [9]. Exceptions in terms of age group were Moldova (15-26 months) and Bosnia and

118 Herzegovina, Costa Rica, Jamaica, North Macedonia and Ukraine (18-29 months), due
119 to the fact that the measles vaccine was given at 15 and 18 months in those countries
120 in the survey year, respectively.

121 The studied outcome was the zero-dose status of the child, defined as not
122 having received any doses of the BCG, polio, DPT and measles vaccines. In case of
123 missing information for a vaccine, the child was considered as not having received it,
124 according to WHO recommendation [9].

125 *Potential indicators of vulnerability*

126 We selected 15 potential indicators that were available in both DHS and MICS
127 surveys, that have been previously used for zero dose investigation in the literature
128 [5,10,11] and that can provide useful information for targeting and tailoring vaccination
129 interventions.

130 The household indicators from the literature were: number of household
131 members, sex of household head, place of residence (urban or rural), wealth quintiles
132 (from poorest to richest: Q1, Q2, Q3, Q4 and Q5), religious affiliation (classified as
133 Christian, Muslim, Hindu, Buddhist, Folk, other or unaffiliated) and ownership of radio
134 or television. The mother indicators were: education (classified as none, primary or
135 secondary+), marital status (classified as never married or in a union, currently married
136 or in a union, or formerly married or in a union) and maternal age in complete years.
137 We also included the number of live born children to the mother that had died to date.
138 The child and pregnancy indicators were: sex of the child, number of antenatal care
139 visits during the pregnancy, place of delivery (classified as noninstitutional, institutional
140 – public, or institutional – private) and the number of tetanus toxoid injections received
141 before and during the pregnancy.

142 Additionally, we included two other household indicators: ownership of a
143 refrigerator and the number of living children, resulting in 17 indicators. Ownership of
144 a refrigerator was included as a form of negative control to radio and television, since
145 the three are indicators of wealth, but, different from the other two, a refrigerator is not
146 a source of information. The number of living children was included as a possible
147 indicator of competition for time and household resources necessary for vaccination.

148 The wealth quintiles are based on a wealth index created by the MICS and DHS
149 teams via principal components analysis (PCA) of household indicators: building
150 materials, ownership of appliances, presence of electricity, water and sanitation,
151 among others [12,13]. Separate PCAs are carried out for urban and rural settings in
152 order to take into account the differences in relevance that household assets can have
153 in those settings. The households are then divided into quintiles according to their
154 wealth index [14].

155 *Analysis methods*

156 We calculated the zero-dose prevalence and its 95% confidence interval
157 (95%CI) for each country taking into account the complex survey design used by DHS
158 and MICS using Stata 16 [7] and created a world map to present those prevalences
159 using R 4.0.2 and publicly available map datasets [15].

160 In order to identify the groups of children with the highest zero-dose prevalence
161 in the studied countries combined and for each individual country, we used a decision
162 tree approach.

163 The most frequently used technique for creating decision trees is the
164 Classification and Regression Tree (CART) method [6]. For the identification of groups
165 with high risk, CART will perform a binary recursive partitioning process. This means
166 that, using the indicators available for the analysis, it will divide the sample into two

167 groups (hence binary), one with a higher zero-dose prevalence and one with a lower,
168 when compared to the original complete sample. The algorithm will test all the
169 indicators. Continuous or polytomous indicators will have all the possible binary splits
170 tested (e.g. belonging to Q1 vs Q2-Q5, belonging to Q1-Q2 vs Q3-Q5, and so on). The
171 best split is the one that maximizes within-group homogeneity. This process eliminates
172 the necessity of previously specifying cut-off points for non-binary indicators and also
173 allows for a specific stratification to happen within different subgroups of the sample,
174 dealing naturally with the complex interactions of indicators.

175 This process will go on, splitting the groups in the previous iteration into smaller
176 groups (hence recursive) until a stopping rule halts the process. The stopping rule can
177 be, for example, a minimum number of children in one subgroup, also called a node.

178 Since the outcome is a binary indicator, CART will create a classification tree,
179 meaning a tree that will classify each child as “zero dose” or “non zero dose” according
180 to the terminal node the child is assigned. However, the zero-dose prevalence in most
181 surveys is much smaller than the prevalence of children who received at least one
182 dose, creating an unbalanced sample for the classification algorithm. This can be
183 thought of as the problem of using a diagnostic tool to identify people with a rare
184 disease. Since most children in the sample have received at least one vaccine dose,
185 the tree can simply classify all children as “non zero dose” and still have high accuracy.
186 The specificity will be of 100% and the sensitivity of 0%, but the overall accuracy –
187 given the low prevalence of zero dose – will be high. In order to mitigate this, we
188 adjusted the misclassification cost (also called loss) used by CART. During the creation
189 of the tree, classifying a “zero dose” child as “non zero dose” costed double than
190 classifying a “non zero dose” child as “zero dose”, reducing the threshold necessary
191 for the children in a node to be classified as zero dose. A more detailed description of

192 how misclassification costs are used in CART can be found elsewhere [16,17]. We
193 made this specification in order to increase the tree sensitivity, even if accompanied
194 by a reduction in specificity. Also, any split leading to a node with less than 50 children
195 was discarded.

196 We created a decision tree for each of the 92 countries separately and one for
197 all countries combined using the CART implementation of the rpart package version
198 4.1-15 [17] in R 4.0.2. All the trees were created using sample weights that consider
199 the complex survey design. The individual trees are presented in country profiles that
200 also include the prevalence of key inequality and health care indicators, selected
201 according to the results of the pooled tree. The pooled analyses sample weights were
202 adjusted in order to take into consideration the number of children aged 12-23 months
203 living in each country.

204 We analyzed the composition of each risk groups identified by the pooled tree
205 in terms of wealth, maternal education, place of residence and sex of the child in order
206 to better characterize inequalities among these groups. This is an important
207 complementary equity analysis to the decision tree since the indicators featured in the
208 tree are the ones best suited for identifying the zero dose groups, but this does not
209 mean that there are no differences in terms of other indicators.

210 *Ethical clearance*

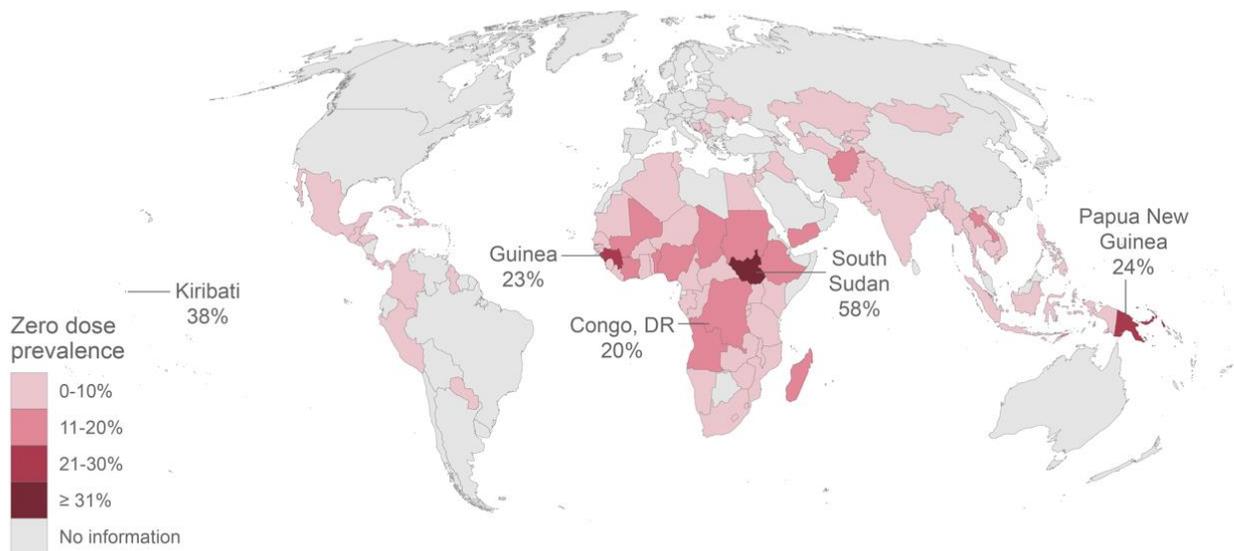
211 All the analyses were carried out using only publicly available datasets. The
212 organizations who administered the surveys were responsible for ethical clearance
213 according to norms of each country.

214 **Results**

215 *Zero-dose prevalence*

216 The pooled zero-dose prevalence was 7.7% (95%CI 7.4-7.9%), but varied
217 widely across all countries, with a median prevalence of 2.9% (interquartile range: 1.5-
218 8.7%). Moldova had the lowest prevalence, with no zero-dose children identified in the
219 sample. Meanwhile, South Sudan had the highest prevalence: 57.6% (95%CI 54.1-
220 61.0%); followed by Kiribati: 38.2% (95%CI 33.4-43.4%); Papua New Guinea: 24.2%
221 (95%CI 20.9-27.9%); Guinea: 23.0% (95%CI 19.8-26.4%); and Congo, DR: 20.5%
222 (95%CI 17.9-23.3%). The national prevalence level for all countries is presented in
223 Figure 1 and the actual values in Table S1 in the supplementary material.

224



225 Figure 1 – Zero dose national prevalence map and top 5 countries with highest
226 prevalence
227

228 *Pooled decision tree*

229 The decision tree created for all countries combined divided the sample into 4
230 groups of children with vastly different zero-dose prevalence (Figure 2). Of all the 17

231 indicators included in the tree analyses, only three were selected to separate the
232 children into risk groups: the number of tetanus vaccine doses received by the mother
233 during or before the pregnancy of the child, the number of antenatal care visits and the
234 child's place of delivery.

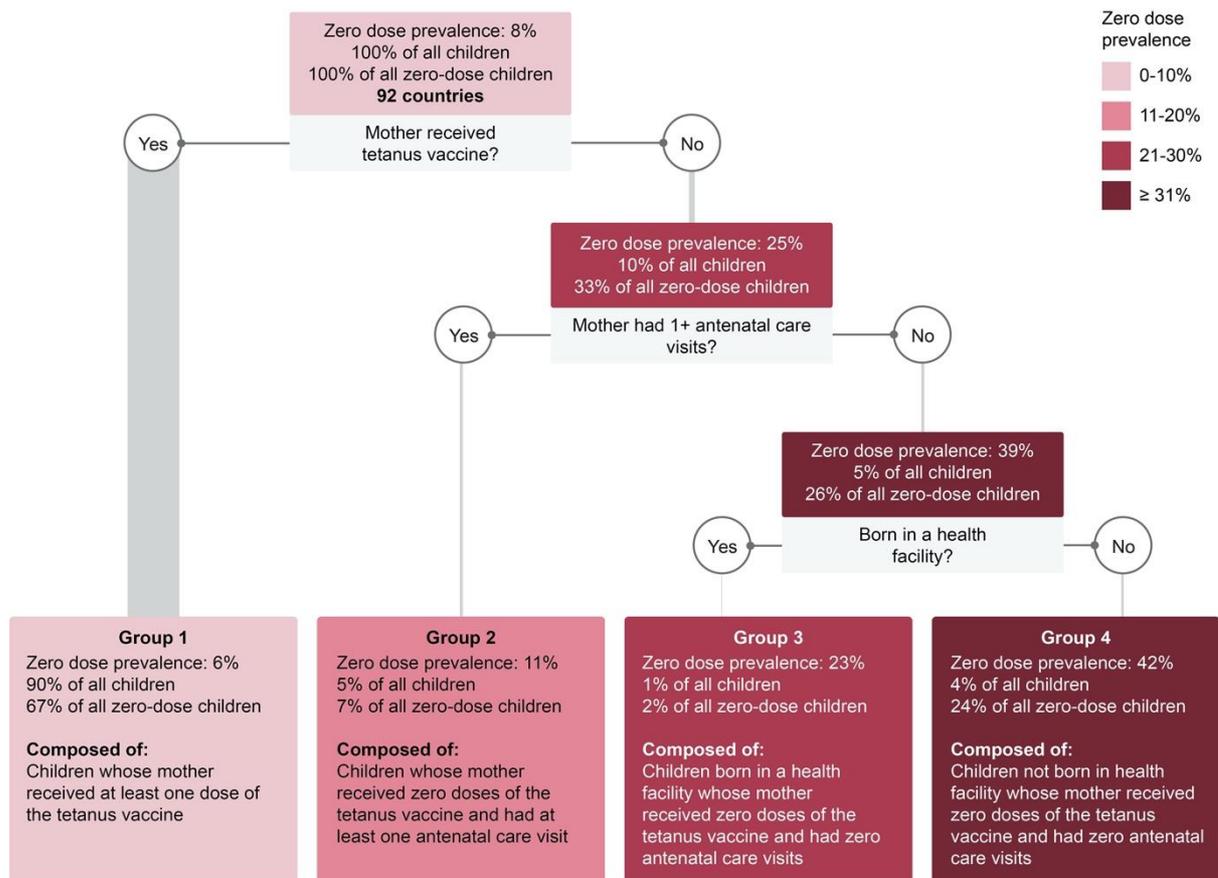
235 Group 1 was composed of children whose mother received at least one dose of
236 the tetanus vaccine. It includes 90% of all children and also the majority of all zero-
237 dose children (67%), given it is a very large group. But it was the group with the lowest
238 zero-dose prevalence: 5.7% (95%CI 5.5-5.9%).

239 Group 2 was composed of children whose mothers received no tetanus vaccine,
240 but who had at least one antenatal care visit during the child's pregnancy. It includes
241 5% of all children and 7% of all zero-dose children. It had almost double the zero-dose
242 prevalence of Group 1: 11.0% (95%CI 9.9-12.3%).

243 Group 3 was composed of children whose mothers received no tetanus vaccine
244 and had no antenatal care visits, but who were born in a health facility. It includes 1%
245 of all children and 2% of all zero-dose children. It had the second highest zero-dose
246 prevalence among the groups: 22.6% (95%CI 18.6-27.2%).

247 Group 4 was composed of children whose mothers received no tetanus vaccine
248 and had no antenatal care visits and whose delivery was not in a health facility. With
249 only 4% of all children in this group, it includes 24% of all zero-dose children and a
250 zero-dose prevalence of 42.2% (95%CI 39.7-44.8%).

251



252

253

Figure 2 – Decision tree for the combined sample of 92 countries

254

255 *Characterization of the risk groups*

256

257

258

259

260

261

262

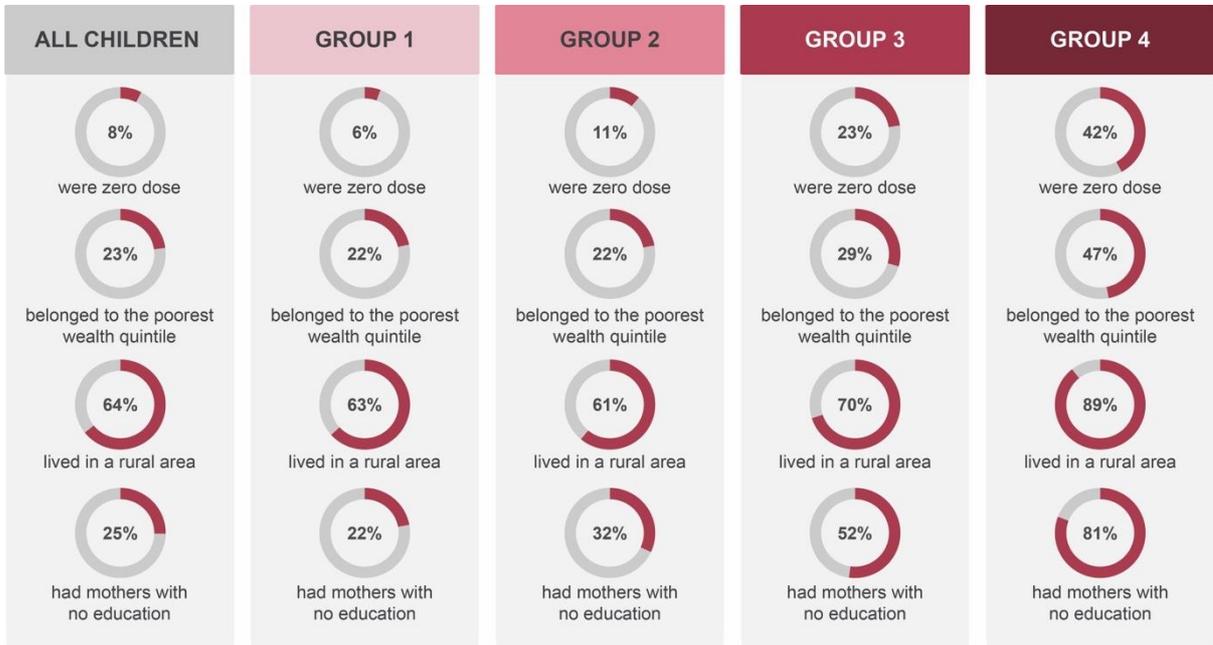
263

264

There was a clear pattern of inequality between the 4 groups: the higher the zero-dose prevalence in a group, the poorer, more rural and with a less educated mother the children in that group. In Groups 1 to 4, respectively, 21.7% (95%CI 21.3-22.1%), 22.0% (95%CI 20.0-24.1%), 29.5% (95%CI 25.2-34.2%) and 47.0% (95%CI 44.1-50.0%) of the children belonged to the poorest wealth quintile; 63.3% (95%CI 62.7-63.8%), 61.0% (95%CI 58.8-63.2%), 70.1% (95%CI 64.2-75.4%) and 89.3% (95%CI 87.7-90.8%) lived in a rural area; and 22.1% (95%CI 21.7-22.5%), 32.0% (95%CI 29.6-34.5%), 52.2 (95%CI 47.1-57.2%) and 80.9% (95%CI 79.0-82.6%) had a mother with no education. A comparison between the 4 groups and all children

265 combined is presented in Figure 3. There were no statistically significant differences in
 266 terms of sex of the child among any groups.

267



268

269 Figure 3 – Description of the 4 zero dose groups according to key inequality
 270 indicators
 271

272 *Country-specific trees*

273 Of the 92 countries studied, in only 25 it was possible to identify high risk groups
 274 with the decision tree approach. The median zero-dose prevalence for the 67 countries
 275 for which a tree was not built was 1.7% (ranging from 0.0% to 16.3%), while the median
 276 for the other 25 countries was 13.0% (ranging from 3.3% to 57.6%). The median
 277 sample size for the countries for which a tree was not built was 1337 children (ranging
 278 from 266 to 5820), while the median for the other countries was 2151 children (ranging
 279 from 453 to 49284). Among the countries with risk groups separation, 15 featured the
 280 number of antenatal care visits in their tree, 10 the place of delivery, 8 the number of
 281 tetanus vaccine doses and also 8 the wealth quintile. The results for all indicators are

282 presented in Table 1. The 92 trees are presented in the country profiles in
283 supplementary figures S1 to S92.

284

285 Table 1 – Indicators featured in national decision trees

Indicator	Number of national trees that featured the indicator (out of 25 trees)
Number of antenatal care visits	15
Place of delivery	10
Mother's tetanus vaccine doses	8
Wealth quintiles	8
Number of household members	5
Mother's age	4
Mother's education	4
Religious group	4
Ownership of radio	2
Place of residence	1
Child's sex	1
Number of live born children that had died	1
Number of living children	1
Ownership of TV	1
Sex of household head	0
Mother's marital status	0
Ownership of refrigerator	0

286

287 Discussion

288 The pooled decision tree indicates a clear message: the children at a higher risk
289 of not being vaccinated are the ones who themselves and whose mothers have also
290 been left out by other health services and interventions, in particular antenatal care.
291 The less health services and interventions the child/mother pair received, the higher
292 the risk of the child not have been vaccinated. The triple exposed children (those
293 whose mother did not received any doses of the tetanus vaccine, who did not have
294 any antenatal care visits and whose delivery was not in a health facility) had an
295 alarming zero-dose prevalence of 42%. But it is important to stress: no causal

296 relationship is established by the decision tree. The analysis can only indicate that both
297 phenomena have happened in the same household.

298 The fact that the triple exposed children were only a small percentage of the
299 sample (4%), but they accounted for one in every four zero-dose children indicates: 1)
300 the opportunity they represent for targeted interventions and 2) how challenging it is to
301 reach them, as they and their mothers have already been left out by vaccination,
302 antenatal care and delivery services. This result corroborates the current
303 recommendation by the IA2030 to “Encourage greater collaboration and integration
304 within and beyond the health sector”, reinforcing the importance of integration between
305 primary health care services in order to increase efficiency and reach those who are
306 being left out by multiple basic interventions [3].

307 It is no surprise that the first indicator selected by the tree was the mother not
308 having received any doses of the tetanus vaccine, which could evidence similar
309 challenges in effectively vaccinating mother and child, involving vaccine supply,
310 logistics, application, monitoring and long-term predictable funding [18]. Both
311 vaccination and antenatal care services failed to reach and vaccinate these mothers
312 during and before the index pregnancy, which contributes to the burden of maternal
313 and neonatal tetanus [19]. Furthermore, one in every 4 children of these mothers were
314 also zero dose, therefore failing to achieve both short and long-term protection against
315 tetanus for the child.

316 The lack of any antenatal care visits can function as a further indicator of families
317 that are harder to reach by health services, but can also have an impact in the future
318 vaccination of the child. Mothers with more antenatal care visits have more
319 opportunities to receive positive messages by healthcare providers regarding the
320 advantages of vaccinating their soon to be born child, increasing awareness of the

321 benefits and safety of vaccines and giving information such as the appropriate place
322 and time for vaccination [20]. Meanwhile, the child not been delivered in a health facility
323 can have a more direct link with the zero-dose status. The WHO recommends that in
324 countries with high burden of tuberculosis a single dose of the BCG vaccine should be
325 given at birth to all healthy neonates [21] and 152 LMICs have a policy of universal
326 neonatal vaccination at birth or at the first week of life [22]. This is in accordance with
327 our results: among children whose mothers have not been vaccinated against tetanus
328 and who had zero antenatal care visits, those who were born in a health facility had a
329 zero-dose prevalence of 23% while those who were not had almost double the
330 prevalence (42%).

331 Our equity analyses showed that the groups of children at higher risk of being
332 zero dose were significant and incrementally poorer, more rural and with less educated
333 mothers. This is line with the literature, as children in those conditions are less likely to
334 be vaccinated [5] and their mothers are also less likely to receive qualified antenatal
335 care [23]. This result further indicates the level of vulnerability of the children in the
336 higher risk groups, as their socioeconomic condition has also been associated with
337 lower medical treatment seeking behavior [24], worst nutritional status [25] and higher
338 mortality among unvaccinated children [26].

339 In 2002, the WHO and the UNICEF launched the “Reach Every District” (RED)
340 strategy, an initiative that promotes the prioritization of districts with poor access and
341 utilization of vaccination services [27], with reports of increased vaccination coverage
342 [28] and detection of vulnerable populations [29,30] after the adoption of the RED
343 strategy. A proposed improvement to the RED strategy is the “reach every community”
344 approach, with a focus on facility level planning, monitoring of community access and
345 utilization of vaccination services, local communication strategies and health networks

346 that should allow for the tailoring of vaccination delivery systems according to the
347 specific necessities of vulnerable communities [29]. Furthermore, those strategies are
348 an opportunity to promote other maternal, newborn and child health (MNCH) services,
349 since the community involvement can help to identify newborn and pregnant women
350 that could be targets for those services [31]. An example is the inclusion of a wider
351 MNCH, environmental health and water and sanitation package by planners involved
352 in the application of the RED strategy in the Byanzurkh District in Mongolia, since they
353 determined that both vaccination and MNCH services shared similar barriers in the
354 district (distance, level of education and poverty) [30].

355 The strengths of this paper include: the multicountry approach, with the
356 inclusion of 92 nationally representative surveys and the resulting large sample size;
357 the choice of zero dose as the studied vaccination outcome, therefore focusing on the
358 most vulnerable children; the national profiles created, with the prevalence of key
359 indicators and national decision trees; and the novel use of decision trees. They are a
360 versatile tool that can be used for identifying risk groups of other outcomes in the
361 context of global and public health. The CART implementation, in particular, can thrive
362 in the large sample sizes that are quite common in the field. Furthermore, it allowed
363 for the inclusion of 17 indicators of vulnerability without previous specification of cut-
364 off points and interactions, dealing naturally with the complex intersections between
365 them.

366 Unfortunately, we were able to identify risk groups at the national level for only
367 25 out of the 92 countries studied. The tree inability to identify risk groups is most likely
368 due to three factors (individually or in combination): the indicators selected for the
369 analyses not corresponding to actual zero dose risk groups in the country, low sample
370 size and low zero-dose prevalence. The last two factors impact the algorithm's ability

371 to create groups that are large enough and with enough zero dose children to be
372 maintained during tree creation, especially with our imposed limitation of at least 50
373 children in each group. This is supported by the fact that countries where risk groups
374 were identified had a median zero-dose prevalence more than 6 times higher and a
375 median sample size more than 60% higher than the countries with no risk groups
376 identified. Furthermore, countries with really low zero dose prevalence may not have
377 well defined risk groups from a population perspective and the remaining children
378 might only be reached by population level interventions. Nevertheless, for the 25
379 countries where risk groups were identified, the result was similar to the pooled tree
380 analysis: antenatal care visits, delivery in a health facility and tetanus vaccination were
381 the most common indicators selected by the trees; the last one together with the wealth
382 indicator.

383 There are several limitations to these analyses. First, since children with missing
384 information in a vaccine are considered as non-vaccinated, the zero-dose prevalence
385 might be overestimated. As a sensitivity analysis, we calculated the pooled zero-dose
386 prevalence removing all zero dose children with missing information for any vaccine,
387 resulting in a prevalence of 7.4% (7.1-7.6%). Among the top five countries, South
388 Sudan and Congo DR had 21% and 18% of all zero dose from children with missing
389 information in at least one vaccine. Although this is far from ideal, it follows the current
390 recommendation by the WHO on how to treat missing information for vaccination
391 records [9] and is a more conservative approach, considering the risks of non-
392 vaccination.

393 Second, although we were successful in identifying zero dose risk groups in the
394 pooled analysis, 67% of all zero dose children in the sample were classified as part of
395 the lowest risk group. This may be due to indicators of vulnerability not included in the

396 analyses or due to the fact that some children do not belong to a specific and well-
397 defined risk group. Further investigation is necessary – possibly using other indicators
398 – as well as a general strengthening of health services, combined with population
399 interventions, in order to ensure that those children are not missed.

400 Third, the highest zero-dose prevalence found in a risk group from the pooled
401 tree was 42%. One could argue that this is a low prevalence for a more traditional
402 application of a classification algorithm, especially if used as a prediction tool.
403 Considering that our main goal was to identify risk groups and not to create predictions
404 and since zero-dose prevalence tend to be really low in most countries, we believe that
405 42% represent a significantly high prevalence from a public health perspective.

406 Fourth, the choice of using a double misclassification cost was arbitrary and the
407 final tree is dependent on that choice. As a sensitivity analysis, we tested other
408 adjustment weights (1, 3, 4 and 5). For 1, no risk groups were identified. For 3 to 5, a
409 very similar tree was created, but the last split (if the child was born in a health facility)
410 was not present. Therefore, the final message remains valid, but with the caveat that
411 the last split should be interpreted with caution.

412

413 **Conclusion**

414 “Leave no one behind” is an opportunity to reach not only those who have been
415 left out by vaccination, but also those who have been left out by other primary health
416 care services. The children at higher risk of being zero dose are also the ones whose
417 mothers were left out by antenatal care, delivery and vaccination services. Those
418 children are also poorer, more rural and their mothers less educated. Further
419 integration of primary health care services and targeted interventions towards the most

420 vulnerable communities are necessary in order to achieve the Sustainable
421 Development Goals and the Immunization Agenda 2030.

422

423 **Declaration of Competing Interest**

424 The authors declare that they have no known competing financial interests or
425 personal relationships that could have appeared to influence the work reported in this
426 paper. All authors attest they meet the ICMJE criteria for authorship.

427

428 **Acknowledgements**

429 This paper was made possible with funds from the Wellcome Trust (Grant
430 Number: 101815/Z/13/Z), Bill & Melinda Gates Foundation (Grant Number:
431 OPP1199234), Associação Brasileira de Saúde Coletiva, and Coordenação de
432 Aperfeiçoamento de Pessoal de Nível Superior (Finance Code 001). We are thankful
433 to Cíntia Borges for her help in the national profiles design.

434

435 **Contributors**

436 TMS and AJDB conceptualized the paper. TMS and BOCP conducted the
437 analyses, with support from AJDB. TMS, BOCP and AJDB interpreted the results, with
438 support from GCV. TMS wrote the manuscript, which was critically reviewed and edited
439 by all other authors. All authors read and approved the final manuscript.

440

441 **References**

442 [1] World Health Organization. The global vaccine action plan 2011-2020: review
443 and lessons learned: strategic advisory group of experts on immunization. World
444 Health Organization; 2019.

- 445 [2] Patel MK, Dumolard L, Nedelec Y, Sodha S V, Steulet C, Gacic-Dobo M, et al.
446 Progress Toward Regional Measles Elimination - Worldwide, 2000-2018. *MMWR*
447 *Morb Mortal Wkly Rep* 2019;68:1105–11. <https://doi.org/10.15585/mmwr.mm6848a1>.
- 448 [3] World Health Organization. *Immunization Agenda 2030: A Global Strategy to*
449 *Leave No One Behind*. Geneva: n.d.
- 450 [4] Cutts FT, Claquin P, Danovaro-Holliday MC, Rhoda DA. Monitoring
451 vaccination coverage: Defining the role of surveys. *Vaccine* 2016;34:4103–9.
452 <https://doi.org/10.1016/j.vaccine.2016.06.053>.
- 453 [5] Bosch-Capblanch X, Banerjee K, Burton A. Unvaccinated children in years of
454 increasing coverage: How many and who are they? Evidence from 96 low- and
455 middle-income countries. *Trop Med Int Health* 2012;17:697–710.
456 <https://doi.org/10.1111/j.1365-3156.2012.02989.x>.
- 457 [6] Wolfson J, Venkatasubramaniam A. Branching Out: Use of Decision Trees in
458 Epidemiology. *Curr Epidemiol Reports* 2018;5:221–9. [https://doi.org/10.1007/s40471-](https://doi.org/10.1007/s40471-018-0163-y)
459 [018-0163-y](https://doi.org/10.1007/s40471-018-0163-y).
- 460 [7] Hancioglu A, Arnold F. Measuring Coverage in MNCH: Tracking Progress in
461 Health for Women and Children Using DHS and MICS Household Surveys. *PLOS*
462 *Med* 2013;10:1–8. <https://doi.org/10.1371/journal.pmed.1001391>.
- 463 [8] The World Bank. *World Bank Country and Lending Groups* n.d.
464 [https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups)
465 [country-and-lending-groups](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups) (accessed October 15, 2020).
- 466 [9] World Health Organization. *Expanded Programme on Immunization (EPI).*
467 *Harmonizing vaccination coverage measures in household surveys: A primer* 2019.
- 468 [10] Shenton L, Wagner A, Bettampadi D, Masters N, Carlson BF, Boulton M.
469 *Factors Associated with Vaccination Status of Children Aged 12–48 Months in India,*

470 2012–2013. *Matern Child Health J* 2017;22. <https://doi.org/10.1007/s10995-017->
471 2409-6.

472 [11] Francis MR, Nohynek H, Larson H, Balraj V, Mohan VR, Kang G, et al.
473 Factors associated with routine childhood vaccine uptake and reasons for non-
474 vaccination in India: 1998-2008. *Vaccine* 2018;36:6559–66.
475 <https://doi.org/10.1016/j.vaccine.2017.08.026>.

476 [12] Filmer D, Pritchett LH. Estimating Wealth Effects Without Expenditure Data—
477 Or Tears: An Application To Educational Enrollments In States Of India*.
478 *Demography* 2001;38:115–32. <https://doi.org/10.1353/dem.2001.0003>.

479 [13] Rutstein S, Johnson K. The DHS Wealth Index. 2004.
480 <https://doi.org/10.13140/2.1.2806.4809>.

481 [14] Rutstein S. The DHS Wealth Index: Approaches for Rural and Urban Areas
482 2008.

483 [15] Kelso N V, Patterson T. Natural Earth Vector 2020.

484 [16] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression
485 Trees. Taylor & Francis; 1984.

486 [17] Therneau T, Atkinson B. rpart: Recursive Partitioning and Regression Trees
487 2019.

488 [18] Duclos P, Okwo-Bele J-M, Gacic-Dobo M, Cherian T. Global immunization:
489 status, progress, challenges and future. *BMC Int Health Hum Rights* 2009;9:S2.
490 <https://doi.org/10.1186/1472-698X-9-S1-S2>.

491 [19] World Health Organization. Protecting all against tetanus: guide to sustaining
492 maternal and neonatal tetanus elimination (MNTE) and broadening tetanus
493 protection for all populations. Geneva PP - Geneva: World Health Organization; n.d.

494 [20] Hajizadeh M. Socioeconomic inequalities in child vaccination in low/middle-

495 income countries: what accounts for the differences? *J Epidemiol Community Health*
496 2018;72:719 LP – 725. <https://doi.org/10.1136/jech-2017-210296>.

497 [21] World Health Organization. *Weekly Epidemiological Record*, 2018, vol. 93, 08
498 [full issue]. *Wkly Epidemiol Rec = Relev Épidémiologique Hebd n.d.*;93:73–96.

499 [22] Roy P, Vekemans J, Clark A, Sanderson C, Harris RC, White RG. Potential
500 effect of age of BCG vaccination on global paediatric tuberculosis mortality: a
501 modelling study. *Lancet Glob Heal* 2019;7:e1655–63. [https://doi.org/10.1016/S2214-](https://doi.org/10.1016/S2214-109X(19)30444-9)
502 109X(19)30444-9.

503 [23] Arroyave L, Saad GE, Victora CG, Barros AJD. Inequalities in Antenatal Care
504 Coverage and Quality: an Analysis From 63 Low and Middle-income Countries Using
505 the ANCq Content-qualified Coverage Indicator. *Int J Equity Health* 2021.
506 <https://doi.org/10.21203/rs.3.rs-120640/v1>.

507 [24] Ayalneh AA, Fetene DM, Lee TJ. Inequalities in health care utilization for
508 common childhood illnesses in Ethiopia: evidence from the 2011 Ethiopian
509 Demographic and Health Survey. *Int J Equity Health* 2017;16:67.
510 <https://doi.org/10.1186/s12939-017-0561-7>.

511 [25] da Silva ICM, França G V, Barros AJD, Amouzou A, Krusevec J, Victora CG.
512 Socioeconomic Inequalities Persist Despite Declining Stunting Prevalence in Low-
513 and Middle-Income Countries. *J Nutr* 2018;148:254–8.
514 <https://doi.org/10.1093/jn/nxx050>.

515 [26] Bishaia D, Koenig M, Khan MA. Measles vaccination improves the equity of
516 health outcomes: evidence from Bangladesh. *Health Econ* 2003;12:415–9.

517 [27] Vandelaer J, Bilous J, Nshimirimana D. Reaching Every District (RED)
518 approach: a way to improve immunization performance. *Bull World Health Organ*
519 2008;86:A-B. <https://doi.org/10.2471/blt.07.042127>.

520 [28] Ryman T, Macauley R, Nshimirimana D, Taylor P, Shimp L, Wilkins K.
521 Reaching every district (RED) approach to strengthen routine immunization services:
522 evaluation in the African region, 2005. J Public Health (Bangkok) 2010;32:18–25.
523 <https://doi.org/10.1093/pubmed/fdp048>.

524 [29] Chan Soeung S, Grundy J, Duncan R, Thor R, Bilous JB. From reaching every
525 district to reaching every community: analysis and response to the challenge of
526 equity in immunization in Cambodia. Health Policy Plan 2013;28:526–35.
527 <https://doi.org/10.1093/heapol/czs096>.

528 [30] Enkhtuya B, Badamusuren T, Dondog N, Khandsuren L, Elbegtuya N, Jargal
529 G, et al. Reaching every district – development and testing of a health micro-planning
530 strategy for reaching difficult to reach populations in Mongolia. Rural Remote Health
531 2008;9:1045.

532 [31] World Health Organization. Microplanning for immunization service delivery
533 using the Reaching Every District (RED) strategy. Geneva PP - Geneva: World
534 Health Organization; 2009.

535

536 **Supplementary materials**537 **Table S1 – National zero-dose prevalence and confidence interval**

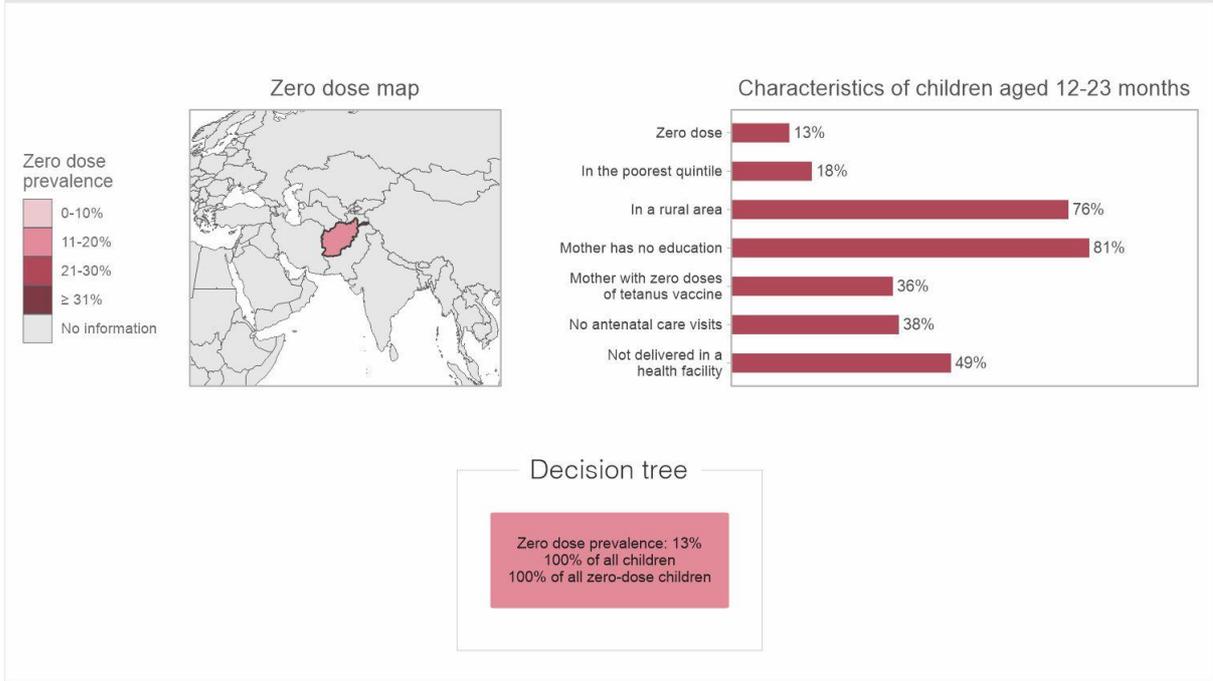
Country	Year	Source	Zero dose prevalence	CI 95%
Afghanistan	2015	DHS	13.1%	(11.4 - 15.0%)
Algeria	2012	MICS	1.6%	(1.0 - 2.4%)
Angola	2015	DHS	19.8%	(17.7 - 22.1%)
Armenia	2015	DHS	0.6%	(0.2 - 1.7%)
Bangladesh	2014	DHS	2.0%	(1.4 - 2.9%)
Belize	2015	MICS	1.6%	(0.5 - 4.9%)
Benin	2017	DHS	11.0%	(9.3 - 12.8%)
Bosnia and Herzegovina	2011	MICS	0.3%	(0.1 - 1.3%)
Burkina Faso	2010	DHS	1.8%	(1.2 - 2.7%)
Burundi	2016	DHS	0.3%	(0.2 - 0.7%)
Cambodia	2014	DHS	2.4%	(1.6 - 3.6%)
Cameroon	2014	MICS	3.9%	(2.8 - 5.3%)
CAR	2010	MICS	9.3%	(7.8 - 11.1%)
Chad	2014	DHS	18.5%	(16.5 - 20.7%)
Colombia	2010	DHS	1.7%	(1.2 - 2.4%)
Comoros	2012	DHS	10.7%	(8.2 - 13.8%)
Congo Brazzaville	2014	MICS	6.3%	(5.0 - 7.9%)
Congo Democratic Republic	2017	MICS	20.5%	(17.9 - 23.3%)
Costa Rica	2011	MICS	0.1%	(0.0 - 0.8%)
Côte d'Ivoire	2016	MICS	11.9%	(9.9 - 14.2%)
Cuba	2014	MICS	0.3%	(0.1 - 0.7%)
Dominican Republic	2014	MICS	4.0%	(3.2 - 5.1%)
Egypt	2014	DHS	0.0%	(0.0 - 0.2%)
El Salvador	2014	MICS	0.4%	(0.1 - 1.9%)
Eswatini	2014	MICS	1.6%	(0.8 - 3.4%)
Ethiopia	2016	DHS	16.2%	(13.5 - 19.3%)
Gabon	2012	DHS	4.8%	(3.3 - 6.9%)
Gambia	2018	MICS	1.5%	(1.0 - 2.3%)
Ghana	2017	MICS	3.2%	(2.1 - 5.0%)
Guatemala	2014	DHS	0.6%	(0.3 - 1.1%)
Guinea	2018	DHS	23.0%	(19.8 - 26.4%)
Guinea Bissau	2014	MICS	3.3%	(2.4 - 4.5%)
Guyana	2014	MICS	2.8%	(1.5 - 5.3%)
Haiti	2016	DHS	9.8%	(7.5 - 12.5%)
Honduras	2011	DHS	0.4%	(0.2 - 1.0%)
India	2015	DHS	6.1%	(5.7 - 6.4%)
Indonesia	2017	DHS	7.0%	(5.9 - 8.1%)
Iraq	2018	MICS	3.0%	(2.4 - 3.9%)
Jamaica	2011	MICS	0.5%	(0.1 - 2.0%)
Jordan	2017	DHS	6.9%	(5.2 - 9.1%)
Kazakhstan	2015	MICS	1.6%	(1.0 - 2.6%)

Country	Year	Source	Zero dose prevalence	CI 95%
Kenya	2014	DHS	1.7%	(1.2 - 2.5%)
Kiribati	2018	MICS	38.2%	(33.4 - 43.4%)
Kosovo	2013	MICS	1.0%	(0.3 - 3.0%)
Kyrgyzstan	2018	MICS	2.4%	(1.2 - 4.6%)
Laos	2017	MICS	13.0%	(11.3 - 14.9%)
Lesotho	2018	MICS	2.5%	(1.3 - 4.6%)
Liberia	2013	DHS	1.7%	(1.0 - 2.7%)
Madagascar	2018	MICS	18.7%	(16.6 - 21.0%)
Malawi	2015	DHS	1.6%	(1.1 - 2.3%)
Maldives	2016	DHS	8.1%	(5.7 - 11.5%)
Mali	2018	DHS	14.5%	(11.9 - 17.4%)
Mauritania	2015	MICS	9.7%	(7.8 - 11.9%)
Mexico	2015	MICS	5.6%	(3.7 - 8.3%)
Moldova	2012	MICS	0.0%	-
Mongolia	2018	MICS	1.3%	(0.7 - 2.4%)
Montenegro	2013	MICS	1.4%	(0.4 - 5.5%)
Mozambique	2015	DHS	5.1%	(3.2 - 8.0%)
Myanmar	2015	DHS	7.9%	(5.4 - 11.3%)
Namibia	2013	DHS	4.4%	(2.9 - 6.7%)
Nepal	2016	DHS	0.8%	(0.4 - 1.8%)
Niger	2012	DHS	4.4%	(3.4 - 5.6%)
Nigeria	2018	DHS	19.5%	(17.9 - 21.3%)
North Macedonia	2011	MICS	2.1%	(0.8 - 5.4%)
Pakistan	2017	DHS	3.7%	(2.6 - 5.3%)
Panama	2013	MICS	1.6%	(0.8 - 3.1%)
Papua New Guinea	2016	DHS	24.2%	(20.9 - 27.9%)
Paraguay	2016	MICS	2.6%	(1.7 - 3.9%)
Peru	2018	DHS	1.1%	(0.7 - 1.6%)
Philippines	2017	DHS	9.5%	(7.9 - 11.4%)
Rwanda	2014	DHS	0.7%	(0.4 - 1.3%)
São Tomé and Príncipe	2014	MICS	1.9%	(0.9 - 4.0%)
Senegal	2018	DHS	2.4%	(1.5 - 3.8%)
Serbia	2014	MICS	0.6%	(0.2 - 1.8%)
Sierra Leone	2017	MICS	2.7%	(1.9 - 3.7%)
South Africa	2016	DHS	5.3%	(3.4 - 8.0%)
South Sudan	2010	MICS	57.6%	(54.1 - 61.0%)
State of Palestine	2014	MICS	0.1%	(0.0 - 0.5%)
Sudan	2014	MICS	12.0%	(10.0 - 14.4%)
Tajikistan	2017	DHS	3.5%	(2.6 - 4.7%)
Tanzania	2015	DHS	2.3%	(1.6 - 3.4%)
Thailand	2015	MICS	3.3%	(1.9 - 5.7%)
Timor-Leste	2016	DHS	19.2%	(16.7 - 22.0%)
Togo	2017	MICS	4.6%	(3.2 - 6.7%)
Tunisia	2018	MICS	0.3%	(0.1 - 1.3%)

Country	Year	Source	Zero dose prevalence	CI 95%
Turkmenistan	2015	MICS	0.1%	(0.0 - 0.9%)
Uganda	2016	DHS	1.5%	(1.0 - 2.1%)
Ukraine	2012	MICS	3.4%	(2.2 - 5.2%)
Vietnam	2013	MICS	1.5%	(0.9 - 2.7%)
Yemen	2013	DHS	16.3%	(14.6 - 18.1%)
Zambia	2018	DHS	1.3%	(0.8 - 2.1%)
Zimbabwe	2019	MICS	5.1%	(3.6 - 7.2%)

538

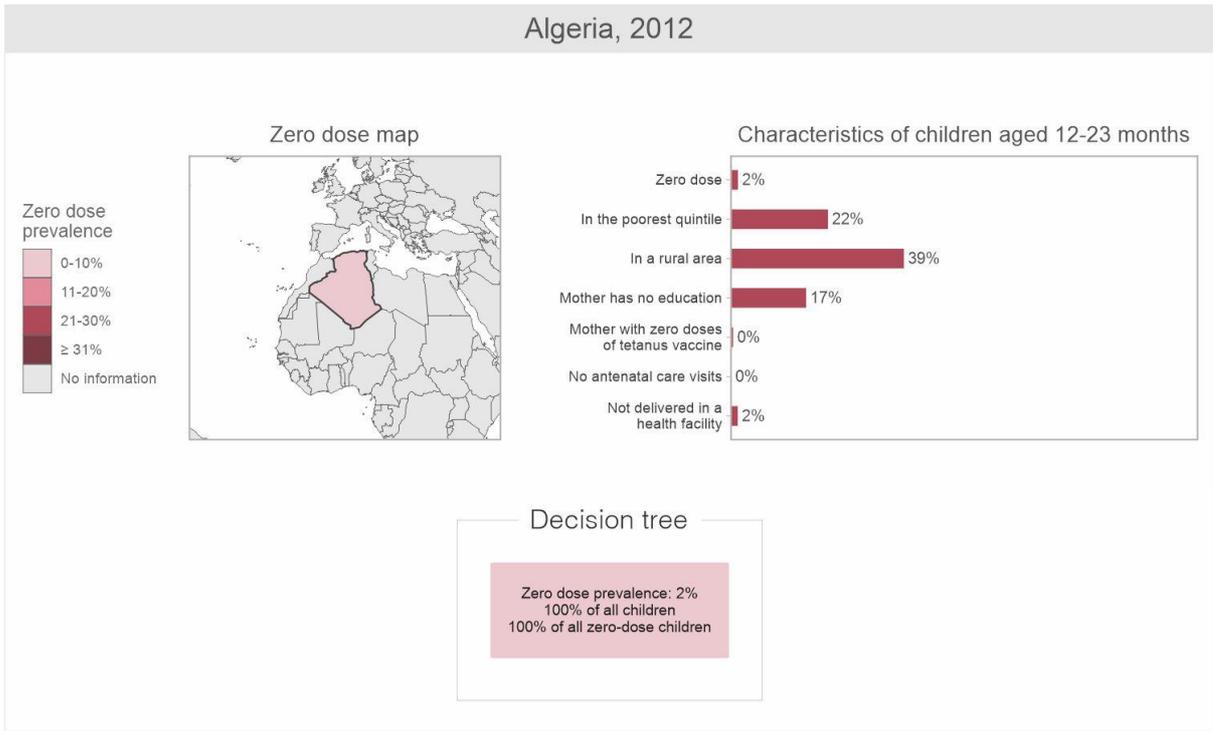
Afghanistan, 2015



539

540

Figure S1 – Afghanistan’s country profile

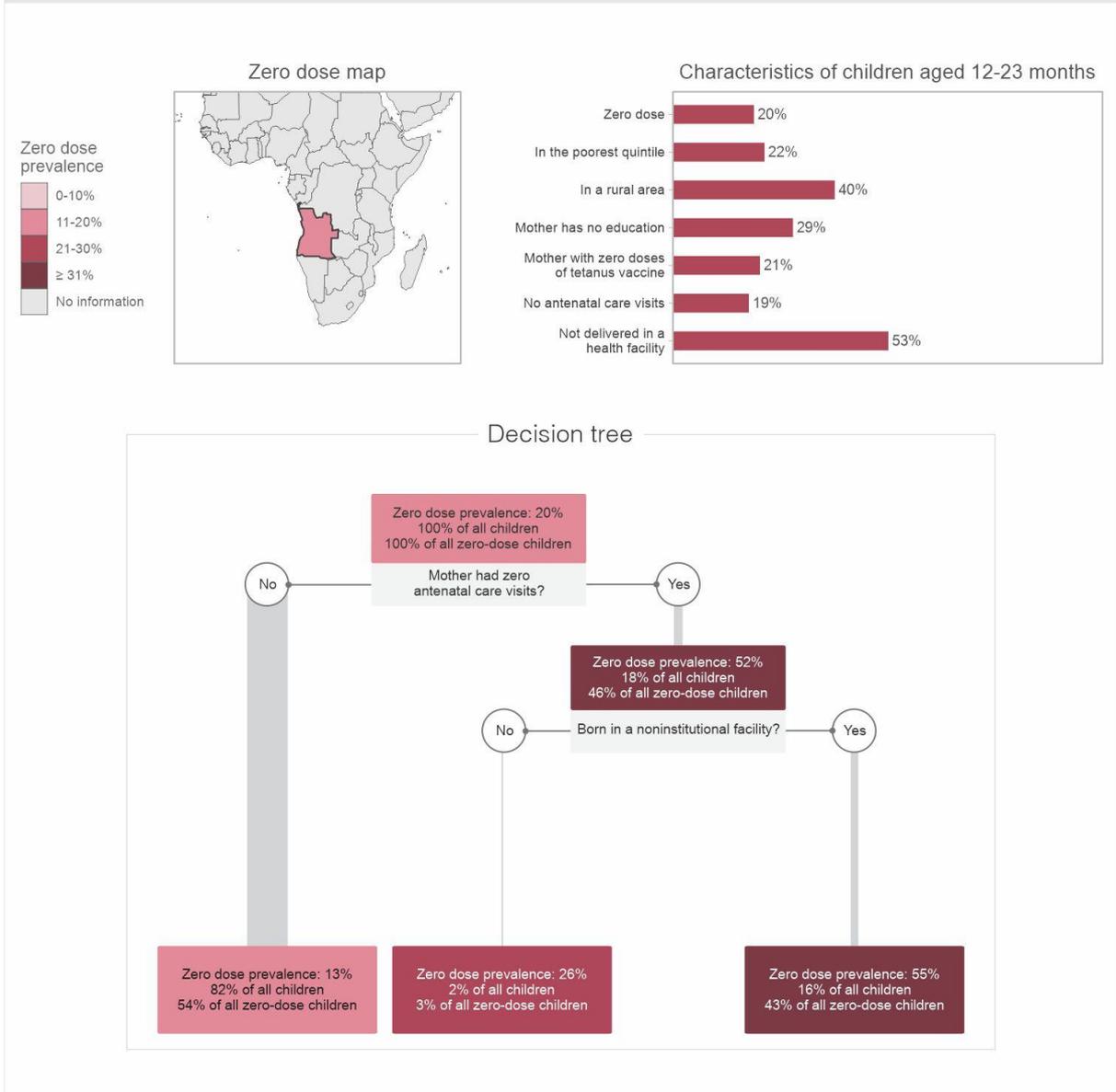


541

542

Figure S2 – Algeria’s country profile

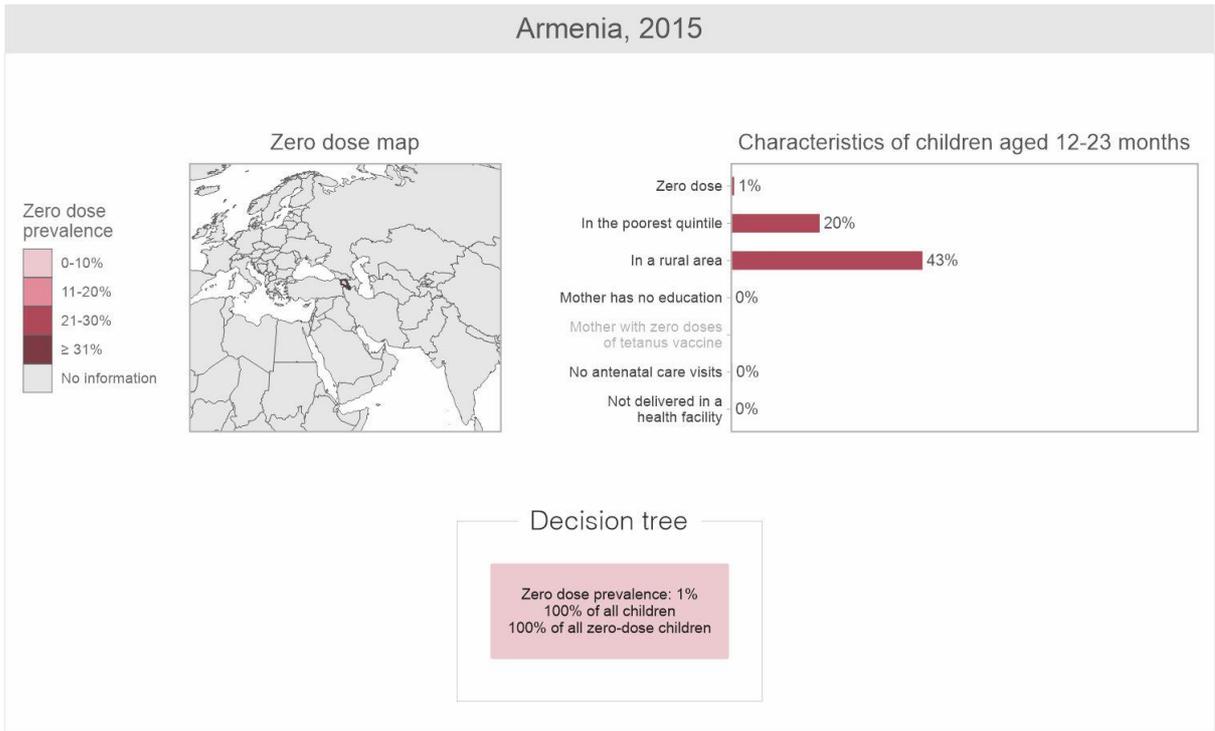
Angola, 2015



543

544

Figure S3 – Angola's country profile

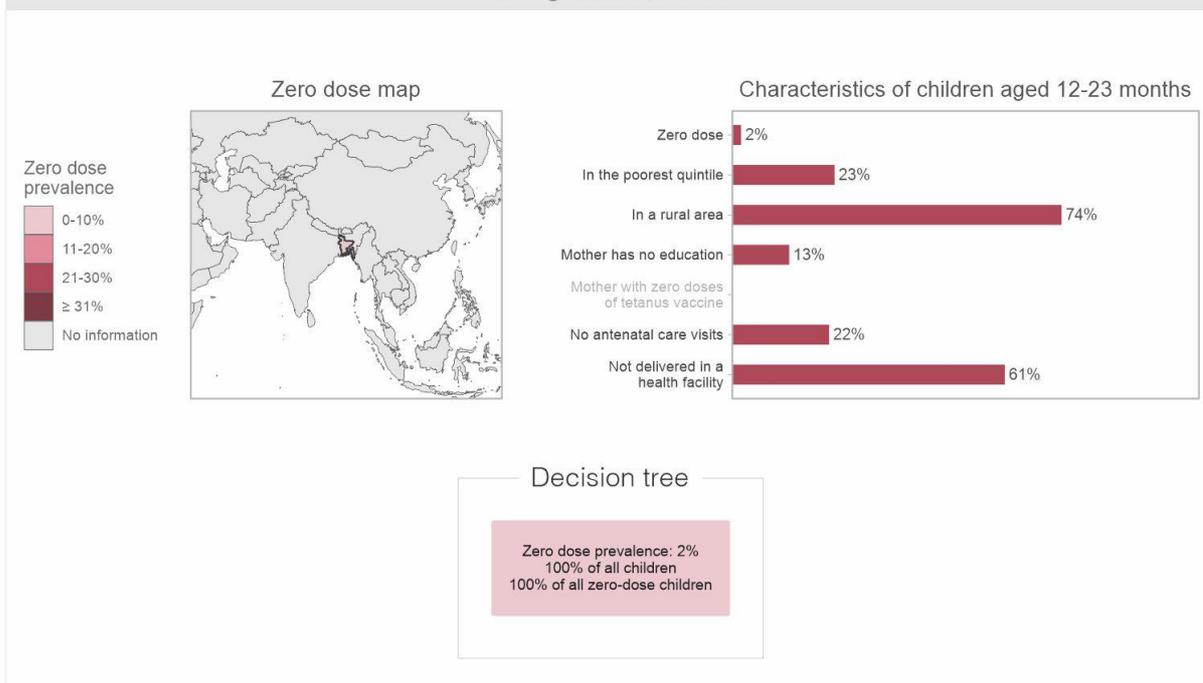


545

546

Figure S4 – Armenia’s country profile

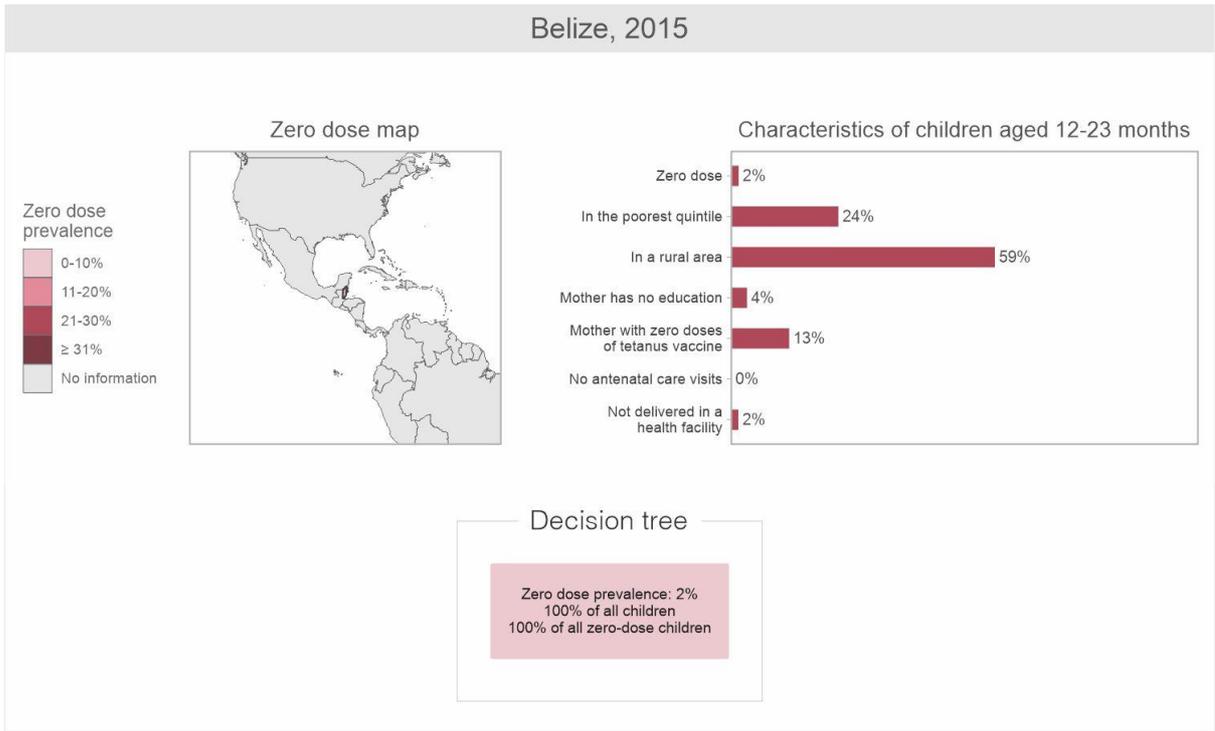
Bangladesh, 2014



547

548

Figure S5 – Bangladesh's country profile

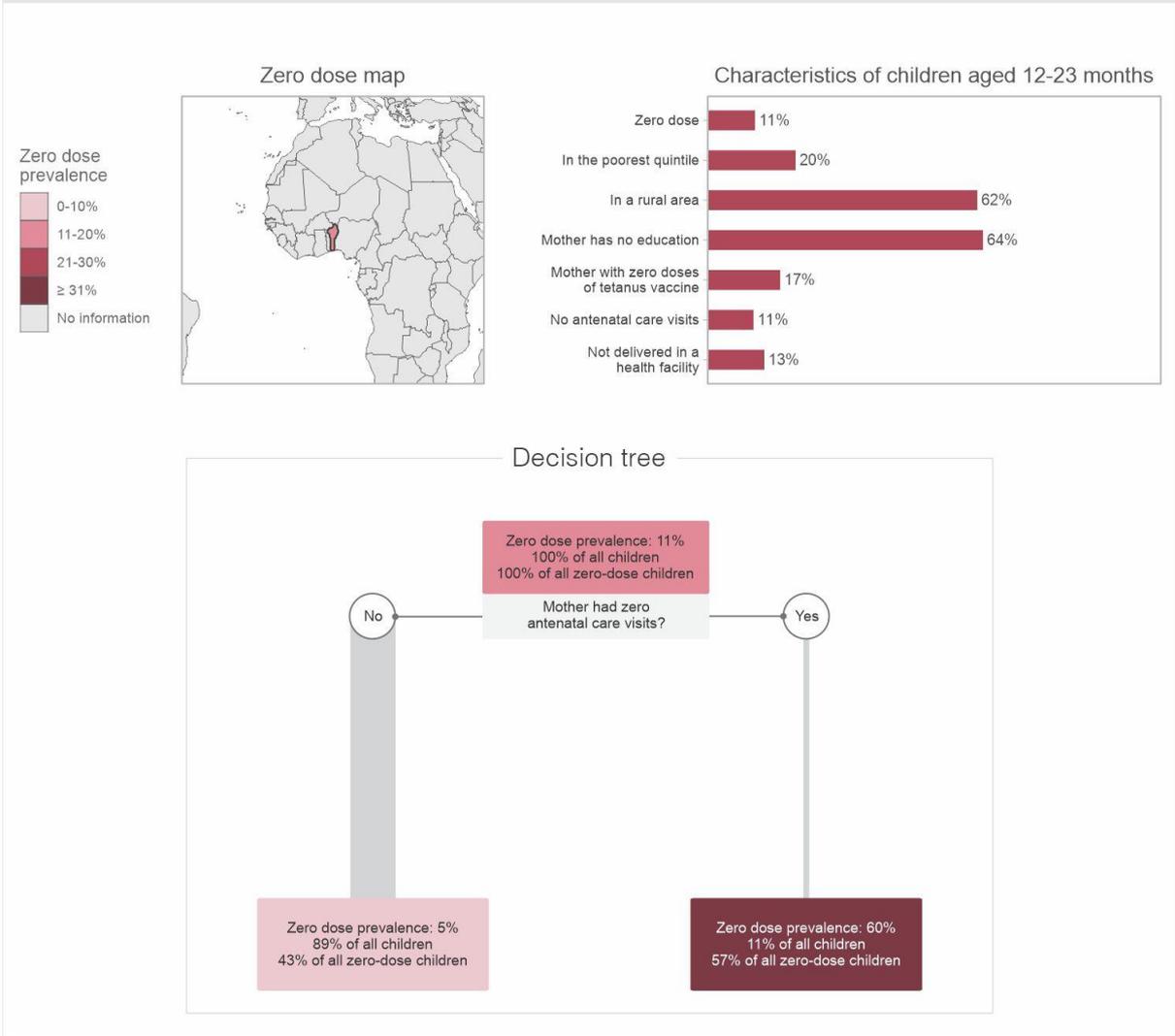


549

550

Figure S6 – Belize’s country profile

Benin, 2017

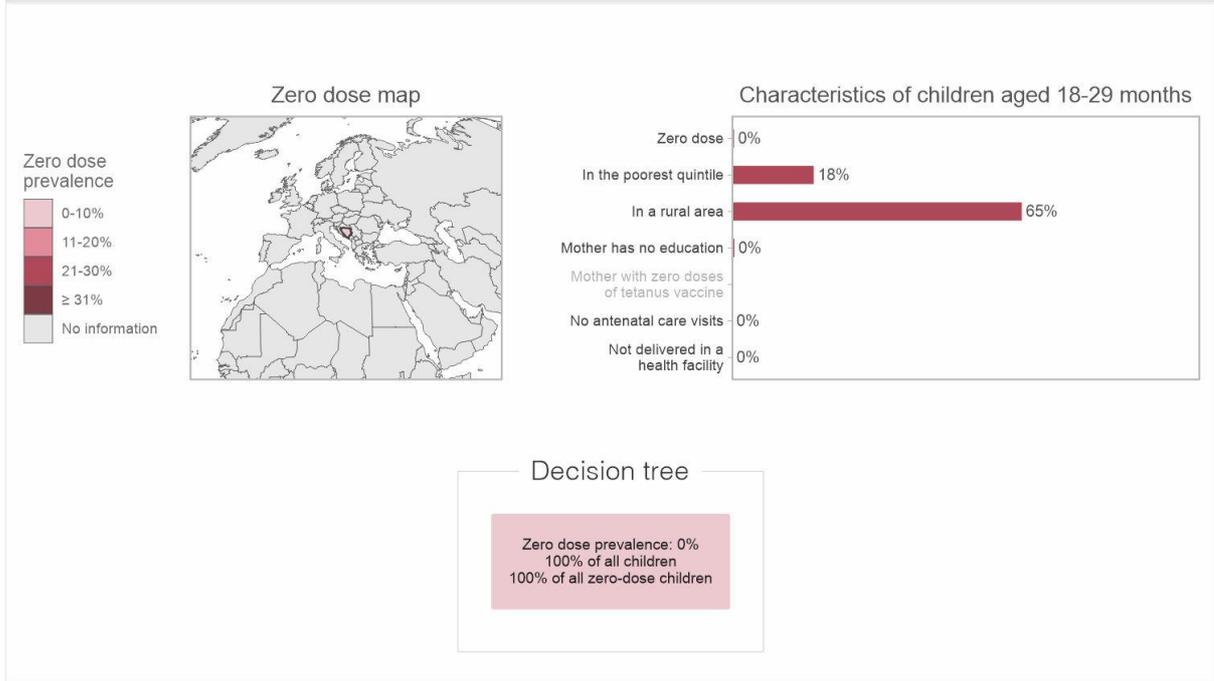


551

552

Figure S7 – Benin's country profile

Bosnia and Herzegovina, 2011

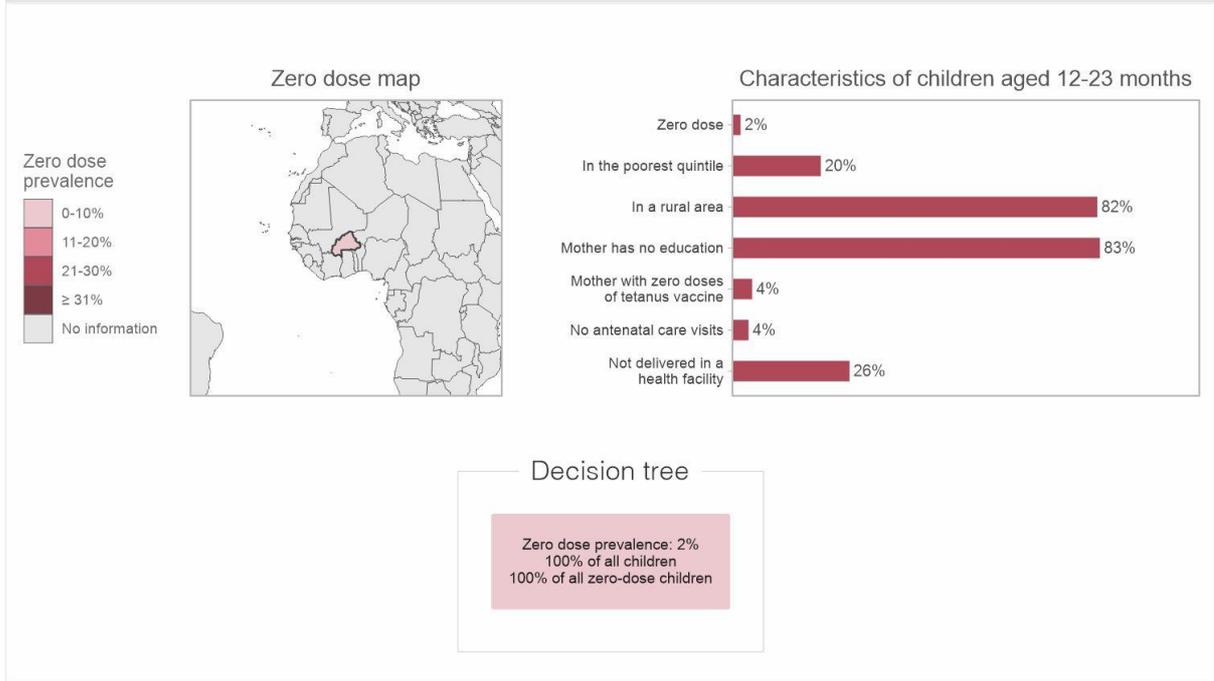


553

554

Figure S8 – Bosnia and Herzegovina’s country profile

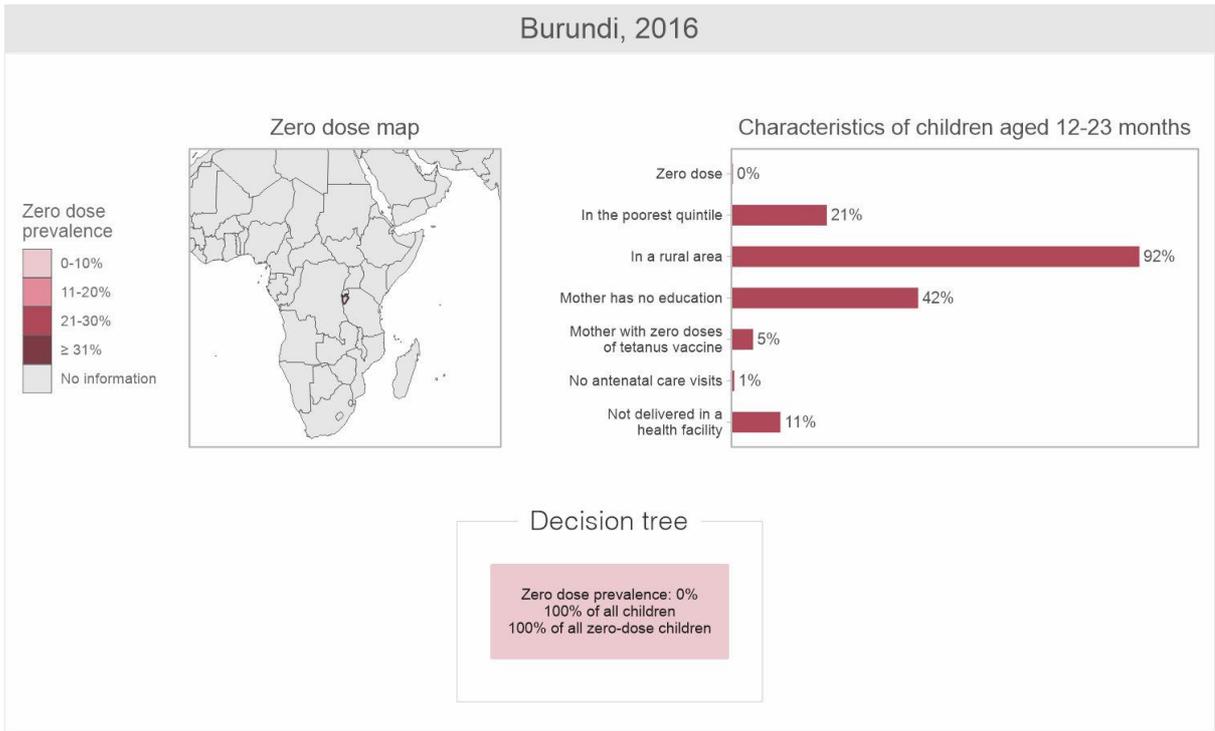
Burkina Faso, 2010



555

556

Figure S9 – Burkina Faso’s country profile

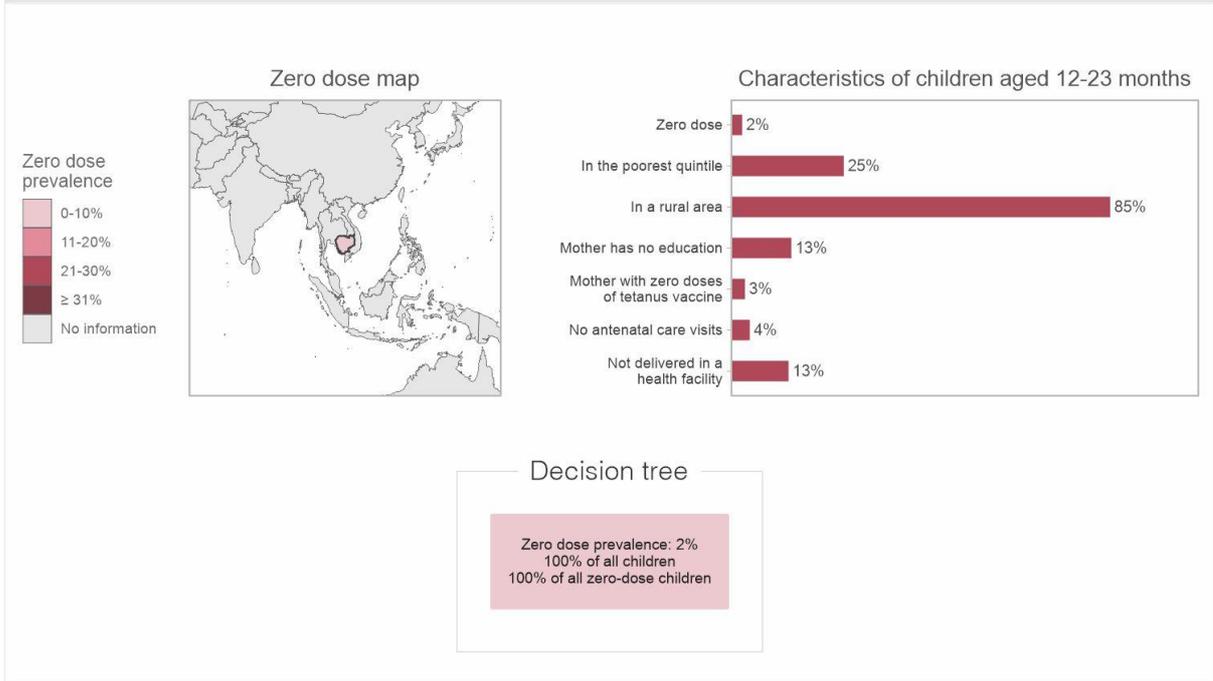


557

558

Figure S10 – Burundi’s country profile

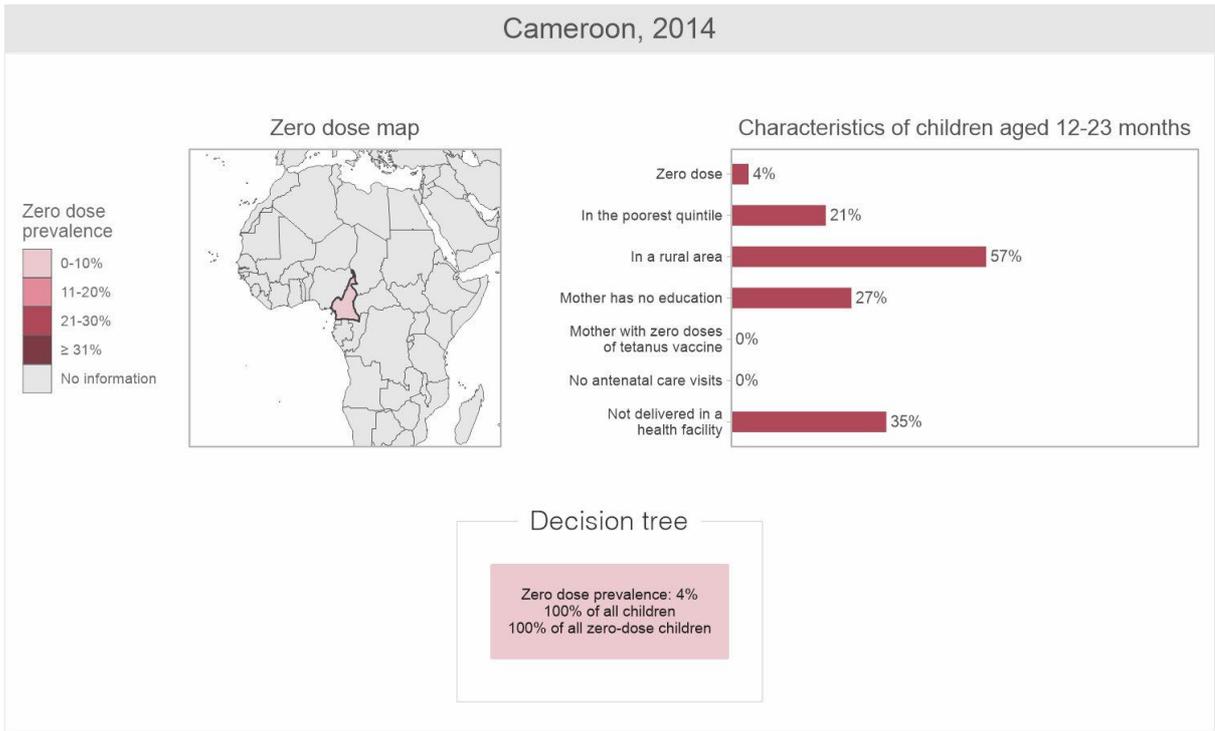
Cambodia, 2014



559

560

Figure S11 – Cambodia's country profile

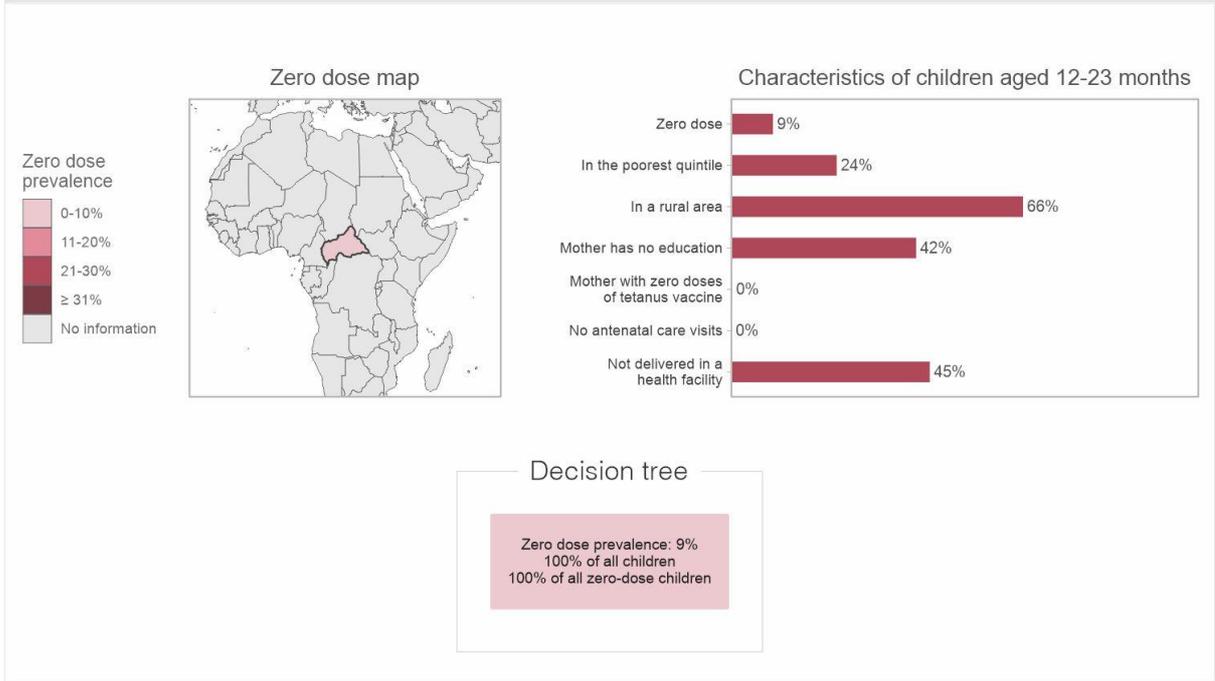


561

562

Figure S12 – Cameroon’s country profile

CAR, 2010

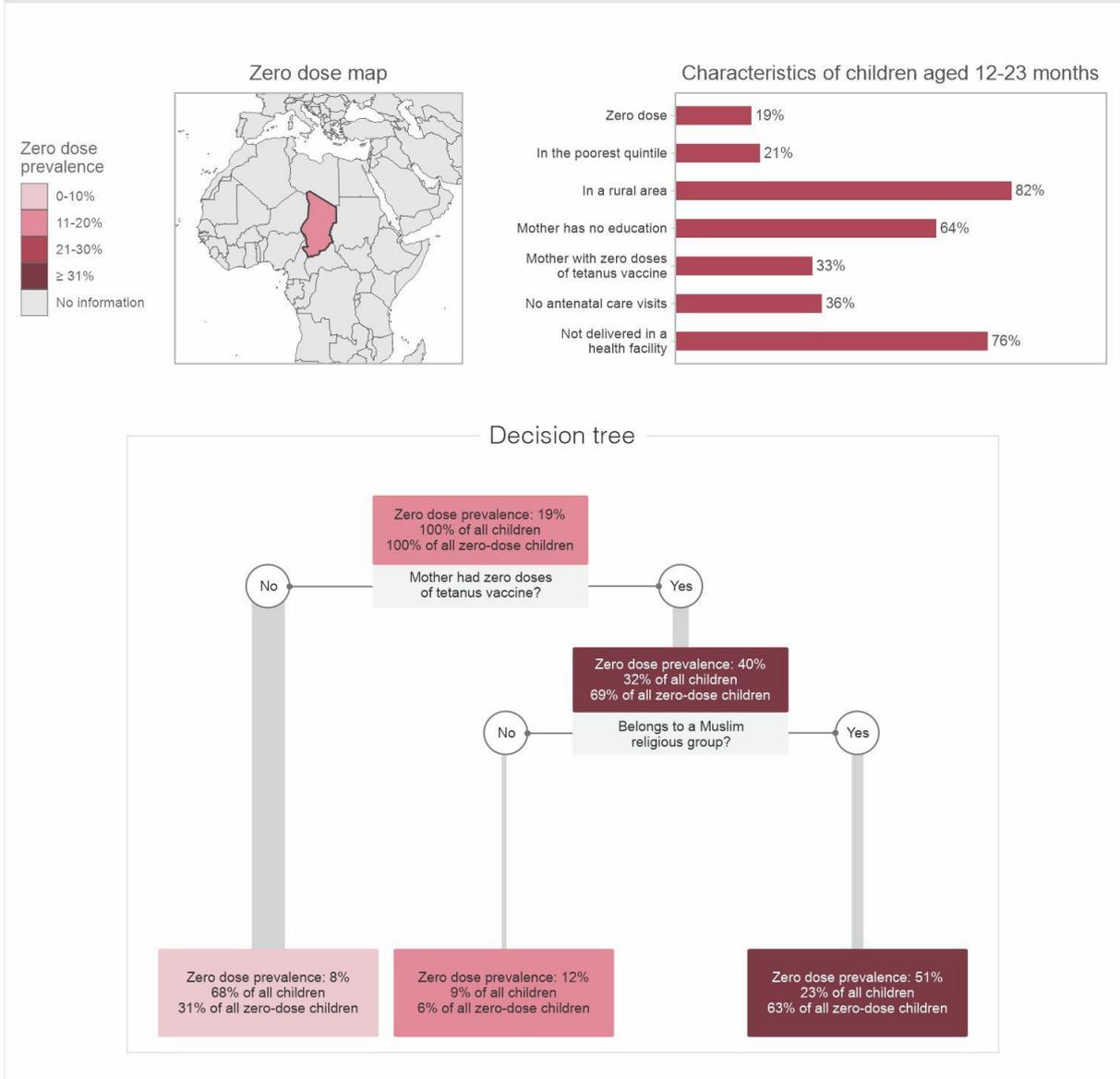


563

564

Figure S13 – Central African Republic's country profile

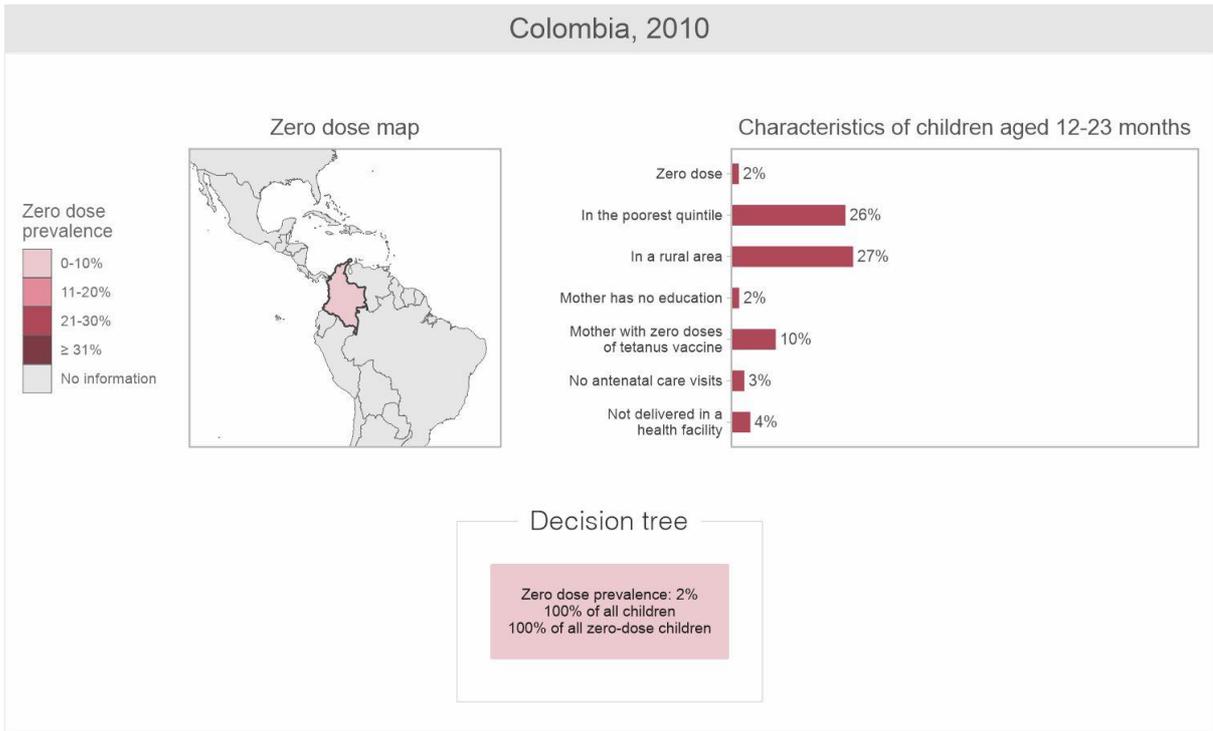
Chad, 2014



565

566

Figure S14 – Chad's country profile



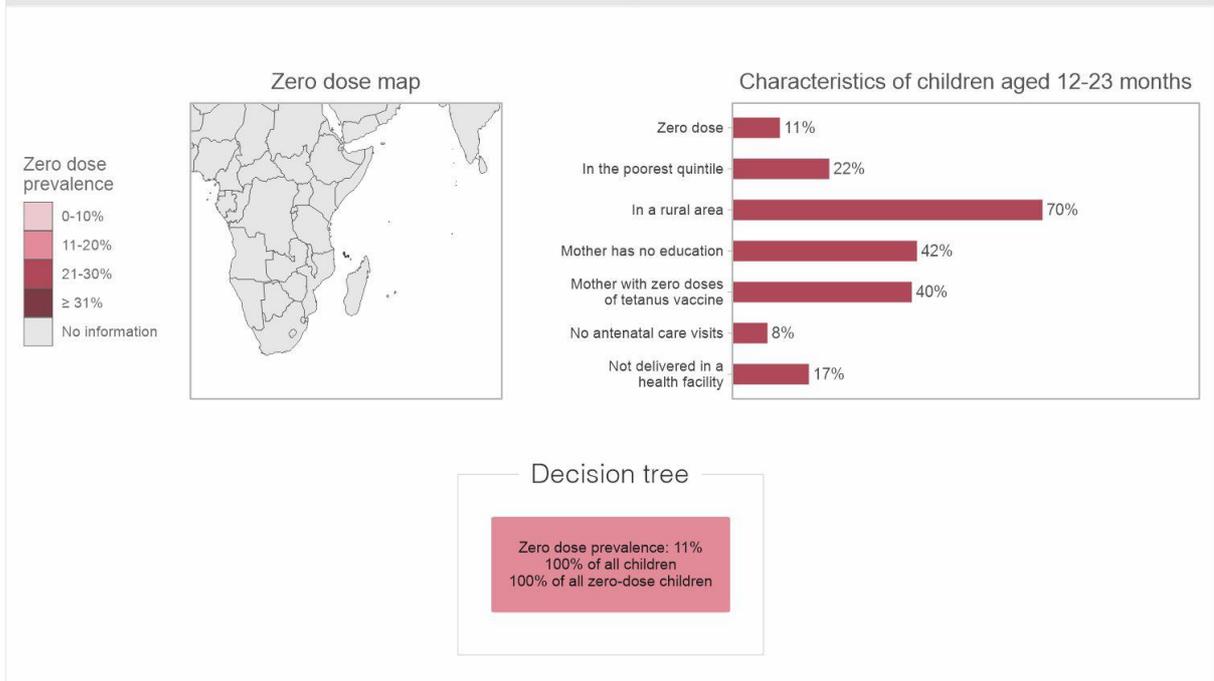
567

568

Figure S15 – Colombia’s country profile

569

Comoros, 2012



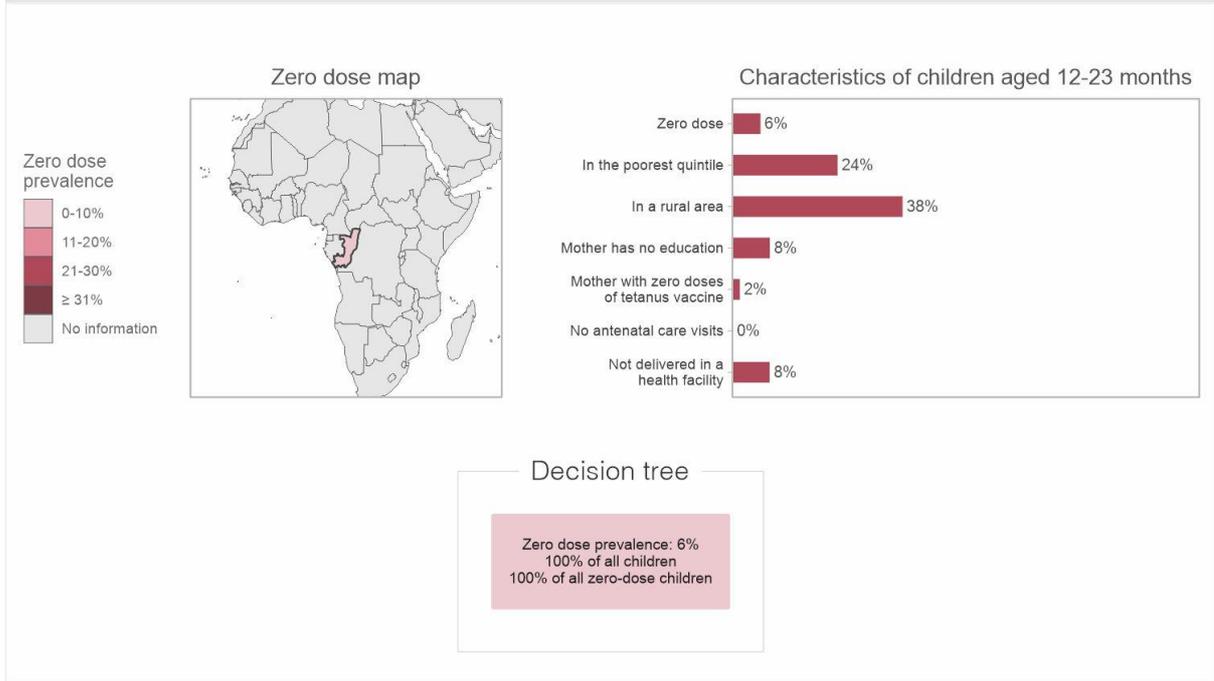
570

571

572

Figure S16 – Comoro's country profile

Congo Brazzaville, 2014



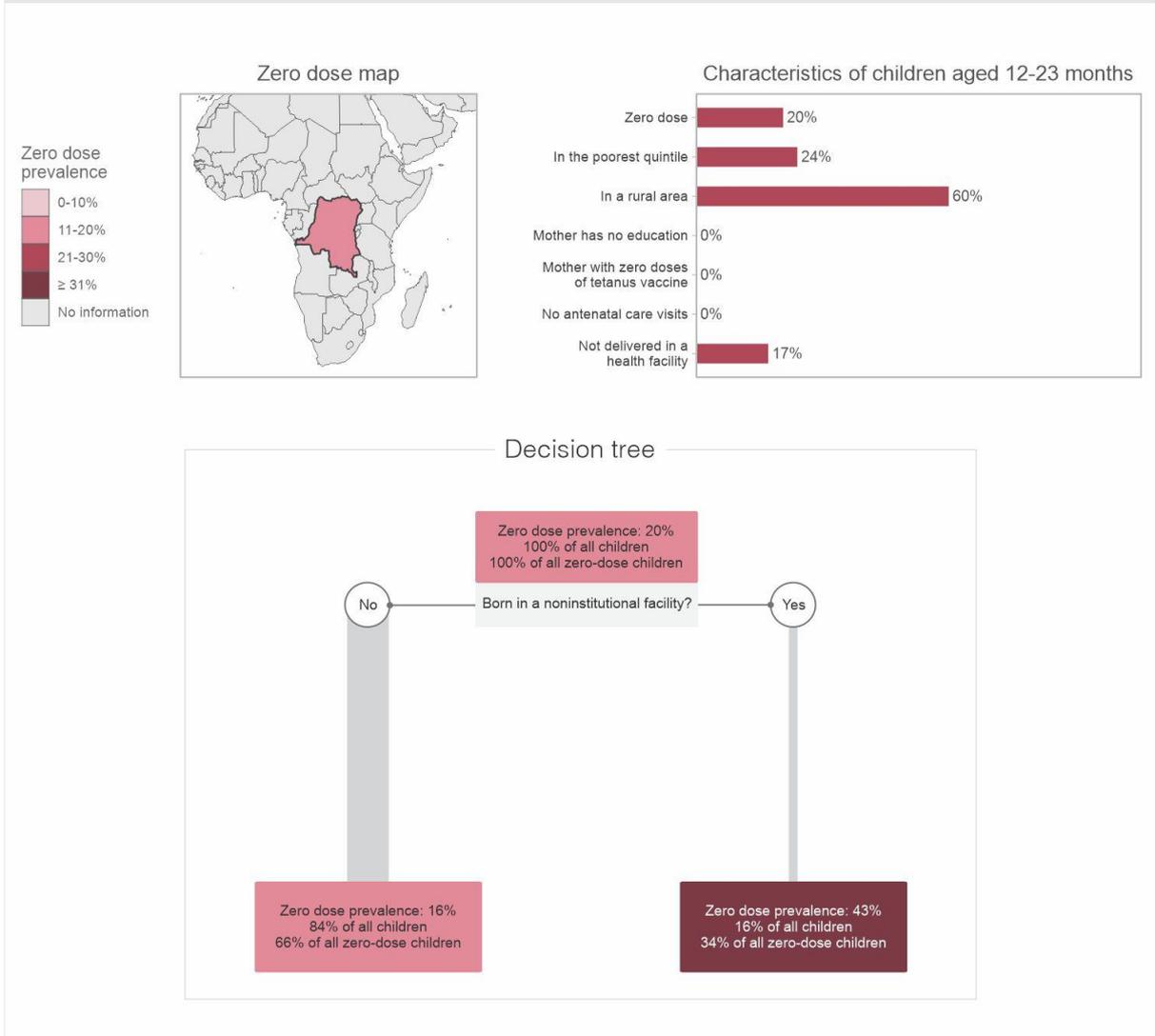
573

574

Figure S17 – Congo Brazzaville’s country profile

575

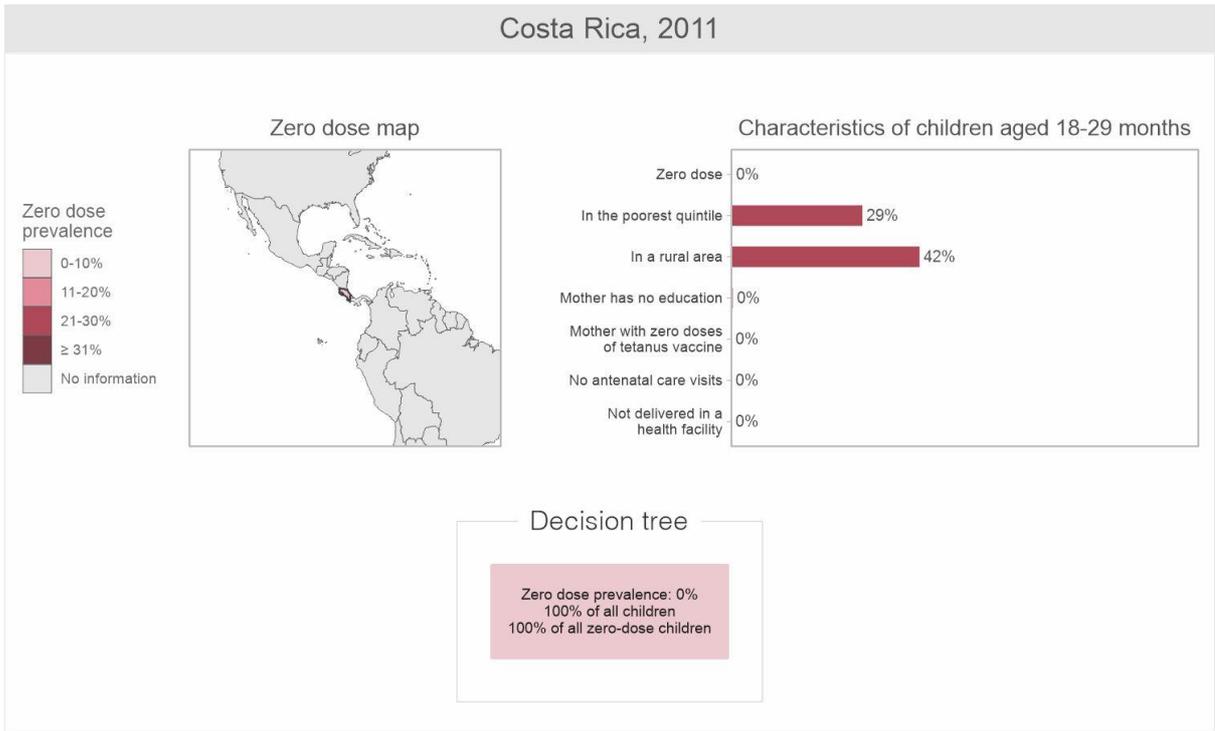
Congo, DR, 2017



576

577

Figure S18 – Congo, Democratic Republic’s country profile

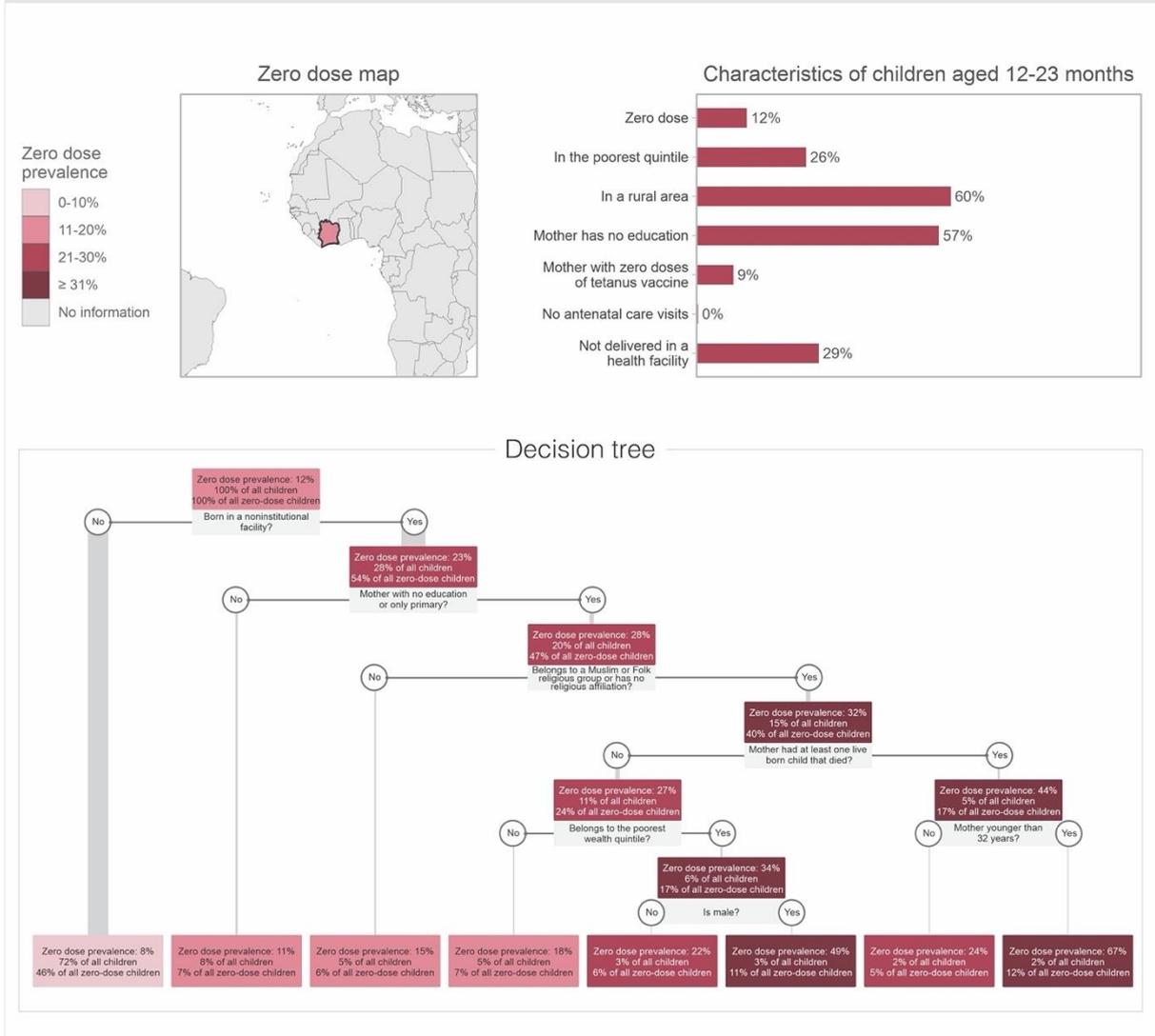


578

579

Figure S19 – Costa Rica’s country profile

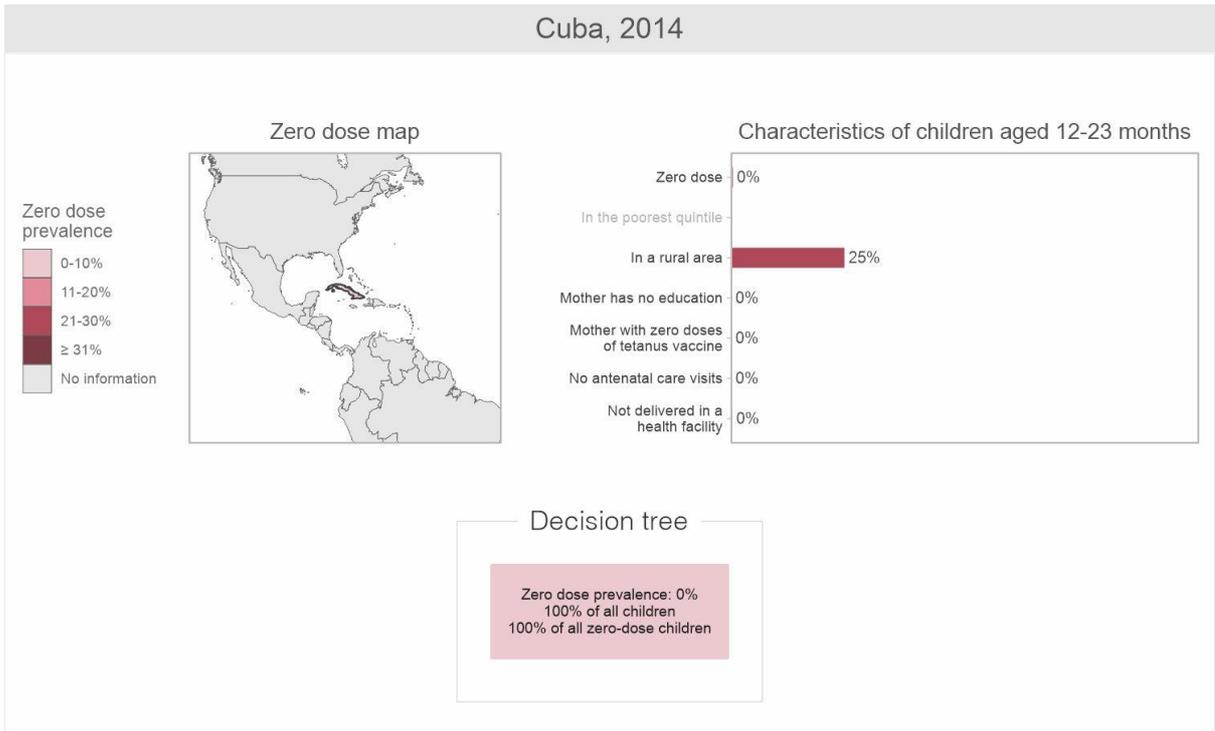
Côte d'Ivoire, 2016



580

581

Figure S20 – Côte d'Ivoire country profile

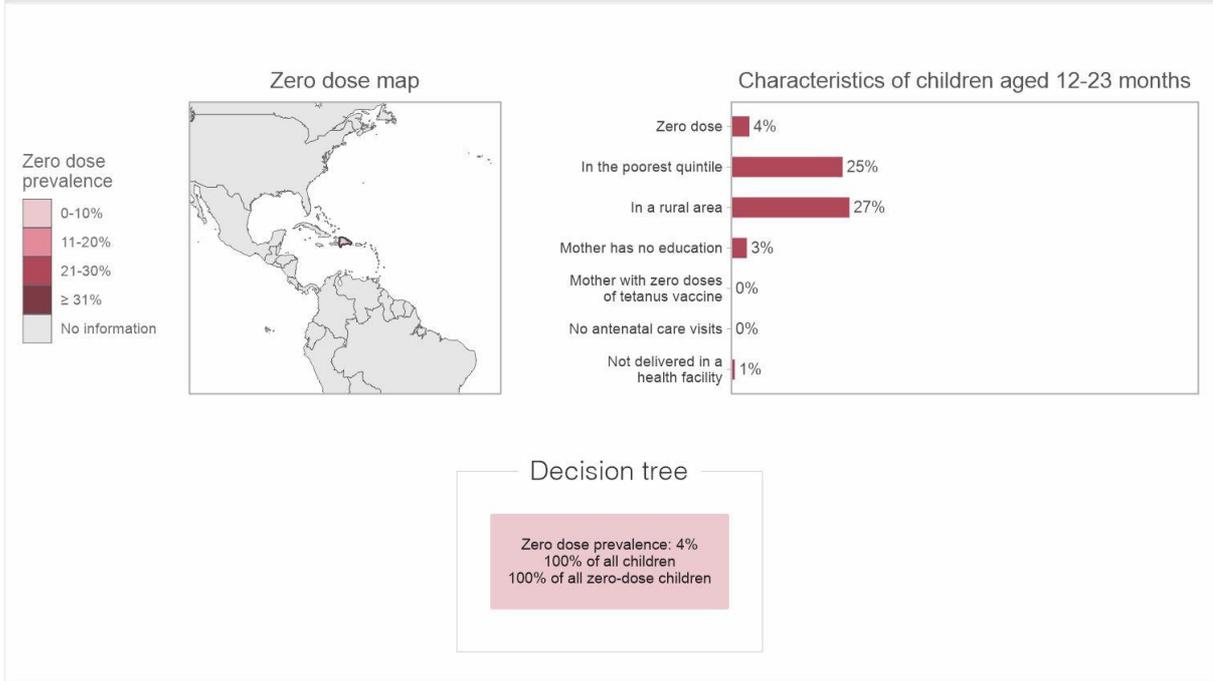


582

583

Figure S21 – Cuba’s country profile

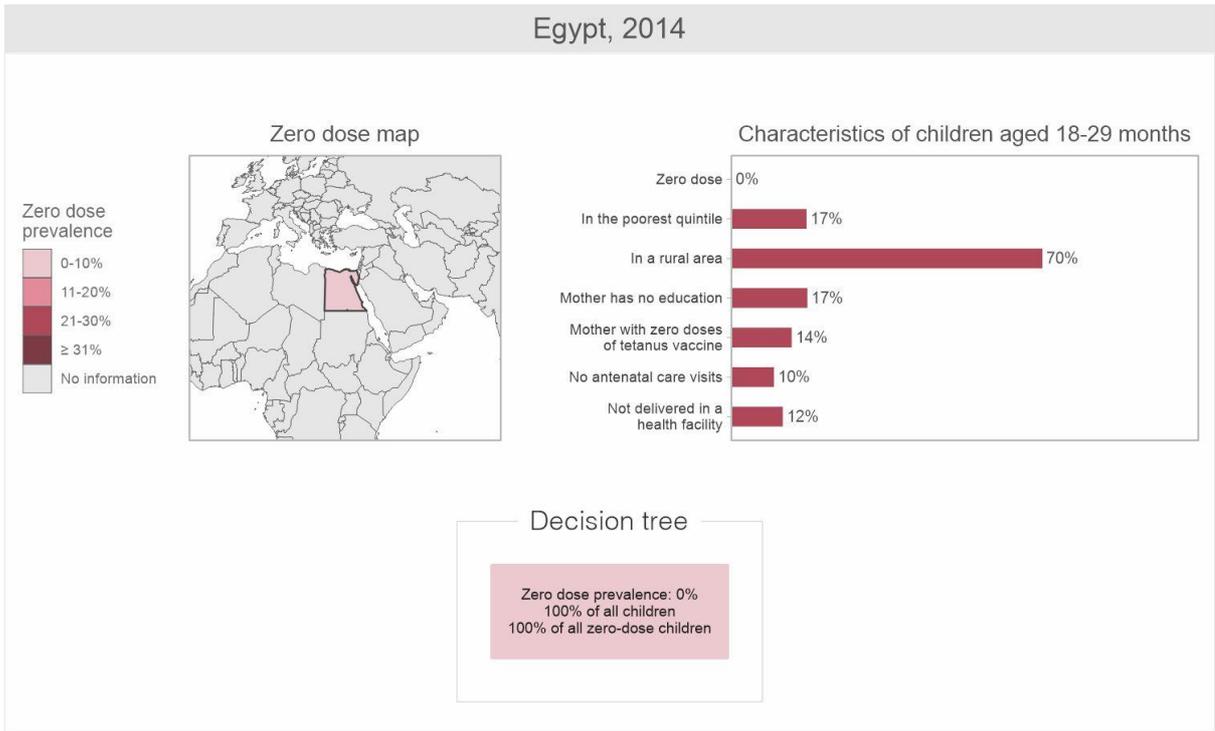
Dominican Republic, 2014



584

585

Figure S22 – Dominican Republic's country profile

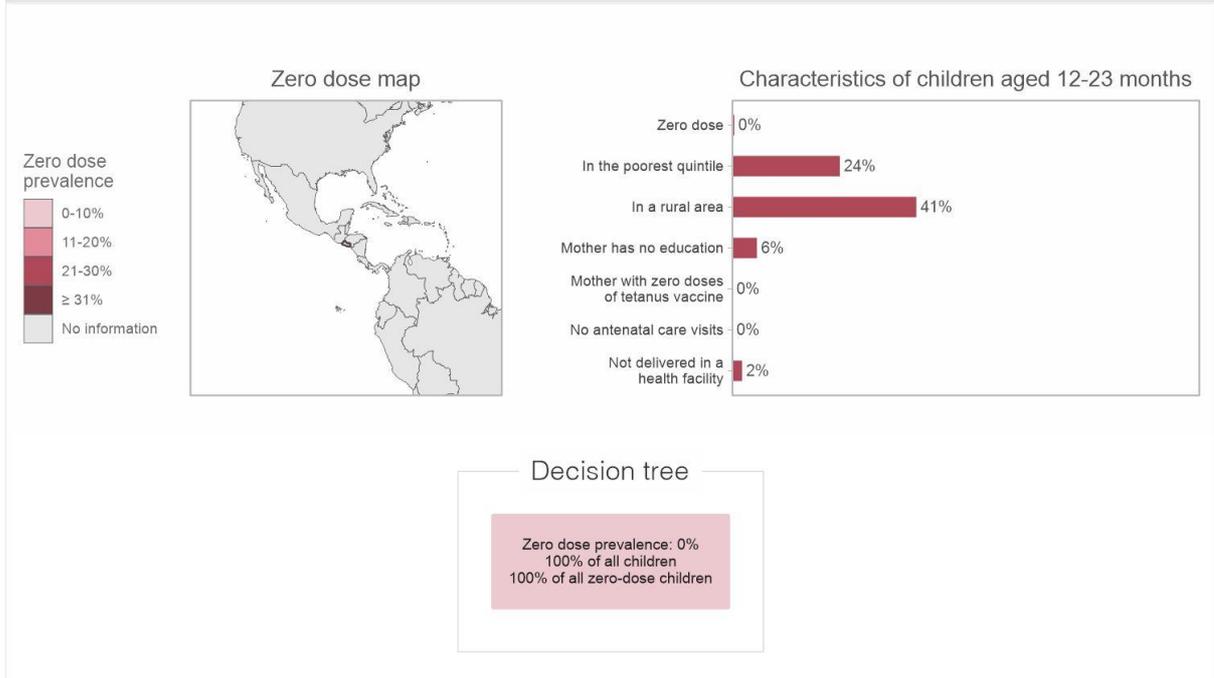


586

587

Figure S23 – Egypt's country profile

El Salvador, 2014

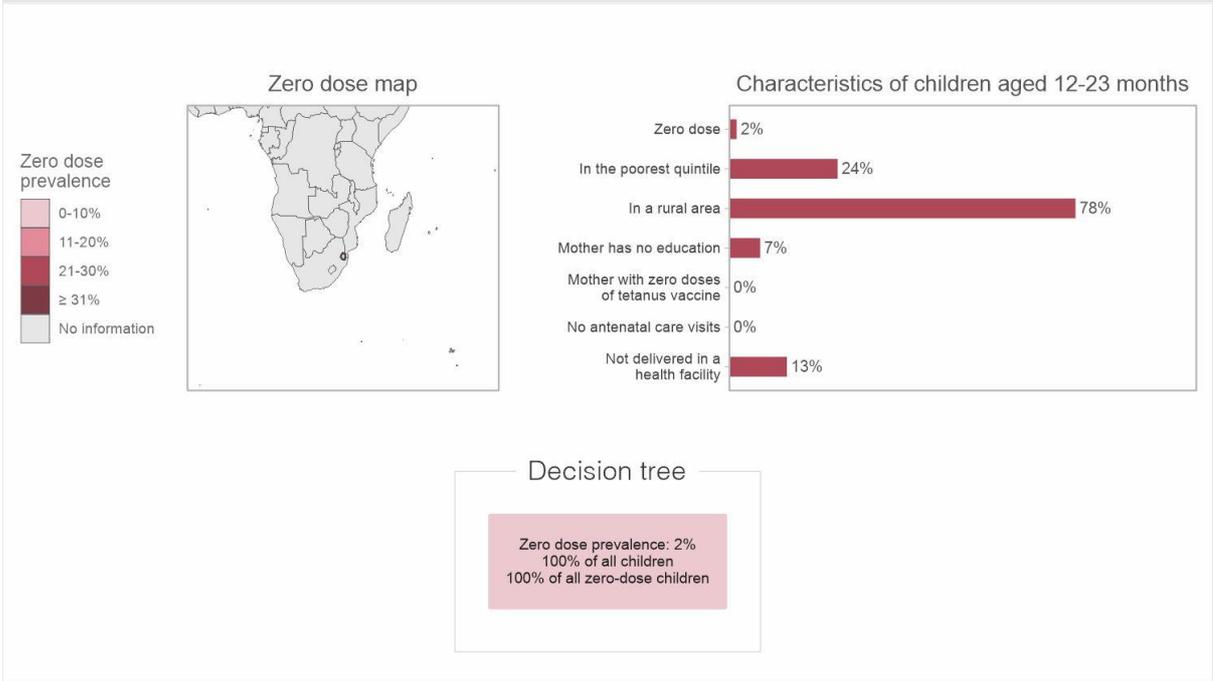


588

589

Figure S24 – El Salvador's country profile

Eswatini, 2014

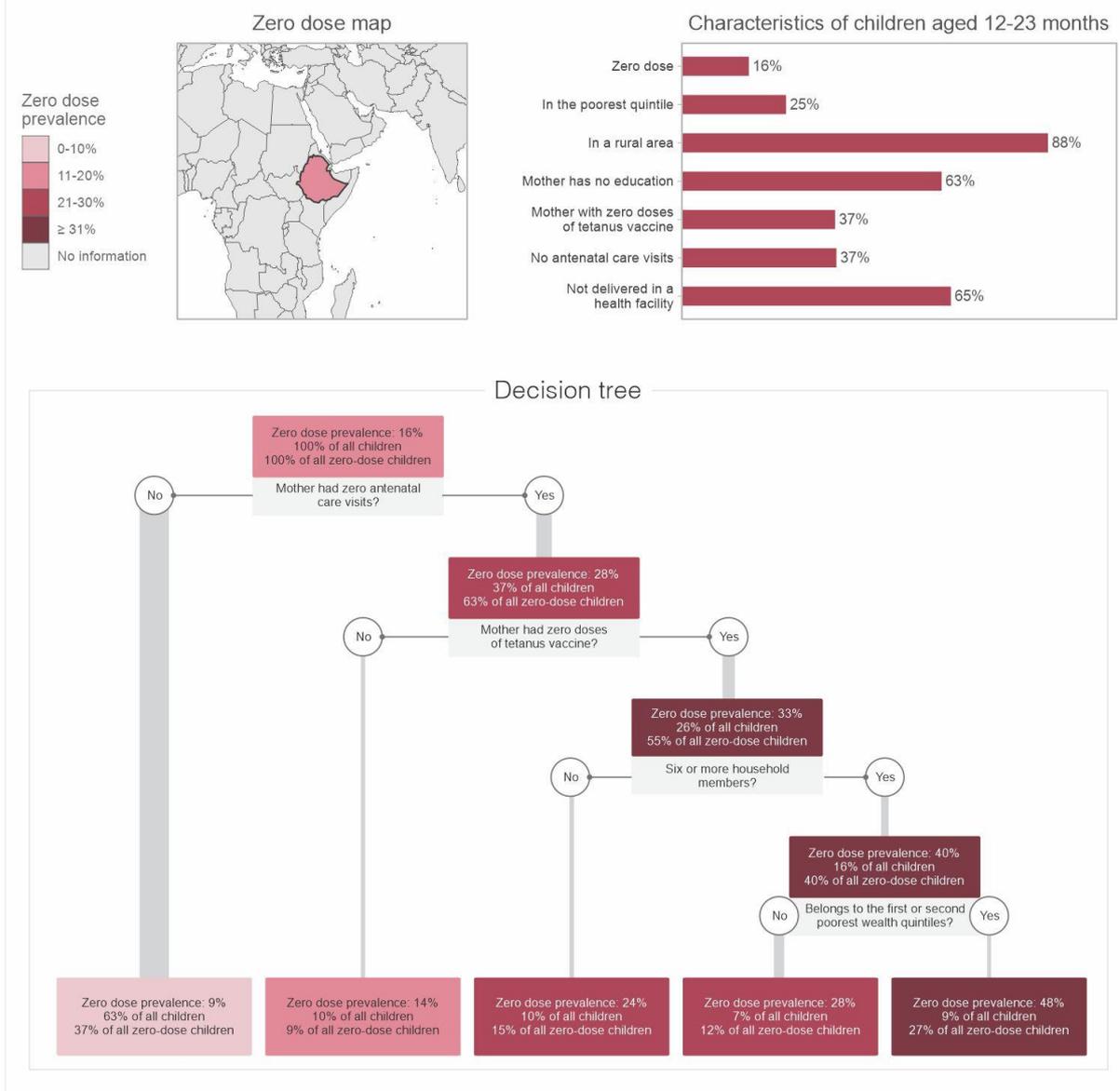


590

591

Figure S25 – Eswatini's country profile

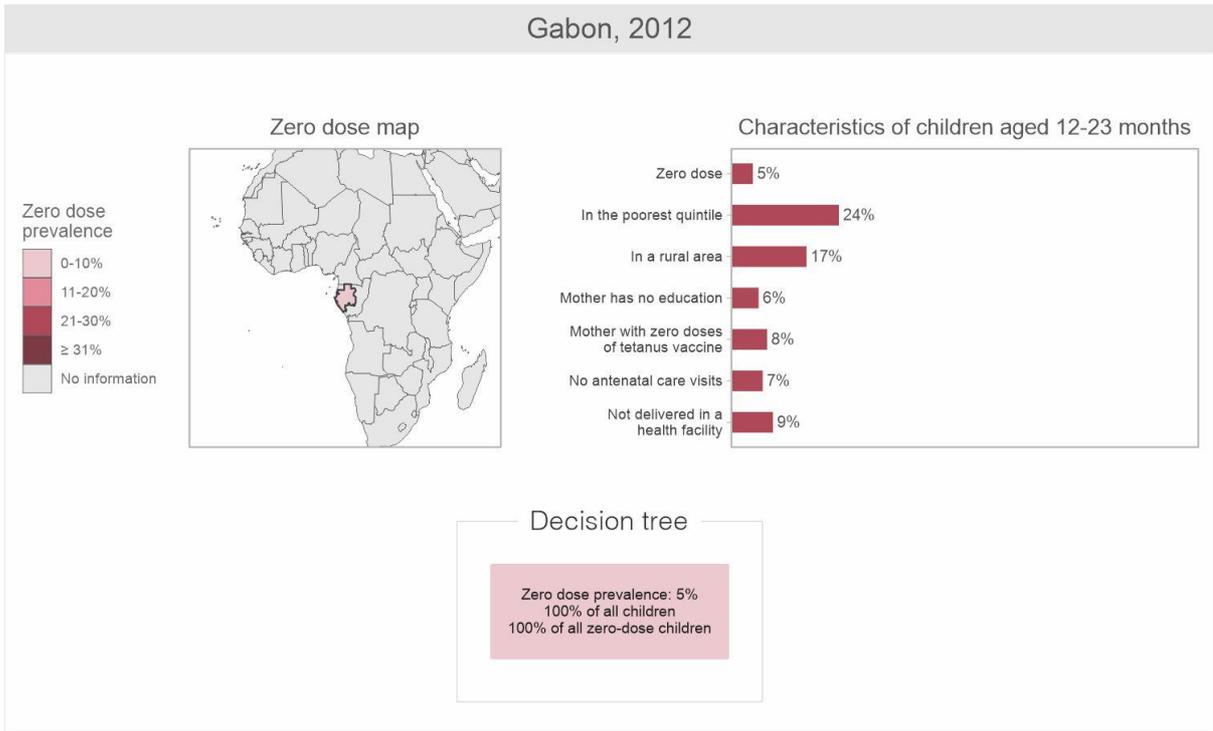
Ethiopia, 2016



592

593

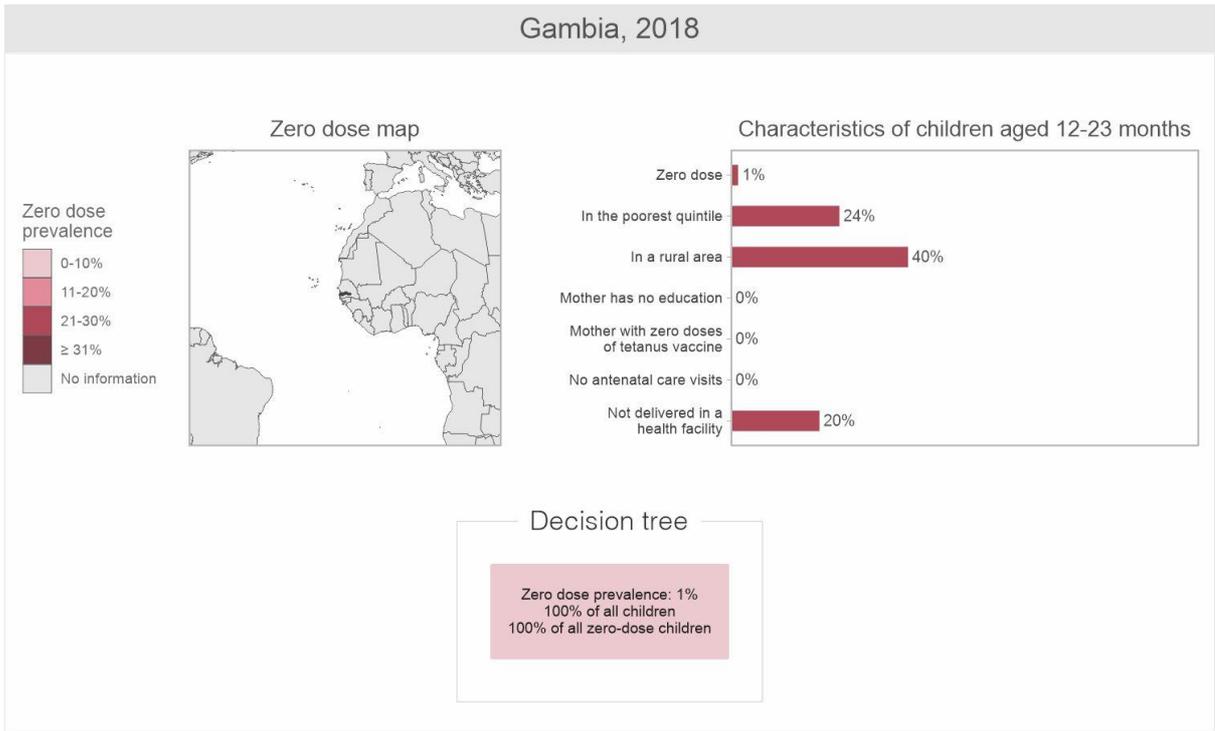
Figure S26 – Ethiopia’s country profile



594

595

Figure S27 – Gabon’s country profile

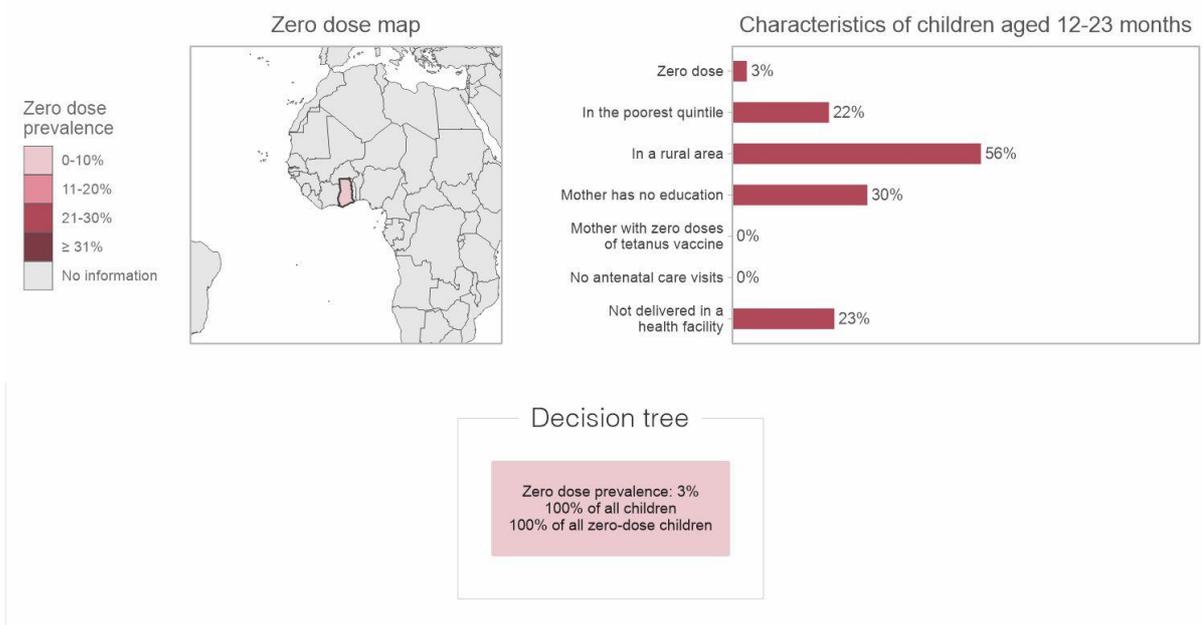


596

597

Figure S28 – Gambia’s country profile

Ghana, 2017

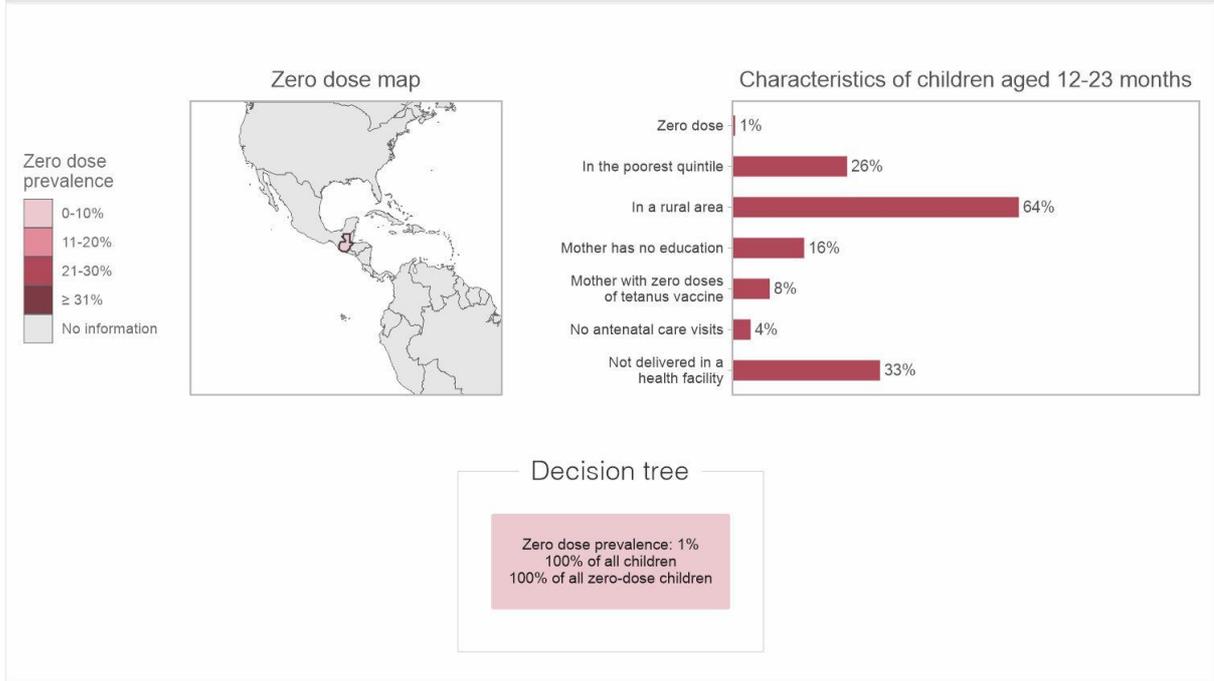


598

599

Figure S29 – Ghana’s country profile

Guatemala, 2014

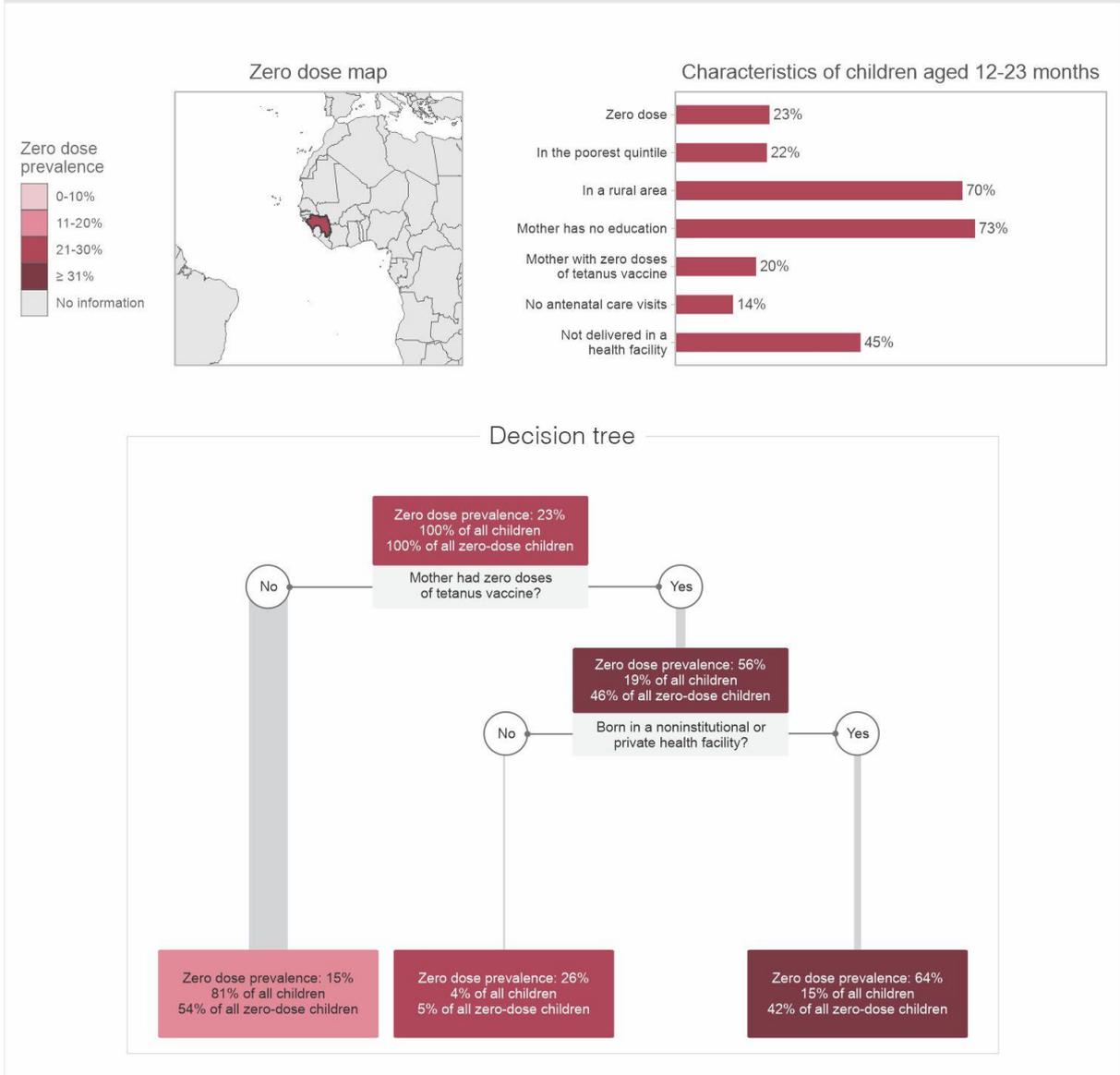


600

601

Figure S30 – Guatemala's country profile

Guinea, 2018

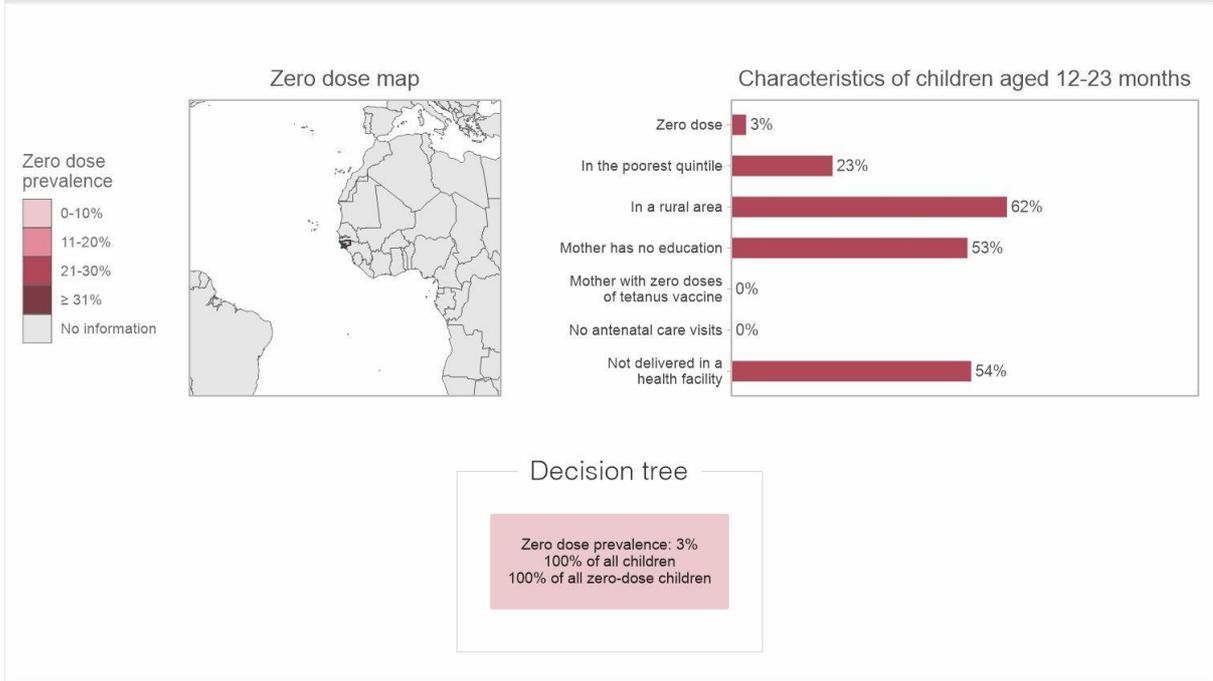


602

603

Figure S31 – Guinea’s country profile

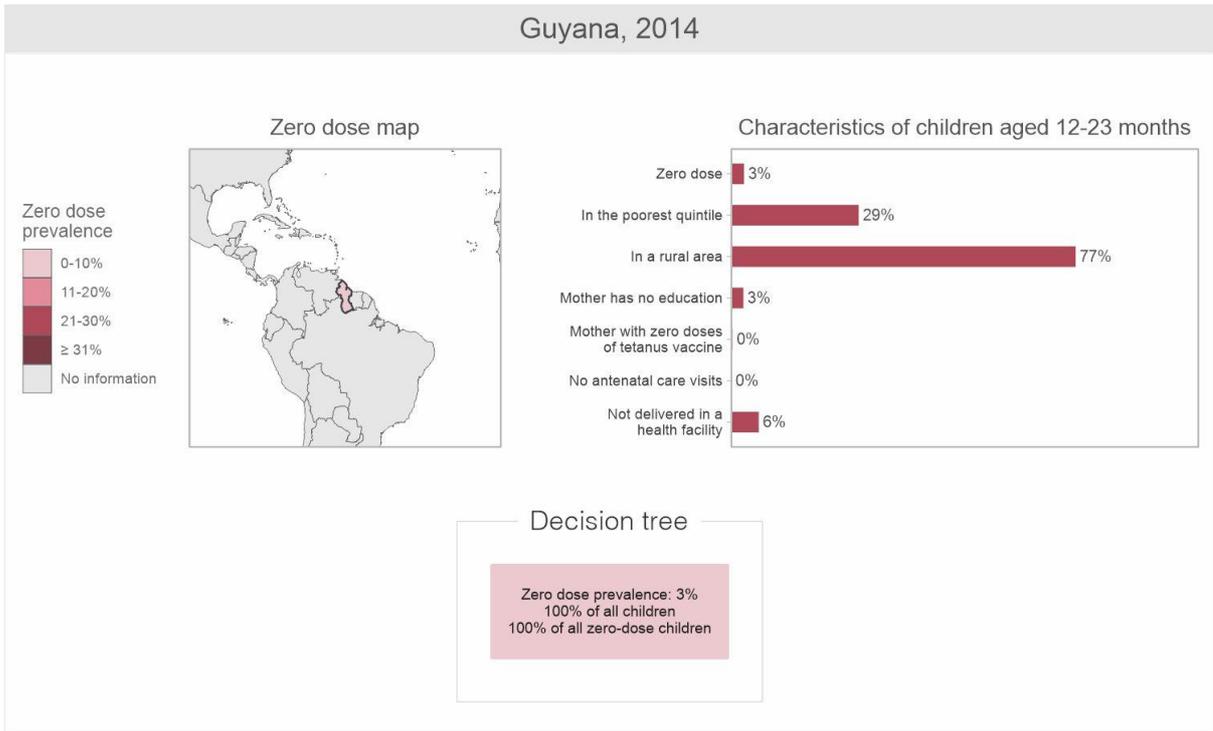
Guinea Bissau, 2014



604

605

Figure S32 – Guinea Bissau’s country profile



606

607

Figure S33 – Guyana’s country profile

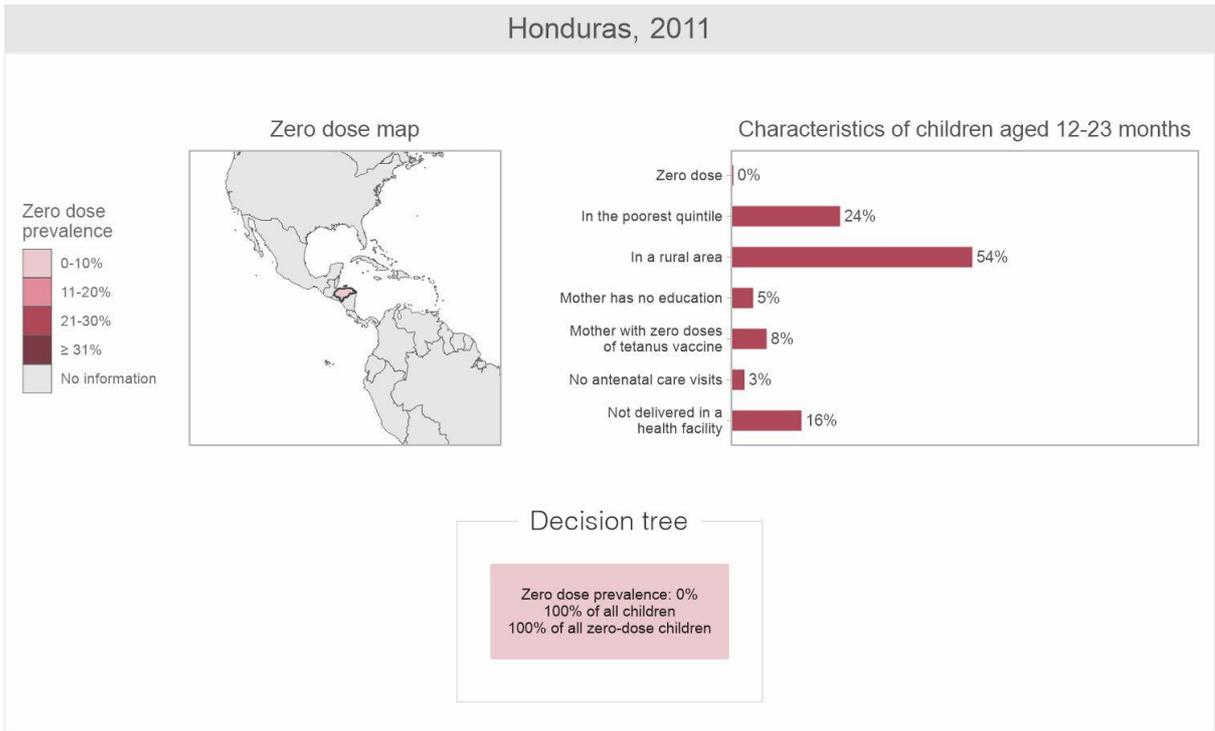
Haiti, 2016



608

609

Figure S34 – Haiti’s country profile

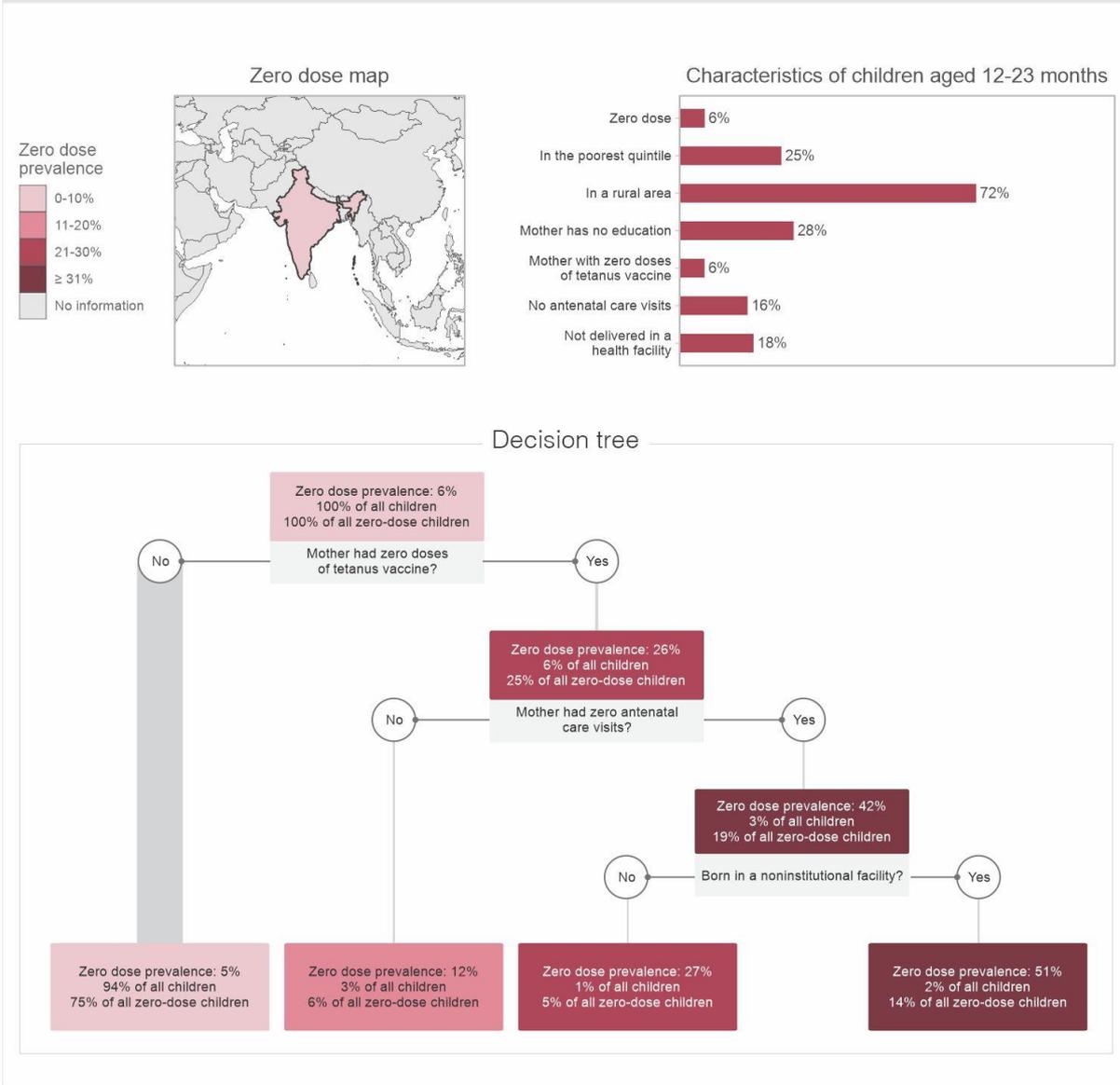


610

611

Figure S35 – Honduras's country profile

India, 2015

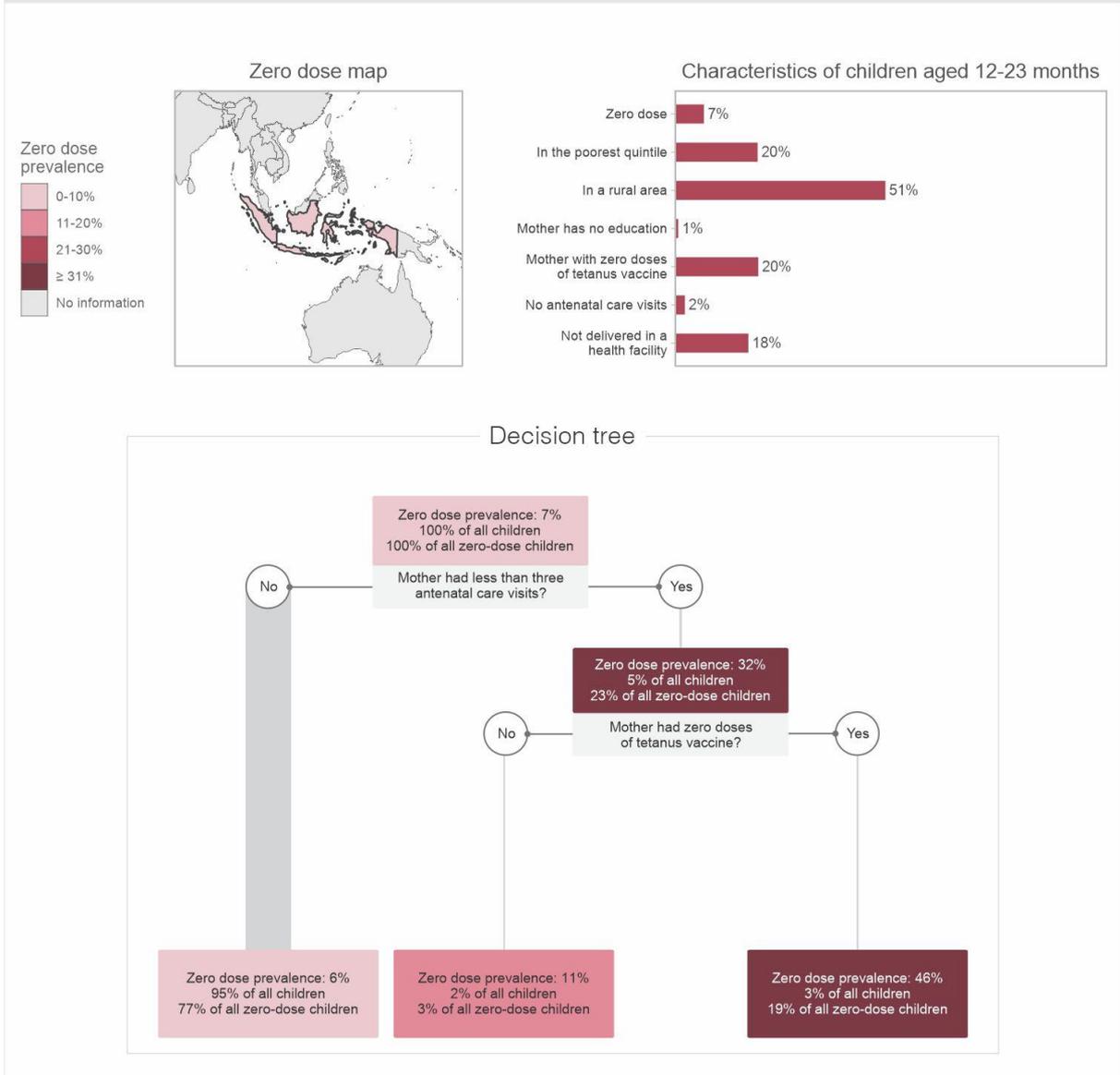


612

613

Figure S36 – India’s country profile

Indonesia, 2017

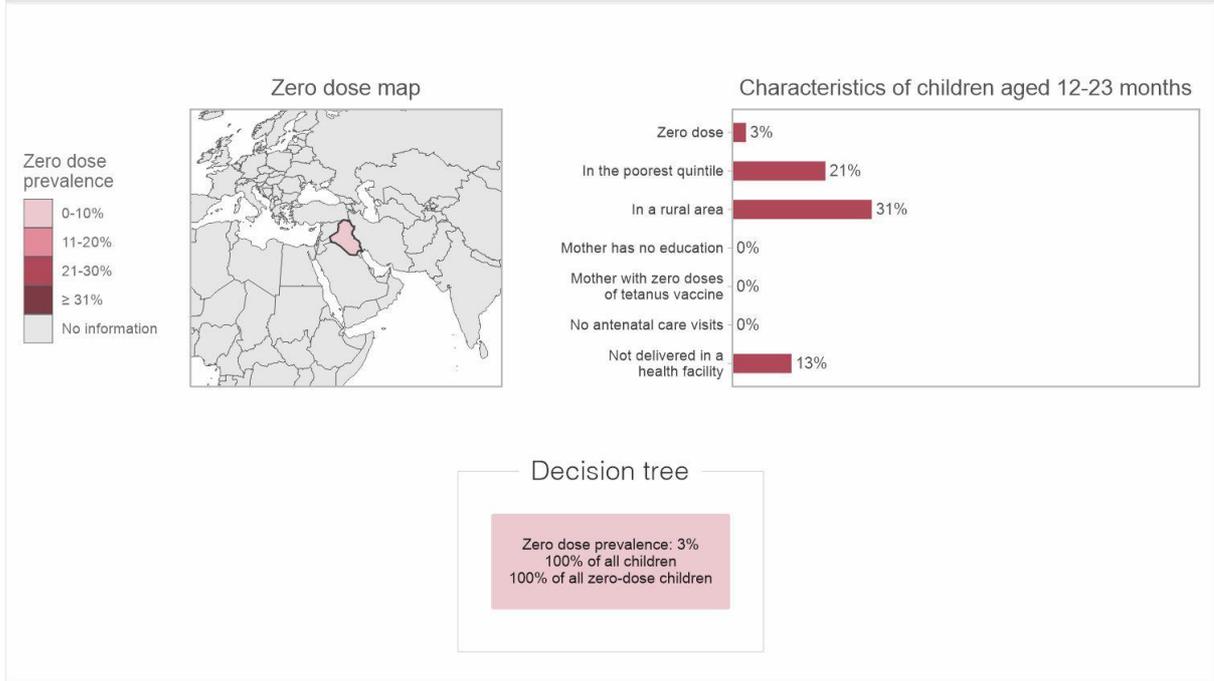


614

615

Figure S37 – Indonesia’s country profile

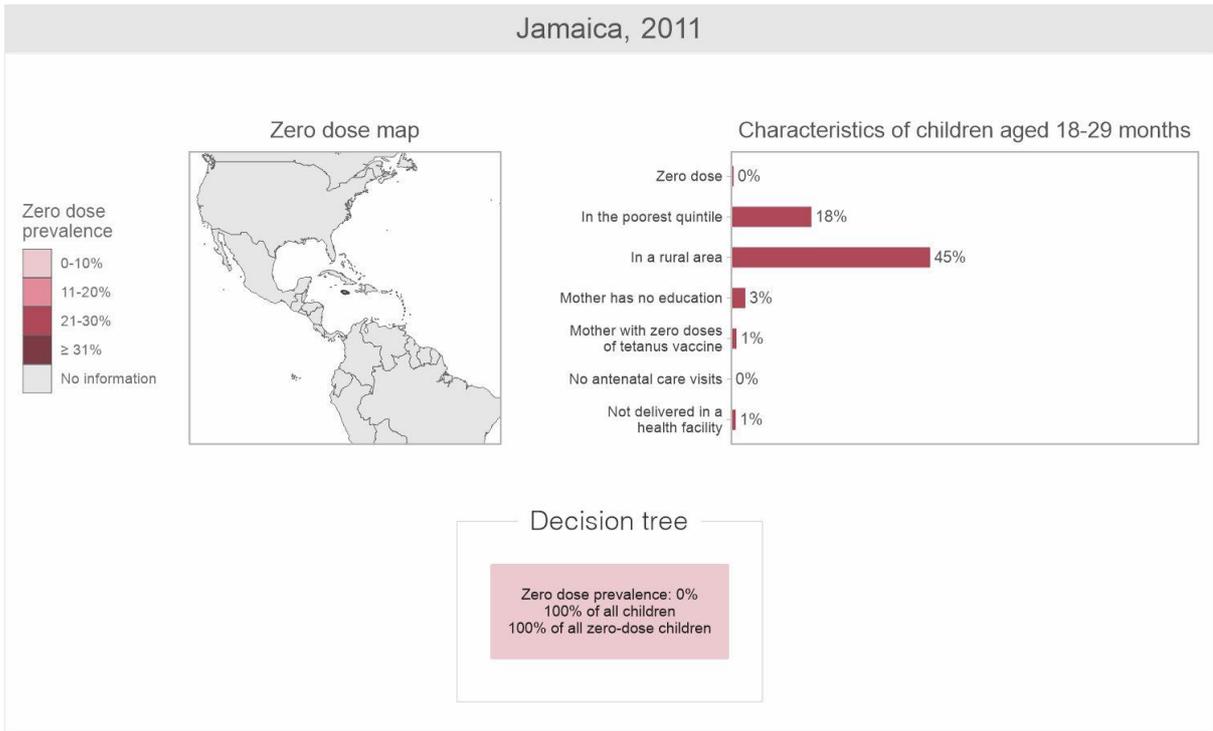
Iraq, 2018



616

617

Figure S38 – Iraq’s country profile

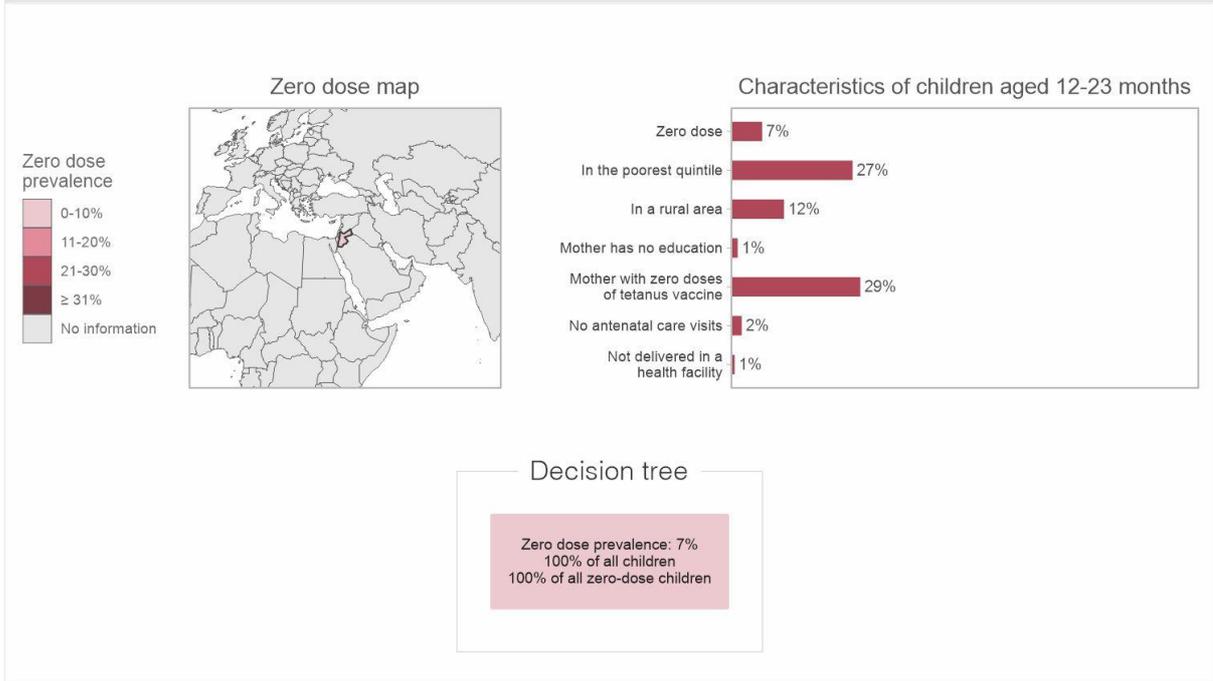


618

619

Figure S39 – Jamaica’s country profile

Jordan, 2017

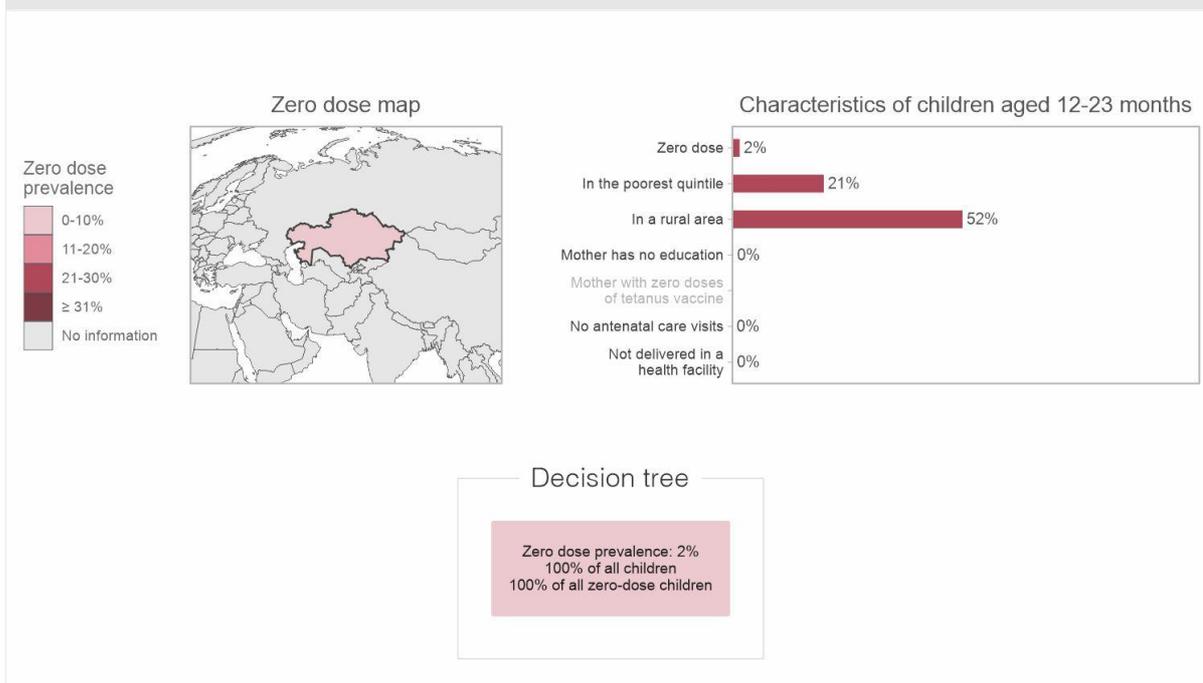


620

621

Figure S40 – Jordan’s country profile

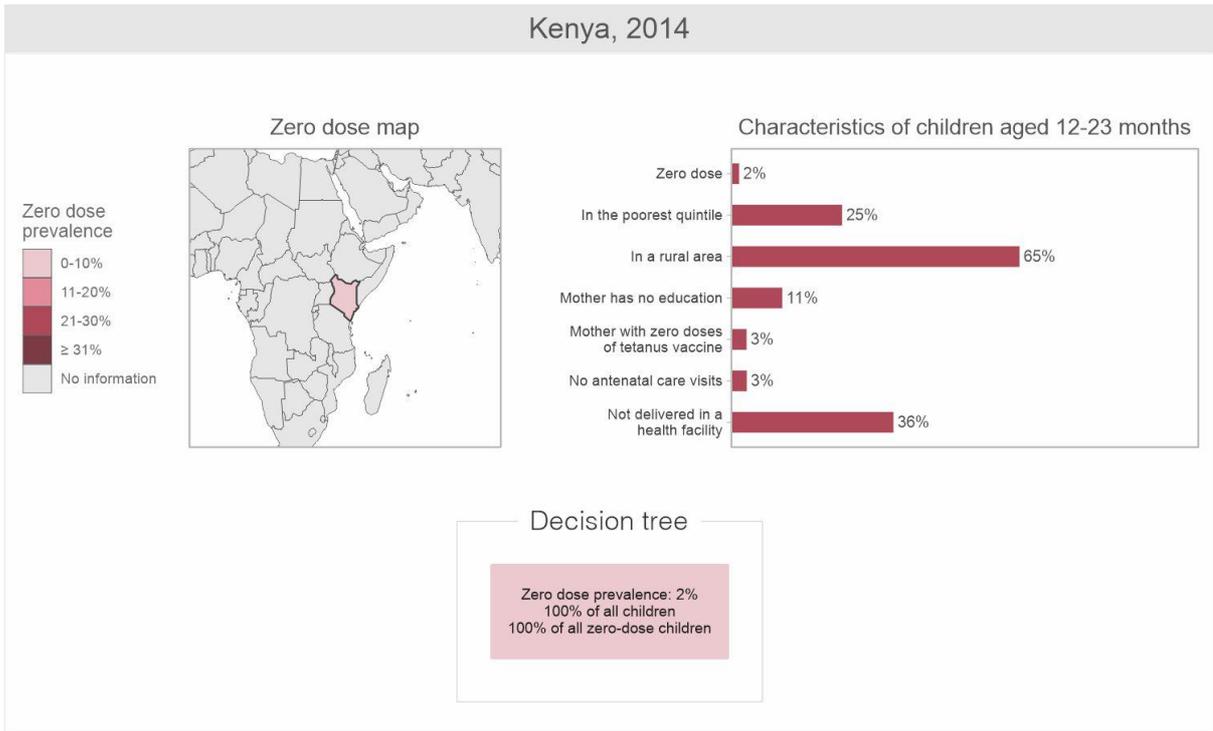
Kazakhstan, 2015



622

623

Figure S41 – Kazakhstan’s country profile

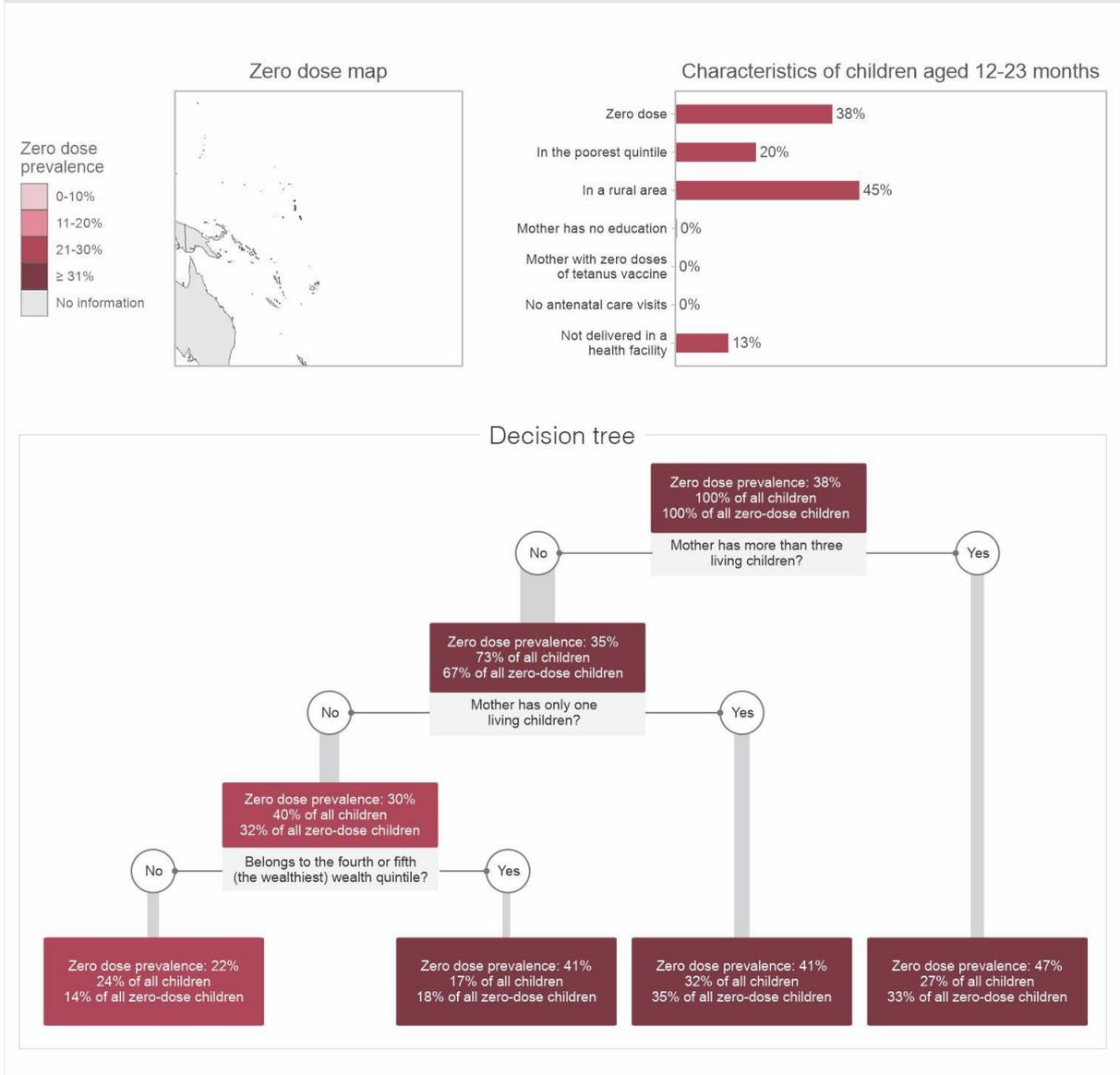


624

625

Figure S42 – Kenya’s country profile

Kiribati, 2018

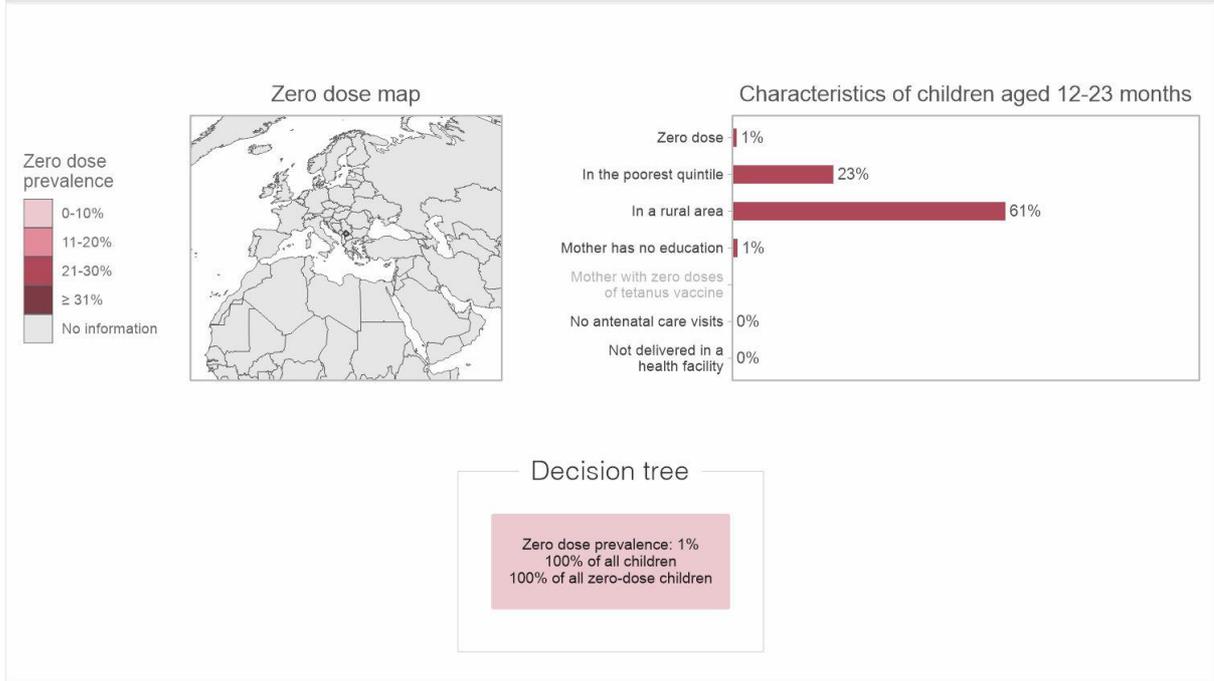


626

627

Figure S43 – Kiribati’s country profile

Kosovo, 2013

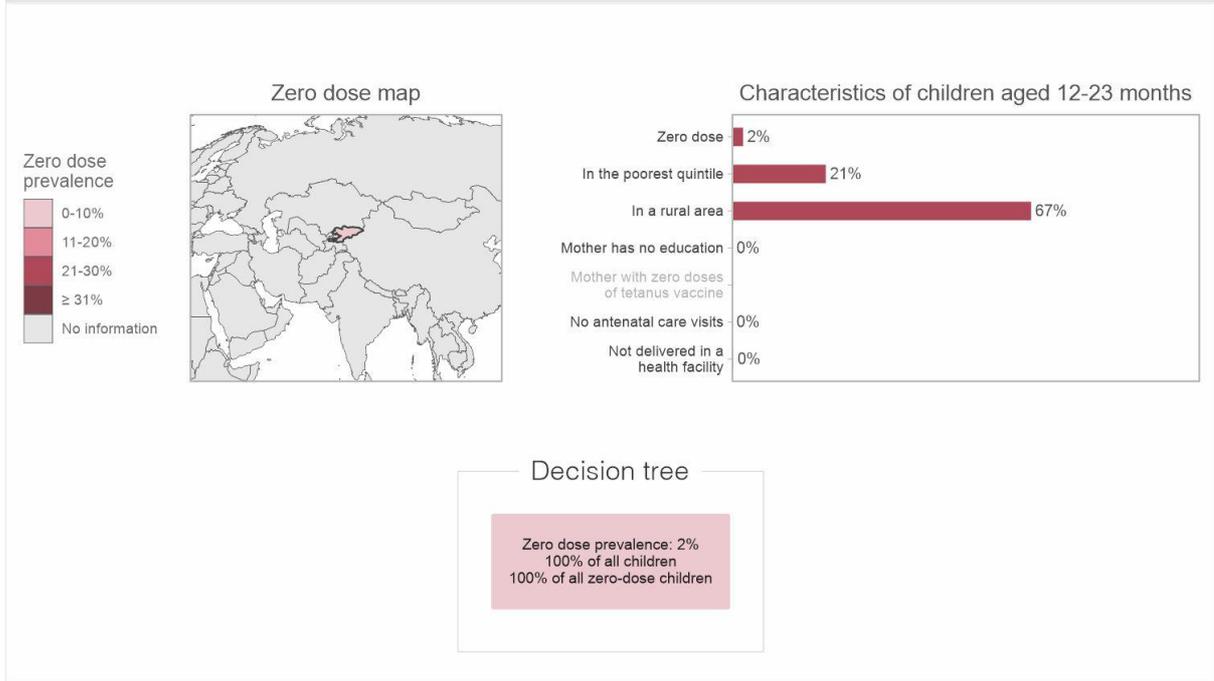


628

629

Figure S44 – Kosovo’s country profile

Kyrgyzstan, 2018

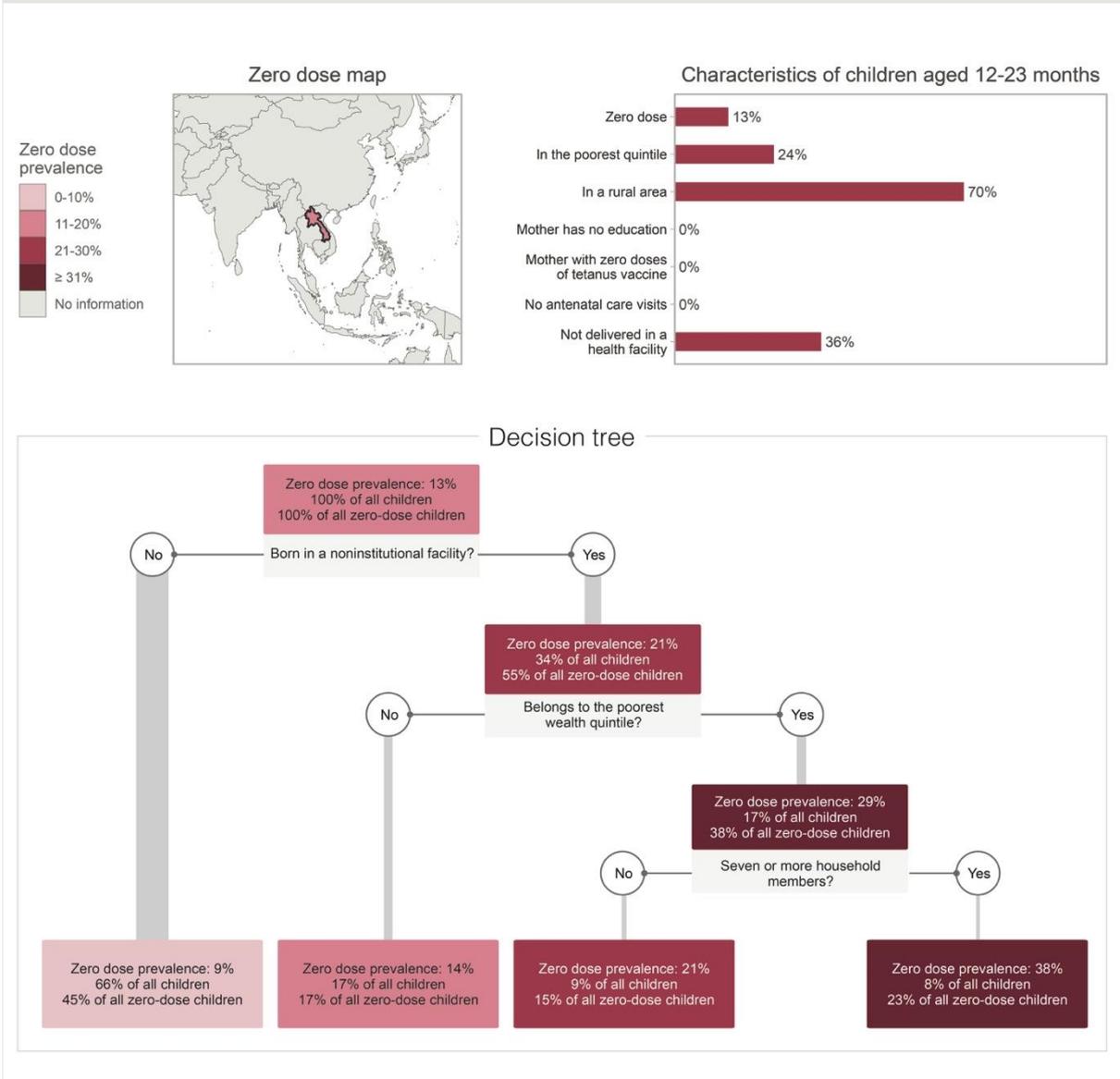


630

631

Figure S45 – Kyrgyzstan’s country profile

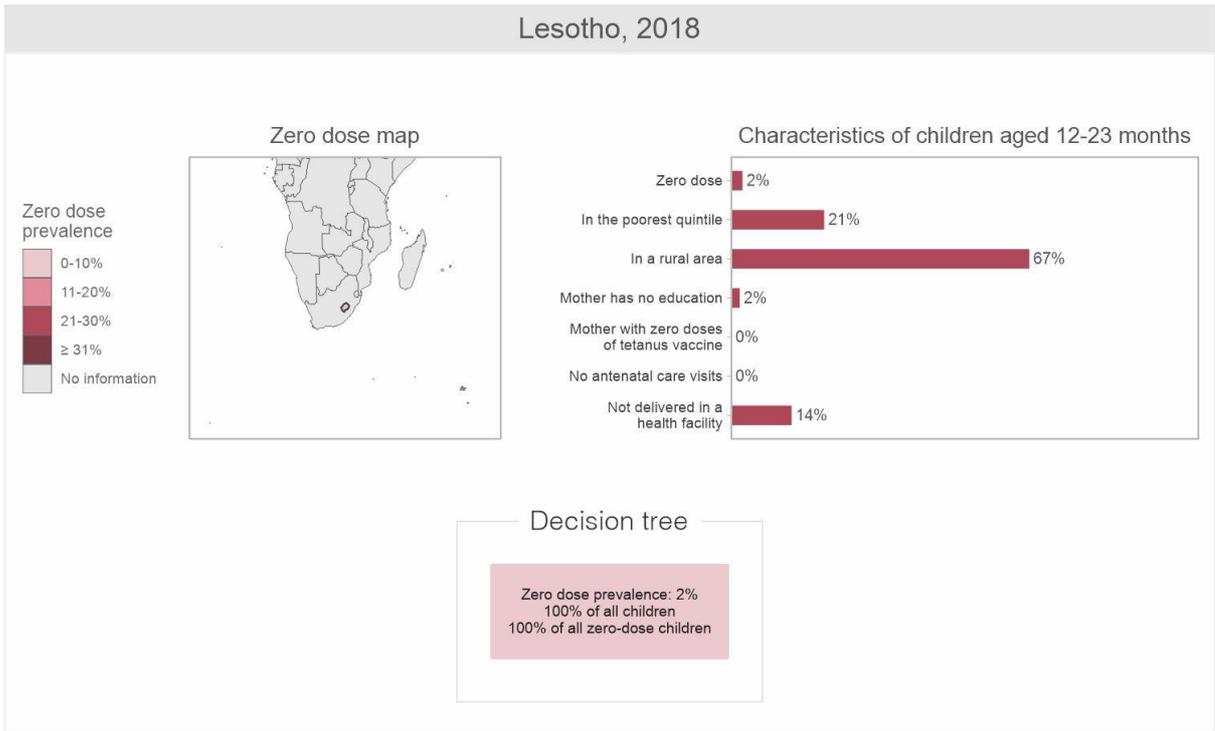
Laos, 2017



632

633

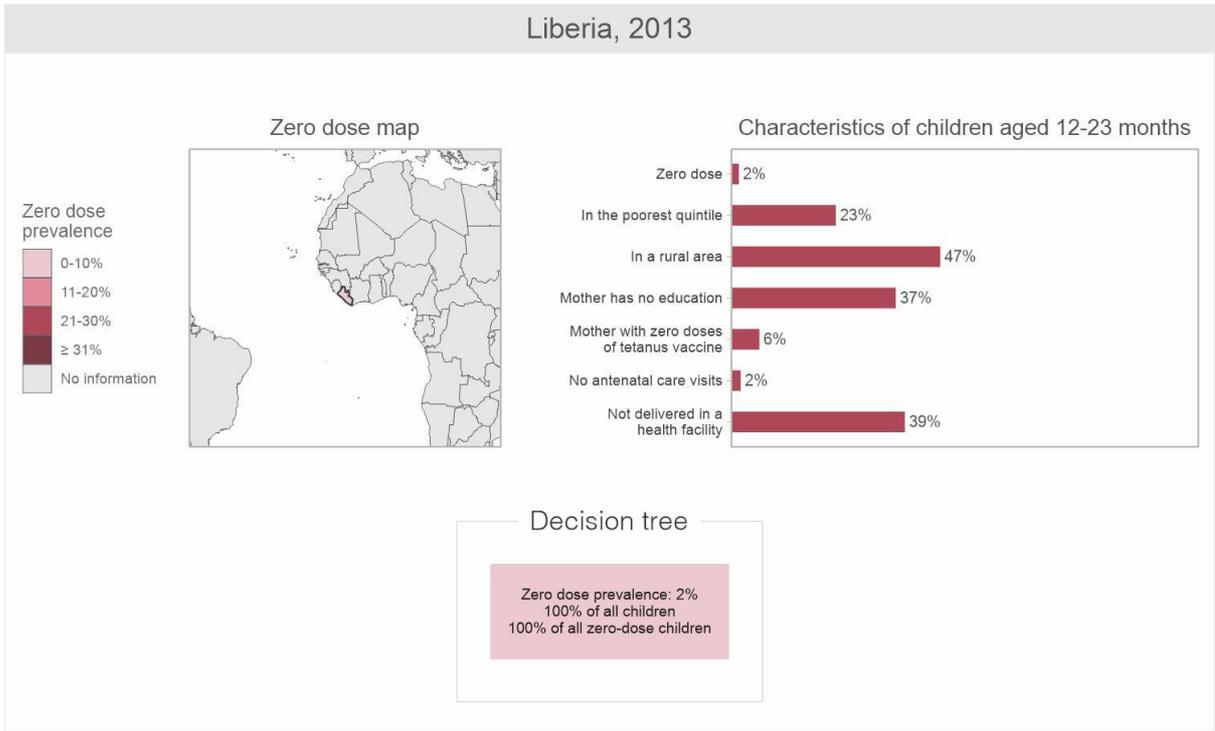
Figure S46 – Laos' country profile



634

635

Figure S47 – Lesotho’s country profile

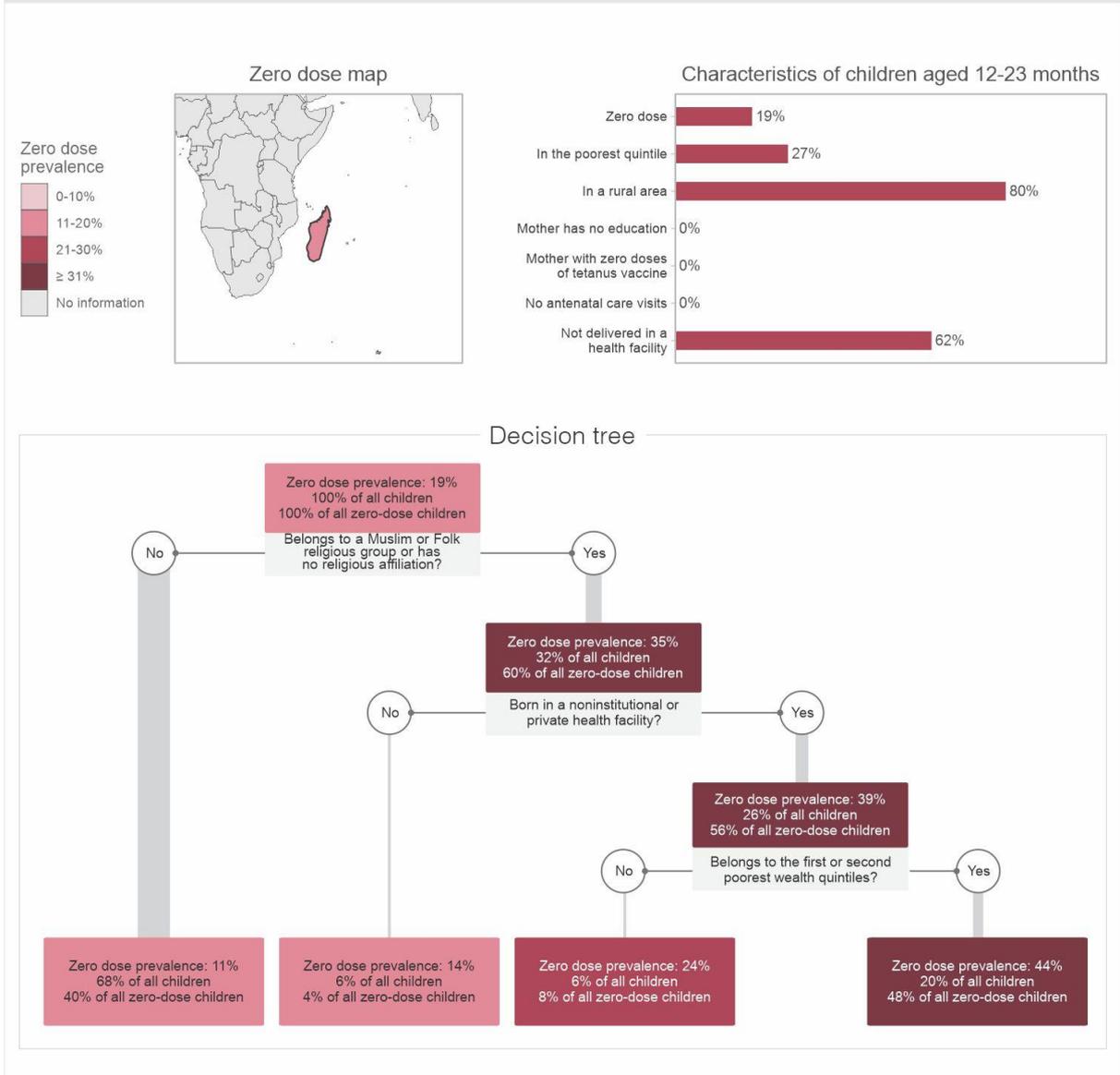


636

637

Figure S48 – Liberia’s country profile

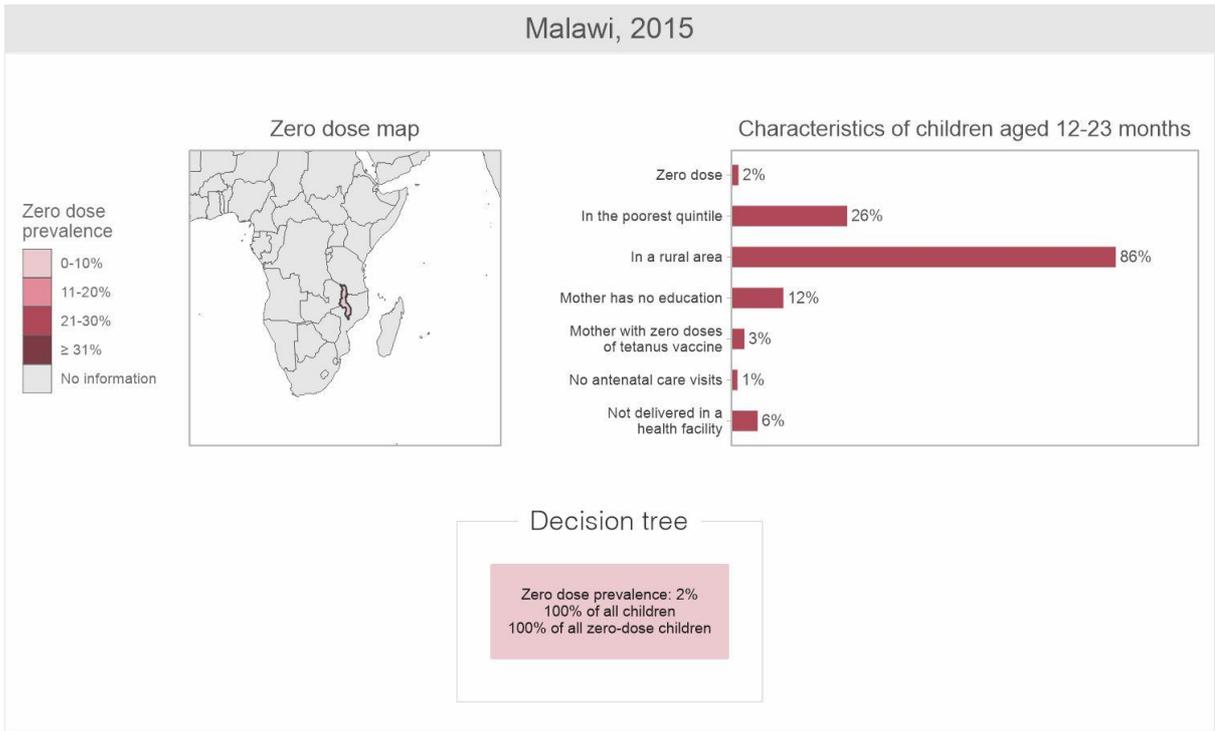
Madagascar, 2018



638

639

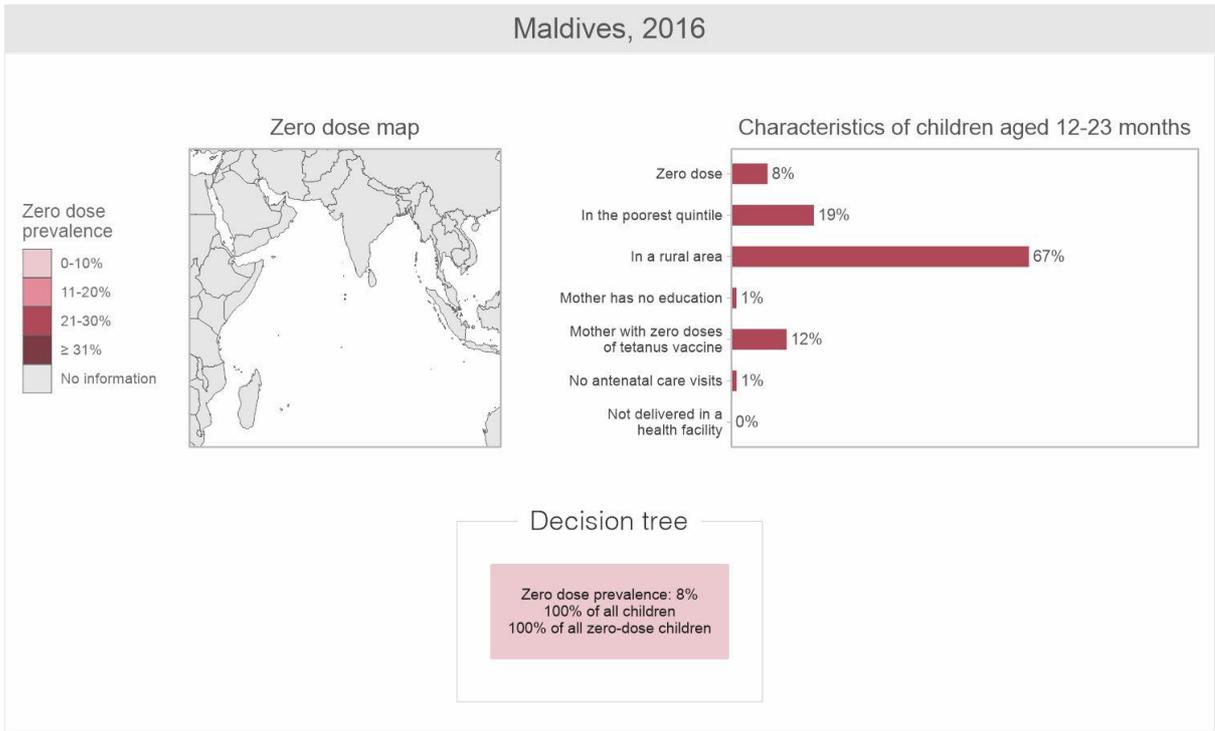
Figure S49 – Madagascar’s country profile



640

641

Figure S50 – Malawi’s country profile

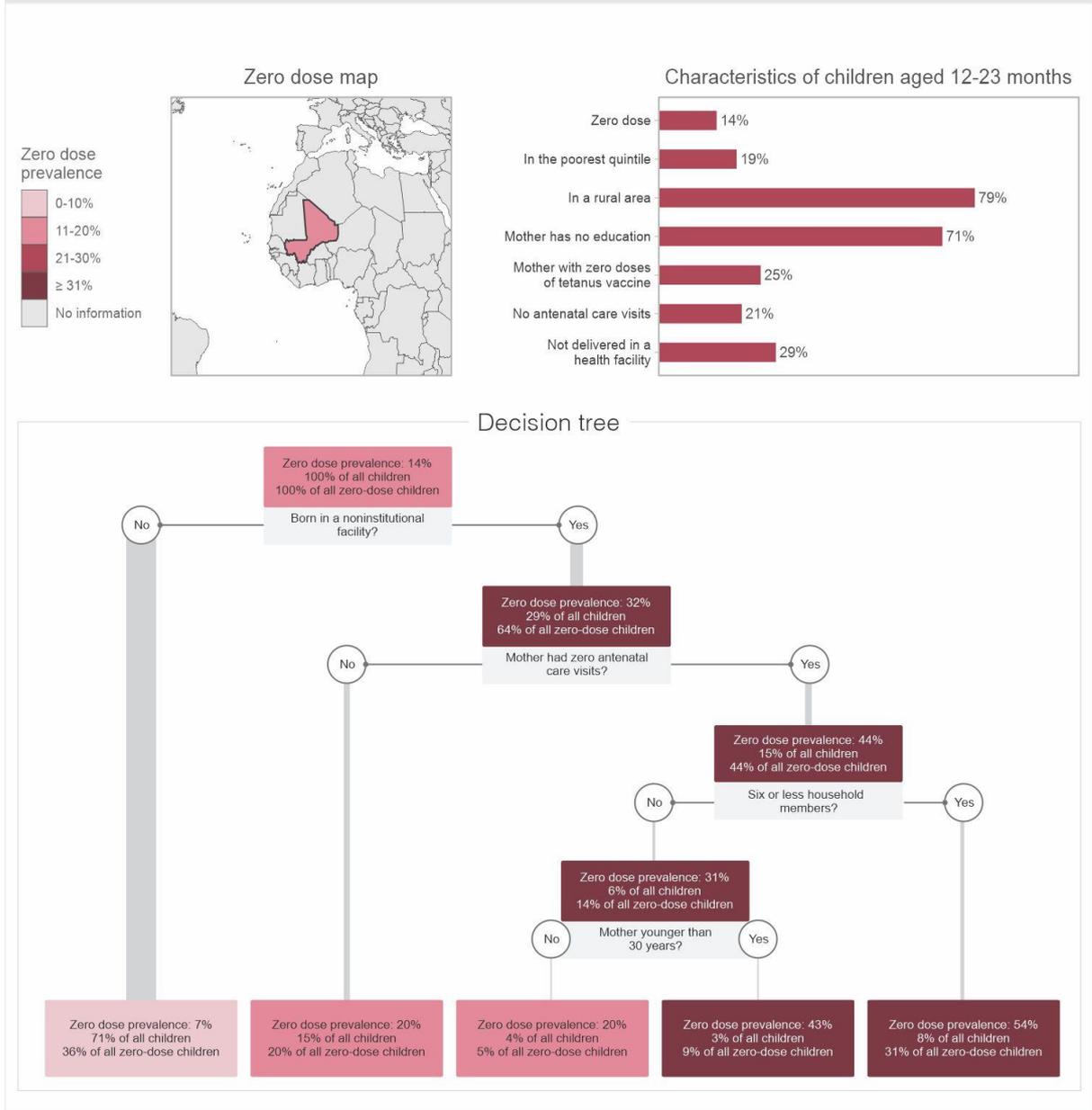


642

643

Figure S51 – Maldives' country profile

Mali, 2018

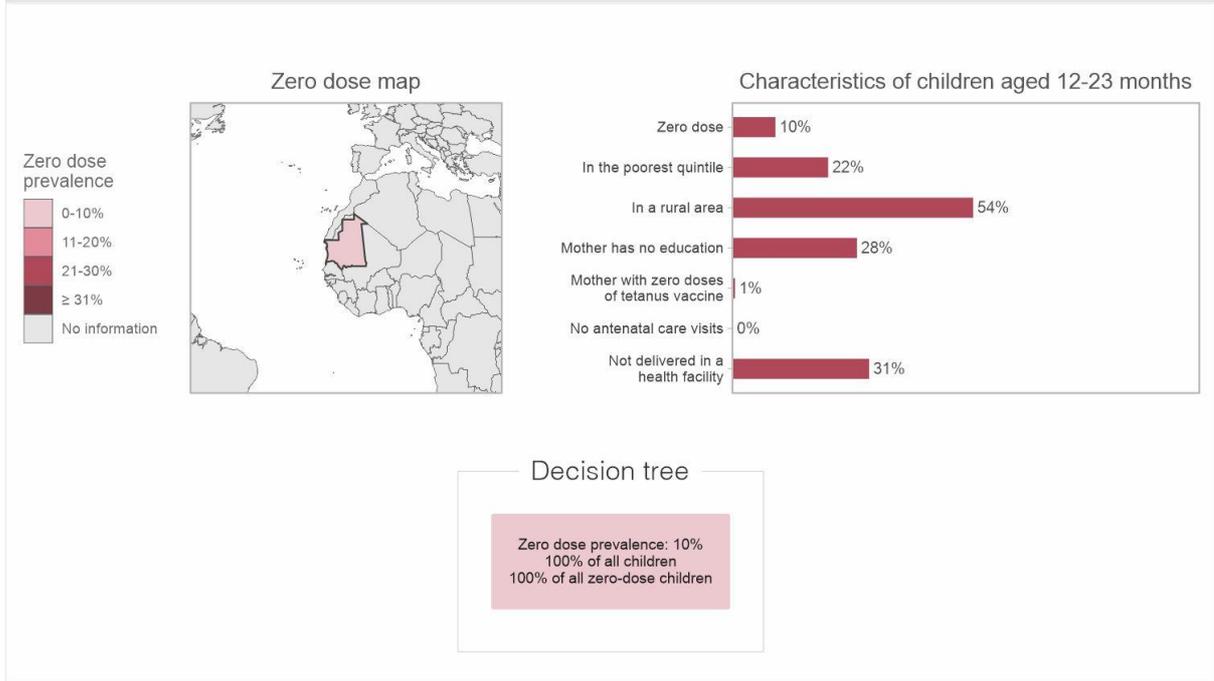


644

645

Figure S52 – Mali’s country profile

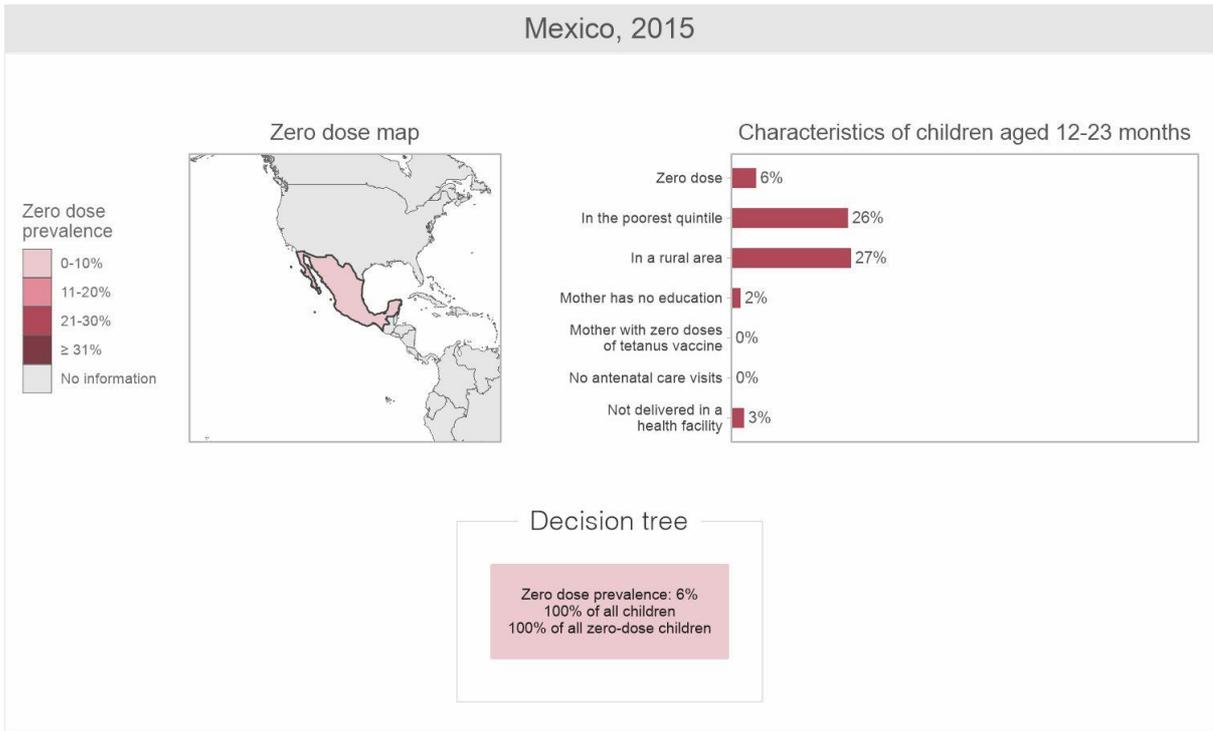
Mauritania, 2015



646

647

Figure S53 – Mauritania's country profile

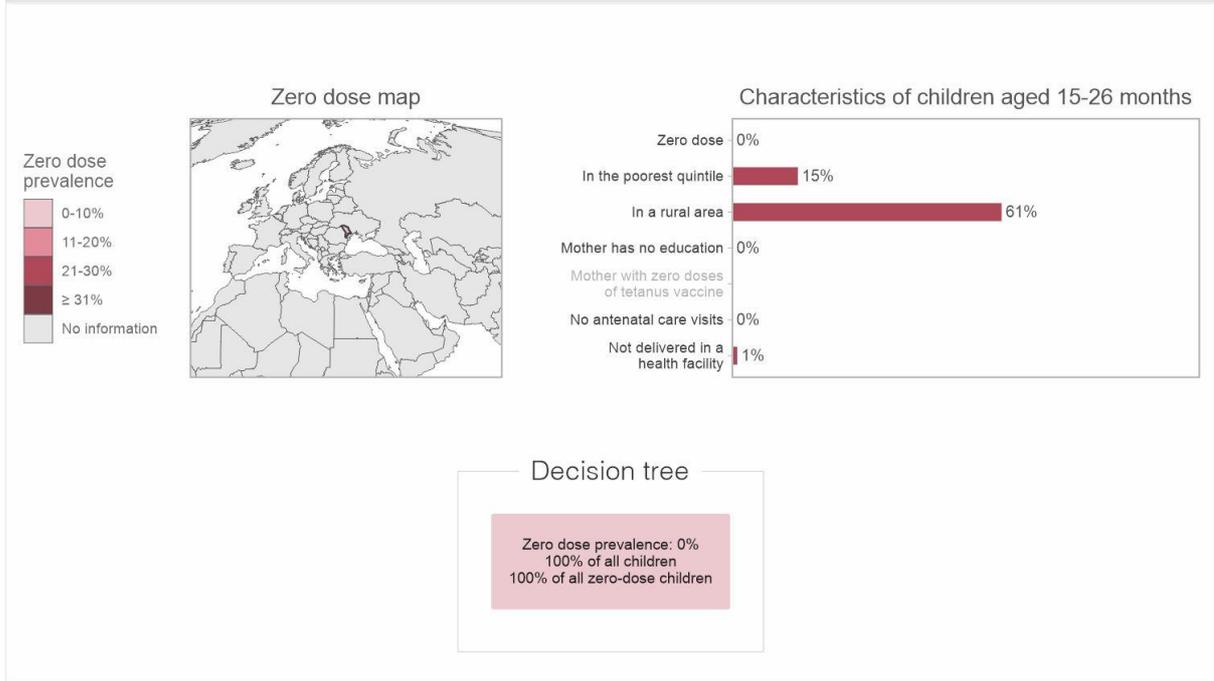


648

649

Figure S54 – Mexico’s country profile

Moldova, 2012

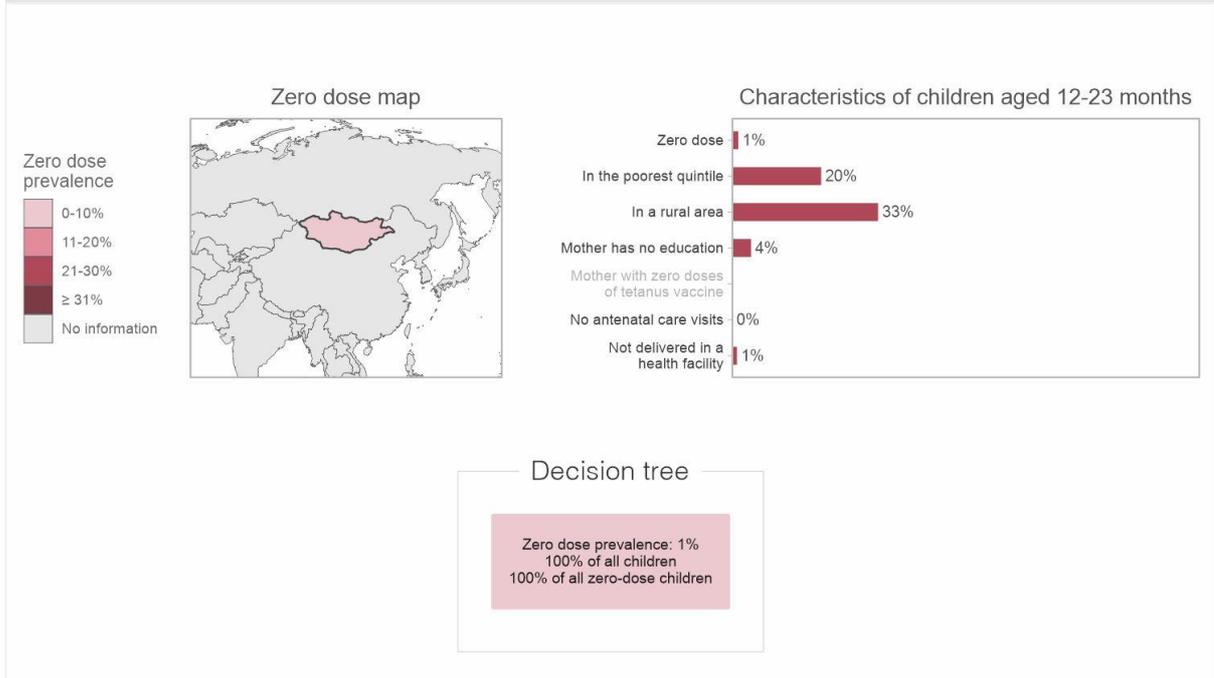


650

651

Figure S55 – Moldova’s country profile

Mongolia, 2018

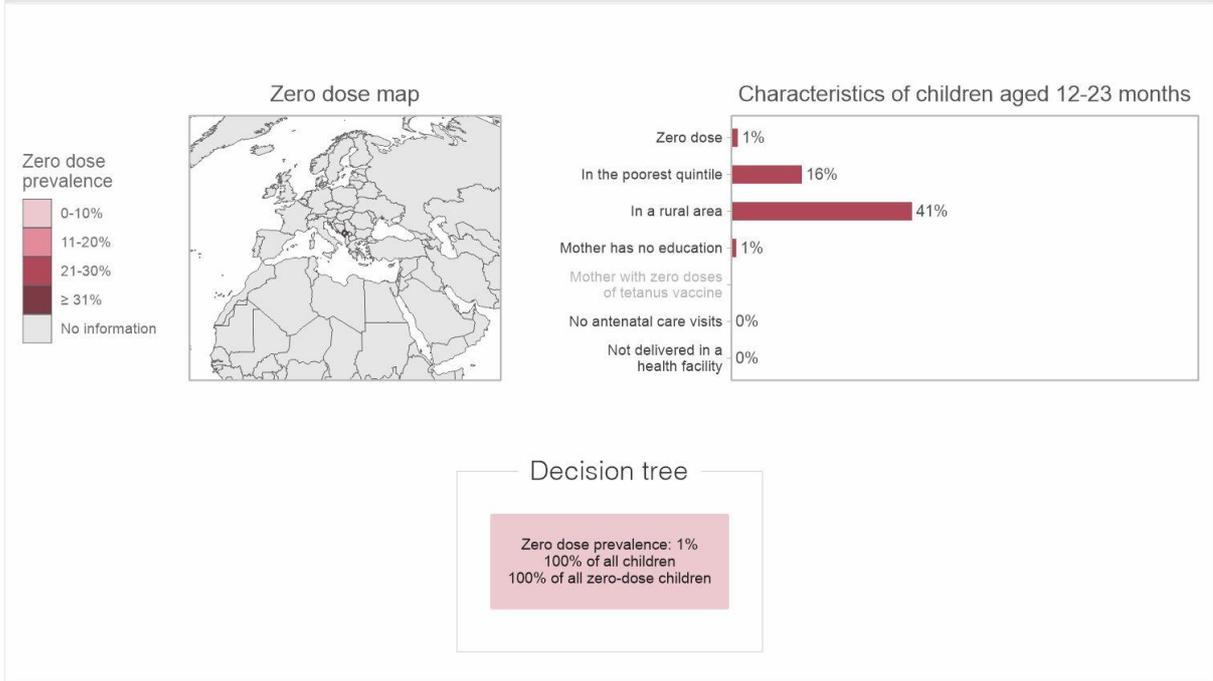


652

653

Figure S56 – Mongolia’s country profile

Montenegro, 2013

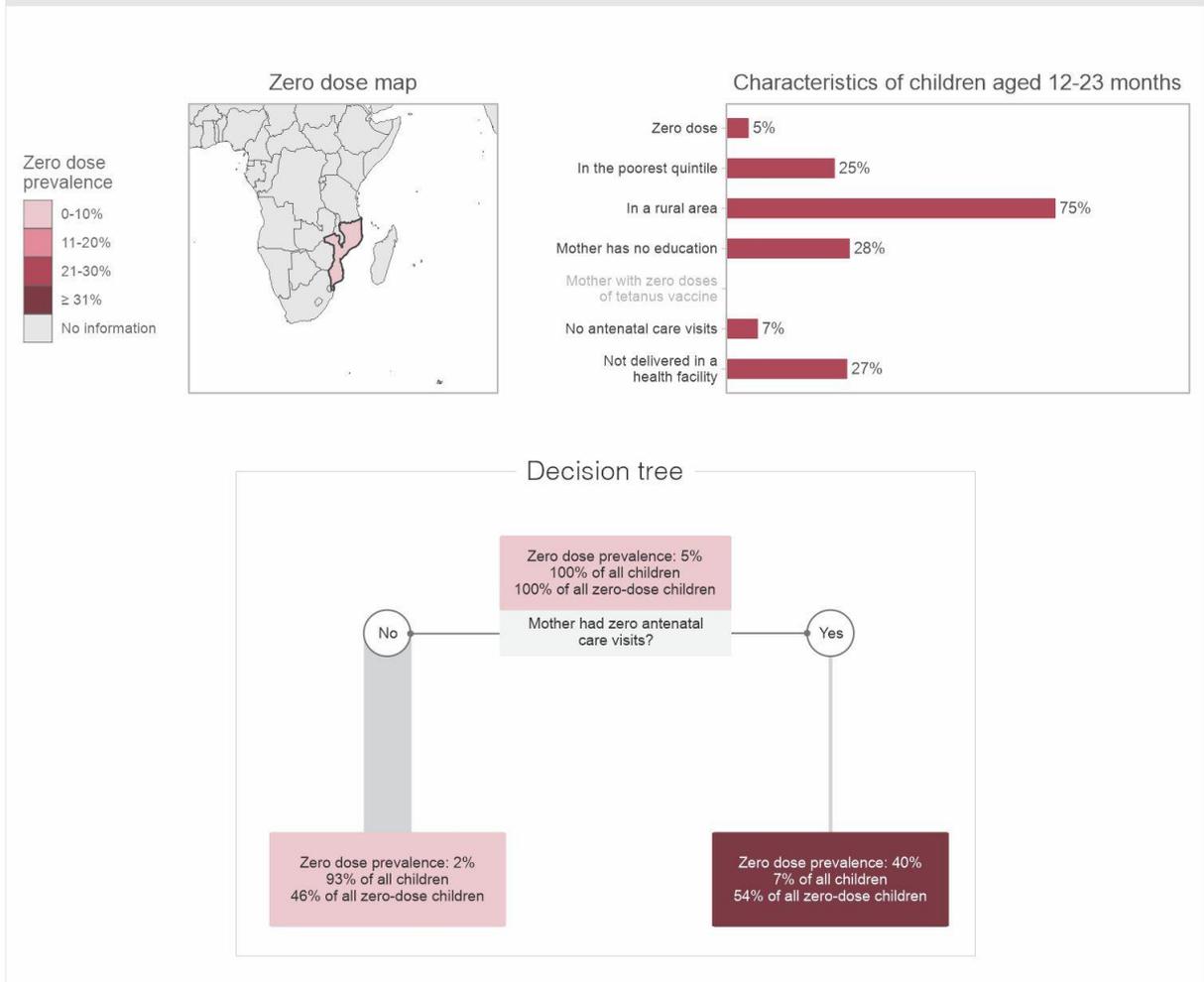


654

655

Figure S57 – Montenegro's country profile

Mozambique, 2015

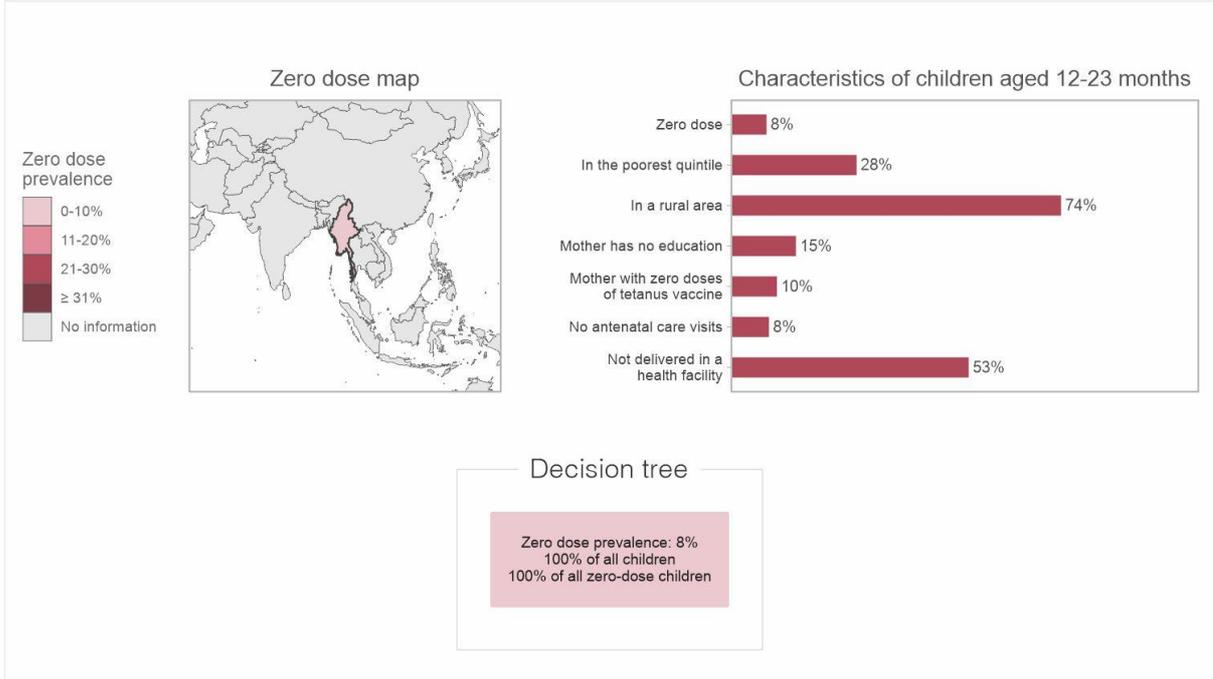


656

657

Figure S58 – Mozambique's country profile

Myanmar, 2015

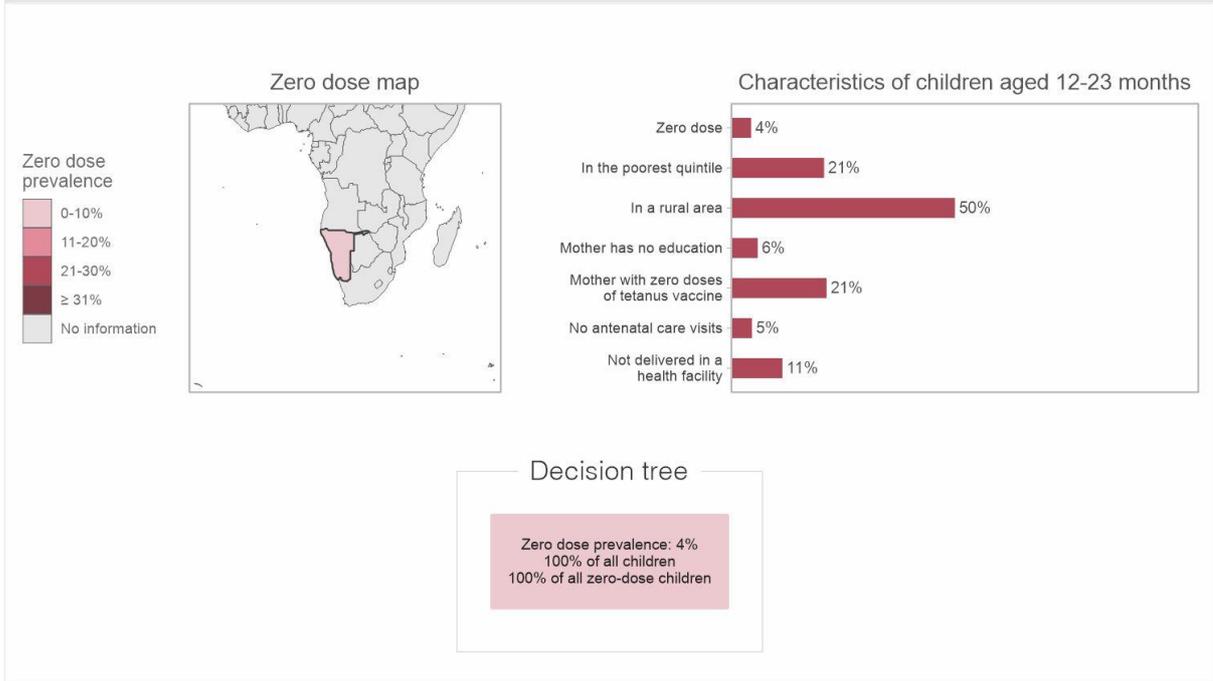


658

659

Figure S59 – Myanmar’s country profile

Namibia, 2013

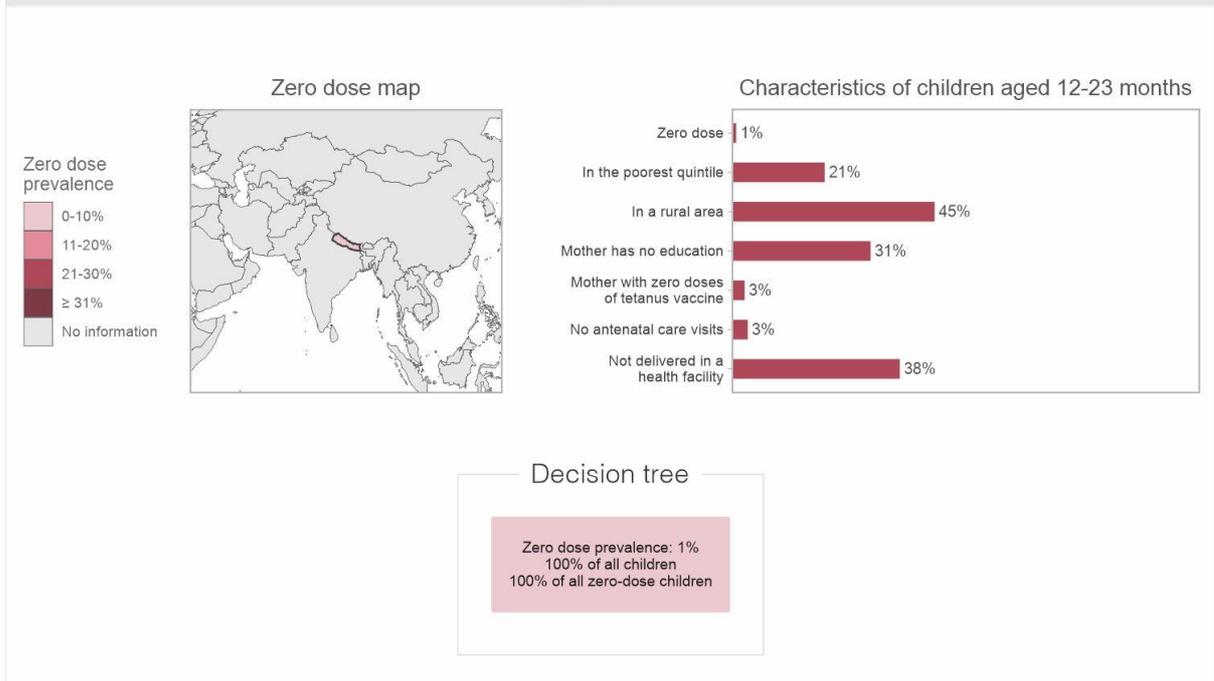


660

661

Figure S60 – Namibia's country profile

Nepal, 2016

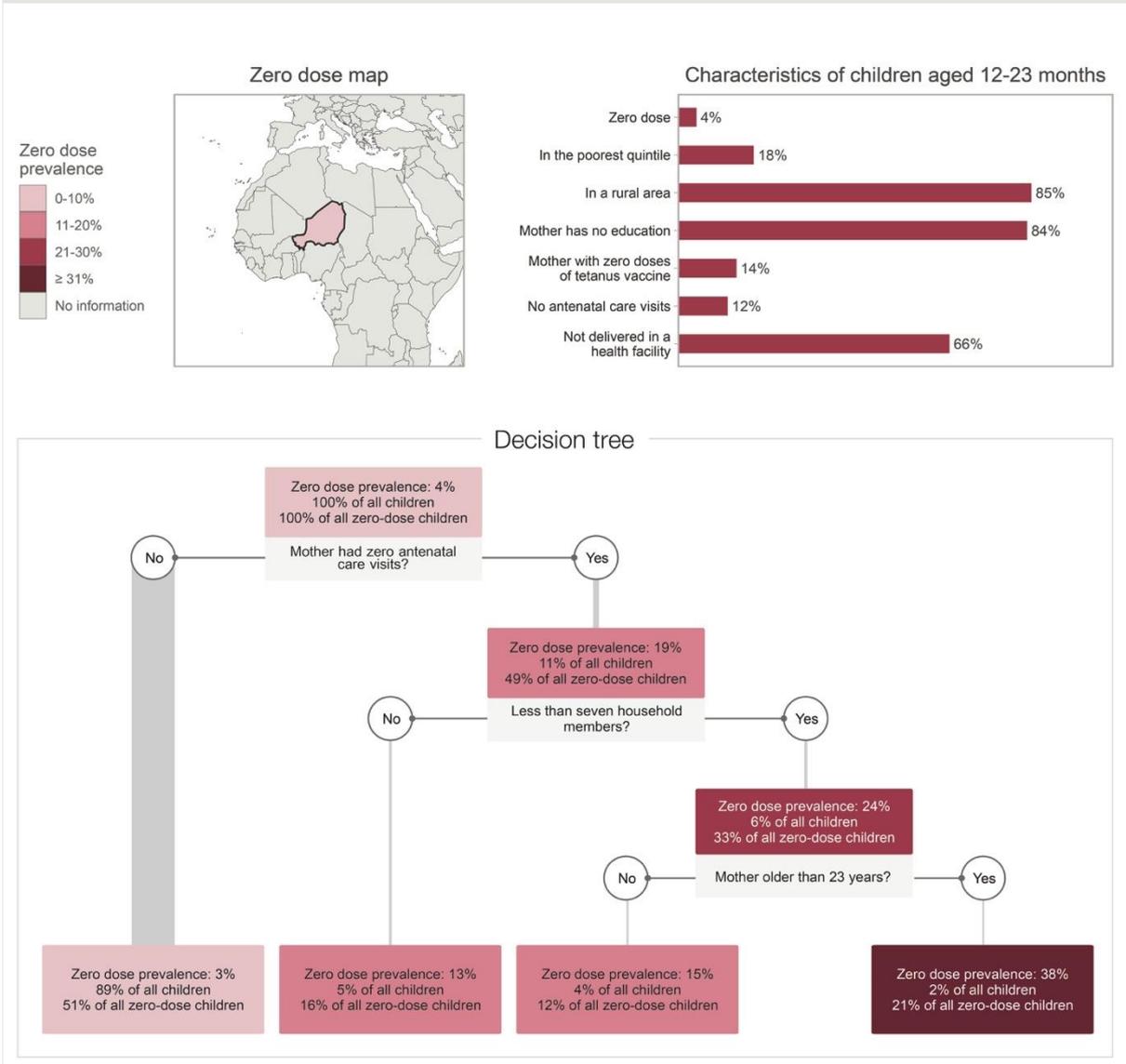


662

663

Figure S61 – Nepal’s country profile

Niger, 2012



664

665

Figure S62 – Niger’s country profile

Nigeria, 2018

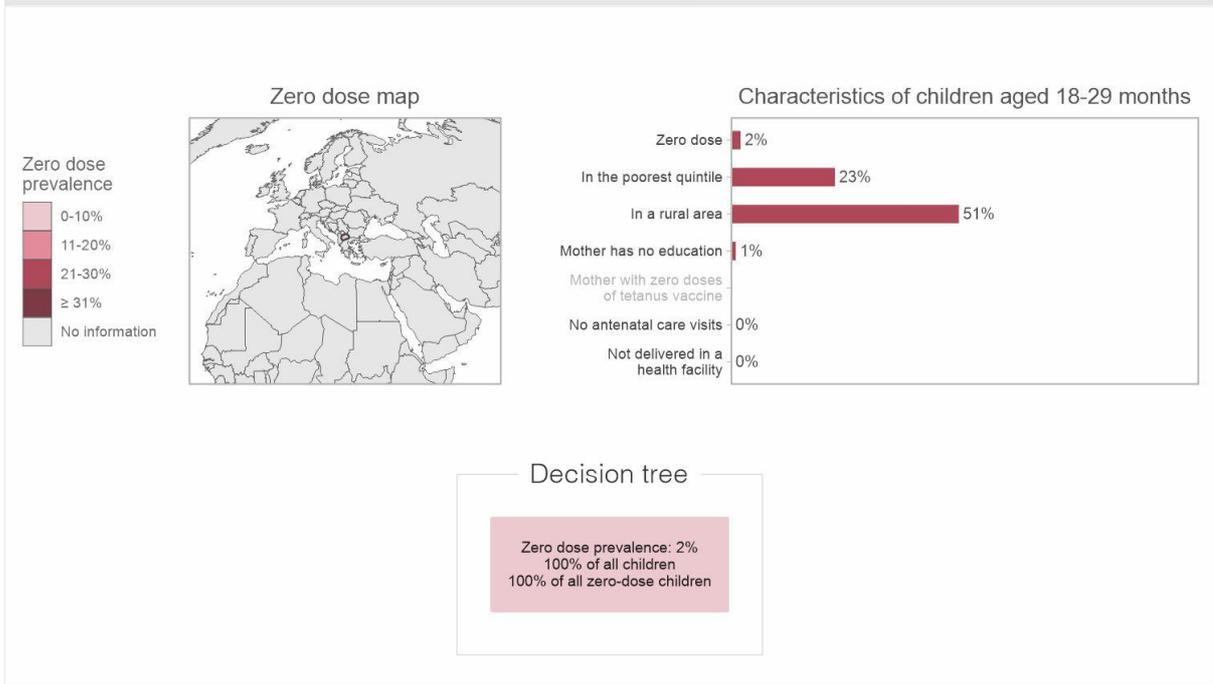


666

667

Figure S63 – Nigeria’s country profile

North Macedonia, 2011

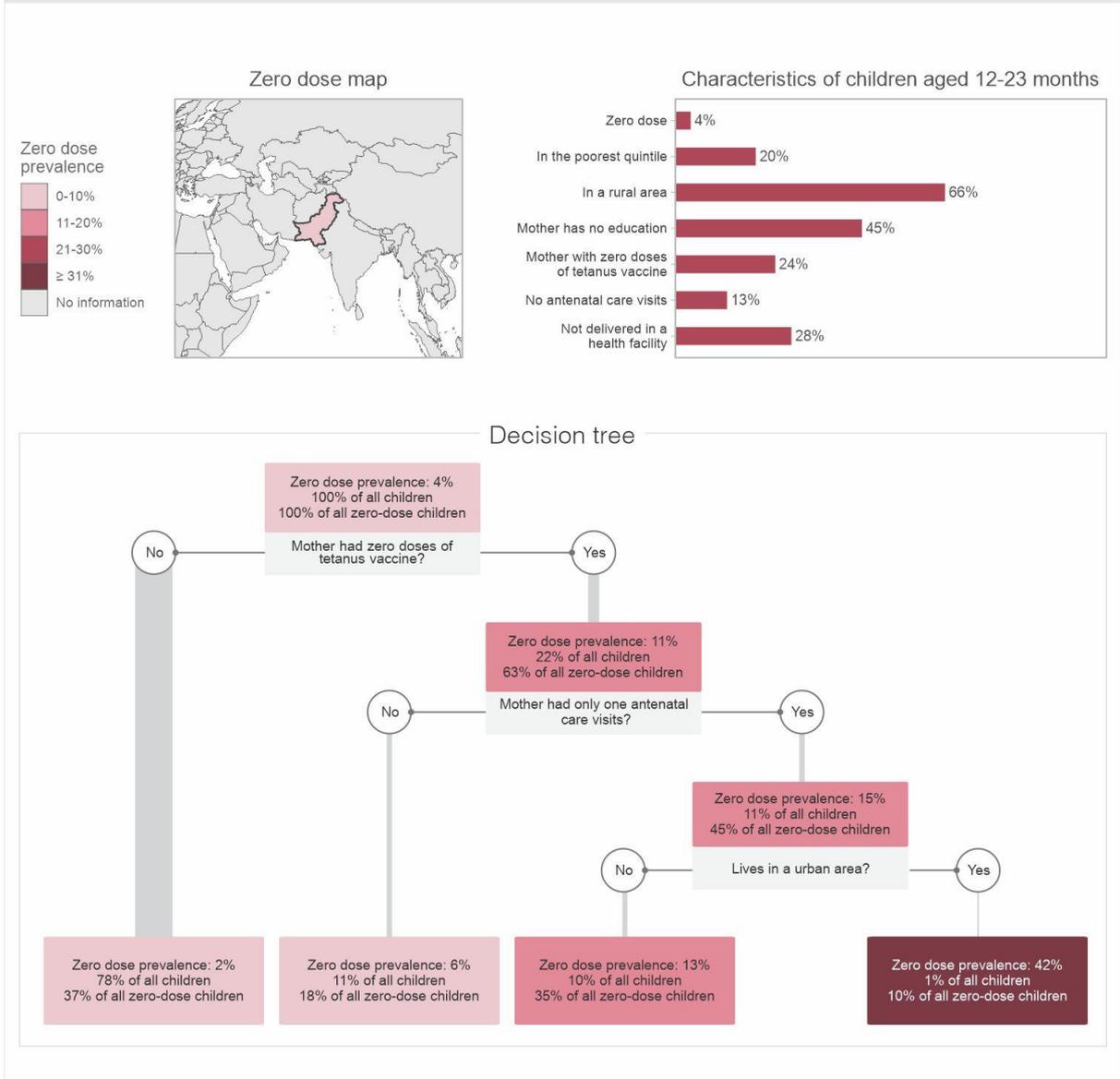


668

669

Figure S64 – North Macedonia's country profile

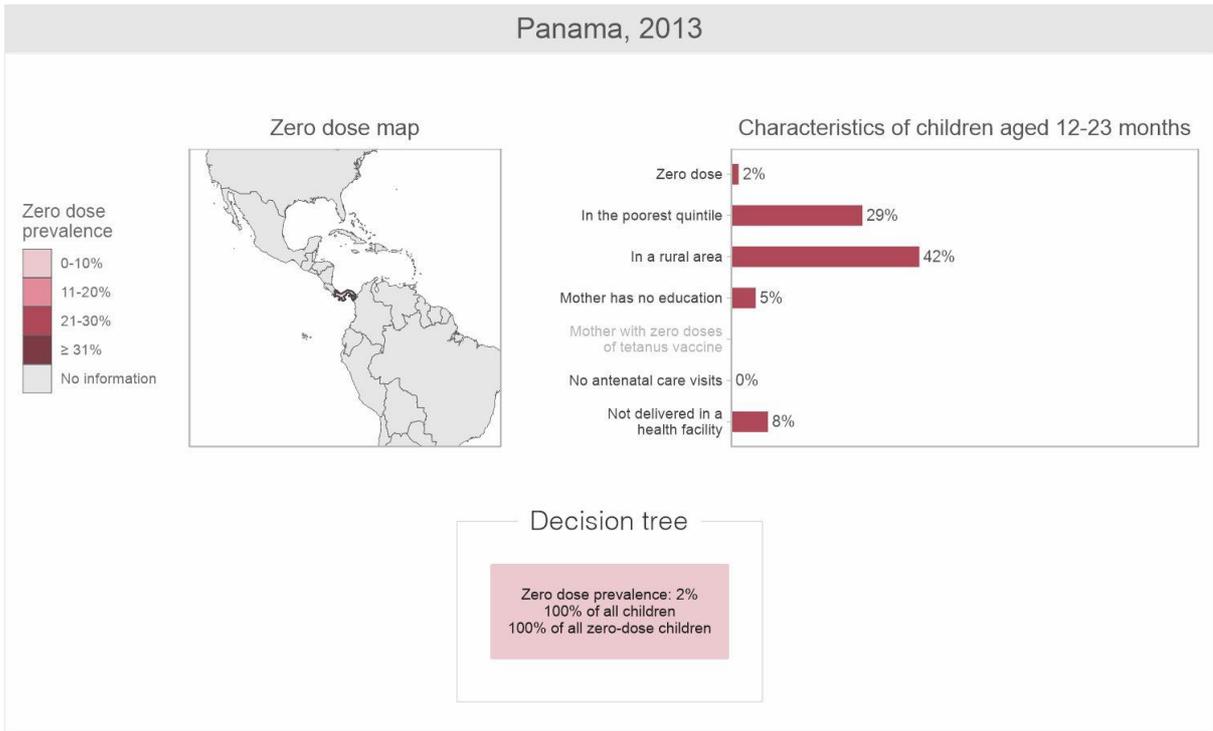
Pakistan, 2017



670

671

Figure S65 – Pakistan’s country profile



672

673

Figure S66 – Panama’s country profile

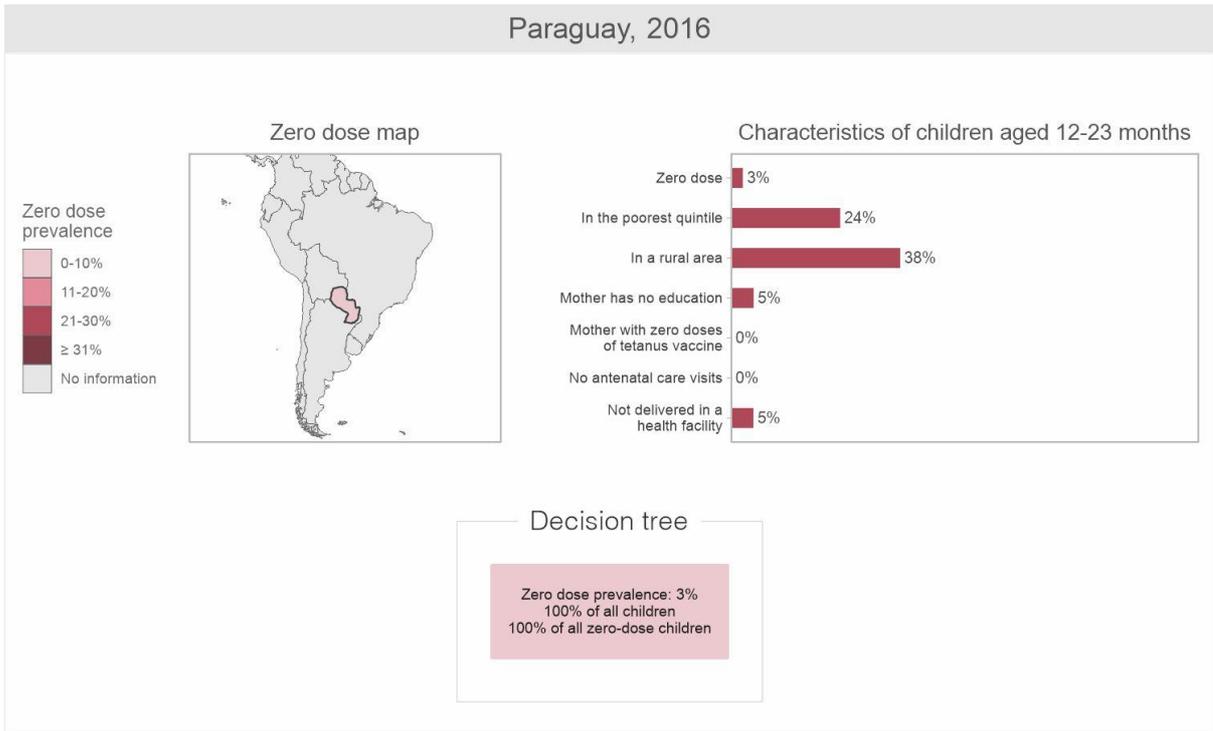
Papua New Guinea, 2016



674

675

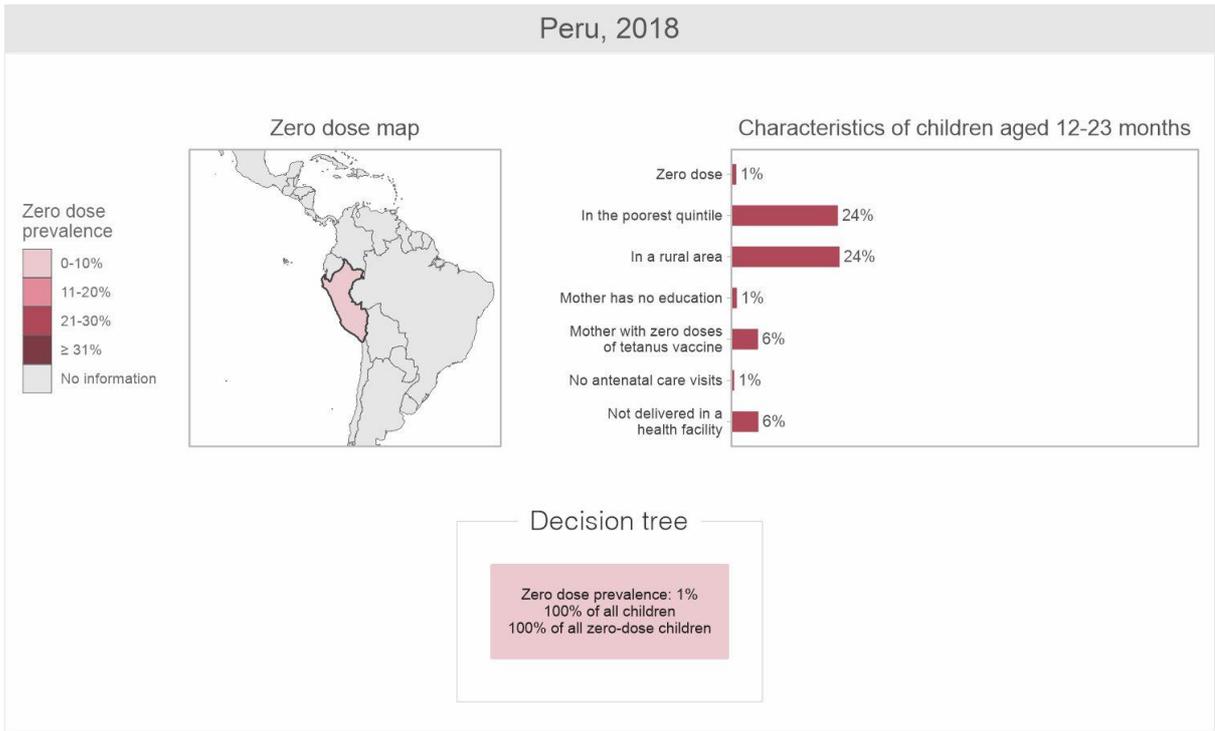
Figure S67 – Papua New Guinea’s country profile



676

677

Figure S68 – Paraguay’s country profile

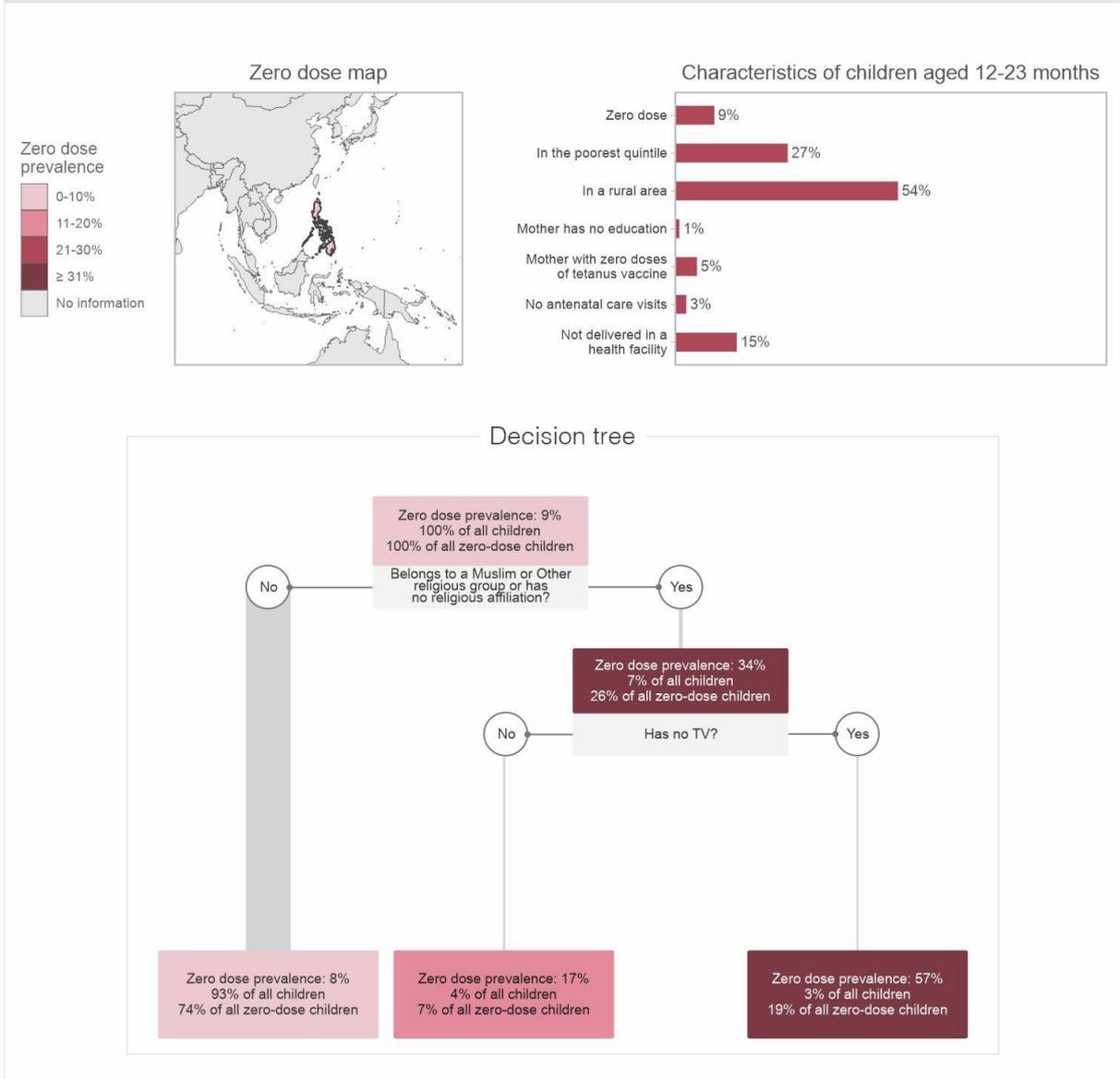


678

679

Figure S69 – Peru’s country profile

Philippines, 2017

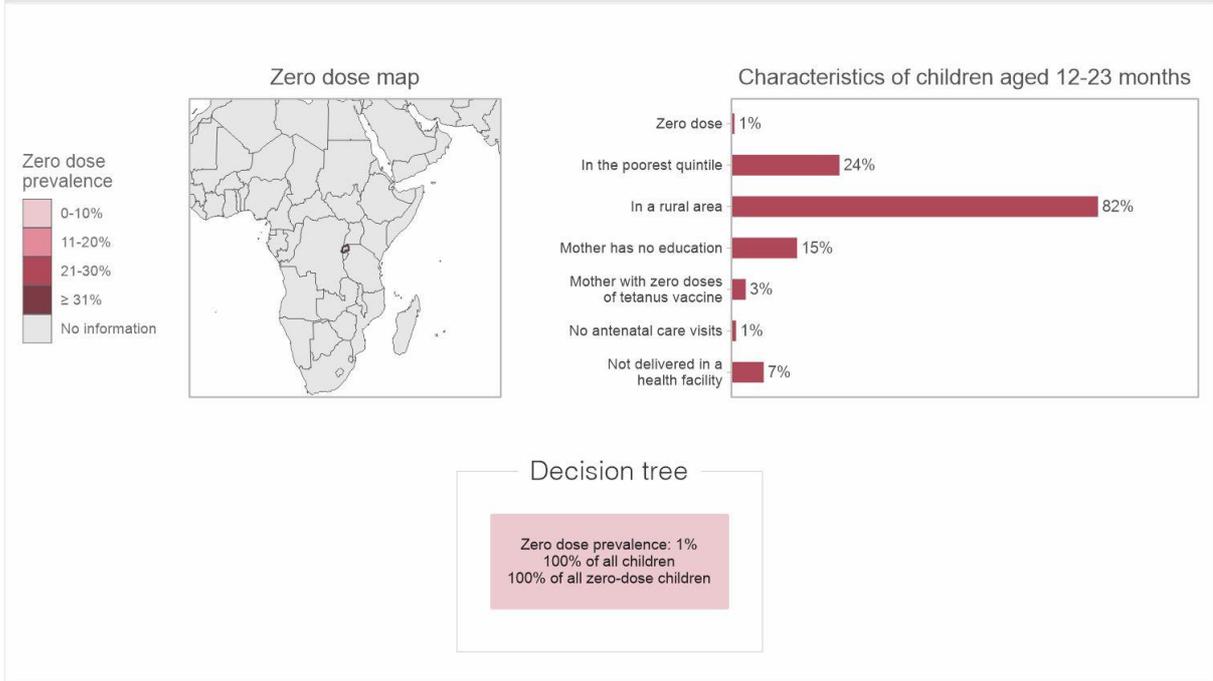


680

681

Figure S70 – Philippines’ country profile

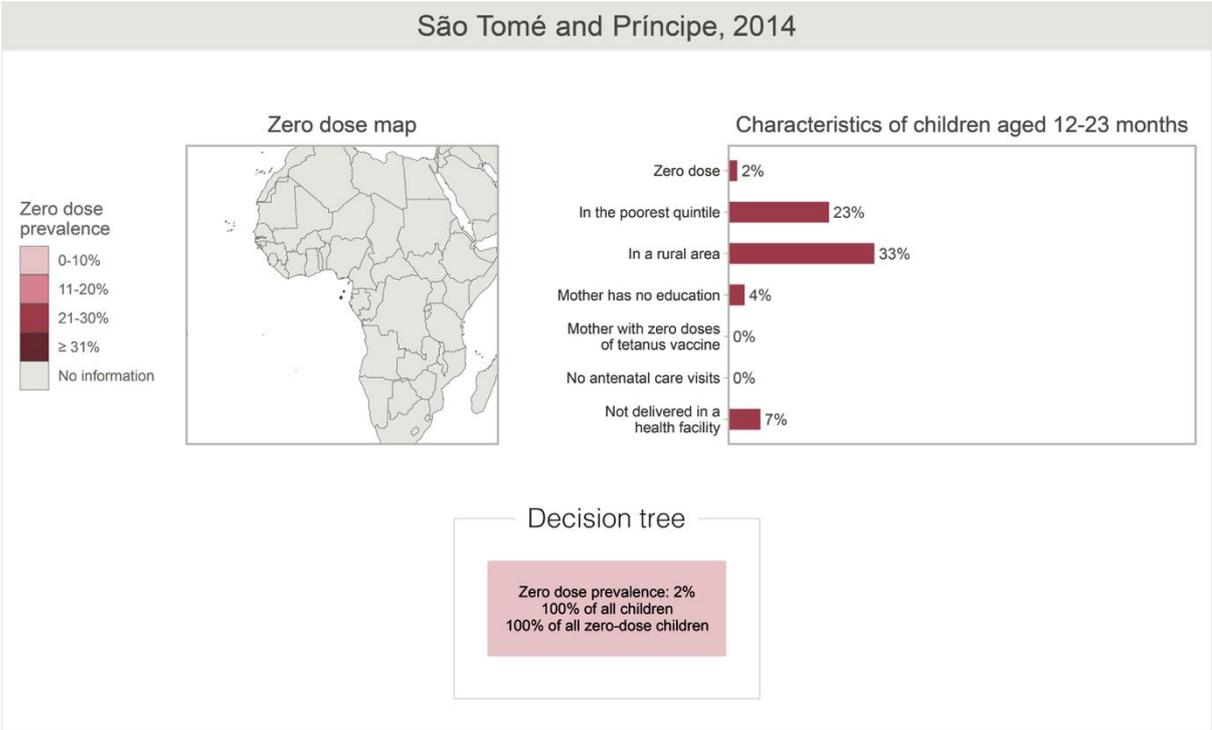
Rwanda, 2014



682

683

Figure S71 – Rwanda’s country profile

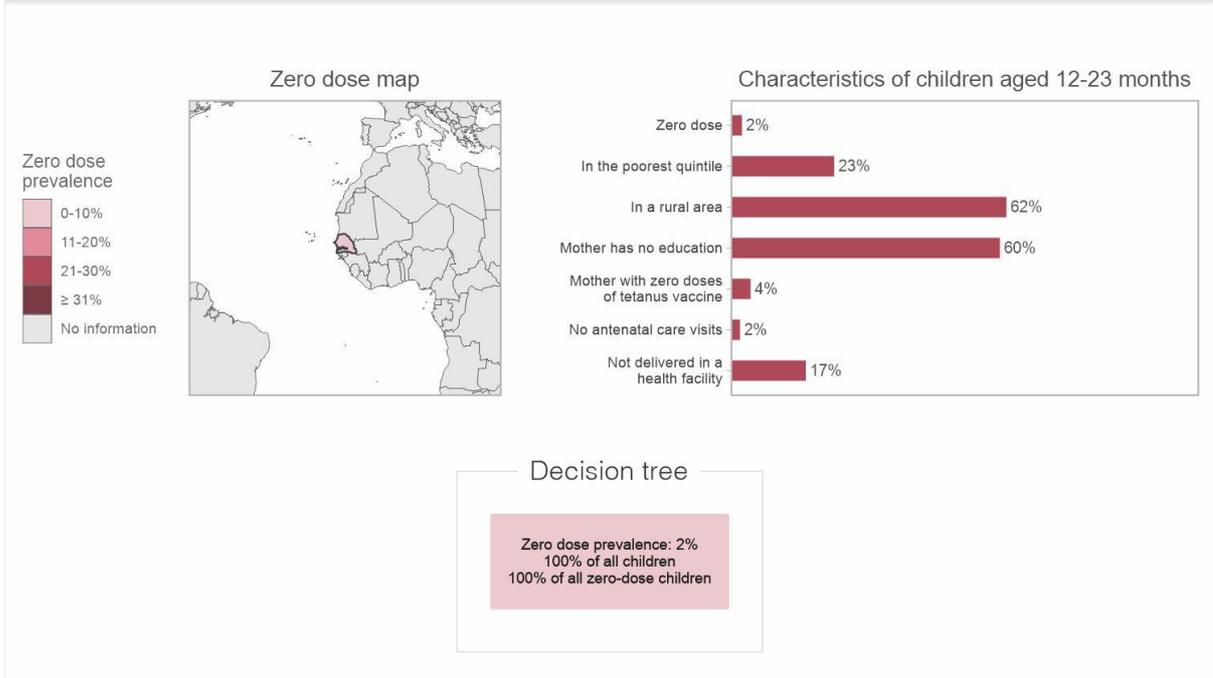


684

685

Figure S72 – São Tomé and Príncipe’s country profile

Senegal, 2018

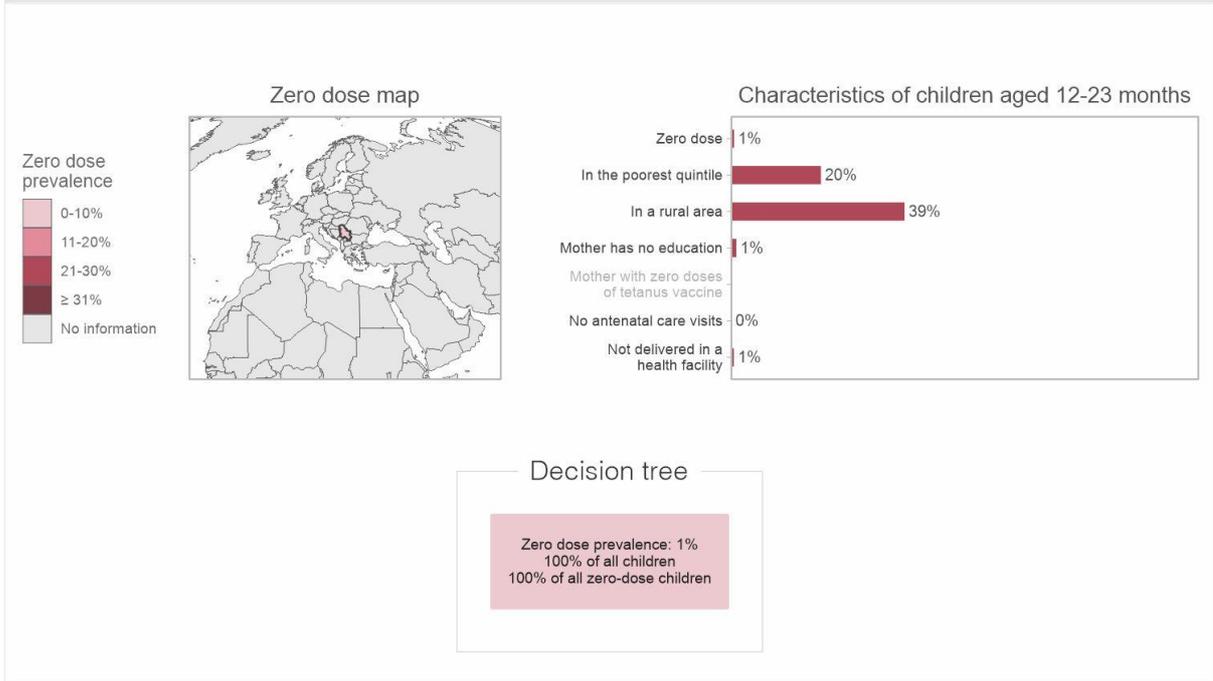


686

687

Figure S73 – Senegal’s country profile

Serbia, 2014

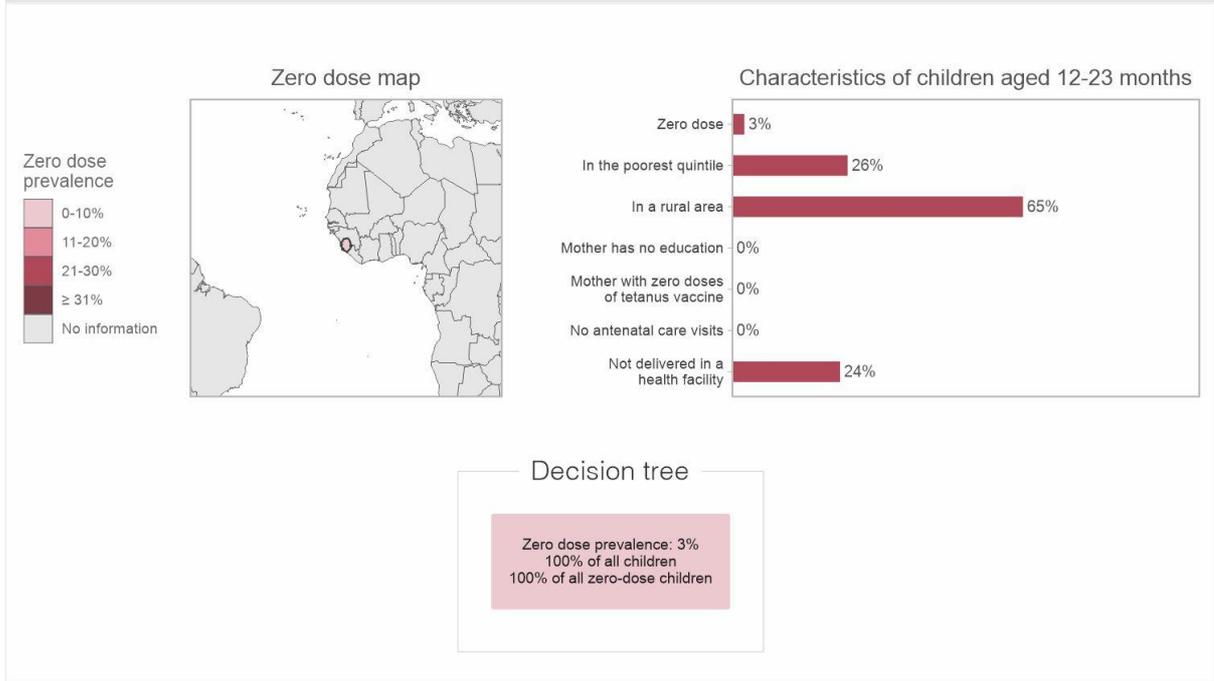


688

689

Figure S74 – Serbia's country profile

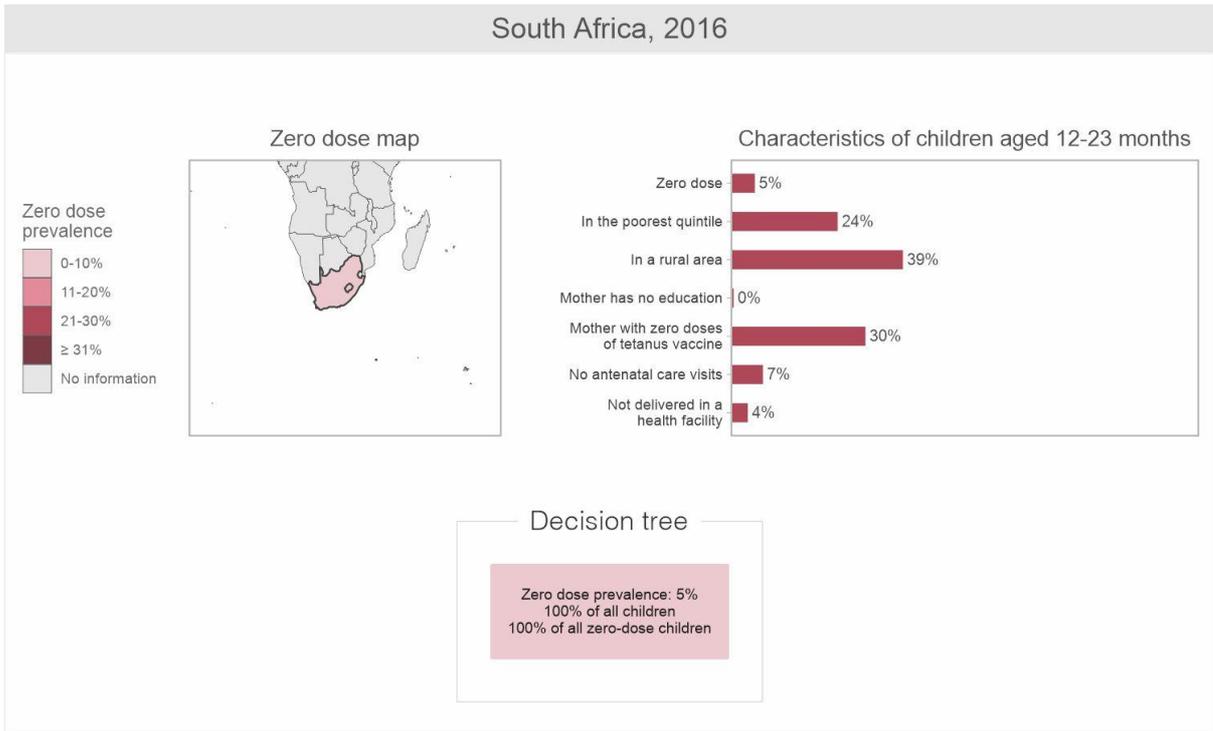
Sierra Leone, 2017



690

691

Figure S75 – Sierra Leone’s country profile

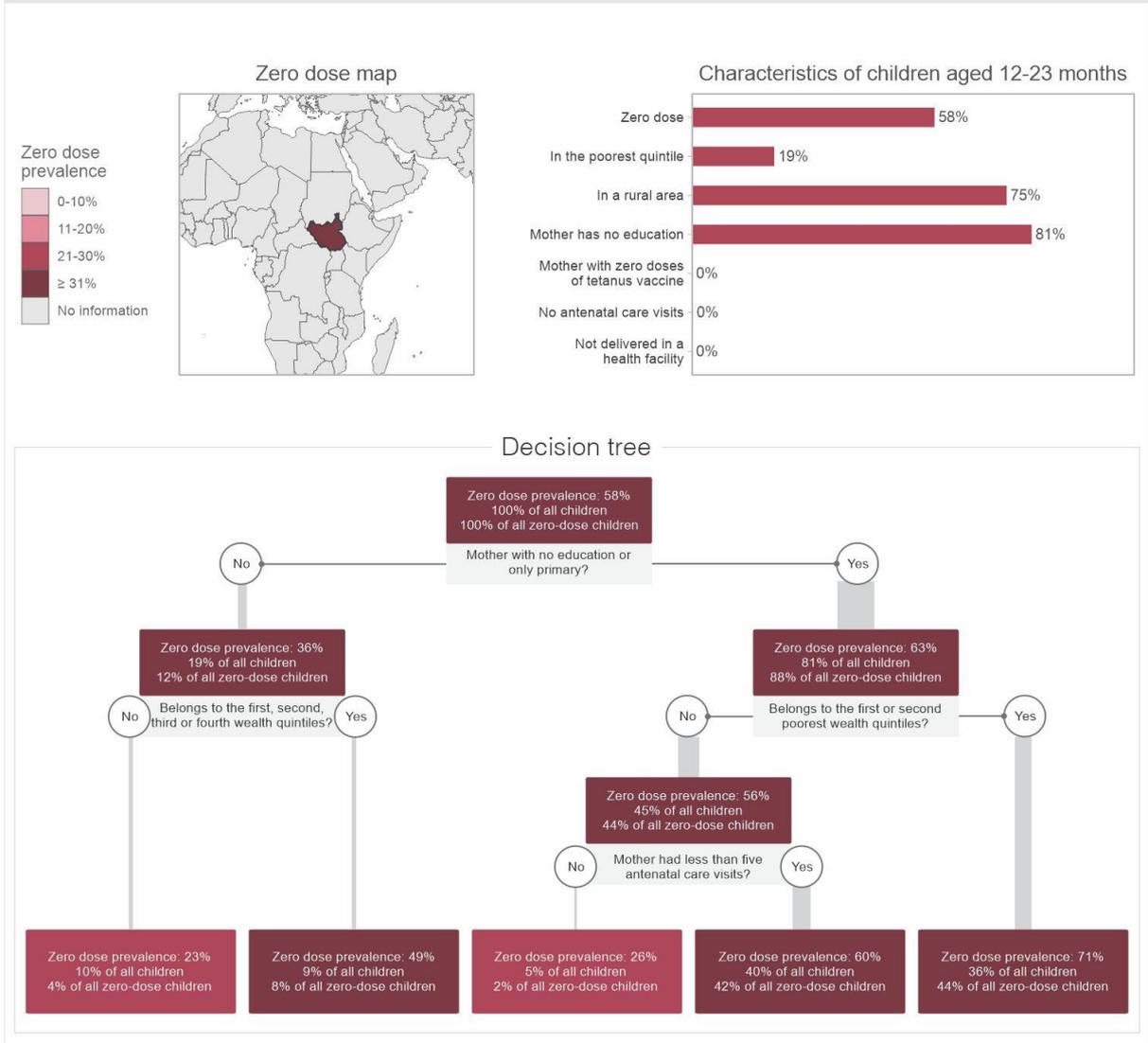


692

693

Figure S76 – South Africa’s country profile

South Sudan, 2010

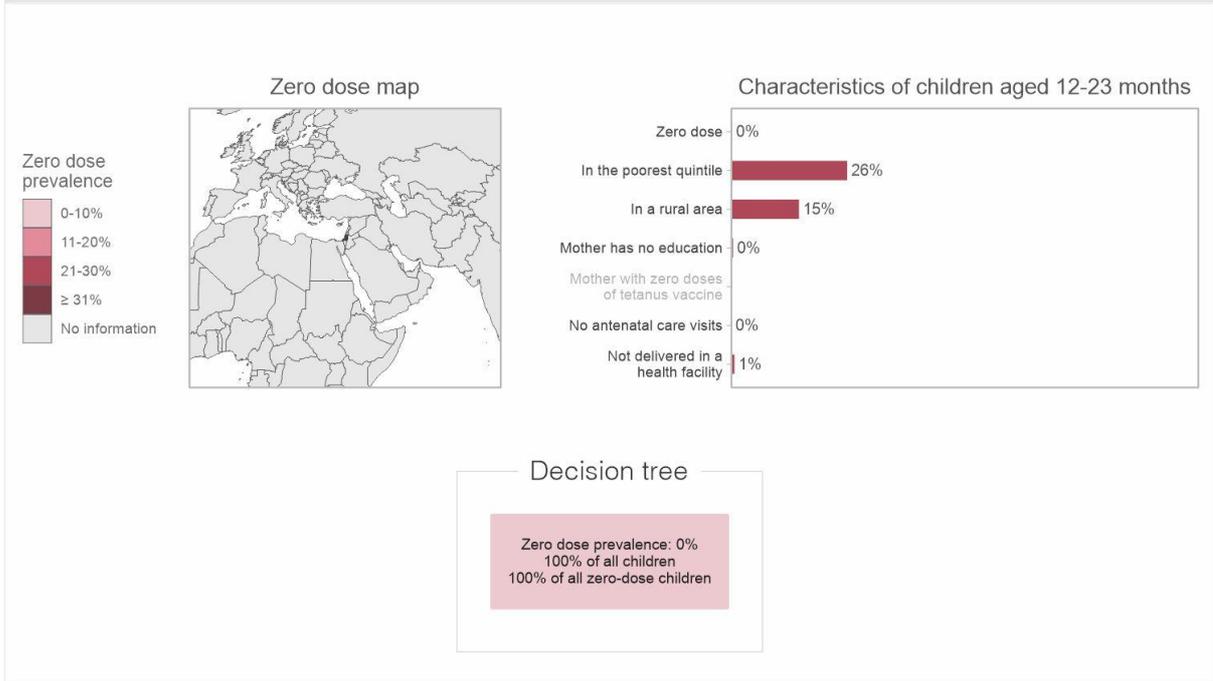


694

695

Figure S77 – South Sudan's country profile

State of Palestine, 2014

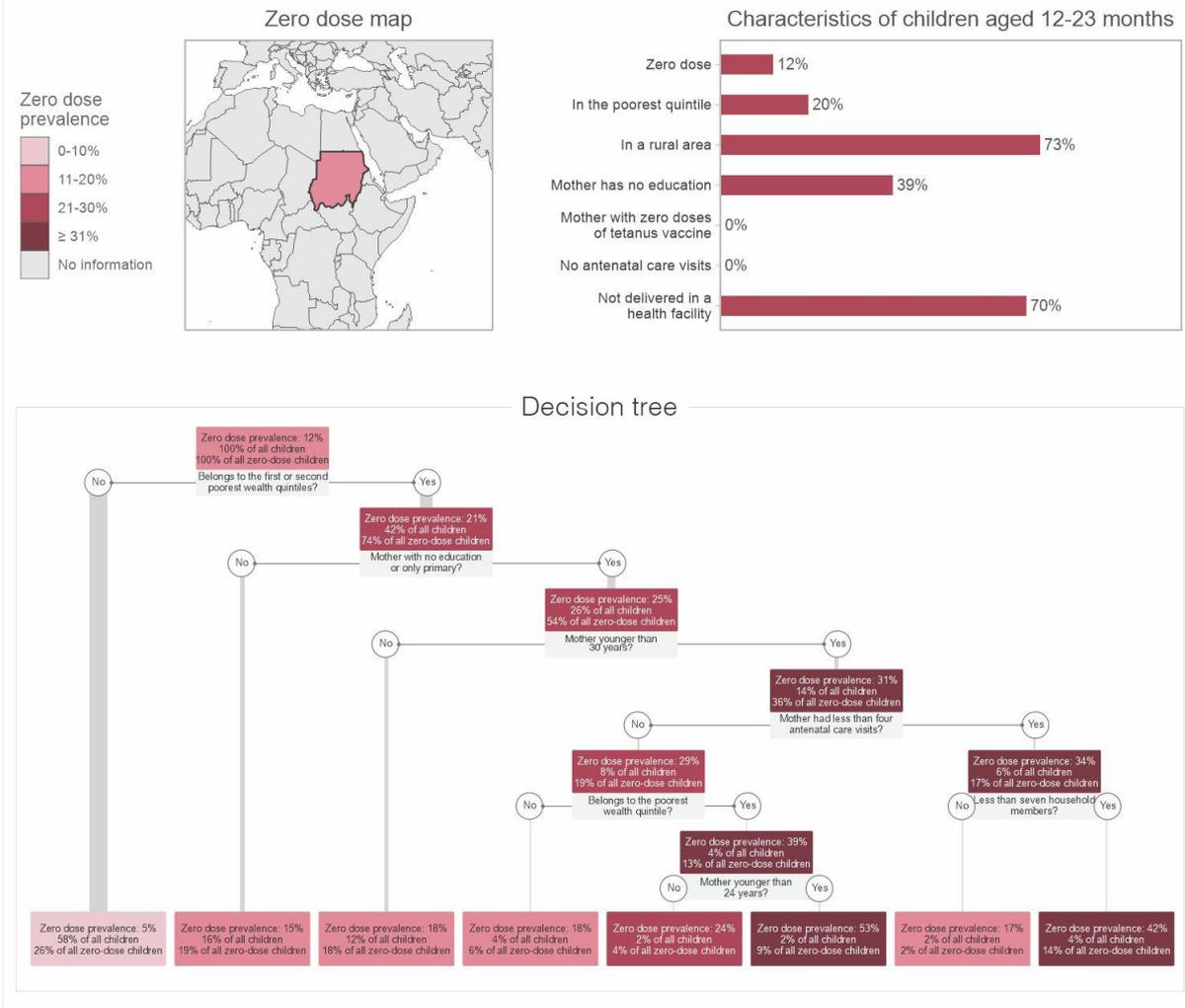


696

697

Figure S78 – State of Palestine’s country profile

Sudan, 2014

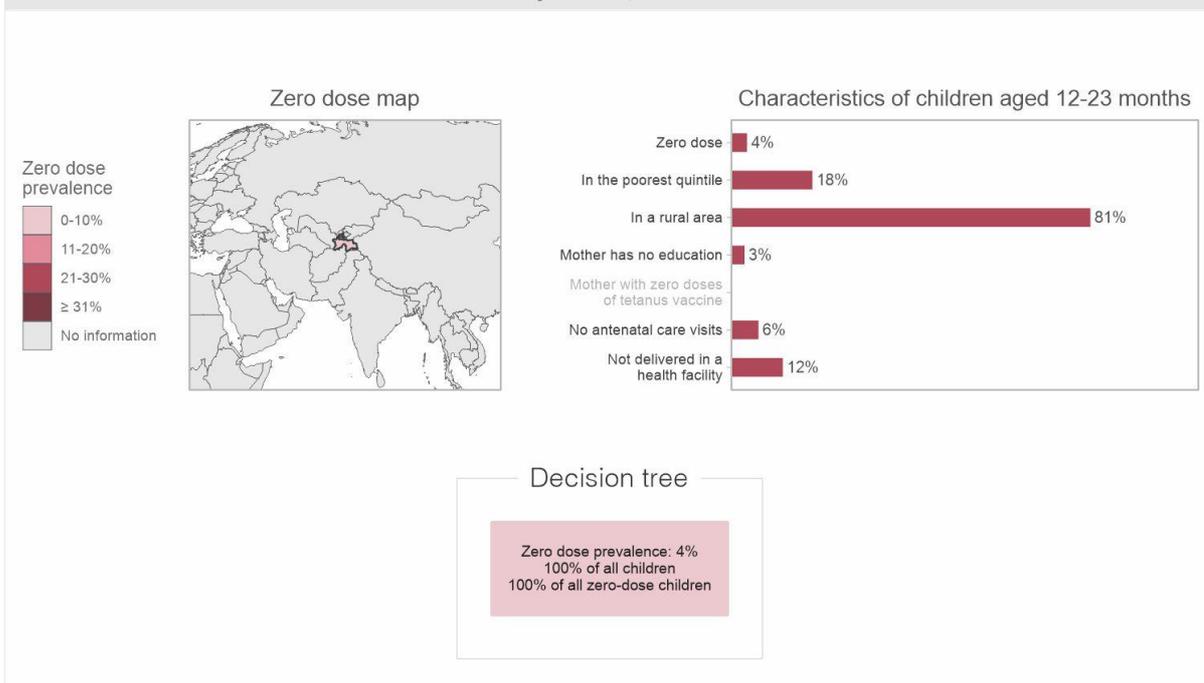


698

699

Figure S79 – Sudan’s country profile

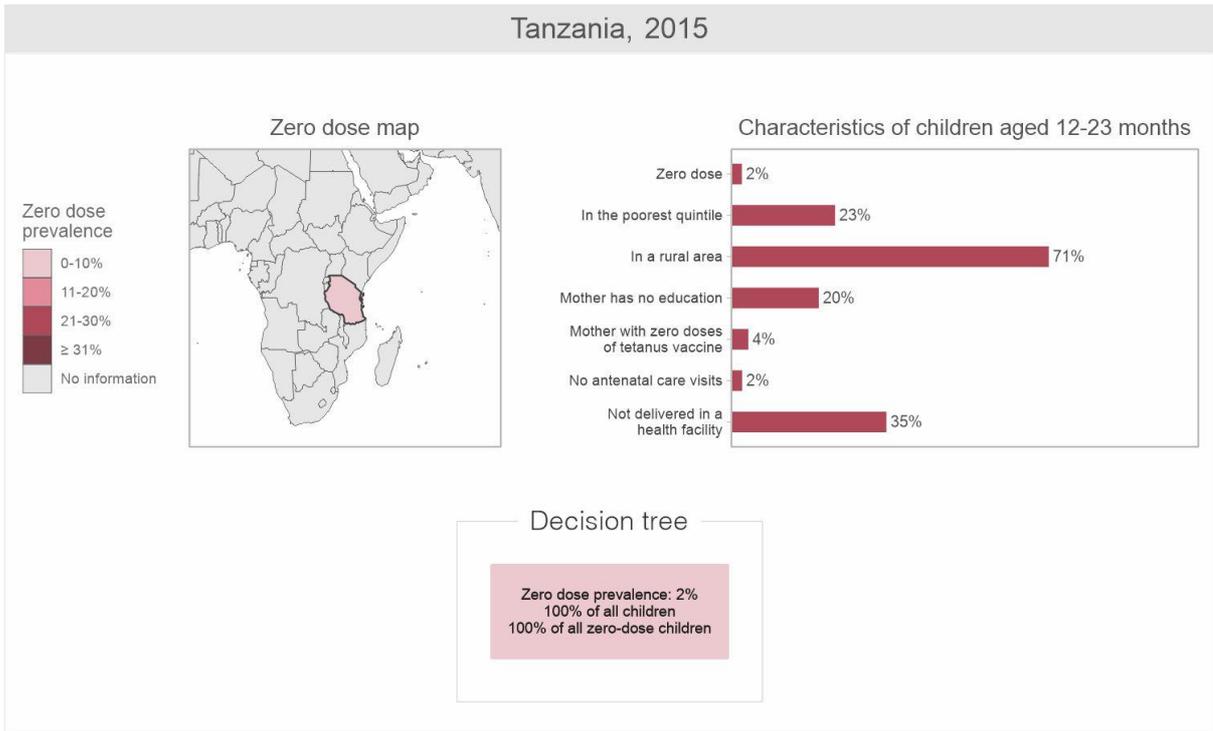
Tajikistan, 2017



700

701

Figure S80 – Tajikistan’s country profile

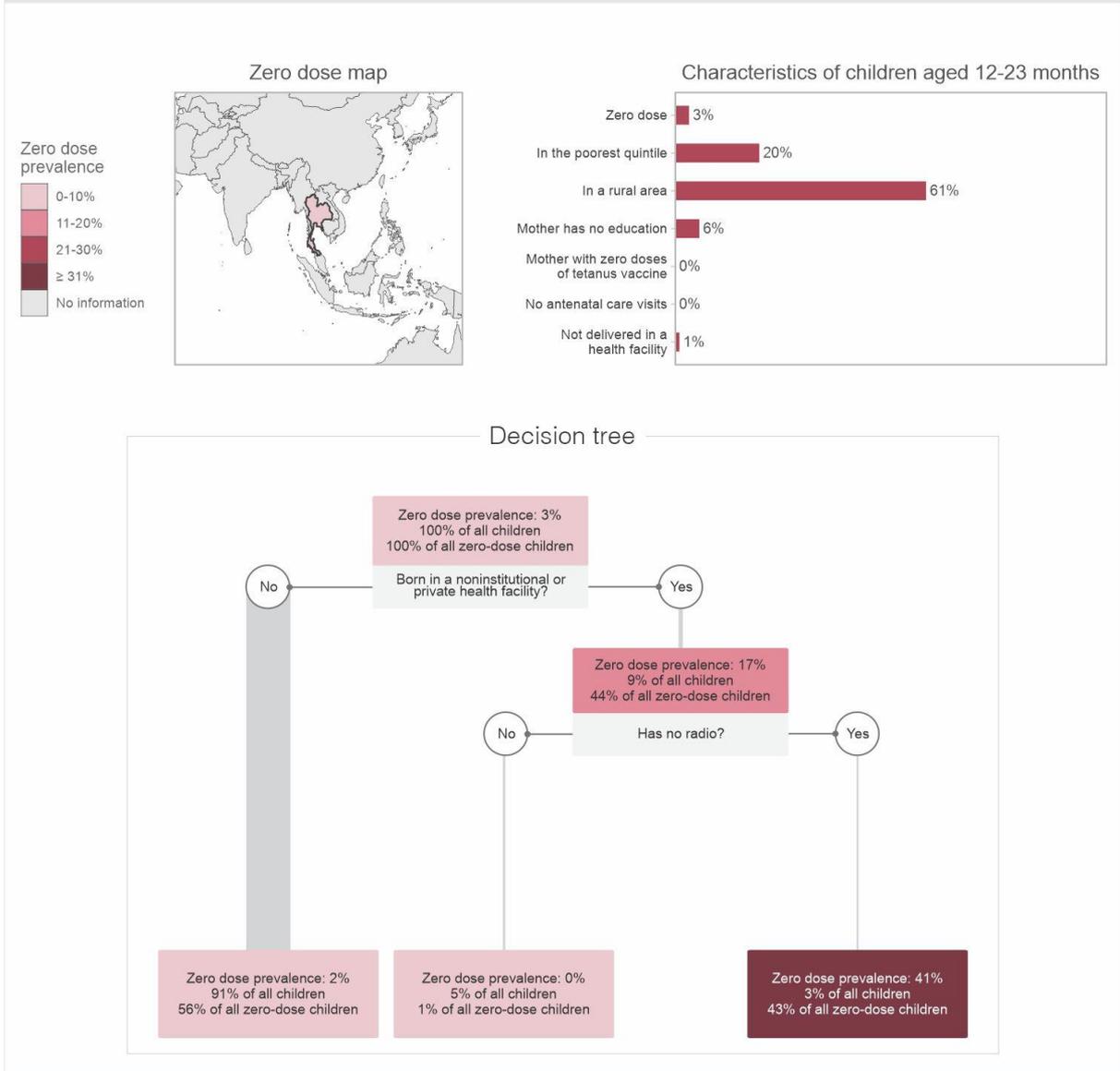


702

703

Figure S81 – Tanzania’s country profile

Thailand, 2015

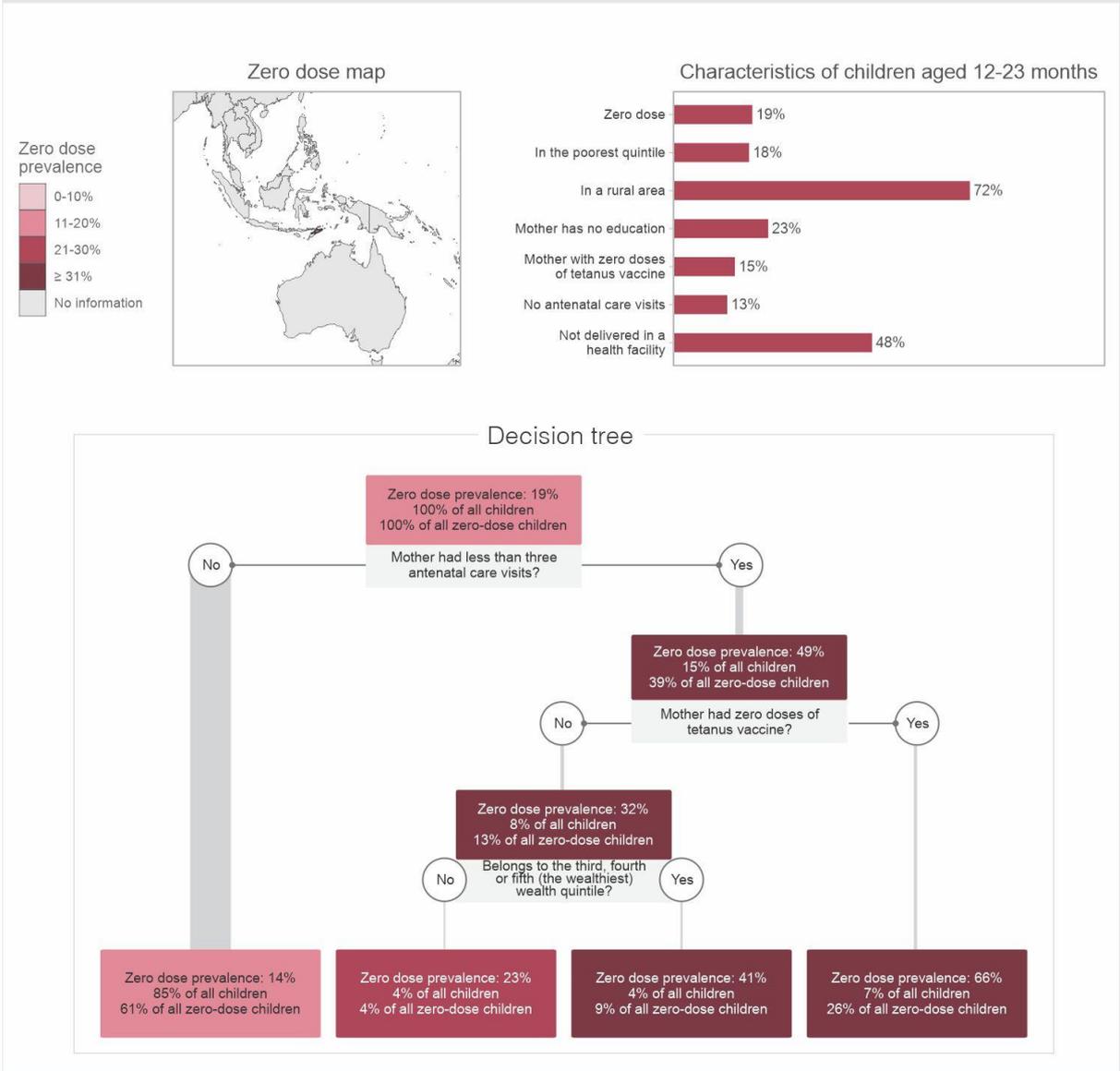


704

705

Figure S82 – Thailand's country profile

Timor-Leste, 2016

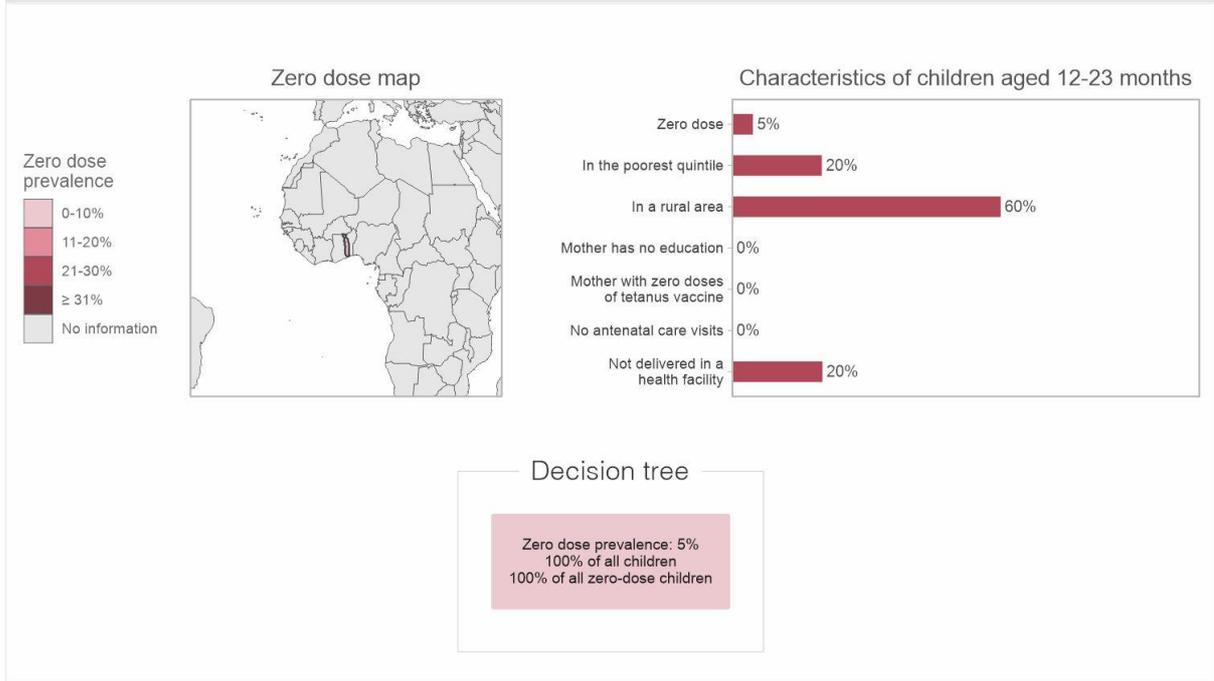


706

707

Figure S83 – Timor-Leste's country profile

Togo, 2017

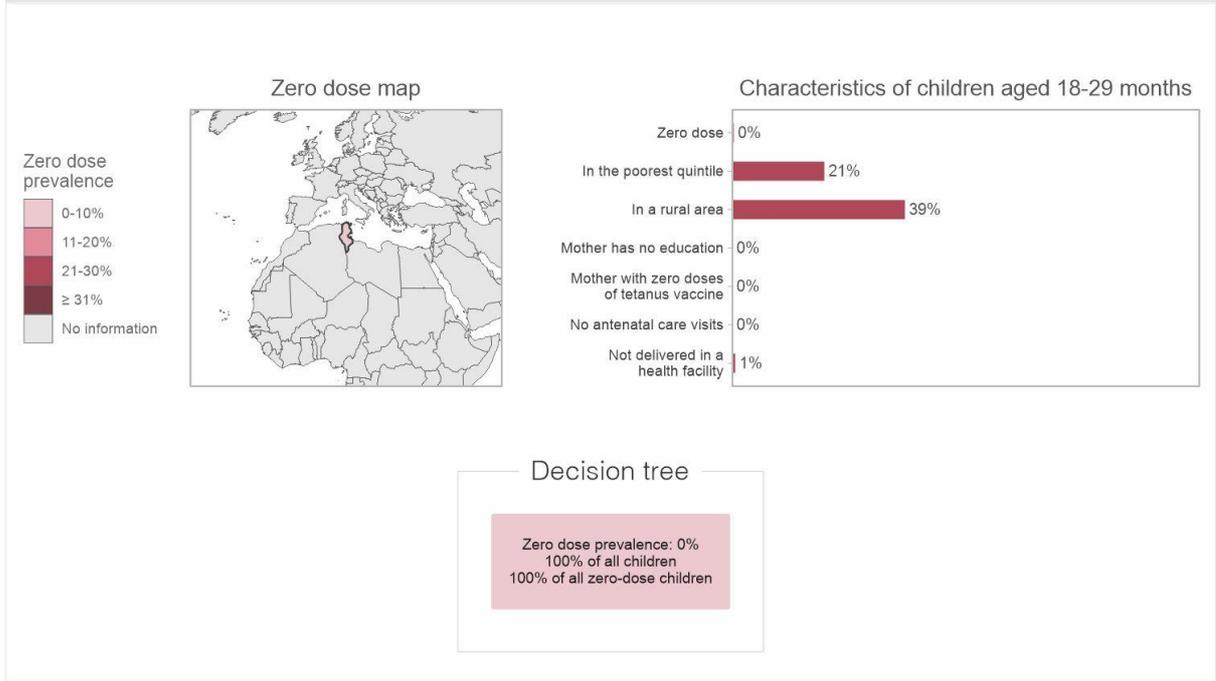


708

709

Figure S84 – Togo's country profile

Tunisia, 2018

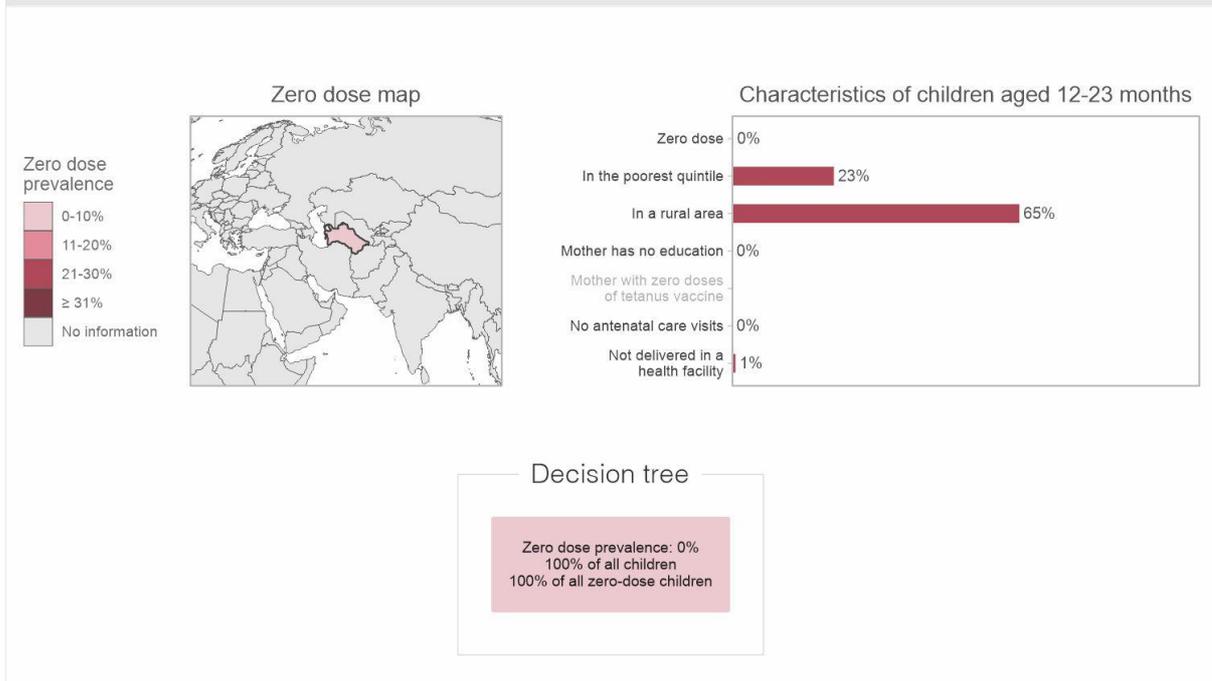


710

711

Figure S85 – Tunisia’s country profile

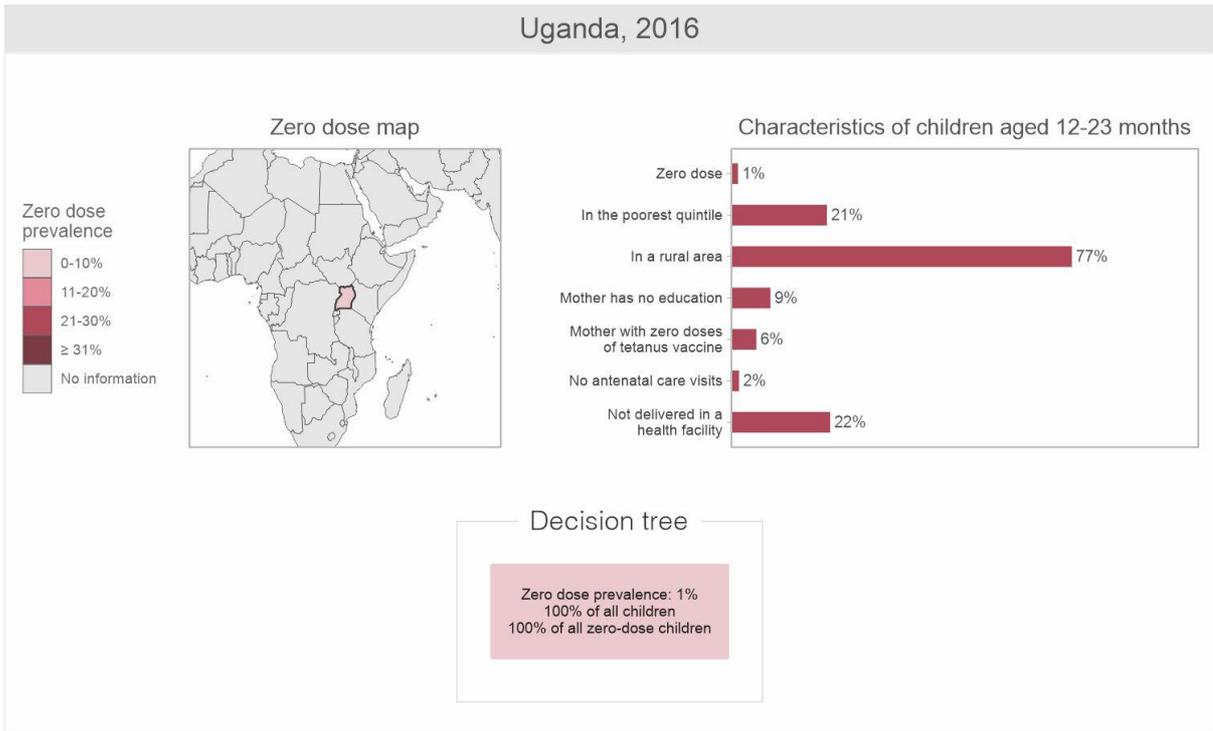
Turkmenistan, 2015



712

713

Figure S86 – Turkmenistan’s country profile

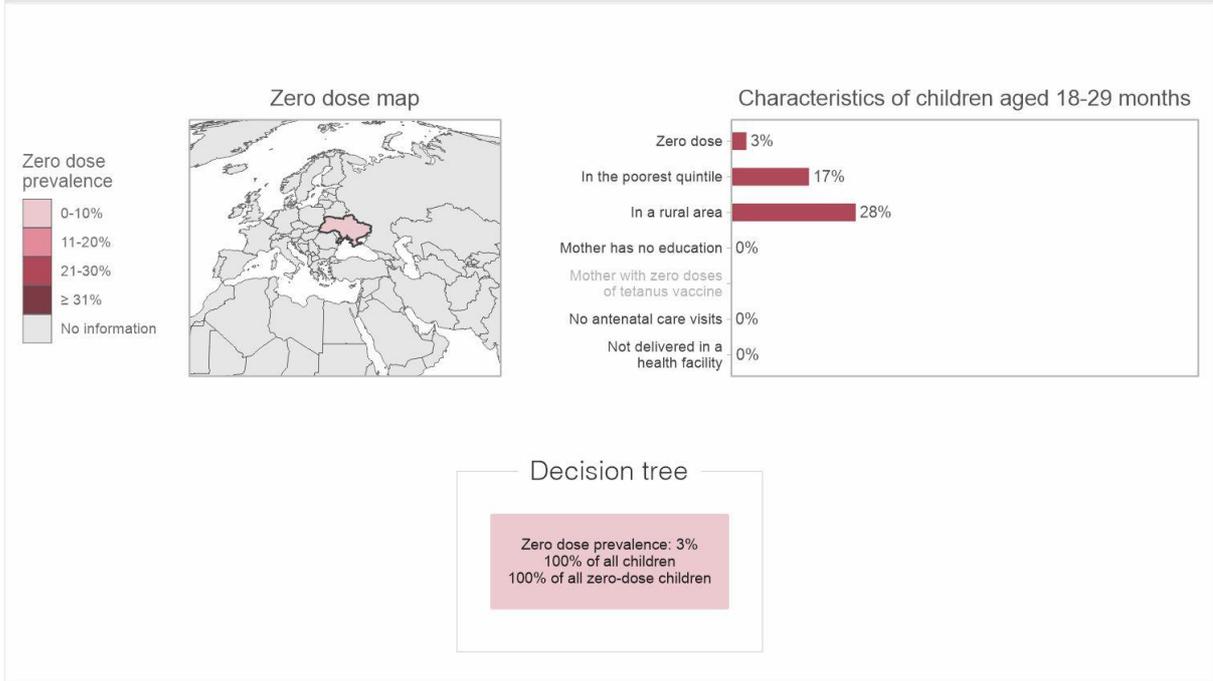


714

715

Figure S87 – Uganda’s country profile

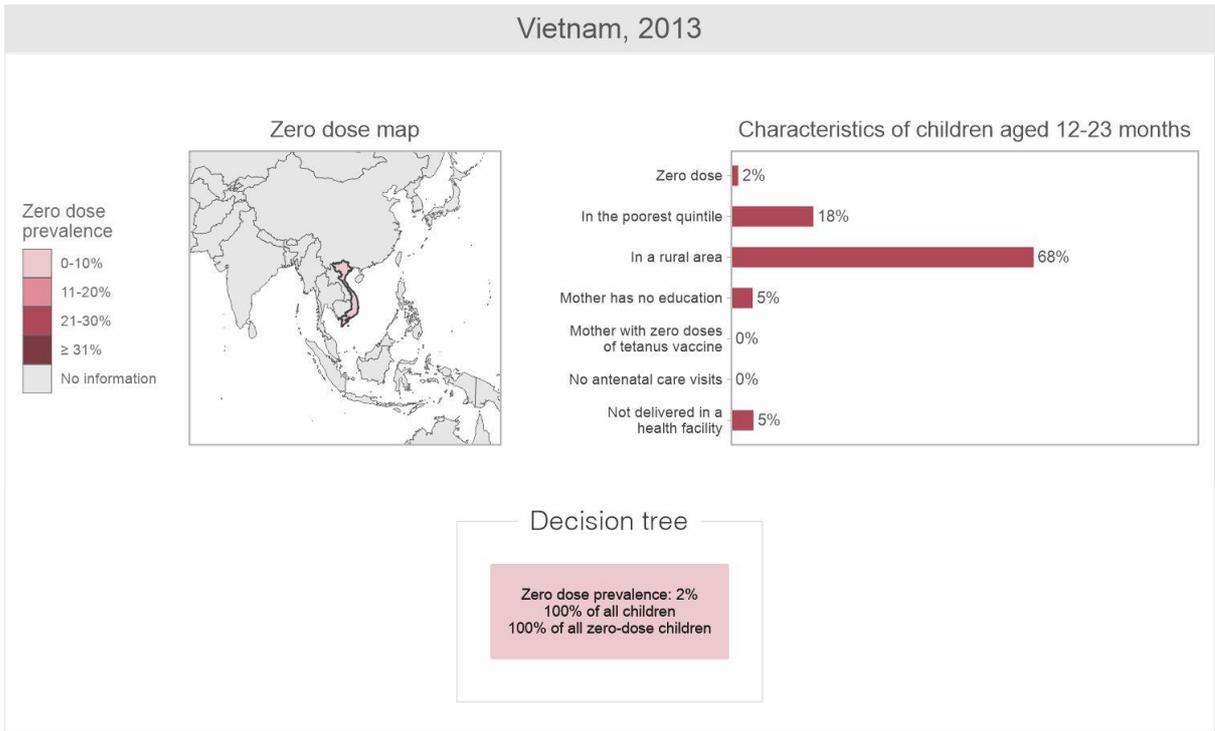
Ukraine, 2012



716

717

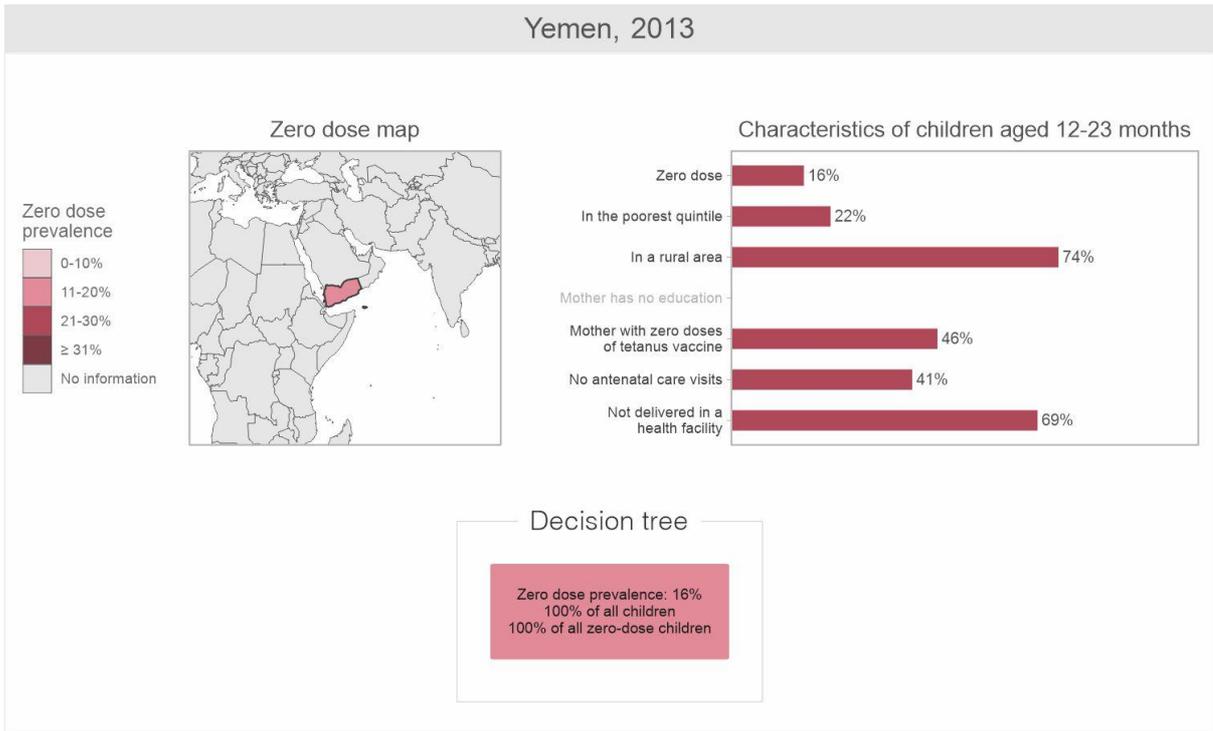
Figure S88 – Ukraine’s country profile



718

719

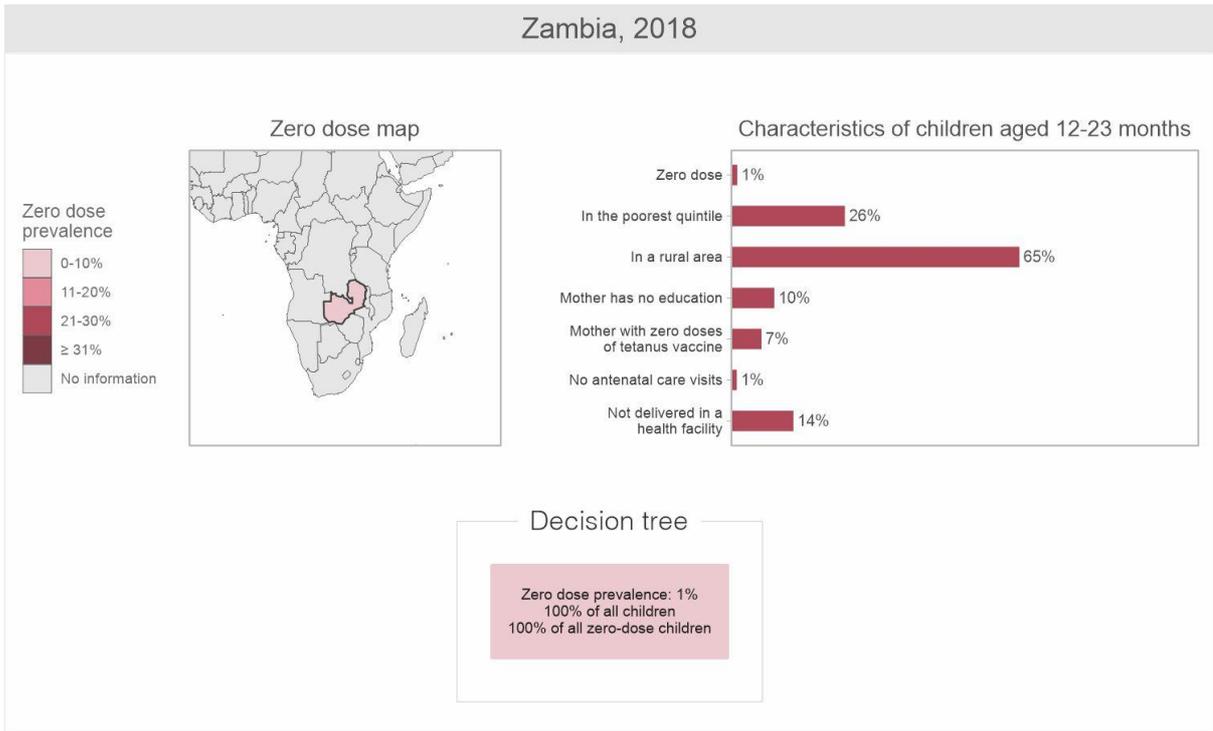
Figure S89 – Vietnam’s country profile



720

721

Figure S90 – Yemen’s country profile

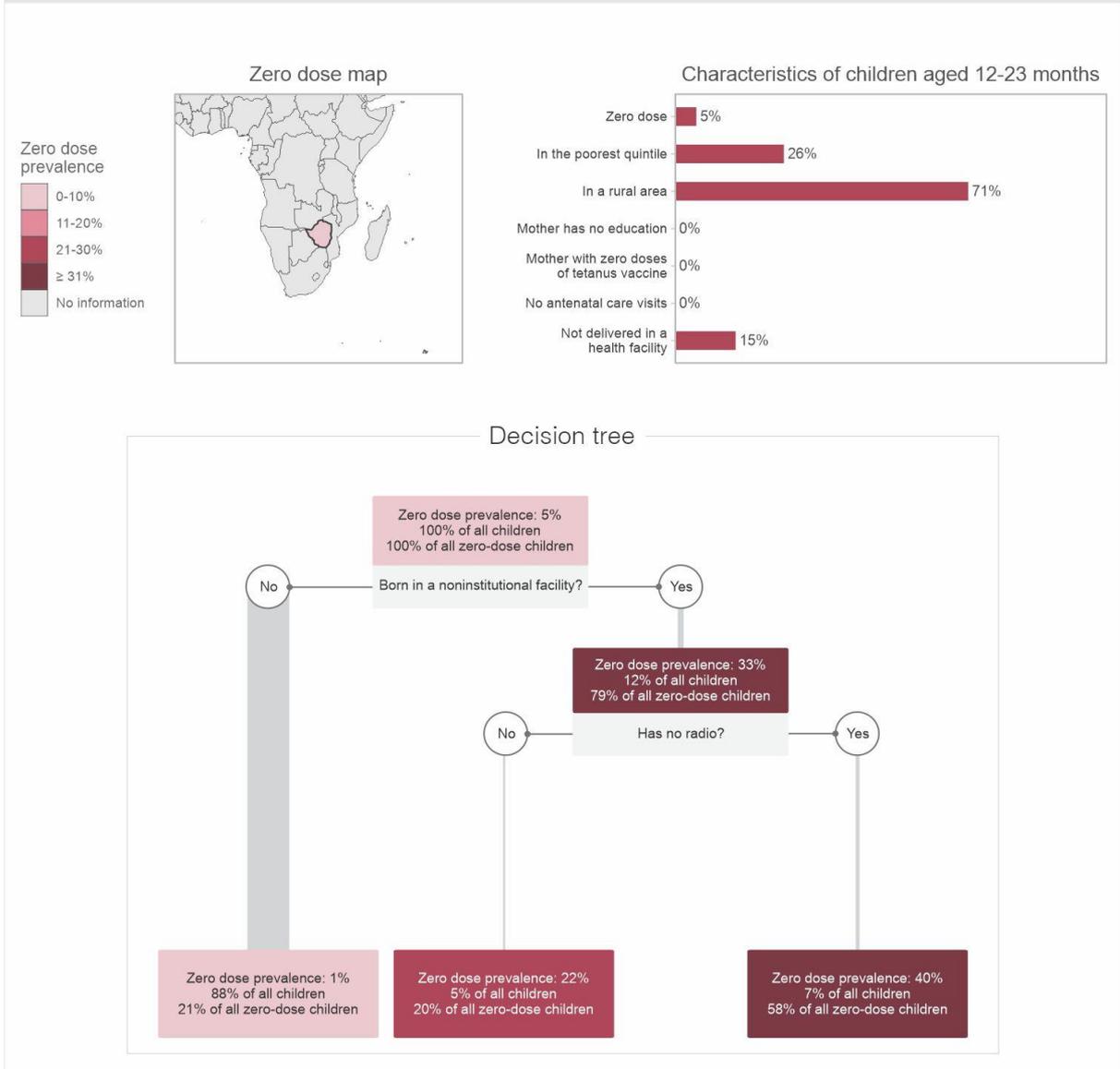


722

723

Figure S91 – Zambia’s country profile

Zimbabwe, 2019



724

725

Figure S92 – Zimbabwe’s country profile

6. COMUNICAÇÃO PARA A IMPRENSA

Quem são as crianças que não estão sendo vacinadas?

Estima-se que, anualmente, 2 a 3 milhões de mortes sejam evitadas pela aplicação em massa de vacinas. Mas elas não atingem toda a população mundial igualmente: 13 milhões de crianças nunca foram vacinadas em todo o mundo. Encontrar tais crianças é o objetivo da dissertação de mestrado intitulada "Identificação e descrição de grupos de risco de crianças nunca vacinadas em países de baixa e média renda por meio de árvores de decisão", desenvolvida por Thiago Melo no Programa de Pós-Graduação em Epidemiologia da UFPel, sob orientação do Professor Aluísio Barros.

Identificar crianças com alto risco de não serem vacinadas é crucial para o alcançar os Objetivos do Desenvolvimento Sustentável das Nações Unidas e seu lema central: não deixar ninguém para trás. Segundo os resultados do estudo – que avaliou mais de 200 mil crianças em 92 países – as crianças em maior risco de não serem vacinadas são aquelas cujas mães não receberam os serviços mais básicos de vacinação, atendimento pré-natal e parto em uma unidade de saúde. São crianças em situação de risco não só para as doenças que essas vacinas devem prevenir, mas também em risco de não terem acesso a todo um conjunto de serviços de saúde pública que deveria assegurar seu direito à saúde, bem-estar, cuidado e assistência especiais garantidos na Declaração Universal dos Direitos Humanos. “Percebemos que as mesmas famílias estão sendo deixadas para trás por vários serviços de saúde”, diz o autor, “e campanhas de vacinação infantil desconectadas de um cuidado mais amplo são uma oportunidade perdida. É necessária maior integração entre os serviços básicos de saúde nesses países.”

O estudo não investigava apenas a vacinação, mas também o método estatístico conhecido como “árvores de decisão”. Apesar de ainda ser pouco usado em pesquisas na área da saúde, os resultados são promissores: “Já estamos aplicando o mesmo método em outros importantes problemas de saúde pública, como a violência contra mulher. As árvores de decisão são um método bem flexível na busca por grupos de risco, seja de vacinação ou violência. E o resultado é bem simples de ser interpretado. Ainda temos muito a investigar, mas as árvores podem ser uma ferramenta importante para ajudar gestores a criar programas de saúde mais direcionados e adaptados às diferentes realidades de cada população”, conclui.