Title

The role of human mobility and land cover change in the epidemiology of malaria in the Peruvian Amazon

Author

Carrasco-Escobar, Gabriel

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

The role of human mobility and land cover change in the epidemiology of malaria in the Peruvian Amazon

A Dissertation submitted in partial satisfaction of the requirements

for the degree Doctor of Philosophy

in

Public Health (Global Health)

by

Gabriel Carrasco Escobar

Committee in charge:

University of California San Diego

Tarik Benmarhnia, Chair
Kimberly Brouwer
Gordon McCord

San Diego State University

John Elder
Karen Ferran
Eyal Oren

2023

The Dissertation of Gabriel Carrasco Escobar is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

_____

_____

_____

_____

_____

_____

Chair

University of California San Diego

San Diego State University

2023

Dedication

Para Santiago y Salvador

Table of contents

List of Figures

List of Tables

Acknowledgements

Mi más profundo agradecimiento a Marita Escobar y Bertha Alegría por su fortaleza y amor, y porque todo lo que pasamos valió la pena. A mi esposa Kelly Tello, por nuestra hermosa familia y el hermoso ejemplo que le das a nuestros hijos. A Víctor Carrasco por enseñarme *cuanto control y cuanto amor tiene que haber en una casa*. Y a Abraham Escobar, Raúl (Lin) Carrasco y Martin Tello para que desde donde estén, sepan que lo logramos.

I would also like to express my deep gratitude to Tarik Benmarhnia, my mentor, dissertation chair, and role model of what I want to build in my career. Thanks for all your support, encouragement, and challenges during the last 5 years.

I acknowledge the support for this research from Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (162–2016-FONDECYT), National Institutes of Health (NIH, R01CA228147), National Institute of Allergy and Infectious Diseases (U19AI089681), and NIH/Fogarty International Center Global Infectious Diseases Training Program (D43 TW007120) for the work presented in Chapter 3. The research in Chapter 4 was supported by the Institute of Tropical Medicine Alexander von Humboldt at Universidad Peruana Cayetano Heredia and the University of Alabama at Birmingham (UAB) with funding received for this project through the GORGAS grant-2017 and the National Cancer Institute (R01CA228147).

Chapter 2, in full, has been submitted for publication in Nature Communications. Gabriel Carrasco-Escobar, Diego Villa, Antony Barja, Rachel Lowe, Mercedes Pascual, Alejandro

Llanos-Cuentas, Tarik Benmarhnia. The paper title is “The role of connectivity on malaria dynamics across areas with contrasting control coverage in the Peruvian Amazon”. The dissertation author was the primary researcher and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Scientific Reports 2022. Gabriel Carrasco-Escobar, Jason Rosado, Oscar Nolasco, Michael T. White, Ivo Mueller, Marcia C. Castro, Hugo Rodriguez-Ferruci, Dionicia Gamboa, Alejandro Llanos-Cuentas, Joseph M. Vinetz, Tarik Benmarhnia. The paper title is “Effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon using parametric g-formula”. The dissertation author was the primary researcher and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in Royal Society Open Science 2022. Gabriel Carrasco-Escobar, Jose Matta-Chuquisapon, Edgar Manrique, Jorge Ruiz-Cabrejos, Jose Luis Barboza, Daniel Wong, German Henostroza, Alejandro Llanos-Cuentas, Tarik Benmarhnia. The paper title is “Quantifying the effect of human population mobility on malaria risk in the Peruvian Amazon”. The dissertation author was the primary researcher and author of this paper.

Vita

Education

2023	Doctor of Philosophy in Public Health (Global Health) University of California San Diego & San Diego State University San Diego, California
2016	Master of Science (Epidemiology) Universidad Peruana Cayetano Heredia Lima, Peru
2014	Bachelor of Science (Biology) Universidad Peruana Cayetano Heredia Lima, Peru

Publications

Carrasco-Escobar G, Moreno M, Fornace K, Herrera-Varela M, Manrique E, Conn JE. The use of drones for mosquito surveillance and control. Parasit Vectors. 2022;15(1):1-21. doi:10.1186/s13071-022-05580-5

Navarro R, Paredes JL, Tucto L, Medina C, Angles-Yanqui E, Nario JC, Ruiz-Cabrejos J, Quintana JL, Turpo-Espinoza K, Mejia-Cordero F, Aphang-Lam M, Florez J, **Carrasco-Escobar G**, Ochoa TJ. Bovine lactoferrin for the prevention of COVID-19 infection in health care personnel: a double-blinded randomized clinical trial (LF-COVID). BioMetals. Published online December 7, 2022. doi:10.1007/s10534-022-00477-3

Carrasco-Escobar G, Rosado J, Nolasco O, White MT, Mueller I, Castro MC, Rodriguez-Ferruci H, Gamboa D, Llanos-Cuentas A, Vinetz JM, Benmarhnia T. Effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon using parametric g-formula. Sci Rep. 2022;12(1):19144. doi:10.1038/s41598-022-23528-8

Nottmeyer L, Armstrong B, Lowe R, Abbott S, Meakin S, O'Reilly K, von Borries R, Schneider R, Royé D, Hashizume M, Pascal M, Tobias A, Vicedo-Cabrera AM, Lavigne E, Correa PM, Ortega NV, Kynčl J, Urban A, Orru H, Ryti N, Jaakkola J, Dallavalle M, Schneider A, Honda Y, Ng CFS, Alahmad B, **Carrasco G**, Holobâc IH, Kim H, Lee W, Íñiguez C, Bell ML, Zanobetti A, Schwartz J, Scovronick N, Coélho M de SZS, Saldiva PHN, Diaz MH, Gasparrini A, Sera F. The association of COVID-19 incidence with temperature, humidity, and UV radiation – A global multi-city analysis. Sci Total Environ. Published online September 7, 2022:158636. doi:10.1016/j.scitotenv.2022.158636

Carrasco-Escobar G, Matta-Chuquisapon J, Manrique E, Ruiz-Cabrejos J, Barboza JL, Wong D, Henostroza G, Llanos-Cuentas A, Benmarhnia T. Quantifying the effect of human population mobility on malaria risk in the Peruvian Amazon. R Soc Open Sci. 9(7):211611. doi:10.1098/rsos.211611

Matta-Chuquisapon J, Gianella C, **Carrasco-Escobar G**. Missed opportunities for vaccination in Peru 2010–2020: A study of socioeconomic inequalities. Lancet Reg Health – Am. 2022;14. doi:10.1016/j.lana.2022.100321

Moreno M, Torres K, Tong C, Castillo SSG, **Carrasco-Escobar G**, Guedez G, Torres L, Herrera-Varela M, Guerra L, Guzman M, Wong D, Ramirez R, Llanos-Cuentas A, Conn JE, Gamboa D, Vinetz JM. Insights into Plasmodium vivax Asymptomatic Malaria Infections and Direct Skin-Feeding Assays to Assess Onward Malaria Transmission in the Amazon. Am J Trop Med Hyg. 2022;1(aop). doi:10.4269/ajtmh.21-1217

Wu Y, Li S, Zhao Q, Wen B, Gasparrini A, Tong S, Overcenco A, Urban A, Schneider A, Entezari A, Vicedo-Cabrera AM, Zanobetti A, Analitis A, Zeka A, Tobias A, Nunes B, Alahmad B, Armstrong B, Forsberg B, Pan SC, Íñiguez C, Ameling C, Valencia CD la C, Åström C, Houthuijs D, Dung DV, Royé D, Indermitte E, Lavigne E, Mayvaneh F, Acquaotta F, de'Donato F, Rao S, Sera F, **Carrasco-Escobar G**, Kan H, Orru H, Kim H, Holobaca IH, Kyselý J, Madureira J, Schwartz J, Jaakkola JJK, Katsouyanni K, Diaz MH, Ragettli MS, Hashizume M, Pascal M, Coélho M de SZS, Ortega NV, Ryti N, Scovronick N, Michelozzi P, Correa PM, Goodman P, Saldiva PHN, Abrutzky R, Osorio S, Dang TN, Colistro V, Huber V, Lee W, Seposo X, Honda Y, Guo YL, Bell ML, Guo Y. Global, regional, and national burden of mortality associated with short-term temperature variability from 2000–19: a three-stage modelling study. Lancet Planet Health. 2022;6(5):e410-e421. doi:10.1016/S2542-5196(22)00073-0

Rosado J, **Carrasco-Escobar G**, Nolasco O, Garro K, Rodriguez-Ferruci H, Guzman-Guzman M, Llanos-Cuentas A, Vinetz JM, Nekkab N, White MT, Mueller I, Gamboa D. Malaria transmission structure in the Peruvian Amazon through antibody signatures to Plasmodium vivax. PLoS Negl Trop Dis. 2022;16(5):e0010415. doi:10.1371/journal.pntd.0010415

Alvarez MVN, Alonso DP, Kadri SM, Rufalco-Moutinho P, Bernardes IAF, de Mello ACF, Souto AC, **Carrasco-Escobar G**, Moreno M, Gamboa D, Vinetz JM, Conn JE, Ribolla PEM. Nyssorhynchus darlingi genome-wide studies related to microgeographic dispersion and blood-seeking behavior. Parasit Vectors. 2022;15(1):106. doi:10.1186/s13071-022-05219-5

Wu Y, Wen B, Li S, Gasparrini A, Tong S, Overcenco A, Urban A, Schneider A, Entezari A, Vicedo-Cabrera AM, Zanobetti A, Analitis A, Zeka A, Tobias A, Alahmad B, Armstrong B, Forsberg B, Íñiguez C, Ameling C, Valencia CD la C, Åström C, Houthuijs D, Dung DV, Royé D, Indermitte E, Lavigne E, Mayvaneh F, Acquaotta F, de'Donato F, Sera F, **Carrasco G**, Kan H, Orru H, Kim H, Holobaca IH, Kyselý J, Madureira J, Schwartz J, Katsouyanni K, Hurtado-Diaz M, Ragettli MS, Hashizume M, Pascal M, Coélho M de SZS, Scovronick N, Michelozzi P, Goodman P, Saldiva PHN, Abrutzky R, Osorio S, Dang TN, Colistro V, Huber V, Lee W, Seposo X, Honda Y, Bell ML, Guo Y. Fluctuating temperature modifies heat-mortality association in the globe. The Innovation. 2022;0(0). doi:10.1016/j.xinn.2022.100225

Letellier N, Zamora S, Spoon C, Yang JA, Mortamais M, **Carrasco-Escobar G**, Sears DD, Jankowska MM, Benmarhnia T. Air pollution and metabolic disorders: Dynamic versus static measures of exposure among Hispanics/Latinos and non-Hispanics. Environ Res. 2022;209:112846. doi:10.1016/j.envres.2022.112846

Schwarz L, Castillo EM, Chan TC, Brennan JJ, Sbiroli ES, **Carrasco-Escobar G**, Nguyen A, Clemesha RES, Gershunov A, Benmarhnia T. Heat Waves and Emergency Department Visits Among the Homeless, San Diego, 2012–2019. Am J Public Health. 2022;112(1):98-106. doi:10.2105/AJPH.2021.306557

Byrne I, Chan K, Manrique E, Lines J, Wolie RZ, Trujillano F, Garay GJ, Del Prado Cortez MN, Alatrista-Salas H, Sternberg E, Cook J, N'Guessan R, Koffi A, Ahoua Alou LP, Apollinaire N, Messenger LA, Kristan M, **Carrasco-Escobar G**, Fornace K. Technical Workflow Development for Integrating Drone Surveys and Entomological Sampling to Characterise Aquatic Larval Habitats of Anopheles funestus in Agricultural Landscapes in Côte d'Ivoire. J Environ Public Health. 2021;2021:e3220244. doi:10.1155/2021/3220244

Sera F, Armstrong B, Abbott S, Meakin S, O'Reilly K, von Borries R, Schneider R, Royé D, Hashizume M, Pascal M, Tobias A, Vicedo-Cabrera AM, Gasparrini A, Lowe R. A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries. Nat Commun. 2021;12(1):5968. doi:10.1038/s41467-021-25914-8

Carrasco-Escobar G, Qquellon J, Villa D, Cava R, Llanos-Cuentas A, Benmarhnia T. Time-Varying Effects of Meteorological Variables on Malaria Epidemiology in the Context of Interrupted Control Efforts in the Amazon Rainforest, 2000–2017. Front Med. 2021;8:1649. doi:10.3389/fmed.2021.721515

Tobías A, Hashizume M, Honda Y, Sera F, Ng CFS, Kim Y, Roye D, Chung Y, Dang TN, Kim H, Lee W, Íñiguez C, Vicedo-Cabrera A, Abrutzky R, Guo Y, Tong S, Coelho M de SZS, Saldiva PHN, Lavigne E, Correa PM, Ortega NV, Kan H, Osorio S, Kyselý J, Urban A, Orru H, Indermitte E, Jaakkola JJK, Ryti NRI, Pascal M, Huber V, Schneider A, Katsouyanni K, Analitis A, Entezari A, Mayvaneh F, Goodman P, Zeka A, Michelozzi P, de'Donato F, Alahmad B, Diaz MH, De la Cruz Valencia C, Overcenco A, Houthuijs D, Ameling C, Rao S, Di Ruscio F, **Carrasco G**, Seposo X, Nunes B, Madureira J, Holobaca I-H, Scovronick N, Acquaotta F, Forsberg B, Åström C, Ragettli MS, Guo Y-LL, Chen B-Y, Li S, Colistro V, Zanobetti A, Schwartz J, Dung DV, Armstrong B, Gasparrini A. Geographical Variations of the Minimum Mortality Temperature at a Global Scale: A Multicountry Study. Environ Epidemiol. 2021;5(5):e169. doi:10.1097/EE9.0000000000000169

Chen G, Guo Y, Yue X, Tong S, Gasparrini A, Bell ML, Armstrong B, Schwartz J, Jaakkola JJK, Zanobetti A, Lavigne E, Nascimento Saldiva PH, Kan H, Royé D, Milojevic A, Overcenco A, Urban A, Schneider A, Entezari A, Vicedo-Cabrera AM, Zeka A, Tobias A, Nunes B, Alahmad B, Forsberg B, Pan S-C, Íñiguez C, Ameling C, De la Cruz Valencia C, Åström C, Houthuijs D, Van Dung D, Samoli E, Mayvaneh F, Sera F, **Carrasco-Escobar G**, Lei Y, Orru H, Kim H, Holobaca I-H, Kyselý J, Teixeira JP, Madureira J, Katsouyanni K, Hurtado-Díaz M, Maasikmets M, Ragettli MS, Hashizume M, Stafoggia M, Pascal M, Scortichini M, de Sousa Zanotti Stagliorio Coêlho M, Valdés Ortega N, Ryti NRI, Scovronick N, Matus P, Goodman P, Garland RM, Abrutzky R, Garcia SO, Rao S, Fratianni S, Dang TN, Colistro V, Huber V, Lee W, Seposo X, Honda Y, Guo YL, Ye T, Yu W, Abramson MJ, Samet JM, Li S. Mortality risk attributable to wildfire-related PM2·5 pollution: a global time series study in 749 locations. Lancet Planet Health. 2021;5(9):e579-e587. doi:10.1016/S2542-5196(21)00200-X

Carrasco-Escobar G, Fornace K, Benmarhnia T. Mapping socioeconomic inequalities in malaria in Sub-Sahara African countries. Sci Rep. 2021;11(1):15121. doi:10.1038/s41598-021-94601-x

Zhao Q, Guo Y, Ye T, Gasparrini A, Tong S, Overcenco A, Urban A, Schneider A, Entezari A, Vicedo-Cabrera AM, Zanobetti A, Analitis A, Zeka A, Tobias A, Nunes B, Alahmad B, Armstrong B, Forsberg B, Pan S-C, Íñiguez C, Ameling C, De la Cruz Valencia C, Åström C,

Houthuijs D, Dung DV, Royé D, Indermitte E, Lavigne E, Mayvaneh F, Acquaotta F, de'Donato F, Di Ruscio F, Sera F, **Carrasco-Escobar G**, Kan H, Orru H, Kim H, Holobaca I-H, Kyselý J, Madureira J, Schwartz J, Jaakkola JJK, Katsouyanni K, Hurtado Diaz M, Ragettli MS, Hashizume M, Pascal M, de Sousa Zanotti Stagliorio Coélho M, Valdés Ortega N, Ryti N, Scovronick N, Michelozzi P, Matus Correa P, Goodman P, Nascimento Saldiva PH, Abrutzky R, Osorio S, Rao S, Fratianni S, Dang TN, Colistro V, Huber V, Lee W, Seposo X, Honda Y, Guo YL, Bell ML, Li S. Global, regional, and national burden of mortality associated with non-optimal ambient temperatures from 2000 to 2019: a three-stage modelling study. Lancet Planet Health. 2021;5(7):e415-e425. doi:10.1016/S2542-5196(21)00081-4

Marks C, Abramovitz D, Donnelly CA, **Carrasco-Escobar G**, Carrasco-Hernández R, Ciccarone D, González-Izquierdo A, Martin NK, Strathdee SA, Smith DM, Bórquez A. Identifying counties at risk of high overdose mortality burden during the emerging fentanyl epidemic in the USA: a predictive statistical modelling study. Lancet Public Health. 2021;0(0). doi:10.1016/S2468-2667(21)00080-3

Vicedo-Cabrera AM, Scovronick N, Sera F, Royé D, Schneider R, Tobias A, Astrom C, Guo Y, Honda Y, Hondula DM, Abrutzky R, Tong S, Coelho M de SZS, Saldiva PHN, Lavigne E, Correa PM, Ortega NV, Kan H, Osorio S, Kyselý J, Urban A, Orru H, Indermitte E, Jaakkola JJK, Ryti N, Pascal M, Schneider A, Katsouyanni K, Samoli E, Mayvaneh F, Entezari A, Goodman P, Zeka A, Michelozzi P, de'Donato F, Hashizume M, Alahmad B, Diaz MH, Valencia CDLC, Overcenco A, Houthuijs D, Ameling C, Rao S, Di Ruscio F, **Carrasco-Escobar G**, Seposo X, Silva S, Madureira J, Holobaca IH, Fratianni S, Acquaotta F, Kim H, Lee W, Iniguez C, Forsberg B, Ragettli MS, Guo YLL, Chen BY, Li S, Armstrong B, Aleman A, Zanobetti A, Schwartz J, Dang TN, Dung DV, Gillett N, Haines A, Mengel M, Huber V, Gasparrini A. The burden of heat-related mortality attributable to recent human-induced climate change. Nat Clim Change. Published online May 31, 2021:1-9. doi:10.1038/s41558-021-01058-x

Rosas-Aguirre A, Moreno M, Moreno-Gutierrez D, Llanos-Cuentas A, Saavedra M, Contreras-Mancilla J, Barboza J, Alava F, Aguirre K, **Carrasco-Escobar G**, Prussing C, Vinetz J, Conn JE, Speybroeck N, Gamboa D. Integrating Parasitological and Entomological Observations to Understand Malaria Transmission in Riverine Villages in the Peruvian Amazon. J Infect Dis. 2021;223(Supplement_2):S99-S110. doi:10.1093/infdis/jiaa496

Lowe R, Lee SA, O'Reilly KM, Brady OJ, Bastos L, **Carrasco-Escobar G**, Catão R de C, Colón-González FJ, Barcellos C, Carvalho MS, Blangiardo M, Rue H, Gasparrini A. Combined effects of hydrometeorological hazards and urbanisation on dengue risk in Brazil: a spatiotemporal modelling study. Lancet Planet Health. 2021;5(4):e209-e219. doi:10.1016/S2542-5196(20)30292-8

Rufalco-Moutinho P, Kadri SM, Alonso DP, Moreno M, **Carrasco-Escobar G**, Prussing C, Gamboa D, Vinetz JM, Sallum MAM, Conn JE, Ribolla PEM. Ecology and larval population dynamics of the primary malaria vector Nyssorhynchus darlingi in a high transmission setting dominated by fish farming in western Amazonian Brazil. PLOS ONE. 2021;16(4):e0246215. doi:10.1371/journal.pone.0246215

Marks C, **Carrasco-Escobar G**, Carrasco-Hernandez R, Johnson D, Ciccarone D, Strathdee SA, Smith D, Bórquez A. Methodological approaches for the prediction of opioid use-related

epidemics in the United States: a narrative review and cross-disciplinary call to action. Transl Res. Published online March 31, 2021. doi:10.1016/j.trsl.2021.03.018

Bussalleu A, Di-Liberto A, Carcamo C, **Carrasco-Escobar G**, Zavaleta-Cortijo C, King M, Berrang-Ford L, Maurtua D, Llanos-Cuentas A, Garcia P, Harper SL, Edge V, Ford J, Lwasa S, Namanya DB, Indigenous Health and Adaption to Climate Change Research Group (IHACC). Cultural Values and the Coliform Bacterial Load of "Masato," an Amazon Indigenous Beverage. EcoHealth. Published online November 20, 2020. doi:10.1007/s10393-020-01498-5

Carrasco-Escobar G, Fornace K, Wong D, Padilla-Huamantinco PG, Saldaña-Lopez JA, Castillo-Meza OE, Caballero-Andrade AE, Manrique E, Ruiz-Cabrejos J, Barboza JL, Rodriguez H, Henostroza G, Gamboa D, Castro MC, Vinetz JM, Llanos-Cuentas A. Open-Source 3D Printable GPS Tracker to Characterize the Role of Human Population Movement on Malaria Epidemiology in River Networks: A Proof-of-Concept Study in the Peruvian Amazon. Front Public Health. 2020;8. doi:10.3389/fpubh.2020.526468

Carrasco-Escobar G, Manrique E, Tello-Lizarraga K, Miranda JJ. Travel Time to Health Facilities as a Marker of Geographical Accessibility Across Heterogeneous Land Coverage in Peru. Front Public Health. 2020;8. doi:10.3389/fpubh.2020.00498

Fletcher IK, Stewart-Ibarra AM, Sippy R, **Carrasco-Escobar G**, Silva M, Beltran-Ayala E, Ordoñez T, Adrian J, Sáenz FE, Drakeley C, Jones KE, Lowe R. The Relative Role of Climate Variation and Control Interventions on Malaria Elimination Efforts in El Oro, Ecuador: A Modeling Study. Front Environ Sci. 2020;8. doi:10.3389/fenvs.2020.00135

Rosas-Aguirre A, Guzman-Guzman M, Chuquiyauri R, Moreno M, Manrique P, Ramirez R, **Carrasco-Escobar G**, Rodriguez H, Speybroeck N, Conn JE, Gamboa D, Vinetz JM, Llanos-Cuentas A. Temporal and micro-spatial heterogeneity in transmission dynamics of co-endemic Plasmodium vivax and Plasmodium falciparum in two rural cohort populations in the Peruvian Amazon. J Infect Dis. Published online August 24, 2020. doi:10.1093/infdis/jiaa526

Curioso WH, **Carrasco-Escobar G**. Collaboration in times of COVID-19: the urgent need for open-data sharing in Latin America. BMJ Health Care Inform. 2020;27(1). doi:10.1136/bmjhci-2020-100159

Loyola S, Valle A, Montero S, **Carrasco-Escobar G**. Recomendaciones para describir de forma adecuada una curva epidémica de COVID-19. Rev Peru Med Exp Salud Pública. 2020;37(2):378-380. doi:10.17843/rpmesp.2020.372.5461

Carrasco-Escobar G, Schwarz L, Miranda JJ, Benmarhnia T. Revealing the air pollution burden associated with internal Migration in Peru. Sci Rep. 2020;10(1):1-12. doi:10.1038/s41598-020-64043-y

Carrasco-Escobar G, Schwalb A, Tello-Lizarraga K, Vega-Guerovich P, Ugarte-Gil C. Spatio-temporal co-occurrence of hotspots of tuberculosis, poverty and air pollution in Lima, Peru. Infect Dis Poverty. 2020;9. doi:10.1186/s40249-020-00647-w

Castro MC, Baeza A, Codeço CT, Cucunubá ZM, Dal'Asta AP, Leo GAD, Dobson AP, **Carrasco-Escobar G**, Lana RM, Lowe R, Monteiro AMV, Pascual M, Santos-Vega M. Development, environmental degradation, and disease spread in the Brazilian Amazon. PLOS Biol. 2019;17(11):e3000526. doi:10.1371/journal.pbio.3000526

Manrique P, Miranda-Alban J, Alarcon-Baldeon J, Ramirez R, **Carrasco-Escobar G**, Herrera H, Guzman-Guzman M, Rosas-Aguirre A, Llanos-Cuentas A, Vinetz JM, Escalante AA, Gamboa D. Microsatellite analysis reveals connectivity among geographically distant transmission zones of Plasmodium vivax in the Peruvian Amazon: A critical barrier to regional malaria elimination. PLoS Negl Trop Dis. 2019;13(11):e0007876. doi:10.1371/journal.pntd.0007876

Saavedra MP, Conn JE, Alava F, **Carrasco-Escobar G**, Prussing C, Bickersmith SA, Sangama JL, Fernandez-Miñope C, Guzman M, Tong C, Valderrama C, Vinetz JM, Gamboa D, Moreno M. Higher risk of malaria transmission outdoors than indoors by Nyssorhynchus darlingi in riverine communities in the Peruvian Amazon. Parasit Vectors. 2019;12(1):374. doi:10.1186/s13071-019-3619-0

Prussing C, Saavedra MP, Bickersmith SA, Alava F, Guzmán M, Manrique E, **Carrasco-Escobar G**, Moreno M, Gamboa D, Vinetz JM, Conn JE. Malaria vector species in Amazonian Peru co-occur in larval habitats but have distinct larval microbial communities. PLoS Negl Trop Dis. 2019;13(5):e0007412. doi:10.1371/journal.pntd.0007412

Carrasco-Escobar G, Castro MC, Barboza JL, Ruiz-Cabrejos J, Llanos-Cuentas A, Vinetz JM, Gamboa D. Use of open mobile mapping tool to assess human mobility traceability in rural offline populations with contrasting malaria dynamics. PeerJ. 2019;7:e6298. doi:10.7717/peerj.6298

Carrasco-Escobar G, Manrique E, Ruiz-Cabrejos J, Saavedra M, Alava F, Bickersmith S, Prussing C, Vinetz JM, Conn JE, Moreno M, Gamboa D. High-accuracy detection of malaria vector larval habitats using drone-based multispectral imagery. PLoS Negl Trop Dis. 2019;13(1):e0007105. doi:10.1371/journal.pntd.0007105

Moreno-Gutierrez D, Llanos-Cuentas A, Luis Barboza J, Contreras-Mancilla J, Gamboa D, Rodriguez H, **Carrasco-Escobar G**, Boreux R, Hayette M-P, Beutels P, Speybroeck N, Rosas-Aguirre A. Effectiveness of a Malaria Surveillance Strategy Based on Active Case Detection during High Transmission Season in the Peruvian Amazon. Int J Environ Res Public Health. 2018;15(12). doi:10.3390/ijerph15122670

Moreno M, Tong-Rios C, Orjuela-Sanchez P, **Carrasco-Escobar G**, Campo B, Gamboa D, Winzeler EA, Vinetz JM. Continuous Supply of Plasmodium vivax Sporozoites from Colonized Anopheles darlingi in the Peruvian Amazon. ACS Infect Dis. Published online February 21, 2018. doi:10.1021/acsinfecdis.7b00195

Carrasco-Escobar G, Miranda-Alban J, Fernandez-Miñope C, Brouwer KC, Torres K, Calderon M, Gamboa D, Llanos-Cuentas A, Vinetz JM. High prevalence of very-low Plasmodium falciparum and Plasmodium vivax parasitaemia carriers in the Peruvian Amazon: insights into local and occupational mobility-related transmission. Malar J. 2017;16:415. doi:10.1186/s12936-017-2063-x

Carrasco-Escobar G, Gamboa D, Castro MC, Bangdiwala SI, Rodriguez H, Contreras-Mancilla J, Alava F, Speybroeck N, Lescano AG, Vinetz JM, Rosas-Aguirre A, Llanos-Cuentas A. Micro-epidemiology and spatial heterogeneity of P . vivax parasitaemia in riverine communities of the Peruvian Amazon: A multilevel analysis. Sci Rep. 2017;7(1):8082. doi:10.1038/s41598-017-07818-0

Serra-Casas E, Manrique P, Ding XC, **Carrasco-Escobar G**, Alava F, Gave A, Rodriguez H, Contreras-Mancilla J, Rosas-Aguirre A, Speybroeck N, González IJ, Rosanas-Urgell A, Gamboa D. Loop-mediated isothermal DNA amplification for asymptomatic malaria detection in challenging field settings: Technical performance and pilot implementation in the Peruvian Amazon. PLOS ONE. 2017;12(10):e0185742. doi:10.1371/journal.pone.0185742

Rosas-Aguirre A, Guzman-Guzman M, Gamboa D, Chuquiyauri R, Ramirez R, Manrique P, **Carrasco-Escobar G**, Puemape C, Llanos-Cuentas A, Vinetz JM. Micro-heterogeneity of malaria transmission in the Peruvian Amazon: a baseline assessment underlying a population-based cohort study. Malar J. 2017;16(1):312. doi:10.1186/s12936-017-1957-y

Rovira-Vallbona E, Contreras-Mancilla JJ, Ramirez R, Guzmán-Guzmán M, **Carrasco-Escobar G**, Llanos-Cuentas A, Vinetz JM, Gamboa D, Rosanas-Urgell A. Predominance of asymptomatic and sub-microscopic infections characterizes the Plasmodium gametocyte reservoir in the Peruvian Amazon. PLoS Negl Trop Dis. 2017;11(7):e0005674. doi:10.1371/journal.pntd.0005674

Ruiz-Grosso P, Miranda JJ, Gilman RH, Walker BB, **Carrasco-Escobar G**, Varela-Gaona M, Diez-Canseco F, Huicho L, Checkley W, Bernabe-Ortiz A, CRONICAS Cohort Study Group. Spatial distribution of individuals with symptoms of depression in a periurban area in Lima: an example from Peru. Ann Epidemiol. 2016;26(2):93-99.e2. doi:10.1016/j.annepidem.2015.11.002

Rosas-Aguirre A, Speybroeck N, Llanos-Cuentas A, Rosanas-Urgell A, **Carrasco-Escobar G**, Rodriguez H, Gamboa D, Contreras-Mancilla J, Alava F, Soares IS, Remarque E, D Alessandro U, Erhart A. Hotspots of Malaria Transmission in the Peruvian Amazon: Rapid Assessment through a Parasitological and Serological Survey. PloS One. 2015;10(9):e0137458. doi:10.1371/journal.pone.0137458

Rosas-Aguirre A, Ponce OJ, **Carrasco-Escobar G**, Speybroeck N, Contreras-Mancilla J, Gamboa D, Pozo E, Herrera S, Llanos-Cuentas A. Plasmodium vivax malaria at households: spatial clustering and risk factors in a low endemicity urban area of the northwestern Peruvian coast. Malar J. 2015;14:176. doi:10.1186/s12936-015-0670-y

Abstract of the dissertation

The role of human mobility and land cover change in the epidemiology of malaria in the Peruvian Amazon

by

Gabriel Carrasco Escobar

Doctor of Philosophy in Public Health (Global Health)

University of California San Diego, 2023

San Diego State University, 2023

Professor Tarik Benmarhnia, Chair

Recent estimates show significant human population expansion in the Amazon region that in turn implies dramatic changes in natural landscapes and also in human behaviors such as human mobility which may further expose these communities to malaria risk. The current path toward eliminating malaria is severely threatened by these environmental and human activity changes, yet there is scarce evidence of how changing land cover affects human mobility dynamics and ultimately malaria transmission, especially in the Amazon region.

This dissertation includes three studies examining the relationship of human mobility and land cover change on the epidemiology of malaria in the Peruvian Amazon. The first study assesses the connectivity structure and centrality between cities and villages as malaria transmission drivers in rural Amazonia using novel network analysis on granular passive case detection data. The second study determines the effect of out-of-village working activities on recent malaria exposure using two population-based studies and a g-computation approach to simulate multiple scenarios of mobility restrictions (by proportion of travelers, gender, and age) to quantify the impact of such restriction policies on malaria exposure reduction. The third study quantifies the effect of human population mobility on malaria risk using GPS data and fine-scale mobility metrics computed by a novel movement ecology non-parametric Bayesian framework.

The first study showed that localities with high connectivity consistently have higher malaria endemicity that was exacerbated in regions with the highest baseline malaria transmission rates. The second study presented the crucial significance of human mobility in supporting malaria transmission in the Peruvian Amazon. This study demonstrated the importance of targeting key subpopulations when creating occupational interventions by simulating the incidence of out-of-village employment activities to represent different policy scenarios. Targeting males and adults (18 years and older) groups has the greatest influence on malaria seropositivity. Finally, the third study showed that the high interaction between Amazon villages for reasons such as labor, commerce, or recreation may sustain such endemicity levels by increasing exposure to the malaria parasites and eventually increasing the importation risk.

The findings of this research can be used to inform control strategies and policies in the Amazon region to shift towards approaches that incorporate village connectivity structure and

human mobility patterns to prioritize connected areas instead of single villages and intensify malaria screening in sub-populations defined by their mobility profile.

Chapter 1: Introduction

1.1. Overview of Malaria

Malaria is the most important arthropod-borne disease in developing countries. According to the World Malaria Report by the World Health Organization (WHO) [1], there were 247 million malaria cases and an estimated 619,000 deaths in 2021. This vector-borne disease affects other impoverished populations in tropical and subtropical areas due to the suitable environmental conditions that favor the development of the vectors and spread of the parasites, in addition to the lack of access to quality healthcare [2] (***Figure 1.1***)

The transmission cycle involves protozoan parasites of the genus *Plasmodium spp.* transmitted to humans through the bite of mosquitoes of the genus *Anopheles spp.* There are six species within the genus *Plasmodium* that cause human malaria: *Plasmodium vivax*, *P. falciparum*, *P. malariae*, *P. ovale curtisi*, *P. ovale wallikeri*, and *P. knowlesi* [3]. There are differences among these pathogens, including their life cycle, symptoms, and treatment. From an epidemiological point of view, the main difference is mortality risk; untreated *P. falciparum* cases result in high mortality rates [3,4]. On the other hand, there are about 440 anopheline species, of which only approximately 70 are vectors of *Plasmodium spp.* and of these, about 41 are important vectors that transmit the disease to humans [5,6].

Malaria vectors are sensitive to the characteristics (i.e., temperature and precipitation) of the micro-environment and the physicochemical composition of the water bodies where they breed. These sensitivities determine the spatio-temporal distribution of anopheline species across landscape gradients such as in the Amazon region[7]. Landscape modification due to urban

expansion alters human-mosquito interactions, in particular, due to the adaptability of the dominant vector species in the area, *Nyssorhynchus (Anopheles) darlingi* [8–10]. Importantly, as the flight range of *Anopheles spp.* in this area is 400-500m [11,12], the most relevant source of parasite flow between regions that can jeopardize current malaria control efforts is Human Population Movement (HPM) [13–15].

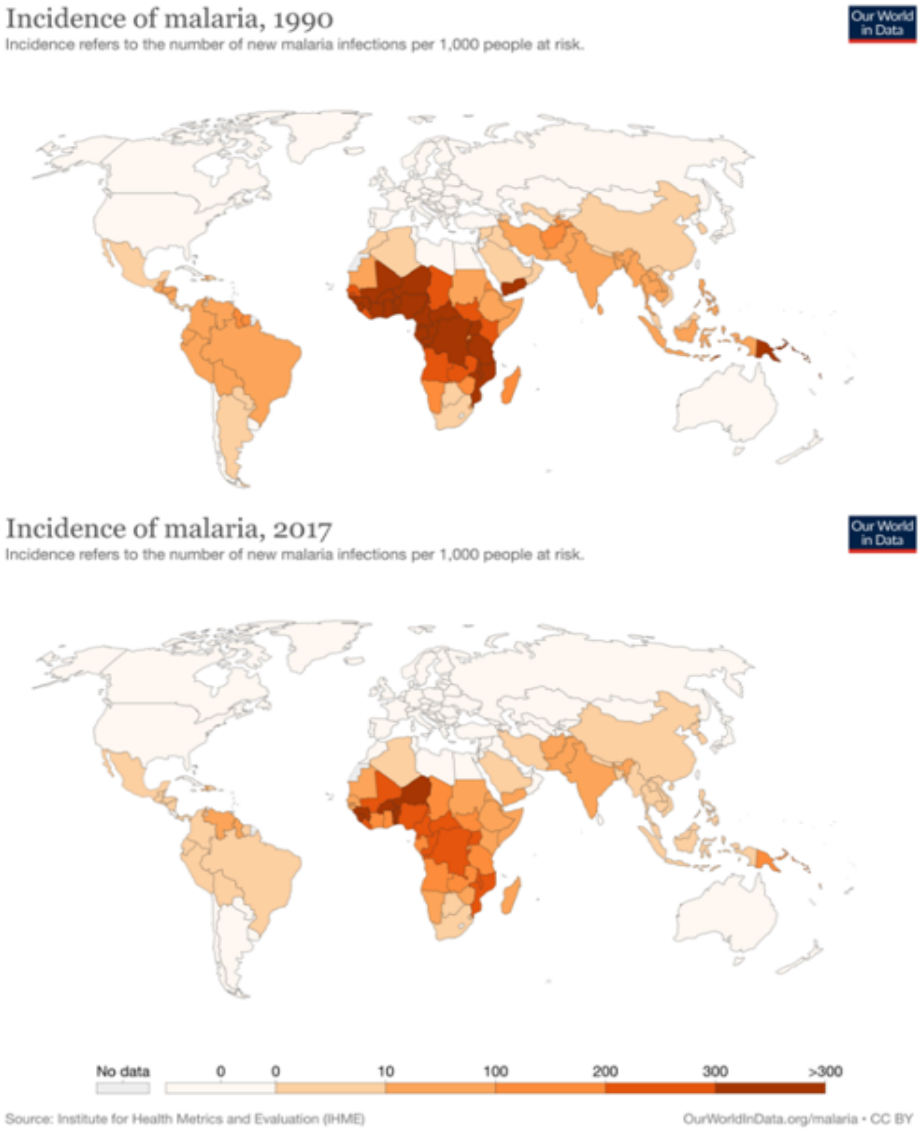


Figure 1.1. Worldwide distribution of malaria incidence in 1990 and 2017. Source: Our World in Data, 2020.

The following sections will cover: i) a general overview of malaria epidemiology in Latin America and Peru, ii) the population and urban expansion in these areas and frameworks that link landscape modification with malaria transmission, iii) the evidence of the role of HPM on the transmission of infectious diseases with particular emphasis on malaria and highlighting where the literature has or has not considered urban development as a driver for increased connectivity due to HPM, and finally, iv) the evidence for the relationship between occupational-related HPM and malaria infection.

1.2. Malaria in Latin America

1.2.1. Region of the Americas

From 2005 to 2014, malaria reduction efforts were successful in the WHO Region of the Americas. However, since 2015, this region experienced an increase in the total number of cases due to outbreaks in the Bolivarian Republic of Venezuela along with increased transmission in endemic areas of countries such as Brazil, Colombia, Guyana, Nicaragua, and Panama, as well as outbreaks in countries that are moving towards elimination (Costa Rica, the Dominican Republic, and Ecuador) [16,17]. Most of the cases in this part of the world are located in the Amazon region (***Figure 1.2***) and are mainly caused by *P. vivax.*

The primary vector in the Americas is *Ny. darlingi* which, in comparison with *An. gambiae* in the African continent, is more flexible in its selection of breeding sites and is more strongly anthropophilic (attracted to biting humans) and exophagic (preferring to bite outdoors) [9,18]. *Ny. darlingi* is highly susceptible to human plasmodium and capable of transmitting the

parasite inside and outside houses, even when its density is low [19]. Breeding sites for *Ny. darlingi* are characteristically represented by collections of clear, shallow water, shaded, with vegetation, and low salt concentration [8,18,20]. Since this species is both anthropophilic and opportunistic [8,10], as the natural environment becomes more altered or deforested, the local population of *Ny. darlingi* tends to cohabit with humans, invading their homes, and increasing its importance as a vector [21]. In the Amazon, it is the anopheline vector that most quickly and efficiently benefits from human-driven land-use change to the pristine environment [7,12,21,22]. The rural, riverine context that characterizes the Amazon rainforest provides suitable breeding sites for *Anopheles* mosquitos, allowing large heterogeneity in spatial patterns and less marked seasonal patterns of malaria transmission in comparison to SSA settings.

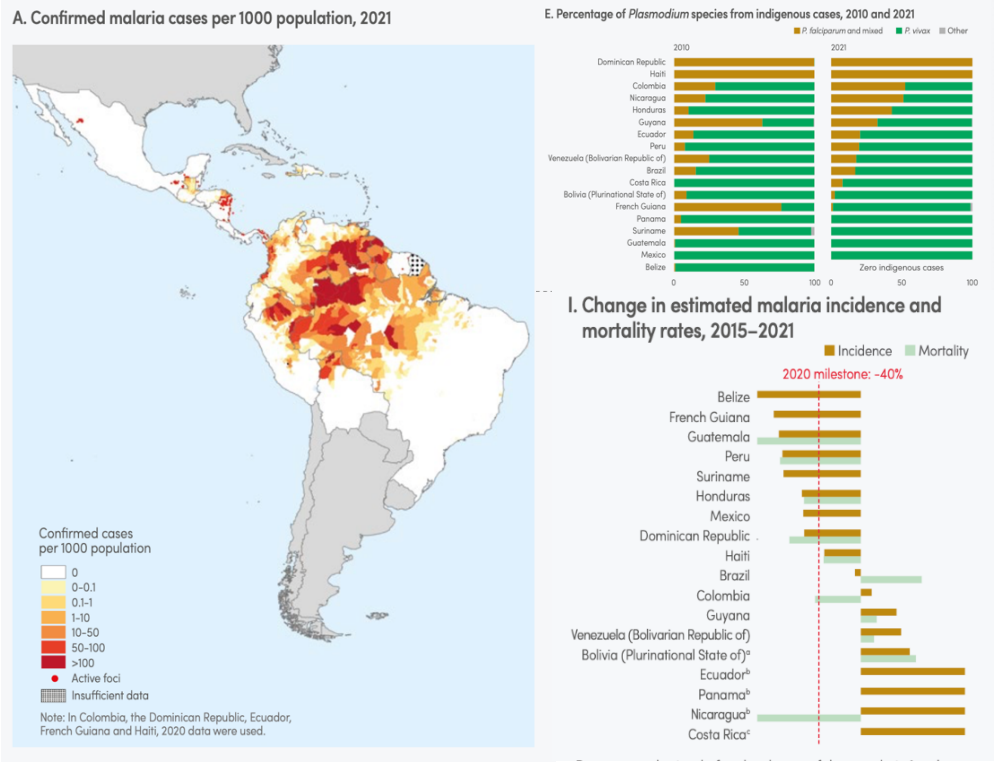


Figure 1.2. Malaria cases per 1000 population in 2021 in the Region of Americas. Percentage of Plasmodium species by country and change in estimated malaria incidence and mortality 2015-2021. Source: World Malaria Report, 2021.

1.2.2. Peru

The Peruvian Amazon is a region undergoing epidemiological changes as a result of complex land alterations/deforestation and climate change. Malaria epidemiology in the Peruvian Amazon is dominated by *P. vivax* (80%), with the remaining 20% of cases attributed to *P. falciparum* [23]. Currently, the province of Loreto accounts for around 90% of all malaria cases reported in Peru [24] and the main vector is *Ny. darlingi* [8,9].

The trends of malaria in Peru show that important reductions in the malaria burden can be obtained with intensive and comprehensive standard control measures, but also that this progress can quickly be lost if there is not a long-term country wide plan to sustain control activities and prevent malaria resurgence [23,25]. In this context, new challenges such as rapid urban expansion and changes in human mobility remain elusive in malaria control plans, which may deter current efforts. For instance, the Global Fund-sponsored "Plan de Malaria en Áreas de Frontera" (PAMAFRO) project, which ran between 2005 and 2010 in the Peruvian Amazon, supported the 1) strengthening of microscopic diagnosis, 2) implementation of active test-and-treat interventions by training community health workers (CHW), community-based larval source management (LSM), and 4) the distribution of long-lasting insecticidal mosquito nets (LLINs) with a community-based intercultural approach. During this period, malaria declined drastically in the most affected department of Loreto from 54,291 reported cases (25% due to *P. falciparum*) in 2005 to 10,504 cases (20% due to *P. falciparum*) in 2010 [23,26]. Regrettably, the marked reduction of international donors' support and the unusually heavy rains that caused floods in riverine villages starting in 2012 led to a five-fold increase in cases between 2010 and 2015 (60,302 cases) [23,24] (***Figure 1.3***).

Political and financial commitment from the Peruvian government was crucial not only to respond to this malaria resurgence, but also to establish a long-term initiative called Plan Malaria Cero started in 2017, with the ambitious aim of eliminating malaria in the country by 2036 [24]. Intensive malaria control interventions based on the PAMAFRO program in villages at high risk of malaria contributed to a substantial reduction in malaria incidence in Loreto, which reported 22,037 cases in 2019 [24]. As transmission decreases, malaria strategies need to be more focused and tailored to the local malaria epidemiology determined by the complex interactions between *Plasmodium* parasites, human behavior, and highly variable environments driving changes in vector behavior and habitat suitability [23,27,28].

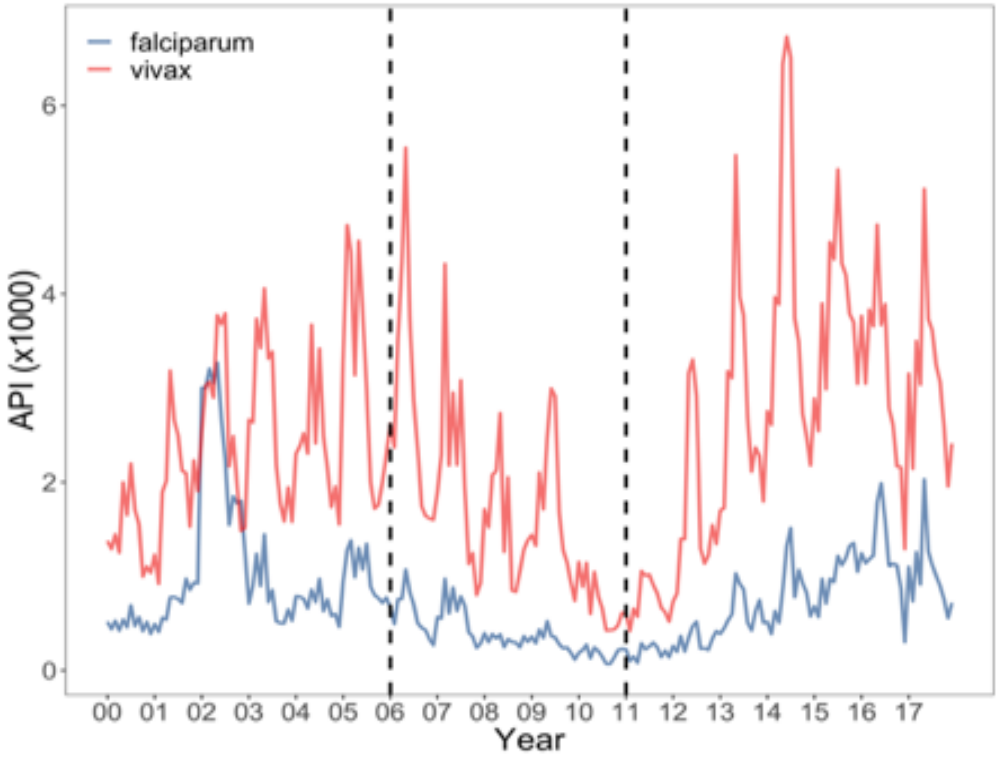


Figure 1.3. Time series of annual parasite index (API) of *P. falciparum* and *P. vivax* in the Loreto Region, Peru, 2000-2017. Source: Own elaboration, Ministry of Health data, 2018.

Rural riverine villages, characterized by poverty and limited access to healthcare facilities [29] (***Figure 1.4***), represent an important proportion of reported malaria cases in Loreto [30]; however, the transmission dynamics in these villages remain understudied [13,31]. Previous studies in the Peruvian Amazon have examined the association between malaria cases and climatic factors using the scarce and scattered network of meteorological stations in this area [32–34] or indices that describe the El Niño Southern Oscillation (ENSO) [35–37]. These studies concluded that higher temperatures, river levels and more abundant rainfall during El Niño conditions were associated with increased malaria risk in the Peruvian Amazon [32,34,35].

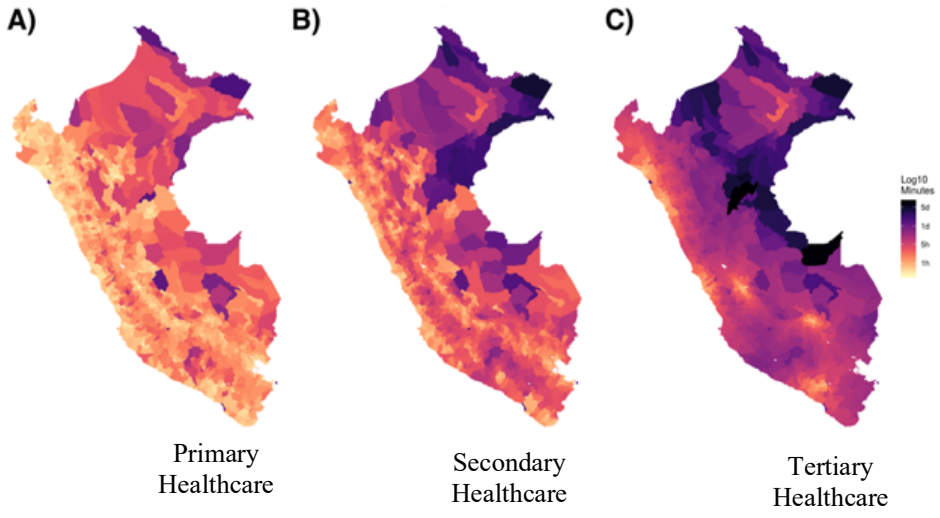


Figure 1.4. District-level travel time (in minutes) from villages to health facilities in Peru, by type of healthcare facility. Source: Carrasco-Escobar, 2020.

1.3. Population growth and land cover change

Population growth and land occupation follow parallel trends particularly in the early stages of urban development. A shift in human populations from rural to urban environments will change global patterns of disease and mortality [38,39]; for example, some studies suggest that urban African populations are healthier and have reduced malaria transmission rates compared

with their rural counterparts [40]. These improved health conditions in urban areas reflect enhanced access to preventative and curative services that might be related to wealth, education, and geographical location. Additionally, with increased human density in urban areas, malaria exposure per capita also decreases since most natural breeding sites remain constant or reduced due to land use and land cover (LULC) change [40]. However, while it remains challenging to determine whether increased urbanization results in decreased transmission, or if malaria reductions promote urban development, the results of multiple studies point to a close relationship between the two, irrespective of national wealth [40,41]. Markedly, prior literature on the subject has focused on the African continent, with few studies evaluating these interactions in Latin America (LA), where different dynamics of the parasite, vectors, and hosts are in place.

Human population size is hypothesized to influence the abundance and distribution of malaria-transmitting mosquitoes in rural settings [42,43], but also in the transition to urban environments. Population density also affects within-city dynamics (i.e., occupational activities, impoverished and vulnerable populations) that cause uneven exposure to mosquitoes. The combination of vector (*Ny. darlingi*) and parasite (*P. vivax*) characteristics and population dynamics in the neglected region of LA creates a situation that potentially jeopardizes global efforts of malaria control, though it also provides motivation to investigate such an understudied region.

In addition, anthropogenic environmental change is a significant driver of infectious disease dynamics in the Amazon rainforest [44]. Increased human population and LULC influence the biological community, including *Anopheles* mosquitoes, particularly those with some degree

of competence to transmit *Plasmodium* sp. that circulate in the Amazon region [45]. Importantly, changes in weather patterns also alter the ecological and spatial distributions of the malaria parasite, *Plasmodium spp.* and its vector, *Anopheles spp.* [46,47]. As a result, the interaction between LULC change, such as deforestation, and climate change is altering the landscape of infectious diseases in the Amazon. These factors are contributing to the re-emergence of malaria and other vector-borne diseases in some areas where it had previously been on the decline, threatening the success of malaria elimination programs [48,49].

Two types of LULC in the Amazon, forest degradation and land clearance, are associated with environmental and ecological changes that increase the breeding habitat of mosquito vectors such as *Ny. darlingi* [7]. In addition, movement of workers into deforested areas increases human exposure to the mosquito vector, elevating mosquito biting rates [50]. These conditions facilitate the spread of malaria vectors and disease transmission, particularly in the early stages of forest clearance, such as in Amazonian frontier settlements, where forest edge habitat is high [44]. This framework is called "frontier malaria" [51] and consists of initial high rates of malaria in newly deforested areas, followed by a decline due to improvement in public services, such as healthcare infrastructure. In the final stage of frontier malaria, disease transmission is reduced and becomes stable [52] (***Figure 1.5***). The complex dynamics of malaria transmission in deforested areas is a challenge for control programs in the Amazon region.

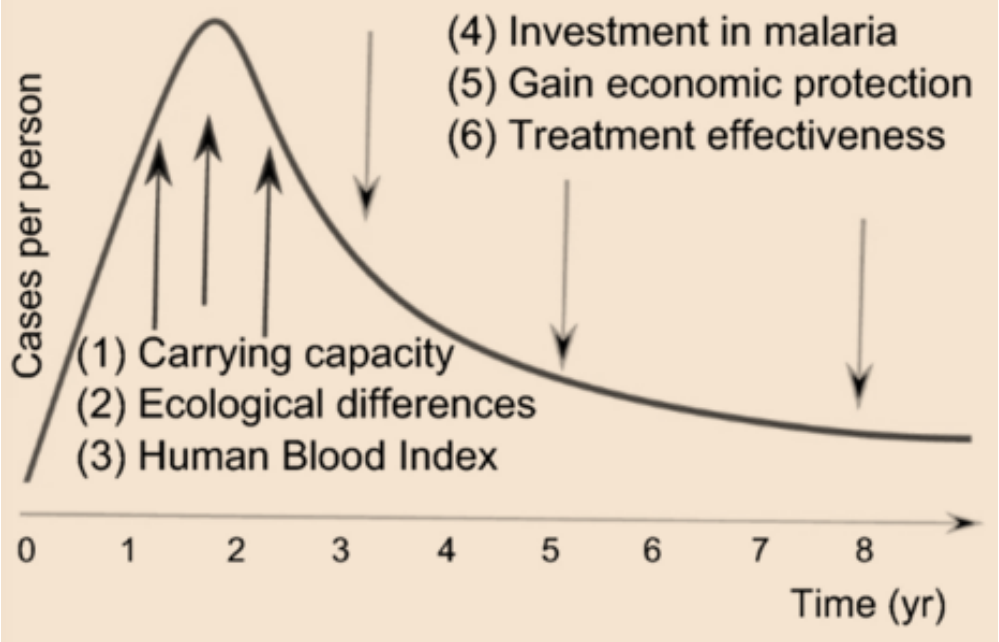


Figure 1.5. Conceptual framework of 'Frontier Malaria'. Source: de Oliveira Padilha, 2019.

Malaria is not the only infectious disease affected by the modification of the environment. The conversion of natural habitats to agricultural or urban ecosystems is widely recognized to influence the risk and emergence of zoonotic disease in humans [53,54]. Gibb et al. [55] analyzed large-scale dataset of ecological assemblages and host species and showed that LULC has global and systematic effects on local zoonotic host communities. Notably, LULC intensity has positive effects on community zoonotic potential both within and between LULC types, with the largest increases in zoonotic potential seen for sites recovering from past disturbances ("substantial-use secondary"), managed sites (i.e., cropland, pasture, or plantation) (18–21% host proportion richness, 21–26% proportion abundance), and urban sites (62–72% proportion richness, 136–144% proportion abundance). Their results suggest that global changes in the mode and the intensity of LULC are expanding and creating new hazardous interfaces between people, livestock, and wildlife reservoirs of zoonotic disease.

Remote sensing has been widely used to help identify the environmental determinants of malaria transmission and how these determinants have changed spatio-temporally concurrent with the LULC trajectories in the Amazon region [56]. These development trajectories in the Amazon rainforest fit within the frontier malaria framework as they involve tight interactions between environmental and contextual variables pertaining to malaria risk [48,51,57–59]. A recent study tested the hypothesis that deforestation, as a marker of LULC, and malaria incidence influence each other in bidirectional causal relationships governed by socio-ecological mechanisms. They found that deforestation leads to an increase in malaria incidence and the increase in malaria burden reduces forest clearing events [44]. In these settings, environmental changes can trigger a change in vector species composition. Some malaria vector species thrive in the new landscapes resulting from changes in deforestation rates, land use, behavioral patterns, or economic activities, whereas other species are replaced [60].

1.4. Networks and connectivity

As villages and cities grow, transit to and from these locations increases. From a malaria control perspective, this flow, also referred to as connectivity, may influence the endemicity level in the system (group of villages/cities). In other words, if there is no population interaction between two places, their malaria endemicity and risk remains independent. In contrast, if there is an intense population flow between two places, their connectivity influences, to some degree, the malaria endemicity and risk in both locations. Most of the literature on this subject uses the disease elimination framework [61], that is, once control strategies are successful and a country achieves the malaria elimination phase (< 1 case / 1000 population at risk) [62], particular attention is paid to the prevention of malaria re-introduction (***Figure 1.6***).

Key concepts to sustain malaria elimination were first defined in the previous Global Malaria Eradication Programme (GMEP, 1955–1969), directed by the World Health Organization. These concepts include: 1) vulnerability (the rate of malaria importation), 2) receptivity (the potential for ongoing local transmission), and 3) malariogenic potential (the expected number of cases that could occur as a product of vulnerability and receptivity). These concepts were revised in 2018 by the WHO Evidence Review Group on the assessment of malariogenic potential to inform elimination strategies and plans to prevent re-establishment of malaria [63], see ***text box 1*** for updated terminology.

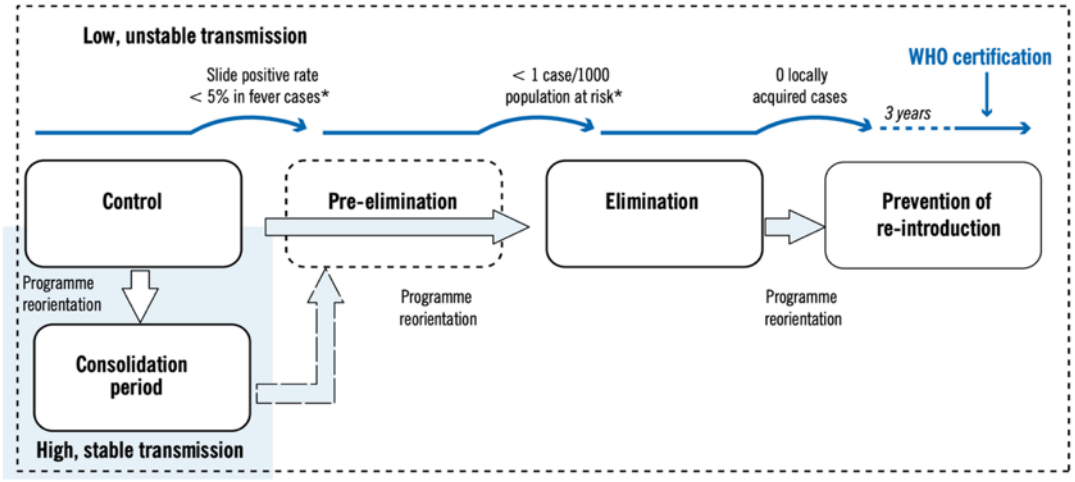


Figure 1.6. Programme phases from malaria control to elimination. Source: Mendis, 2009.

The new version of the WHO guideline "A framework for malaria elimination" [64] expanded these definitions to subnational levels. However, in some contexts, this framework could be further expanded to micro-scale levels. In rural Amazonia, where most commuting occurs by river, the connectivity between villages is shaped by the river network or watershed. Villages with varying levels of malaria endemicity could be located in the same watershed, potentially allowing the circulation of the parasite between eliminated and endemic regions [65].

Text Box 1. Updated terminology by the Malaria Policy Advisory Committee.

Malariogenic potential: Likelihood of local transmission that is the product of receptivity, risk of importation of malaria parasites and infectivity of imported parasites. Note: The concept of malariogenic potential is most relevant for elimination and prevention of reestablishment when indigenous transmission is mostly or entirely eliminated.

Receptivity: Degree to which an ecosystem in a given area at a given time allows for the transmission of Plasmodium spp. from a human through a vector mosquito to another human. Note: This concept reflects vectorial capacity, susceptibility of the human population to malaria infection, and the strength of the health system, including malaria interventions. Receptivity can be influenced by ecological and climatic factors.

Vulnerability: Likelihood of malaria infection based on living conditions or behavioral risk factors, or likelihood of increased risk of severe morbidity and mortality from malaria infection.

Importation risk: Risk or potential influx of parasites via infected individuals or infected Anopheles spp. mosquitoes. Note: "Infected individuals" includes residents infected while visiting endemic areas as well as infected immigrants.

Infectivity: Ability of a given Plasmodium strain to establish an infection in an Anopheles mosquito species and undergo development until the mosquito has sporozoites in its salivary glands.

1.4.1. Parasite flow and malariogenic potential

For a given area, the malariogenic potential is the likelihood that an imported infection establishes onward local malaria transmission due to characteristics of the host, the parasite, the vector, and the ecosystem (see ***text box 1*** for further details). In elimination settings, the connectivity between areas of contrasting endemicity, due to HPM, is considered key to prevent the re-establishment of transmission [61,65,66].

Multiple studies have analyzed the role of HPM in the transportation of parasites and the generation of local subsequent outbreaks [61,65–69]. Importantly, most of these studies identified

multiple types of HPM (i.e., labor, recreational, forced migration, etc.); however, none of these studies quantified of the role of urbanization (or development) of the cities in attracting or generating these types of HPM.

An important, and related, task is the identification of the "source areas" where humans may become infected and subsequently, through travel, introduce the parasite into new "sink areas" [69,70]. Recent access to mobile phone data has improved the estimation of these source and sink areas in Africa [70–73]; however, regions such as the Amazon rainforest lack the mobile phone infrastructure to conduct this type of analysis [74].

1.4.2. Connectivity and centrality applied to infectious diseases

Networks can be used to represent any type of relationship between entities [75,76]. Network theory has been used to represent relationships in scientific fields such as physics, informatics, chemistry, and biology, and more recently in social sciences [75,77]. Networks could represent relations of cells, organisms, information, ideas, or any other study subject. Importantly, network analysis is a branch of applied mathematics derived from graph theory, and as such, multiple properties may be derived from the graph representation [75,78]. These properties could be measured at the entity (herein referred to as nodes or vertices), the relationships (herein referred to as edges or links), or at the overall level. In the study of social science or ecology, these networks are in constant flux due to of the addition or removal of nodes and edges to the system [77]. This property, called "scale-free" (represented in ***Figure 1.7a***), is particularly important in the study of infectious diseases due to their dynamic behavior.

Nodes can be defined as populations (in cities or villages) and the edges as the flow of parasites or susceptible hosts due to HPM. As the size of the nodes increases (due to population growth and/or urbanization in cities/villages), new edges could be created (as people start traveling between villages/cities) and ultimately shape malaria distribution in the system. As more edges are added to the network, the relative importance of the nodes in a network could be measured by centrality properties (***Figure 1.7b***). The most popular measure is the "betweenness centrality", that is computed in a weighted network as follows:

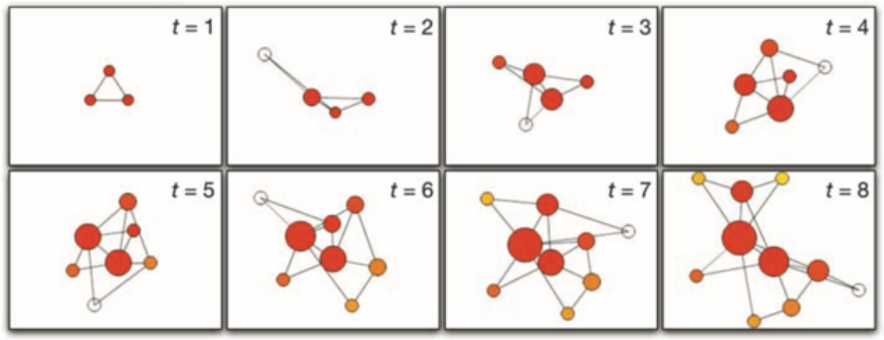
$$C_B(v) = \sum_{s \neq v \neq t} \frac{\delta_{st}(v)}{\delta_{st}}$$

where δ_{st} is the total number of shortest paths from node *s* to node *t* and $\delta_{st}(v)$ is the number of those paths that pass through node v [79,80]. C_B ranges from 0 to 1 and high values indicate the relative high importance of a specific node in a network. Betweenness centrality differs from other centrality measures such as degree, eigenvector, and closeness centrality by considering nodes who play a "bridge spanning" role inside a network.

Network analysis and connectivity have been previously used in the malaria literature. Buckee et al. analyzed the malaria parasite population structure from serological networks [81]. Tatem et al. estimated the role of international population movements on *P. falciparum* malaria elimination strategies [82]. Pindolia et al. further analyzed regional connectivity and the mobility of different demographic groups in in East Africa and showed that demographically-stratified HPM and malaria movement estimates using network analysis can provide quantitative evidence to inform the design of more efficient malaria interventions [83]. Huang et al. expanded this analysis to understand global malaria connectivity through air travel and showed that both

malaria-free areas and other endemic regions are strongly connected, particularly in Africa and Southeast Asia [84].

a)



b)

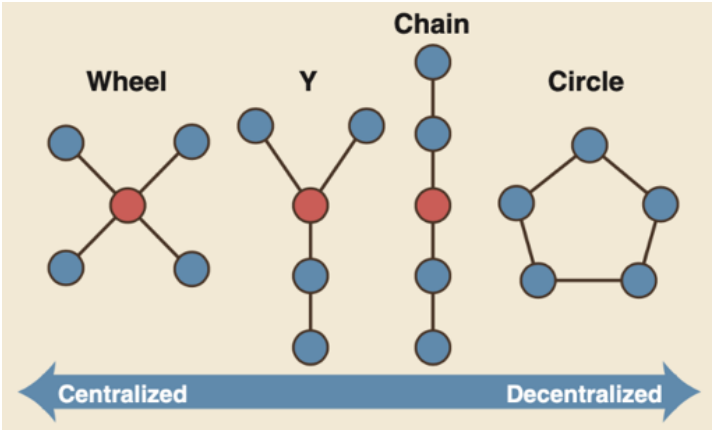


Figure 1.7. Networks representations. A) Scale-free model representation in time (t =1-8). Source: Barabási, 2009 [93]. B) Four network structures in relation to the centrality degree. The most central node in each network is colored red. Source: Borgatti, 2009.

In the mobility and migration literature, the strength of the links between nodes is commonly computed using 'gravity models' [85]. These models are used to understand the connectivity between areas as a consequence of HPM and as a function of the population size in the origin and destination locations. Analogous to Newton's law of gravity, a gravity model assumes that the number of individuals F_{ij} that move between locations i and j per unit time is

proportional to some power of the population of the source (m_i) and destination (n_j) locations, and decays with the distance r_{ij} between them:

$$F_{ij} = \frac{m_i^{\alpha} n_j^{\beta}}{f(r_{ij})}$$

where α and β are adjustable exponents and the deterrence function $f(r_{ij})$ is chosen to fit the empirical data [85]. This model and some variations were recently used to analyze the connectivity between areas due to HPM in malaria-endemic countries [61,70,86,87]. It is important to mention that network analyses (i.e., connectivity, centrality, gravity, etc.) are models where HPM is used to represent a link between two nodes (villages/cities), and taken together with urban expansion, are used in this study to understand the dynamics of the cities and villages in developing settings.

The main mechanism establishing malaria transmission from external sources is the importation of parasites from HPM. As cities grow, HPM intensifies, and in consequence, the probability of malaria importation increases. The magnitude of the connectivity in these areas is conditioned by the size of the cities as a proxy for the number of services and commerce in place.

1.5. Human Mobility

Human travel (referred to as HPM) impacts the spread of pathogens (i.e., viruses, bacteria, parasites, etc.), increases exposure to high-risk or hazardous environments (i.e., forest fringe, air pollution, UV radiation) [88–92], and affects urban planning decisions [89,90]. Prothero et al. [88] highlighted in 1977 that "Not only must mobility be identified as a factor in disease

transmission, because disease itself is a factor responsible for movement: areas may become less suitable for occupation if not totally uninhabitable, thereby forcing the movement and redistribution of population". Nowadays, the availability of new streams of data due to ubiquitous technologies (such as mobile phones, the Internet of Things, and Smart Cities Networks) has revolutionized our understanding of HPM patterns [68,70,72,73]. However, the lack of necessary infrastructure in rural areas such as the Amazon rainforest prevents the use of such technologies and ultimately limits our understanding of the role of HPM on the transmission of diseases in this region [74,93].

1.5.1. Impact of human movement on infectious diseases

People can acquire a parasitic infection during transit between regions with different endemicity levels [65,66]. While travelling, individuals may be exposed to environments favorable to mosquito abundance, biting behavior, and ultimately malaria risk [13,31,94,95]. Multiple factors influence particular HPM patterns. Prothero et al. [88] pioneered a characterization of HPM in circulatory and migratory movements, and in relation to a rural-urban gradient. Stoddard et al. [96] described a framework for characterizing HPM relevant for vector-borne diseases across spatial and temporal scales. Recently, Pindolia et al. [61]expanded these typologies in relation to the distance and frequency of travel in the context of malaria elimination (see ***Figure 1.8***).

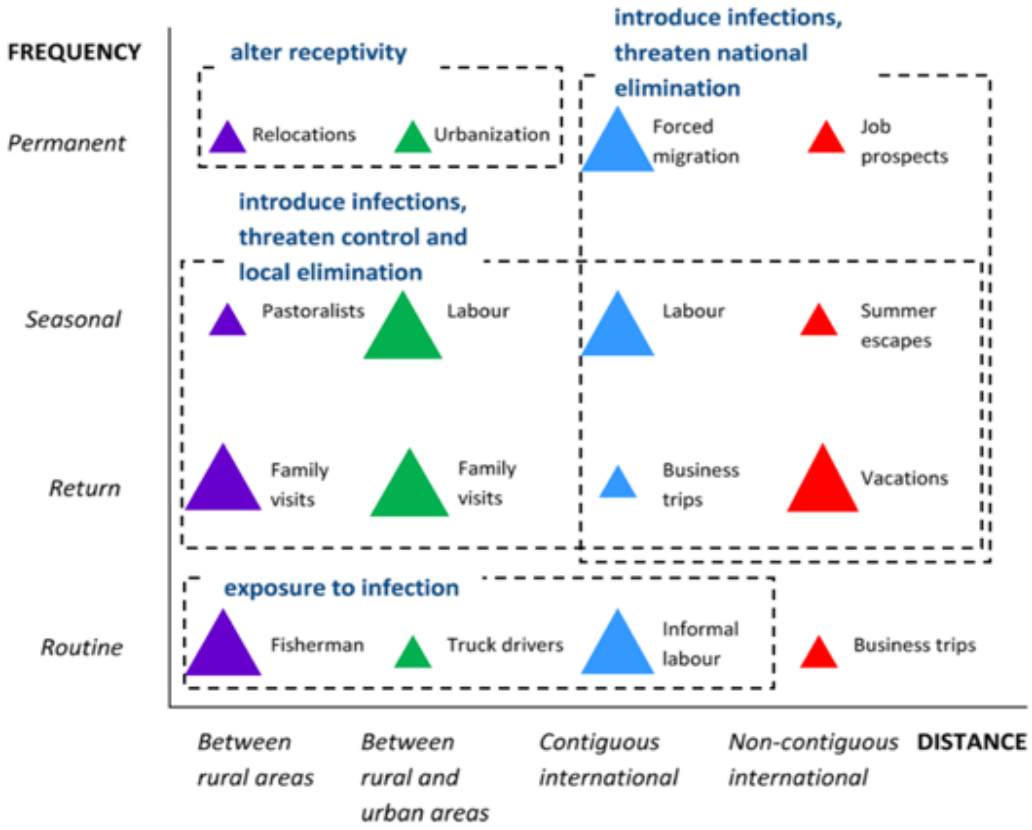


Figure 1.8. Examples of human population movement (HPM) types relevant for malaria control and elimination, stratified by distance and frequency. Source: Pindolia, 2012.

HPM results in increased contact between humans and mosquitos. The extent of this exposure is dependent on human behavior and the type of movement [14]. For instance, sleeping outdoors without bed nets, working certain hours of the day, not using protective clothes, or not taking necessary hygienic measures are all factors that can modulate the spread of malaria while traveling. These tangled interactions have been explored recently by Fornace et al. [94] by using detailed GPS-tracking data to understand the impact of HPM on zoonotic malaria in Malaysia and highlighting the importance of intensified interactions between pathogens, insect vectors, and people around habitat edges.

Previous studies have quantified the role of HPM on malaria transmission in the Amazon region [13,15,97–100]. These studies found that the relationship between HPM and malaria transmission is highly context-dependent, with contrasting malaria exposure levels associated with different HPM patterns. In these studies, consistent risk factors for malaria exposure included travel related to certain occupations, travelling for multiple days at a time, and using multiple means of transportation in a single trip. Most of these studies relied on the use of travel questionaries or participatory mapping [93,98], however, this method has important limitations (i.e., recall bias, lack of accuracy defining transit areas, collecting details on destinations but not on trajectories) in comparison to other methods for measuring HPM [91,96,101], such as telemetry, Radio Frequency Identification (RFID), GPS tracking, and mobile phone tracking (based on cellphone-towers registries). GPS tracking emerged as a suitable approach to measure HPM in the absence of phone infrastructure in regions such as the Amazon. In particular, previous studies on malaria [74] and dengue [102] in the Loreto Region of Peru, showed the feasibility of this technology in rural and urban settings, respectively. Both studies showed the high mobility and connectivity of this population but also how patterns of HPM differed based on factors such as occupational activities and demographic characteristics, with the authors hypothesizing on the potential impact these different HPM patterns might have on the transmission of infectious diseases in the area.

1.5.2. Occupational-related mobility

One of the most important drivers of HPM in malaria-endemic areas is occupational activities like fishing, logging, agriculture, or commerce [83,94,95,98,103]. Previous studies in the Loreto Region have showed consistent evidence suggesting the important role that HPM plays in

the malaria epidemiology of *P. vivax* and *P. falciparum*. Chuquiyauri et al. [104] reported higher odds of several clinical malaria episodes in logging and agricultural workers, and Chuquiyauri et al. [105] showed higher odds of *P. vivax* reinfections in participants with jobs that require travel out of the village. Furthermore, Carrasco-Escobar et al. [13] showed in a longitudinal analysis that loggers, fishers, and farmers had a higher risk of *P. vivax* malaria in riverine communities than occupations carried out inside the villages. Rosas-Aguirre et al. [106] showed that individuals with outdoor occupations (e.g., farmers, guards, loggers, fishers) had higher odds for *P. falciparum* malaria. Finally, Parker et al. [31] reported high measures of vectorial capacity and indices of malaria transmission risk based on direct measurement of anopheline biting rates and sporozoite (dormant stage) rates for several undeveloped, riverine campsites in the Peruvian Amazon, frequented by occupation-related travelers.

1.6. Dissertation overview and specific aims

Despite reductions in malaria incidence and mortality worldwide, malaria remains one of the most important infectious diseases globally in terms of cases and deaths. Malaria transmission occurs mainly in rural areas; however, the expansion of the urban frontier favors complex interactions between hosts, vectors, and parasites in new transient environments. In 2007, the United Nations (UN) estimated that, for the first time, more people in the world lived in urban than in rural areas. Current projections highlight the rapid population growth in urban settings that will in turn involve dramatic changes in natural landscapes and also in human behaviors such as HPM. These increasing changes in the environment and human activities pose a critical challenge to the current malaria elimination track; however, evidence regarding the role

of land cover change on the dynamics of HPM and ultimately on malaria transmission remains scarce, particularly in the Amazon region.

In the Amazon region, bigger cities intensify the HPM between neighboring areas. Increased connectivity, generated due to HPM, also influences the malaria endemicity in a network of cities/villages. HPM provides additional opportunities for environmental exposure to occur in transit between areas, and represents a key, yet neglected factor for tailoring current malaria control strategies, such as household spraying and bed net distribution, which commonly assume stable (not mobile) populations. Using detailed spatio-temporal data at the village level in Peru, this study proposes to unveil the effect of connectivity on malaria transmission in rural settings using network analysis under contrasting transmission levels.

Importantly, different patterns of HPM result in contrasting malaria exposure levels. In rural areas, the highest risk of malaria infections is observed among inhabitants with occupational-related HPM. Logging, fishing, commerce, or agriculture are hypothesized to expose inhabitants to higher mosquito biting rates; however, no etiological assessment has been conducted to understand this relationship. In this study, I propose the use of causal inference methods and detailed individual-level travel data to understand this etiological relationship and simulate potential interventions to optimize the reduction in transmission risk.

Ultimately, studies that examine fine-scale mobility patterns and their role in importing malaria parasites in endemic Amazonian settings remain scarce. GPS data is increasingly being applied to a variety of epidemiological studies to quantitatively characterize mobility patterns at

a fine spatio-temporal resolution. Such GPS-based approaches include time-weighted spatial averaging (TWSA) (e.g., utilization distribution, kernel density, density ranking) and models based on activity spaces (e.g., daily path area, minimum convex polygon, standard deviation ellipse) [107,108]. In this study, I propose quantifying the effect of human population mobility on malaria importation risk by using a densely sampled population cohort tracked with GPS data loggers.

In this dissertation, a multilevel and multiple-method approach is proposed to address the role of urban expansion and HPM in the epidemiology of malaria in the Peruvian Amazon (***Table 1.1***). The following specific aims will be addressed:

Aim 1: Assess connectivity structure and centrality between cities and villages as malaria transmission drivers in rural Amazonia.

Aim 2: Determine the effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon.

Aim 3: Quantify the effect of human population mobility on malaria risk using GPS data and fine-scale mobility metrics.

Evidence generated in this dissertation is aimed to help decision-makers in the Amazon region and other malaria-endemic settings to better prepare for and respond to future changes in malaria risk. By analyzing the effect of environmental modification and HPM on malaria transmission in the Amazon region, the evidence generated in this study will support the adoption of 1) malaria control strategies that account for environmental modification and human

mobility, and 2) urban planning and development policies in rural areas that consider infectious disease dynamics.

Table 1.1. Summary of study area, inference level, temporal scale, and data sources of malaria and covariates for study aims.

Study aims	Study area	Inference level	Temporal scale	Malaria data source	Covariate data sources
Aim 1	Loreto region (Peruvian Amazon)	Village/cities (n = 1608)	Monthly (2009–2018)	Passive case detection (PCD) reports from the Peruvian Ministry of Health	Hansen collection derived deforestation metrics, National watershed network
Aim 2	Rural villages (Peruvian Amazon)	Individuals (n=1790)	Cross-sectional (2018)	Amazonian International Center of Excellence in Malaria Research (ICEMR) data	Sociodemographic and travel surveys
Aim 3	Gamitanacocha village (malaria hotspot)	Trips (GPS logs) (n = 30 individuals)	Weekly (8 weeks, 2018)	Proof-of-concept study, GORGAS study data	Sociodemographic and travel surveys

Chapter 2: The role of connectivity on malaria dynamics across areas with contrasting control coverage in the Peruvian Amazon

2.1. Abstract

Network analysis may improve the understanding of malaria epidemiology in rural areas of the Amazon region by explicitly representing the relationships between villages as a proxy for human population mobility. This study tested a comprehensive set of connectivity metrics and their relationship with malaria incidence across villages with contrasting PAMAFRO (a malaria control initiative) coverage levels in the Loreto department of Peru using data from the passive case detection reports from the Peruvian Ministry of Health between 2011 and 2018 at the village level. A total of 24 centrality metrics were computed and tested on 1,608 nodes (i.e., villages/cities). Based on its consistency and stability, the betweenness centrality type outperformed other metrics. No appreciable differences in the distributions of malaria incidence were found when using different weights, including population, deforested area, Euclidian distance, or travel time. Overall, villages in the top quintile of centrality had a higher malaria incidence in comparison with villages in the bottom quintile of centrality (Mean Difference in cases per 1000 population; *P. vivax* = 165.78 and *P. falciparum* = 76.14). The mean difference between villages at the top and bottom centrality quintiles increased as PAMAFRO coverage increased for both *P. vivax* (Tier 1 = 155.36; Tier 2 = 176.22; Tier 3 = 326.08) and *P. falciparum* (Tier 1 = 48.11; Tier 2 = 95.16; Tier 3 = 139.07). The findings of this study support the shift in current malaria control strategies from targeting specific locations based on malaria metrics to strategies based on connectivity neighborhoods that include influential connected villages.

2.2. Introduction

The Peruvian Amazon is experiencing epidemiological changes in malaria transmission as a result of landscape modifications, climatic factors, malaria control interventions, and other anthropogenic drivers. Regionally, malaria epidemiology is dominated by *P. vivax* (80%), with the remaining 20% of cases attributed to *P. falciparum* [23]. Currently, the Loreto department accounts for an estimated 90% of all malaria cases reported in Peru [24]. In this area, between 2006 and 2010, an intense malaria control program, PAMAFRO, (Project for Malaria Control in Andean Border Areas) was undertaken, supported by the Global Fund [26]. This program was successful and effective, resulting in a sharp reduction in malaria, with cases reaching their lowest number (22,909) in 2011[109]. However, since 2011, this trend has been reversed, with a peak of 61,108 malaria cases reported in 2014 [24]. Re-emergence factors such as asymptomatic reservoirs [110,111], meteorological conditions [112], and changes in the mosquito population [9,31,113] have been previously studied in this area. However, evidence of other factors influencing this uptick in malaria cases, such as human population mobility (HPM) and, as a result, the connectivity between villages with contrasting malaria transmission, remains scarce.

The transit and return of people from locations with contrasting endemicity levels must be addressed to achieve malaria elimination [28]. This flow, also referred to as connectivity, influences the malaria endemicity level in the system (group of villages/cities) and jeopardizes control interventions that focus on targeted villages as isolated from (not connected to) other locations. This human population flow between two areas influences, to some degree, the malaria endemicity and risk in both locations (origin and destination). Under the World Health Organization (WHO) malaria elimination framework, human mobility and connectivity are key

parts of the malariogenic potential, defined as the likelihood that an imported infection establishes local malaria transmission due to characteristics of the host, the parasite, the vector, and the ecosystem [63].

The relationships between two or more entities (i.e., villages) are often analyzed as networks [75,76,78]. Different properties can help capture the level of connectivity between such entities and can be measured at the entities (nodes), the links, or at the overall level. This framework has recently been used in the context of infectious diseases. Buckee et al. analyzed the malaria parasite population structure from serological networks [81]. Tatem et al. estimated the role of international population movements on *P. falciparum* malaria elimination strategies [82]. Pindolia et al. further analyzed regional connectivity and the mobility of different demographic groups in in East Africa and showed that demographically-stratified HPM and malaria movement estimates using network analysis can provide quantitative evidence to inform the design of more efficient malaria interventions [83]. Finally, Huang et al. expanded this analysis to understand the global malaria connectivity through air travel and showed that both malaria-free areas and other endemic regions are strongly connected, particularly in Africa and Southeast Asia [84].

However, no agreement has yet been reached on which network property best captures how HPM affects malaria epidemiology, particularly in areas such as Latin America, which is the region with the most rapid urban growth rate in the world [114,115]. Furthermore, the current projections of population growth in the Amazon region involve dramatic changes in natural landscapes but also in human behaviors such as HPM. This study therefore proposes incorporating LULC changes to reflect the expansion of villages and cities nested in watersheds

that reflect microcircuits of mobility. Taken together, refined metrics of connectivity between villages have the potential to better inform malaria control efforts. In this study, we use data from the passive case detection (PCD) reports from the Peruvian Ministry of Health (MoH) between 2011 and 2018 at the village level to test a comprehensive set of connectivity metrics, including population and environmental (deforestation) weights, and their relationship with malaria incidence across villages with contrasting baseline malaria transmission and PAMAFRO coverage levels in the Loreto department of Peru.

2.3. Methods

2.3.1. Study design

This is an observational ecological study that tests the relationship between connectivity metrics and malaria incidence in the Loreto department of Peru. Connectivity metrics were derived from the combination of multiple centrality measures (i.e., betweenness, strength, eigen, and closeness) and weights such as masses (i.e., population and deforested area) and costs (i.e., distance and travel time). The relative importance of the nodes has been analyzed as a driver for malaria incidence in the area using ten-year (2011–2018) records of the PCD data from the MoH at village/city (node) level. This relationship was further stratified across villages with contrasting baseline malaria transmission and PAMAFRO coverage levels.

2.3.2. Study area

The Loreto department, located in the northeast of Peru, covers 28.7% of the national territory and a total population of 883,510. The political-administrative organization of Loreto is divided into 8 provinces, 53 districts and 31 watersheds (Figure 2.1). Iquitos is the capital city and the most densely populated with 510,000 inhabitants (the 7th most populated city in Peru).

Most inhabitants (69.6%) live in urban areas, 32.2% live in poverty, and 7% live in extreme poverty. Only 39.6% of households have access to basic public services (water, sanitation, electricity, and telephone) [116]. The most common economic activities are agriculture, fishing, and mining [116]. In 2016, Loreto had a total of 521 health care facilities, with a ratio of 1,086 inhabitants per healthcare personnel [116] and important transportation and monetary barriers [29,117] to quality healthcare access. The tropical climate in this area ranges on average from 17°C (between June and July) to 36°C (between December and March) with a rainy season between December and March.

2.3.3. Data sources

2.3.3.1. Malaria passive case detection data

Malaria is a notifiable disease by the Peruvian MoH and the registry of individual-level data started in 2009 [23]. These data are available in both electronic and hardcopy format for the dominant malaria species *P. vivax* and *P. falciparum*. In Peru, malaria diagnosis relies primarily on microscopic inspection of thick and thin blood smears in health facilities. The presence of asexual and sexual stages of *Plasmodium* species is determined after examining 100 high-powered fields [118]. All positive cases are immediately treated according to national guidelines from MoH [119]: chloroquine (CQ) for 3 days and primaquine (PQ) for 7 days in confirmed *P. vivax* malaria infections, and mefloquine (MQ) for 2 days and artesunate (AS) for 3 days in confirmed *P. falciparum* infections. For this study, georeferenced malaria diagnosis data were obtained at the village/city level from 2011 to 2018 for each month and were collapsed for the entire study period. GPS coordinates of the centroids and population size are provided for each

village by the MoH. Malaria endemicity level was computed as the mean 2011–2018 Annual Parasite Index (API), defined as the total number of new cases per 1,000 individuals.

2.3.3.2. Deforestation and watershed data

The Hansen collection [120], a high-spatial resolution (1 arc-second, approximately 30 meters) dataset of yearly forest coverage loss, was used to extract village-level (within 5 km radius from the centroids of villages/cities) mean deforested area (2011–2018; Km^2). The Hansen collection defines forest loss as a stand-replacement disturbance, or a change from a forest to non-forest state, using the year 2000 as reference and bands 3, 4, 5, and 7 of Landsat 7 cloud-free image composites. Hansen collection data were gathered and processed in Google Earth Engine [121], a cloud-based platform for planetary-scale geospatial analysis (https://earthengine.google.com). The watershed boundaries were obtained from the Peruvian National Authority of Water (ANA by its Spanish acronym). ANA provides a division, codification, and systematization of watersheds using two international standard methodologies, the Pfafstetter coding system [122] and a Digital Elevation Model (DEM) such as NASA's SRTM of 30 meters spatial resolution. The final product is a map to a scale of 1:100,000 cm.

2.3.3.3. Distance and travel time estimation

The computation of the distance from each village to all villages analyzed in the entire department of Loreto was performed by calculating the Euclidean distance using the R Statistical Software (v4.2.2; R Core Team 2021). This Euclidean distance is defined as the shortest straight line that exists between two points without considering the type of existing surface. For its calculation the following formula was used:

$$d_{(x_{(i,j)}, y_{(i,j)})} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

Where (x_i, y_i) are the coordinates of the origin and (x_j, y_j) are the coordinates of the destination.

The estimation of travel time was conducted in R Statistical Software (v4.2.2; R Core Team 2021) using the rgee package [123] that bridges R to the Google Earth Engine (GEE) API [121]. We followed travel time estimation procedures described in previous literature [29,124]. To summarize the method, information about land coverage, road infrastructure, and river network was used to create a 30 m resolution grid surface. The speed assigned for each category of land cover was obtained from elsewhere [124] and the Ministry of Transportation provided the speed for the road infrastructure. A friction surface was constructed where each pixel contained the cost (time) to move through the area encompassed in the pixel. Then, a cumulative cost function was applied (least-cost-path algorithm) that examined all potential paths iteratively, and the time-weighted cost was then minimized to calculate the minimum travel time between villages.

2.3.4. Network analysis

From the Euclidean distance and travel time calculations, we obtained two origin-destination datasets of all possible connections between villages in the study area. After data cleaning and harmonization (Supplementary Methods 2.1), we used these interactions (links) to construct graph class objects in R for visualization and calculation of centrality metrics for each community within each of the watersheds (Figure 2.2). Network processing and visualization

was performed using R Statistical Software (v4.2.2; R Core Team 2021). The standardized mean and standard deviation of the computed metrics were estimated overall and by watershed.

2.3.4.1. Network processing

In this study, the 1,608 communities were considered as nodes and the 73,944 possible connections as the edges. The origin-destination dataset containing the connections was formatted as an edge list, as each row represents an edge. The distances and travel times computed for each connection were assigned iteratively as edge weights, which may represent the strength or weakness of the connection between nodes. Additionally, we constructed four different versions of weights based on the gravity model by combining population and average annual forest loss as masses with Euclidean distance and travel time as cost proxies. Finally, we scaled the weights to range from 0 to 1 within each watershed.

Having the nodes and edges with their different weights, we used the tidygraph package in R [125] to create the graphs for each watershed. Since the connections between the villages are unique and do not have directionality, the resulting graphs are undirected graphs. Next, we calculated the centrality metrics. For this study, we considered strength, closeness, betweenness, and eigenvector centrality measures (Supplementary Methods 2.2). The calculations of these centralities were made considering the weight of the edges. Each type of centrality works with a different interpretation of the weights. As mentioned earlier, the weights represent the relative importance of the connection between two nodes. In the case of strength and eigenvector centrality indicators, the weight is interpreted as connection strength and, therefore, a higher weight indicates a stronger connection between the nodes. On the other hand, in the case of

closeness and betweenness centrality indicators, the weight is interpreted as connection weakness and, therefore, a higher weight indicates a weaker connection between the nodes. Therefore, to keep signs consistent and comparable between centrality measures, we use the inverse of these measures for the calculations of the strength and eigenvector centralities. On the other hand, since the weights based on the gravity model represent the "attraction" between communities, we used the inverses of these weights for the calculations of the closeness and betweenness centralities.

We thus obtained six different versions of the centrality indicators depending on the weight used: 1) Euclidean distance, 2) Gravity model with distance and population, 3) Gravity model with distance and forest loss, 4) Travel time, 5) Gravity model with travel time and population, and 6) Gravity model with travel time and forest loss. Within every watershed, we calculated all the versions of the centrality indicators for each village and then scaled them from 0 to 1 using tidygraph. The correlations between all centrality metrics across all villages were computed using Pearson correlation. Dendrograms to cluster centrality metrics were based on a hierarchical cluster analysis using a complete linkage method.

2.3.4.2. Network visualization

Network visualizations were constructed using the ggraph package in R [126]. Two types of visualizations were used for all versions of the centrality metrics. The first consisted of plotting all nodes and edges and distinguishing them by watershed using colors, plotting the opacity of the edges as a function of the weight value (higher weight, less opaque), and setting the node size as a function of the centrality index value (higher centrality, larger size). The second type of

visualization consisted of plotting the networks in different facets for each watershed and following the same settings for edge opacity and node size as in the first type of visualization. For both visualizations, two node layout algorithms were tested: the Kamada-Kawai algorithm and the Stress majorization algorithm. Both algorithms emerge from the same optimization problem, however, the second one uses a more global approximation technique to the problem, resulting in improvements in run time and stability of the resulting node layout [127]. The algorithm used for the final visualizations was chosen by visual inspection of how well the nodes were arranged for our data.

2.3.5. Stratified analysis

Further explorations of the relationship between connectivity metrics and malaria incidence were conducted by stratifying the data across levels of intervention coverage of the PAMAFRO project (2006–2010). Four control activities were recorded per district and year including: i) strengthening of malaria diagnosis, ii) training and supervision of community health workers, iii) community-based larval source management (LSM), and iv) distribution of long-lasting insecticidal nets (LLINs). A more detailed description of control activities carried out during the 2006–2010 intensified malaria control period (PAMAFRO) is found elsewhere [26]. The PAMAFRO intervention coverage was computed as the proportion of intervention-years (maximum of 4 interventions multiplied by 5 years; 20 intervention-years) conducted in each district and assigned to all of that district's villages. To test the trend in the differences in malaria incidence between low and high centrality metrics across the levels of PAMAFRO coverage, an ordinary least squares (OLS) Linear Regression model for interaction between centrality and

PAMAFRO coverage categories were constructed using the watershed units as a fixed effect. (code repository: https://github.com/healthinnovation/network-malaria).

2.4. Results

2.4.1. Baseline characteristics of villages

In total, data from 1,608 nodes (villages/cities) nested within 31 watersheds in the Loreto department were analyzed after data cleaning (Supplementary Figure 2.1). The total malaria cases (2011-2018) was 232,252 *P. vivax* and 60,512 *P. falciparum*. The number of villages in the selected watersheds ranged from 8 to 202, with an average of 51 villages per watershed (Table 2.1). Most populated villages/cities are located on the banks of rivers, mainly close to Iquitos city; in contrast, most highly deforested areas are located on the south-west side of the study area (Figure 2.1).

Important spatial heterogeneity was observed for both *P. vivax* and *P. falciparum* cases. The highest Annual Parasite Index (API) was observed in the watersheds of Pastaza, Tigre, Yavari, and Napo (Figure 2.3). These watersheds showed contrasting epidemiological profiles. Pastaza and Tigre watersheds showed a rapid increase in malaria incidence; on the other hand, Napo and Yavari watersheds, despite their high malaria endemicity levels, were stable during the 2011–2018 period (Figure 2.3). This scattered spatial location of villages and cities is reflected in the contrasting distributions of distance and travel time between village/city dyads (Supplementary Figure 2.2). Both are positively skewed; however, a larger kurtosis is present in the travel time distribution in comparison to Euclidian distance. This pattern is consistent across all the 31 watersheds (Supplementary Figure 2.3).

2.4.2. Centrality estimation

Multiple connectivity metrics were computed from the combination of centrality types and weights. A consensus graph was constructed to represent the network between villages located in the same watershed using multiple iterations of centrality metrics. An example using betweenness centrality with weights based on a gravity model that includes Euclidian distance and population is shown in Figure 2.4. Due to the high density of villages and links, a version divided by watersheds is presented in Supplementary Figure 2.4. The densest networks are Intercuenca 4977, Napo, Medio Bajo Marañon, Pastaza, and Nanay.

A total of 24 centrality metrics were computed. Overall summary statistics are shown in Table 2.2 and watershed-specific statistics are shown in Supplementary Table 2.1. Metrics with the highest standardized mean and variability are those that use distance and travel time as weights. On the other hand, metrics with the lowest standardized mean and variability are those that use a gravity model based on distance and population or travel time and population. Betweenness centrality is consistently the metric with the lowest standardized mean and variability across all weights (Table 2.2) and these trends are consistent across watersheds (Supplementary Table 2.1). Strong correlation patterns were observed between centrality metrics (Figure 2.5). Betweenness centrality is the metric that showed the most consistent clustering pattern in the hierarchical clustering analysis (Supplementary Figure 2.5).

2.4.3. Relationship of centrality measures with malaria incidence

Betweenness centrality indicators using multiple versions of gravity models as weights were selected for further analyses based on the criteria described above. This metric was used to define categories of centrality (low vs. high quintiles) to test their relationship with malaria incidence. Overall, villages in the top quintile of centrality had a higher malaria incidence in comparison with villages in the bottom quintile of centrality (Mean Difference [MD] in cases per 1,000 individuals; *P. vivax* = 165.78 and *P. falciparum* = 76.14) (Supplementary Figure 2.6). When stratifying by levels of PAMAFRO coverage, the mean difference between villages at the top and bottom centrality quintiles increases as PAMAFRO coverage increases (Supplementary Table 2.2) for both *P. vivax* (Tier 1 = 155.36; Tier 2 = 176.22; Tier 3 = 326.08) and *P. falciparum* (Tier 1 = 48.11; Tier 2 = 95.16; Tier 3 = 139.07). Overall distributions are comparable across calculations of weights (i.e., combinations of distance/travel time and population/deforestation); however, the dose-response pattern following the PAMAFRO coverage is consistent across all combinations of weight calculations for *P. falciparum,* in contrast to *P. vivax* where the pattern is more noticeable when using deforestation instead of population (Figure 2.6). The consensus plots and results of the comparison of the malaria incidence between high and low centrality villages using other centrality types (closeness, eigen, and strength) are presented in Supplementary Figure 2.7 – 2.9.

2.5. Discussion

How villages and cities are connected and how these connections influence the transmission of pathogens remains of great interest for global public health. Evidence from this study contributes to the scarce literature on human mobility and its impact on malaria in rural

areas of the Amazon region. This study investigated this relationship in the Peruvian Amazon using a comprehensive set of connectivity metrics and malaria incidence records at a granular spatial resolution. Betweenness centrality outperformed other metrics based on its consistency and stability, and no meaningful differences were detected when using multiple versions of weights such as population, deforested area, Euclidian distance, or travel time. The evidence presented in this study highlights that villages and cities with a high connectivity are consistently locations with higher malaria incidence in the Loreto department of Peru. Further explorations showed that this difference in malaria incidence is exacerbated in areas that received greater coverage of malaria control activities from the PAMAFRO project. These, in turn, are the areas with the greatest baseline malaria transmission. The findings of this study support a shift in current malaria control strategies from targeting particular locations based on their malaria metrics towards strategies based on connectivity neighborhoods that incorporate influential connected villages that share a flow of parasites and hosts.

The main mechanism of malaria transmission reestablishment is the importation of parasites from HPM, and this plays a major role in elimination scenarios. However, it is also meaningful at the micro-geographical level, after malaria control interventions occur and malaria endemicity is expected to be low with only remaining *P. vivax* hypnozoite reservoirs. In these scenarios, interrupting importation pathways may greatly improve the effectiveness of current malaria control efforts [28,61]. However, capturing individual HPM information requires intensive use of resources [73,98,128,129]. This study showed that the use of connectivity metrics between villages contributes to an improved understanding of these complex dynamics in rural areas that are highly connected through river networks.

Importantly, we found that in areas with the greatest malaria endemicity and coverage of PAMAFRO control activities, the influence of connectivity was more prominent. These findings challenge previous literature that highlighted a greater importance of HPM in low-transmission and close-to-elimination settings than in moderate- and high-transmission settings [61–63,82,130]. In fact, less attention was put into the role of HPM and connectivity in high malaria transmission settings [67,131]. Interestingly, areas with high vectorial capacity [31] and parasite genomic diversity [132] are areas with intense HPM in the Peruvian Amazon. These findings are consistent with an emergent body of evidence showing the role HPM in high- to moderate-transmission settings [15,74,98,100,133–135].

In the Amazon region, as cities grow, HPM intensifies, and as a result so does the probability of malaria importation. The magnitude of connectivity in these areas is affected by the size of the cities as a proxy for the number of services and level of commercial activity in place. In addition, these anthropogenic environmental changes impact infectious disease dynamics [44]. Increased human population and environmental modification influence biological communities, including *Anopheles* mosquitoes, particularly those with some degree of competence to transmit *Plasmodium* species that circulate in the Amazon region [45,136]. In this study, centrality metrics computed using population size and deforested area showed comparable performance. These similarities may be leveraged in scenarios with weak vital registration statistics such as rural areas, areas under conflict, or forced displaced populations [137–140] since the collection of deforested area could be conducted using remote sensing tools in comparison to the intense effort involved in a population census.

In this study, connectivity and centrality measures were assessed in relationship to land coverage change using network analysis at the village level in Peru, and their effect on malaria transmission was estimated. This evidence contributes to the understanding of the role of HPM in malaria transmission in rural areas, and secondarily, provides information to optimize the distribution of services or the configuration of networks to reduce the overall flow of malaria infections between cities and villages.

We acknowledge some limitations of this study. First, 221 villages (12%) were excluded from the analysis in the data cleaning process due to missing mass and cost data for the weight calculations (Supplementary Figure 2.1). These exclusions may alter the estimations; however, data was missing completely at random (MCAR). Second, for the connectivity metrics computation, all villages located in the same watershed were assumed to be connected. However, human preferences to avoid villages in the same watershed or travel to villages in another watershed are plausible. We suggest further studies in other settings that consider more complex network structures. Finally, previous studies in the Peruvian Amazon [141–143] reported a high number of sub-clinical infections that are not recorded by the MoH during routine data collection. The findings of this study are relevant only for clinical cases, and caution is suggested when interpreting these results for asymptomatic cases, which can contribute to the maintenance of parasite transmission.

2.6. Conclusion

This study exploited detailed malaria incidence data at the village level to test the influence of a comprehensive set of connectivity metrics. The data in this study show that in the Loreto department of Peru, villages and cities with high connectivity consistently have higher malaria incidence. When stratified by coverage of PAMAFRO control activities, the areas where malaria transmission was the highest are the areas where this difference in malaria incidence is most pronounced. These findings challenge prior research that emphasized the importance of HPM being greater in low-transmission and close-to-elimination settings rather than in moderate- and high-transmission settings. The evidence outlined in this study can be used to tailor malaria control strategies in rural areas by prioritizing influential connected neighborhoods instead of single villages.

2.7. Tables

Table 2.1. Descriptive demographical, epidemiological, and environmental characteristics (2011–2018) in all villages nested in 31 watersheds in the Loreto department, Peru.

Watershed name	Number of villages	Total number of cases		Deforestation	
		P. falciparum	*P. vivax*	Mean	sd
Cuenca Carhuapanas	32	10	664	5.41099	3.35717
Cuenca Itaya	69	2415	21452	4.75289	2.53186
Cuenca Manití	14	327	1228	2.23778	1.79954
Cuenca Morona	46	595	2488	1.29575	0.88265
Cuenca Nanay	91	10276	47046	4.43810	3.29485
Cuenca Napo	169	6333	28177	3.15514	2.35080
Cuenca Paranapura	86	285	5073	9.75672	5.39148
Cuenca Pastaza	95	16825	34343	1.51598	1.24127
Cuenca Potro	8	15	310	1.35012	0.73314
Cuenca Putumayo	42	232	1037	1.07531	1.03147
Cuenca Tahuayo	17	10	213	4.36052	2.01849
Cuenca Tapiche	35	1365	4956	2.42216	2.55880
Cuenca Tigre	75	11830	33253	2.33620	1.84907
Cuenca Yavari	38	2815	10177	2.92976	2.09816
Intercuenca 4977	202	1820	16051	4.27788	3.04429
Intercuenca 49791	12	7	107	4.31276	2.35457
Intercuenca 49793	40	125	1441	5.53734	3.08079
Intercuenca 49795	8	4	34	4.88276	2.69110
Intercuenca 49797	49	82	1580	4.65196	1.88203
Intercuenca 49799	35	6	155	4.14853	1.97210
Intercuenca 49871	11	16	418	6.73127	3.70163
Intercuenca 49873	10	3	32	4.31457	2.41013
Intercuenca 49877	42	130	989	3.49134	2.32782
Intercuenca 49911	19	18	254	3.32678	1.99509
Intercuenca 49913	72	42	618	3.95505	2.86444
Intercuenca 49915	9	1	26	5.73575	2.71946
Intercuenca Bajo Huallga	62	136	500	7.78687	5.19647
Intercuenca Bajo Marañón	48	329	8600	3.89483	2.19509
Intercuenca Medio Bajo Huallaga	36	67	336	10.66703	6.19642
Intercuenca Medio Bajo Marañón	124	4383	10620	2.10576	2.06783
Intercuenca Medio Marañón	12	10	74	3.27035	2.35675

Table 2.2. Descriptive statistics of centrality metrics in all villages in the Loreto department, Peru.

Centrality	Mean (sd)
Distance as weight	
Strength	0.68 (0.26)
Closeness	0.65 (0.28)
Betweenness	0.14 (0.24)
Eigenvector	0.64 (0.28)
Distance and population-based gravity model weight	
Strength	0.11 (0.18)
Closeness	0.19 (0.21)
Betweenness	0.03 (0.15)
Eigenvector	0.12 (0.18)
Distance and forest loss-based gravity model weight	
Strength	0.27 (0.26)
Closeness	0.39 (0.29)
Betweenness	0.10 (0.23)
Eigenvector	0.26 (0.26)
Travel time as weight	
Strength	0.62 (0.28)
Closeness	0.64 (0.28)
Betweenness	0.12 (0.23)
Eigenvector	0.60 (0.30)
Travel time and population-based gravity model weight	
Strength	0.09 (0.17)
Closeness	0.18 (0.20)
Betweenness	0.02 (0.14)
Eigenvector	0.10 (0.18)
Travel time and forest loss-based gravity model weight	
Strength	0.22 (0.26)
Closeness	0.33 (0.28)
Betweenness	0.12 (0.23)
Eigenvector	0.21 (0.26)

2.8. Figures

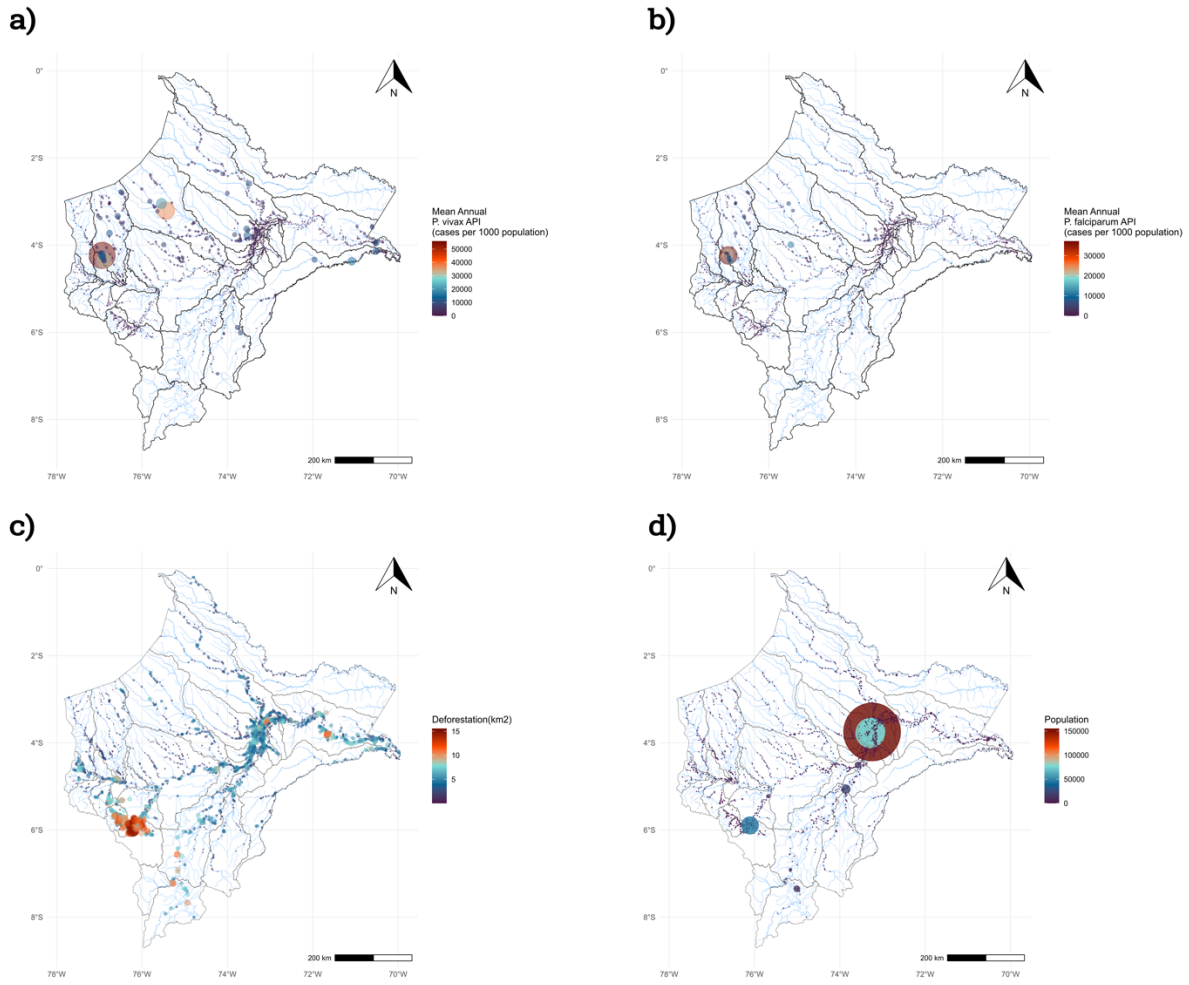


Figure 2.1. Study area and hydro-basins in the Loreto department in the Peruvian Amazon. Each point represent the location of villages and the color and size represent their A) Mean Annual *P. vivax* API, B) Mean Annual *P. falciparum* API, C) Deforested area, and D) Population size. Maps were produced using R v.4.1 (R Development Core Team, R Foundation for Statistical Computing, Australia) based on public geographic data extracted from © OpenStreetMap contributors (www.openstreetmap.org) under Open Data Commons Open Database License (ODbL) 1.0 (http://openstreetmap.org/copyright).

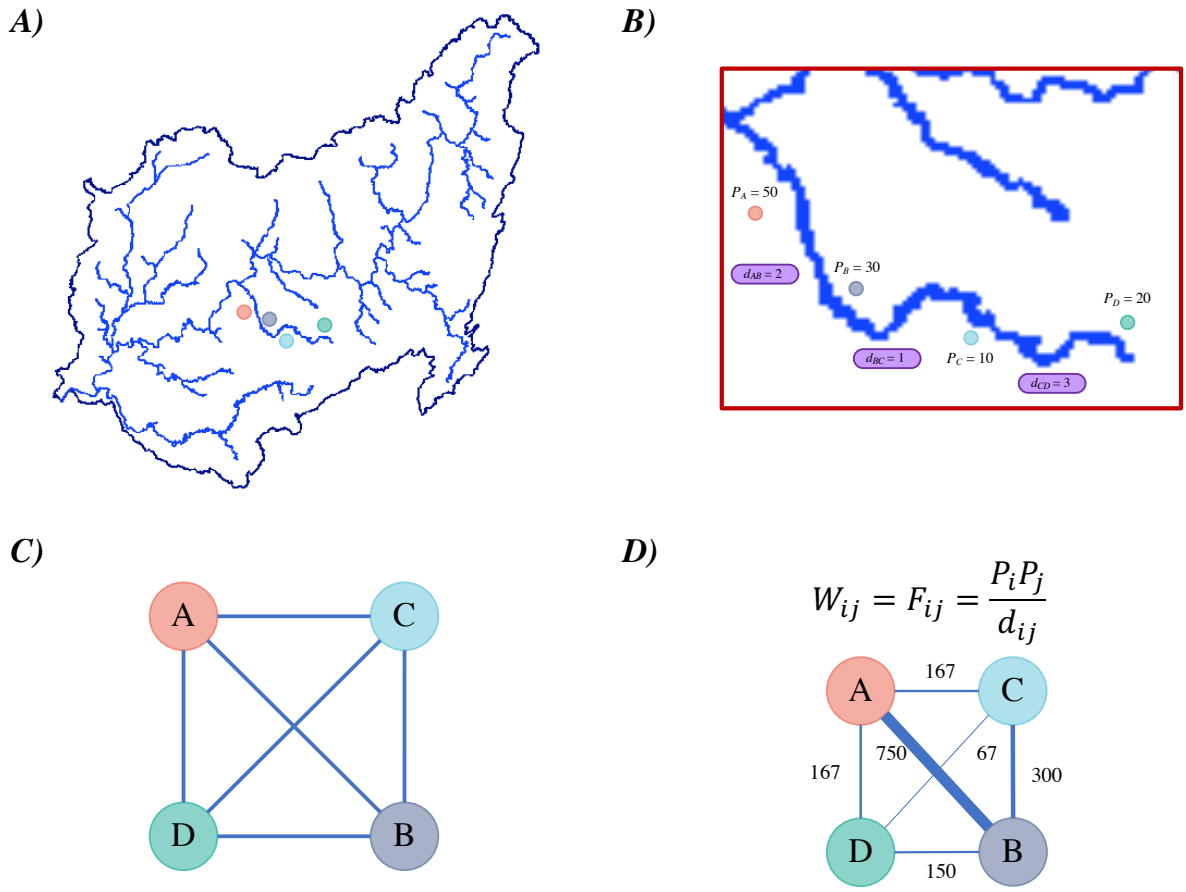


Figure 2.2. Connectivity and centrality estimation workflow. Synthetic example of all the steps to compute the centrality metrics that comprises A) the geolocation of the river network and villages in each watershed area, B) estimating the cost of displacement between villages (i.e., distance and travel time), C) construction of an undirected and unweighted network based on the connections between villages in the same watershed, and finally, D) testing gravity model weights for the links in the network. Weights were computed using multiple masses (i.e., population and deforested areas). Edges widths are relative to weights.

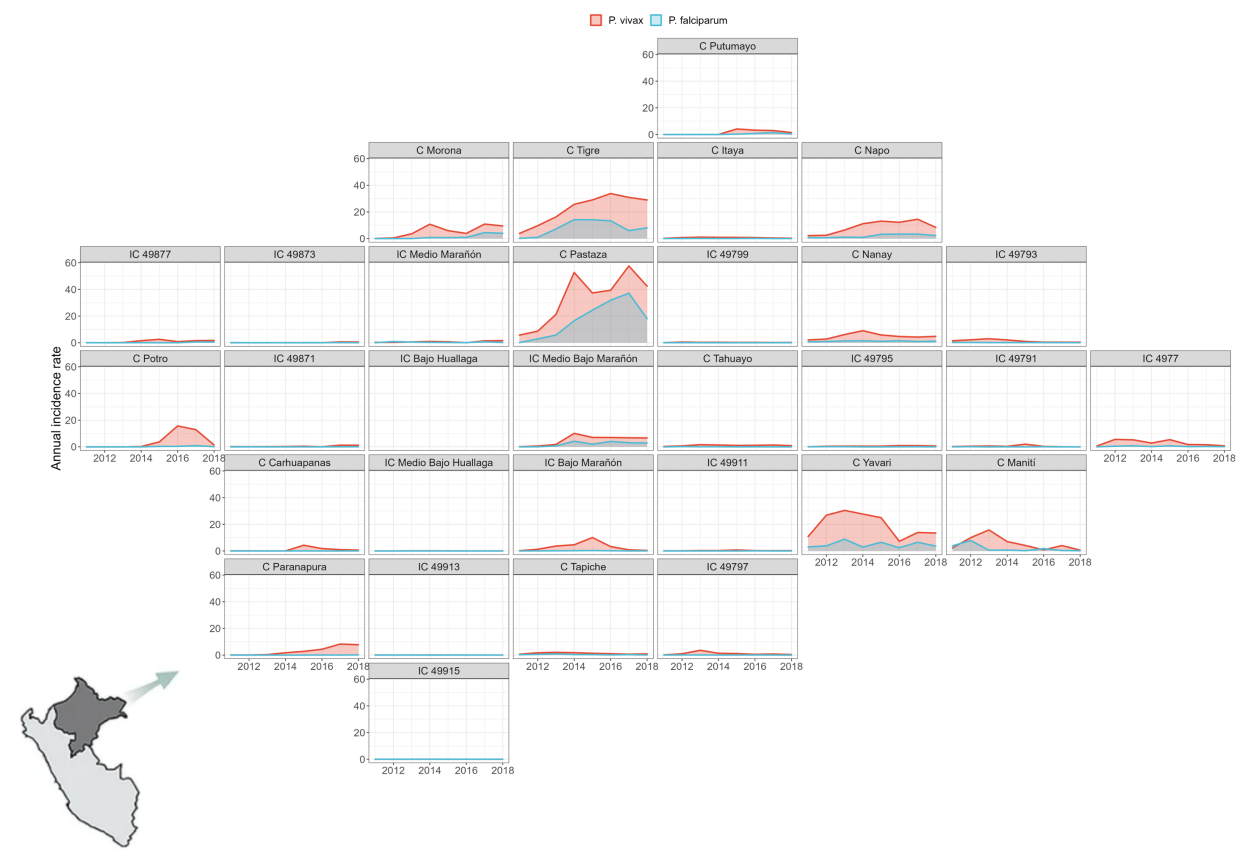


Figure 2.3. Annual malaria incidence rate variation by parasite species. Variation in annual malaria incidence rates due to *P. vivax* (red) and *P. falciparum* (light blue) in 31 watersheds of the Loreto department between 2011 and 2018.

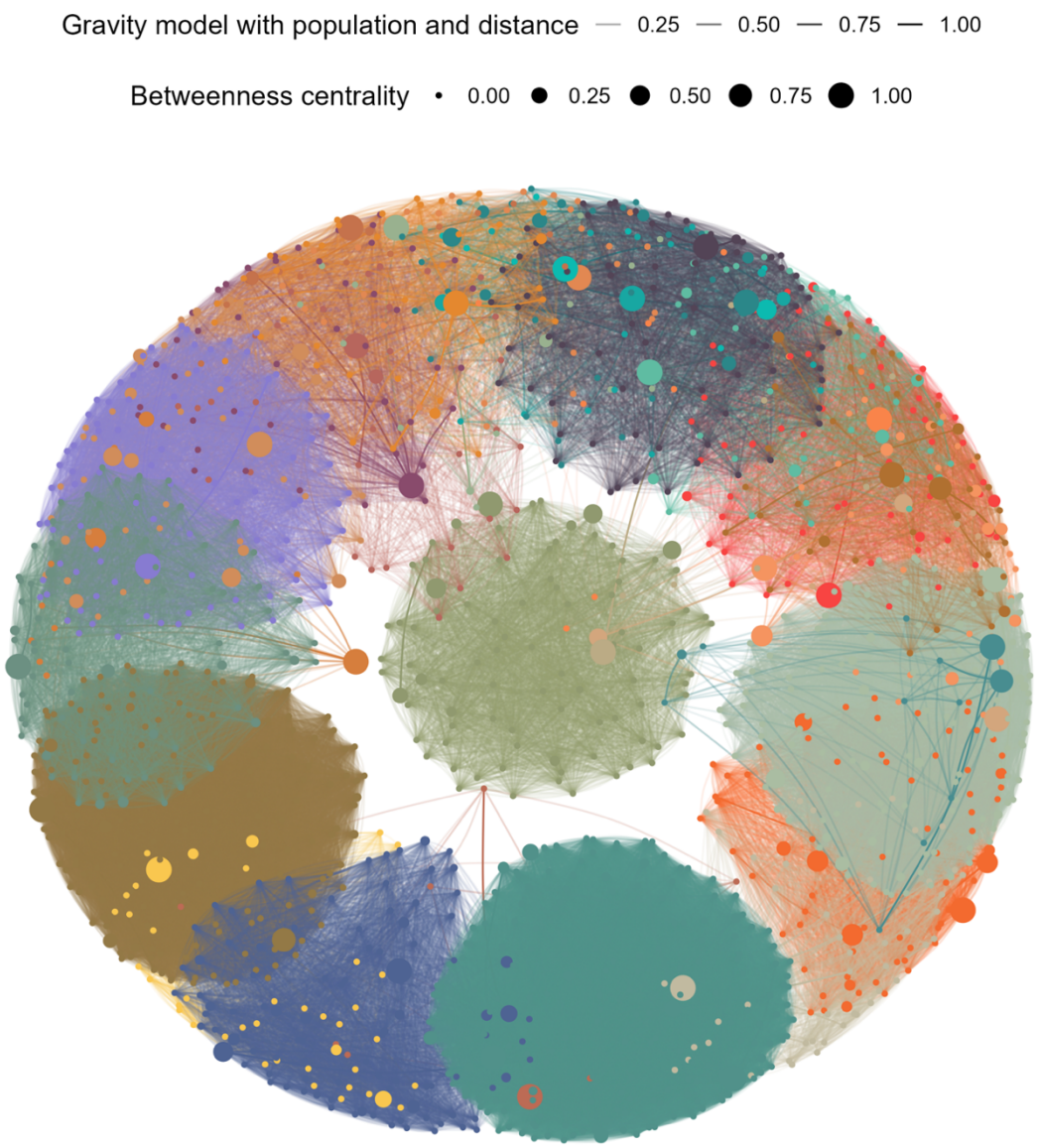


Figure 2.4. Consensus graph of the network of villages in the Loreto department in the Peruvian Amazon.

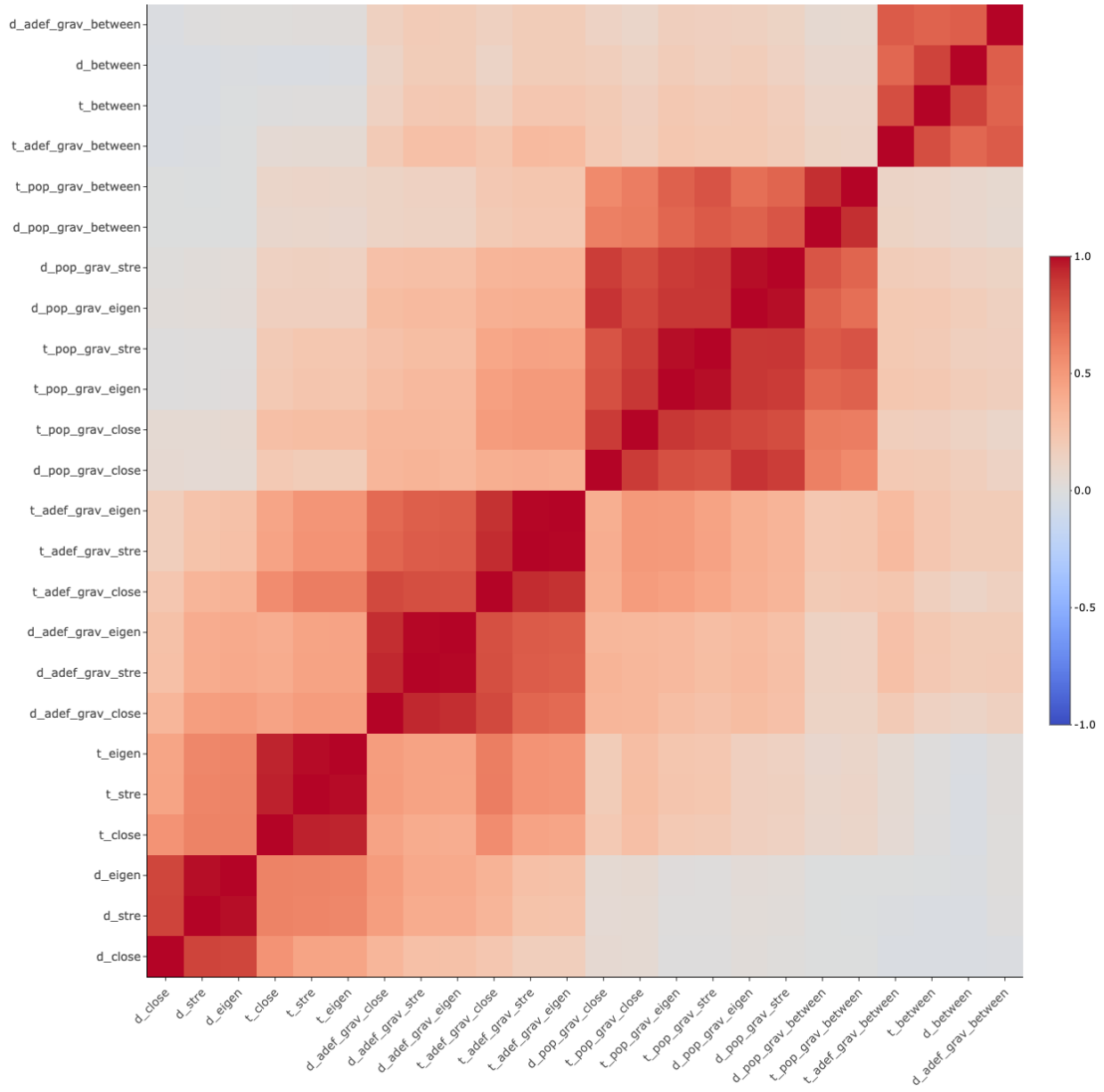
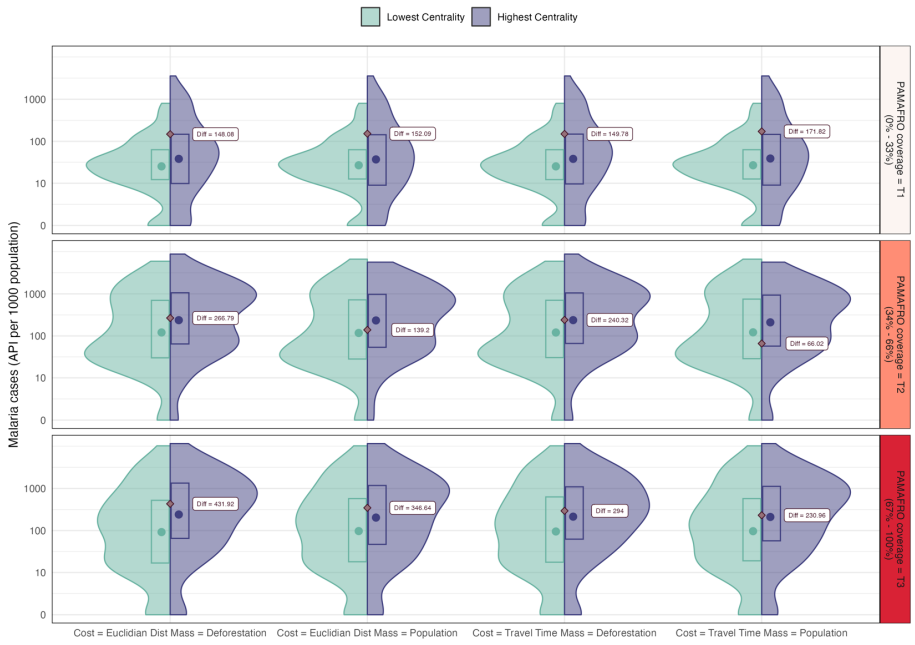


Figure 2.5. Correlation of centrality metrics of villages in the Loreto department in the Peruvian Amazon. Abbreviations: Mass (population [pop], deforested area [adef], and none), cost (distance [d], travel time [t]), and centrality type (betweenness [between], strength [stre], eigen [eigen], and closeness [close]).

A)



B)

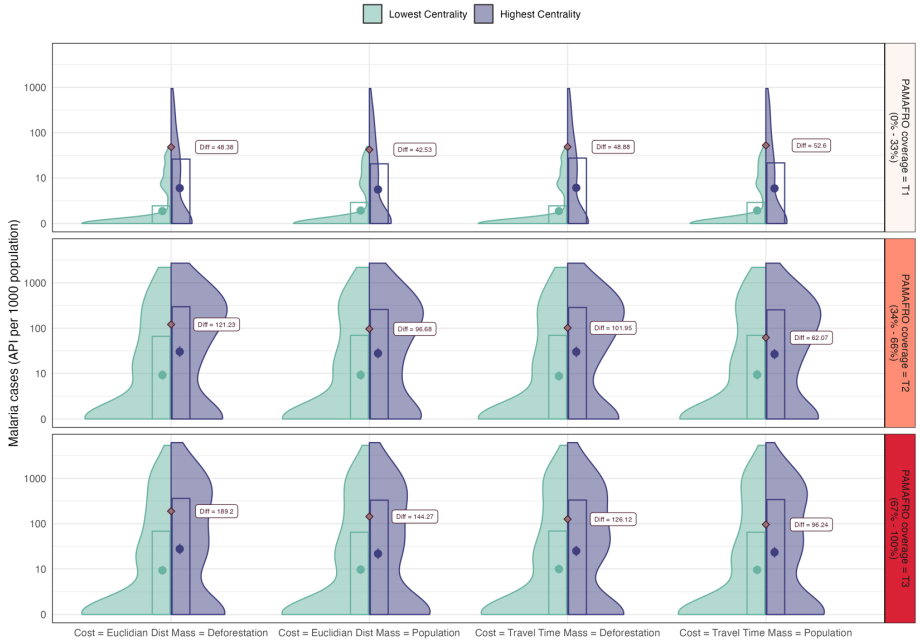


Figure 2.6. Distribution of Total Annual Parasite Index (API) per 1,000 individuals (2011–2018) across high and low centrality villages stratified by levels of PAMAFRO intervention coverage in the Loreto department in the Peruvian Amazon. A) For *P. vivax* and B) for *P. falciparum.* Mean difference between groups are represented as diamonds in each panel. T1 = Tier 1 (Low coverage; Low baseline endemicity), T2 = Tier 2 (Moderate coverage; Moderate baseline endemicity), and T3 = Tier 3 (High coverage; High baseline endemicity).

2.9. Supplementary information

Supplementary Methods 2.1. Network Analysis – Data cleaning

The origin-destination data sets contained duplicates because the calculation of distance and travel time was performed against the entire set. In other words, there could be a connection from village A to B and from B to A, which gave the same distance and travel time. We removed the duplicates and also filtered out those connections that had a calculation equal to zero (connection between the same community). Overall, 1,671,706 distance connections and 1,627,314 travel time connections were computed. The complete case dataset was constructed resulting in a single origin-destination dataset with 1,627,314 connections, having one column for distance and another one for travel time.

The average annual forest loss for each village from 2009 to 2018 was then calculated. Those villages with an average annual forest loss of zero and a population of zero were excluded from further processing. From the total number of villages and connections, 1,634 villages and 73,946 connections remained for analysis. After this, we grouped the connections by watershed and filtered out those that had only one connection, leaving us with 31 watersheds, 73,944 connections and 1,608 communities (Supplementary Figure 2.1).

Supplementary Methods 2.2. Description of centrality metrics

The strength centrality of a particular node is the sum of the weights of all its adjacent edges (Kolaczyk & Csárdi, 2014). The weights here are interpreted as measures of attraction between nodes. On the other hand, the closeness centrality measures how “close” a node is to

other nodes it is connected to, and is defined as the inverse sum of the weights of the adjacent edges of a particular node, given that the weights are measuring the distance between the nodes (Kolaczyk & Csárdi, 2014). Using this same interpretation of the weights, if for every pair of nodes in the graph we find the path or sequence of edges that connects the nodes that minimizes the sum of the edge weights (that is, the shortest path), then the betweenness centrality of a certain node is the proportion of shortest paths that passes through this node (Kolaczyk & Csárdi, 2014). Finally, the eigenvector centrality uses the interpretation of weights as connection strength to give higher scores to the nodes that are more connected to other nodes with high scores (Kolaczyk & Csárdi, 2014). Mathematically, the scores are calculated as the eigenvector components of the weighted adjacency matrix of the graph.

Supplementary Table 2.1. Descriptive centrality metrics by watershed

Centrality	Cuenca CarhuapanaN =321	Cuenca Itaya N = 781	Cuenca Manití N = 151	Cuenca Morona N = 461	Cuenca Nanay N = 841	Cuenca Napo N = 1731	Cuenca Paranapura N = 861	Cuenca Pastaza N = 951	Cuenca Potro N = 81	Cuenca Putumayo N = 421	Cuenca Tahuayo N = 161	Cuenca Tapiche N = 351	Cuenca Tigre N = 761	Cuenca Yavari N = 381
Distance as weight														
Strength	0.61 (0.32)	0.71 (0.22)	0.68 (0.27)	0.53 (0.30)	0.78 (0.19)	0.73 (0.25)	0.74 (0.20)	0.61 (0.21)	0.41 (0.40)	0.64 (0.30)	0.54 (0.34)	0.60 (0.32)	0.66 (0.25)	0.63 (0.29)
Closeness	0.67 (0.29)	0.66 (0.28)	0.61 (0.28)	0.59 (0.33)	0.79 (0.18)	0.75 (0.24)	0.68 (0.22)	0.58 (0.26)	0.58 (0.40)	0.64 (0.33)	0.59 (0.35)	0.62 (0.31)	0.64 (0.29)	0.55 (0.29)
Betweenness	0.03 (0.18)	0.03 (0.16)	0.50 (0.00)	0.02 (0.15)	0.05 (0.21)	0.02 (0.13)	0.02 (0.15)	0.02 (0.14)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.03 (0.17)	0.01 (0.11)	0.08 (0.27)
Eigenvector	0.59 (0.33)	0.68 (0.24)	0.65 (0.28)	0.50 (0.32)	0.76 (0.22)	0.68 (0.29)	0.70 (0.21)	0.56 (0.24)	0.48 (0.38)	0.59 (0.32)	0.54 (0.34)	0.55 (0.34)	0.62 (0.28)	0.57 (0.29)
Distance and population based gravity model weight														
Strength	0.24 (0.23)	0.08 (0.14)	0.30 (0.36)	0.20 (0.19)	0.06 (0.12)	0.05 (0.08)	0.18 (0.16)	0.14 (0.14)	0.42 (0.31)	0.07 (0.15)	0.29 (0.25)	0.08 (0.17)	0.06 (0.13)	0.17 (0.20)
Closeness	0.29 (0.24)	0.28 (0.14)	0.27 (0.27)	0.30 (0.23)	0.11 (0.13)	0.08 (0.11)	0.26 (0.18)	0.21 (0.16)	0.49 (0.31)	0.15 (0.19)	0.40 (0.23)	0.17 (0.22)	0.08 (0.15)	0.22 (0.22)
Betweenness	0.06 (0.25)	0.02 (0.13)	0.07 (0.26)	0.04 (0.16)	0.01 (0.11)	0.01 (0.08)	0.02 (0.12)	0.01 (0.11)	0.17 (0.36)	0.02 (0.15)	0.08 (0.25)	0.03 (0.17)	0.02 (0.12)	0.06 (0.22)
Eigenvector	0.25 (0.23)	0.07 (0.14)	0.28 (0.36)	0.21 (0.20)	0.06 (0.12)	0.05 (0.09)	0.17 (0.16)	0.14 (0.14)	0.46 (0.33)	0.09 (0.16)	0.34 (0.26)	0.10 (0.18)	0.06 (0.13)	0.18 (0.21)
Distance and forest loss based gravity model weight														
Strength	0.40 (0.30)	0.35 (0.24)	0.22 (0.36)	0.24 (0.27)	0.23 (0.23)	0.13 (0.19)	0.44 (0.23)	0.21 (0.24)	0.41 (0.36)	0.21 (0.25)	0.59 (0.35)	0.30 (0.32)	0.26 (0.26)	0.15 (0.19)
Closeness	0.51 (0.31)	0.47 (0.26)	0.35 (0.35)	0.32 (0.28)	0.40 (0.31)	0.21 (0.23)	0.57 (0.24)	0.35 (0.31)	0.51 (0.37)	0.32 (0.28)	0.65 (0.35)	0.38 (0.30)	0.31 (0.25)	0.40 (0.24)
Betweenness	0.03 (0.18)	0.03 (0.16)	0.50 (0.00)	0.02 (0.15)	0.05 (0.21)	0.02 (0.13)	0.02 (0.15)	0.02 (0.14)	0.13 (0.35)	0.02 (0.15)	0.50 (0.00)	0.03 (0.17)	0.01 (0.11)	0.08 (0.27)
Eigenvector	0.38 (0.31)	0.33 (0.24)	0.22 (0.37)	0.21 (0.27)	0.21 (0.23)	0.12 (0.18)	0.40 (0.23)	0.20 (0.24)	0.46 (0.37)	0.20 (0.25)	0.57 (0.35)	0.29 (0.33)	0.26 (0.27)	0.13 (0.20)
Travel time as weight														
Strength	0.64 (0.26)	0.70 (0.25)	0.66 (0.25)	0.59 (0.25)	0.53 (0.32)	0.50 (0.27)	0.65 (0.27)	0.49 (0.27)	0.50 (0.35)	0.54 (0.30)	0.58 (0.35)	0.72 (0.23)	0.62 (0.24)	0.49 (0.32)
Closeness	0.65 (0.26)	0.73 (0.24)	0.66 (0.26)	0.63 (0.24)	0.59 (0.32)	0.54 (0.30)	0.71 (0.26)	0.52 (0.28)	0.49 (0.41)	0.55 (0.29)	0.62 (0.33)	0.66 (0.21)	0.60 (0.27)	0.59 (0.27)
Betweenness	0.03 (0.18)	0.01 (0.11)	0.50 (0.00)	0.02 (0.15)	0.02 (0.13)	0.02 (0.12)	0.02 (0.15)	0.02 (0.14)	0.50 (0.00)	0.02 (0.15)	0.50 (0.00)	0.03 (0.17)	0.01 (0.11)	0.08 (0.27)
Eigenvector	0.58 (0.31)	0.69 (0.26)	0.59 (0.26)	0.58 (0.26)	0.51 (0.33)	0.47 (0.29)	0.63 (0.30)	0.46 (0.29)	0.47 (0.35)	0.54 (0.31)	0.58 (0.35)	0.69 (0.25)	0.59 (0.26)	0.40 (0.36)
Travel time and population based gravity model weight														
Strength	0.24 (0.22)	0.09 (0.14)	0.26 (0.35)	0.14 (0.19)	0.07 (0.12)	0.05 (0.08)	0.10 (0.13)	0.09 (0.13)	0.46 (0.29)	0.06 (0.15)	0.27 (0.25)	0.08 (0.17)	0.07 (0.12)	0.12 (0.17)
Closeness	0.30 (0.24)	0.34 (0.14)	0.29 (0.31)	0.23 (0.21)	0.17 (0.13)	0.10 (0.09)	0.16 (0.17)	0.20 (0.14)	0.49 (0.29)	0.17 (0.20)	0.39 (0.23)	0.15 (0.21)	0.08 (0.14)	0.22 (0.21)
Betweenness	0.06 (0.22)	0.01 (0.11)	0.09 (0.27)	0.03 (0.16)	0.01 (0.11)	0.01 (0.08)	0.01 (0.11)	0.01 (0.10)	0.17 (0.36)	0.02 (0.15)	0.07 (0.25)	0.03 (0.17)	0.01 (0.11)	0.03 (0.16)
Eigenvector	0.24 (0.22)	0.08 (0.14)	0.26 (0.36)	0.16 (0.19)	0.07 (0.12)	0.05 (0.08)	0.10 (0.13)	0.09 (0.13)	0.52 (0.29)	0.09 (0.16)	0.32 (0.26)	0.09 (0.18)	0.07 (0.13)	0.12 (0.18)
Travel time and forest loss based gravity model weight														
Strength	0.42 (0.30)	0.27 (0.27)	0.24 (0.39)	0.24 (0.23)	0.13 (0.18)	0.07 (0.14)	0.18 (0.21)	0.09 (0.15)	0.45 (0.33)	0.18 (0.22)	0.59 (0.35)	0.26 (0.30)	0.16 (0.19)	0.16 (0.23)
Closeness	0.52 (0.31)	0.34 (0.27)	0.33 (0.36)	0.38 (0.29)	0.27 (0.25)	0.14 (0.19)	0.33 (0.28)	0.25 (0.17)	0.49 (0.38)	0.26 (0.23)	0.65 (0.34)	0.37 (0.33)	0.36 (0.28)	0.38 (0.29)
Betweenness	0.03 (0.18)	0.02 (0.13)	0.50 (0.00)	0.03 (0.16)	0.05 (0.20)	0.01 (0.08)	0.02 (0.15)	0.02 (0.12)	0.13 (0.35)	0.02 (0.15)	0.50 (0.00)	0.03 (0.17)	0.01 (0.11)	0.08 (0.27)
Eigenvector	0.40 (0.32)	0.26 (0.26)	0.23 (0.39)	0.23 (0.22)	0.12 (0.18)	0.07 (0.14)	0.17 (0.21)	0.08 (0.15)	0.44 (0.38)	0.17 (0.22)	0.58 (0.35)	0.23 (0.30)	0.15 (0.19)	0.14 (0.23)

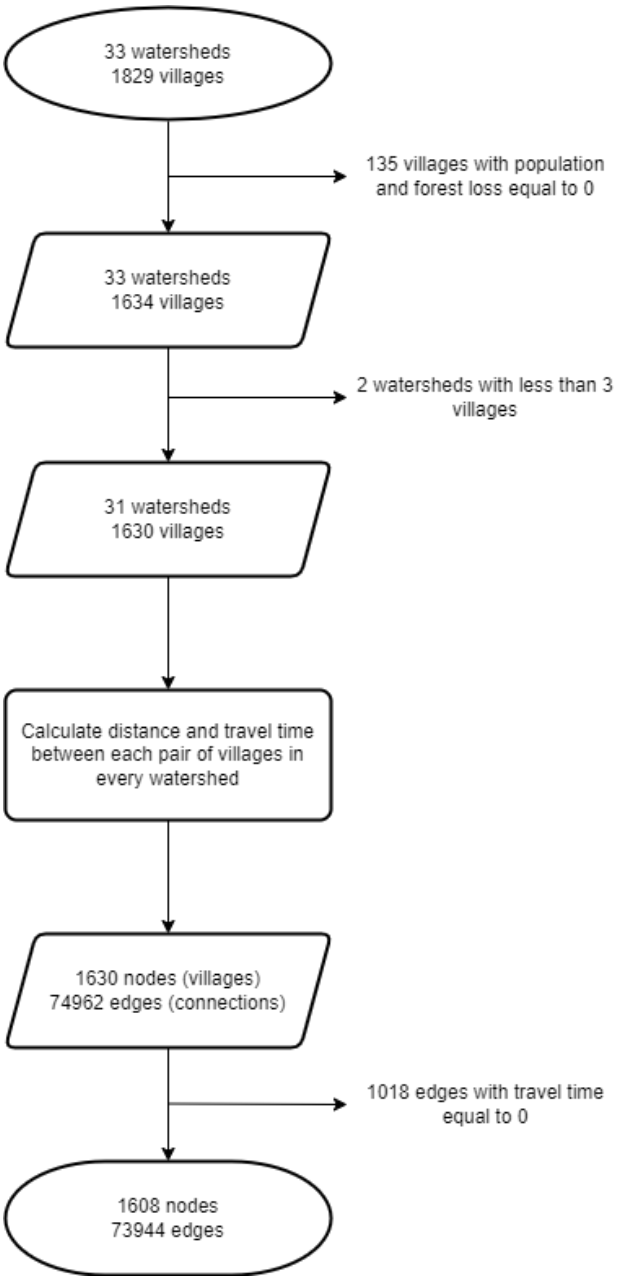
Centrality	Intercuenca 4977 N = 2021	Intercuenca 49791 N = 131	Intercuenca 49793 N = 351	Intercuenca 49795 N = 81	Intercuenca 49797 N = 471	Intercuenca 49799 N = 331	Intercuenca 49871 N = 101	Intercuenca 49873 N = 101
Distance as weight								
Strength	0.75 (0.21)	0.65 (0.30)	0.65 (0.27)	0.49 (0.41)	0.68 (0.28)	0.68 (0.25)	0.47 (0.37)	0.66 (0.34)
Closeness	0.65 (0.26)	0.56 (0.31)	0.57 (0.33)	0.37 (0.41)	0.69 (0.27)	0.60 (0.29)	0.56 (0.38)	0.59 (0.33)
Betweenness	0.02 (0.14)	0.50 (0.00)	0.50 (0.00)	0.40 (0.50)	0.50 (0.00)	0.03 (0.17)	0.50 (0.00)	0.50 (0.00)
Eigenvector	0.72 (0.22)	0.60 (0.31)	0.60 (0.30)	0.48 (0.43)	0.67 (0.29)	0.65 (0.25)	0.47 (0.38)	0.68 (0.37)
Distance and population based gravity model weight								
Strength	0.04 (0.09)	0.32 (0.29)	0.12 (0.16)	0.27 (0.32)	0.07 (0.15)	0.21 (0.21)	0.20 (0.32)	0.36 (0.40)
Closeness	0.06 (0.12)	0.44 (0.32)	0.33 (0.22)	0.33 (0.31)	0.25 (0.21)	0.28 (0.22)	0.31 (0.31)	0.37 (0.37)
Betweenness	0.01 (0.07)	0.16 (0.33)	0.03 (0.17)	0.13 (0.35)	0.02 (0.15)	0.04 (0.18)	0.10 (0.32)	0.12 (0.32)
Eigenvector	0.04 (0.09)	0.36 (0.29)	0.15 (0.17)	0.33 (0.33)	0.10 (0.16)	0.21 (0.22)	0.26 (0.35)	0.38 (0.43)
Distance and forest loss based gravity model weight								
Strength	0.22 (0.17)	0.48 (0.28)	0.42 (0.29)	0.45 (0.39)	0.54 (0.26)	0.39 (0.22)	0.41 (0.36)	0.47 (0.36)
Closeness	0.37 (0.24)	0.65 (0.29)	0.41 (0.25)	0.45 (0.39)	0.54 (0.24)	0.61 (0.25)	0.52 (0.37)	0.63 (0.37)
Betweenness	0.02 (0.14)	0.50 (0.00)	0.50 (0.00)	0.20 (0.38)	0.02 (0.15)	0.03 (0.17)	0.50 (0.00)	0.50 (0.00)
Eigenvector	0.21 (0.17)	0.43 (0.31)	0.39 (0.28)	0.49 (0.42)	0.54 (0.27)	0.36 (0.23)	0.41 (0.37)	0.45 (0.37)
Travel time as weight								
Strength	0.68 (0.24)	0.54 (0.28)	0.55 (0.31)	0.45 (0.34)	0.65 (0.29)	0.57 (0.26)	0.54 (0.39)	0.60 (0.31)
Closeness	0.67 (0.24)	0.63 (0.31)	0.64 (0.30)	0.44 (0.29)	0.70 (0.28)	0.54 (0.27)	0.65 (0.39)	0.62 (0.30)
Betweenness	0.02 (0.14)	0.50 (0.00)	0.50 (0.00)	0.21 (0.40)	0.50 (0.00)	0.03 (0.17)	0.50 (0.00)	0.50 (0.00)
Eigenvector	0.66 (0.24)	0.52 (0.29)	0.52 (0.31)	0.44 (0.36)	0.65 (0.30)	0.53 (0.28)	0.54 (0.39)	0.62 (0.31)
Travel time and population based gravity model weight								
Strength	0.04 (0.11)	0.39 (0.33)	0.18 (0.30)	0.27 (0.32)	0.06 (0.14)	0.15 (0.19)	0.20 (0.32)	0.29 (0.36)
Closeness	0.09 (0.15)	0.47 (0.30)	0.16 (0.22)	0.35 (0.31)	0.18 (0.20)	0.24 (0.23)	0.34 (0.34)	0.33 (0.35)
Betweenness	0.01 (0.08)	0.15 (0.36)	0.04 (0.18)	0.13 (0.35)	0.02 (0.15)	0.03 (0.17)	0.10 (0.32)	0.10 (0.32)
Eigenvector	0.05 (0.11)	0.43 (0.32)	0.17 (0.30)	0.35 (0.32)	0.08 (0.15)	0.16 (0.20)	0.26 (0.35)	0.30 (0.38)
Travel time and forest loss based gravity model weight								
Strength	0.28 (0.23)	0.50 (0.27)	0.17 (0.26)	0.44 (0.43)	0.45 (0.28)	0.30 (0.22)	0.33 (0.37)	0.45 (0.35)
Closeness	0.37 (0.25)	0.62 (0.27)	0.31 (0.33)	0.41 (0.35)	0.47 (0.25)	0.50 (0.26)	0.47 (0.39)	0.53 (0.32)
Betweenness	0.02 (0.14)	0.50 (0.00)	0.50 (0.00)	0.13 (0.35)	0.50 (0.00)	0.03 (0.17)	0.50 (0.00)	0.50 (0.00)
Eigenvector	0.27 (0.22)	0.47 (0.29)	0.15 (0.26)	0.45 (0.43)	0.43 (0.28)	0.27 (0.22)	0.35 (0.39)	0.44 (0.35)

Centrality	Intercuenca 49877 N = 421	Intercuenca 49911 N = 211	Intercuenca 49913 N = 721	Intercuenca 49915 N = 91	Intercuenca Bajo Huallaga N = 621	Intercuenca Bajo Marañón N = 471	Intercuenca Medio Bajo Huallaga N = 371	Intercuenca Medio Bajo Marañón N = 1241	Intercuenca Medio Marañón N = 121	Overall N = 1,608
Distance as weight										
Strength	0.63 (0.27)	0.65 (0.26)	0.69 (0.25)	0.51 (0.37)	0.68 (0.31)	0.55 (0.30)	0.60 (0.29)	0.71 (0.19)	0.60 (0.31)	0.68 (0.26)
Closeness	0.61 (0.26)	0.64 (0.28)	0.68 (0.27)	0.63 (0.40)	0.64 (0.30)	0.59 (0.33)	0.62 (0.30)	0.58 (0.26)	0.50 (0.37)	0.65 (0.28)
Betweenness	0.02 (0.15)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.04 (0.20)	0.03 (0.16)	0.01 (0.09)	0.50 (0.00)	0.14 (0.24)
Eigenvector	0.62 (0.27)	0.60 (0.30)	0.66 (0.26)	0.48 (0.40)	0.64 (0.33)	0.50 (0.33)	0.58 (0.31)	0.67 (0.21)	0.56 (0.31)	0.64 (0.28)
Distance and population based gravity model weight										
Strength	0.12 (0.16)	0.19 (0.23)	0.14 (0.18)	0.22 (0.31)	0.06 (0.14)	0.06 (0.15)	0.06 (0.16)	0.10 (0.14)	0.40 (0.33)	0.11 (0.18)
Closeness	0.33 (0.22)	0.33 (0.25)	0.23 (0.24)	0.47 (0.28)	0.09 (0.15)	0.16 (0.20)	0.27 (0.25)	0.13 (0.15)	0.42 (0.31)	0.19 (0.21)
Betweenness	0.03 (0.15)	0.05 (0.22)	0.03 (0.13)	0.11 (0.33)	0.02 (0.13)	0.02 (0.15)	0.03 (0.16)	0.01 (0.11)	0.14 (0.34)	0.03 (0.15)
Eigenvector	0.15 (0.17)	0.22 (0.25)	0.15 (0.18)	0.32 (0.31)	0.07 (0.14)	0.08 (0.15)	0.10 (0.17)	0.09 (0.14)	0.43 (0.34)	0.12 (0.18)
Distance and forest loss based gravity model weight										
Strength	0.27 (0.24)	0.37 (0.27)	0.27 (0.23)	0.39 (0.39)	0.27 (0.27)	0.34 (0.24)	0.37 (0.26)	0.15 (0.22)	0.34 (0.35)	0.27 (0.26)
Closeness	0.39 (0.25)	0.44 (0.26)	0.40 (0.26)	0.49 (0.37)	0.39 (0.32)	0.50 (0.27)	0.55 (0.29)	0.23 (0.25)	0.46 (0.33)	0.39 (0.29)
Betweenness	0.02 (0.15)	0.08 (0.26)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.04 (0.20)	0.03 (0.16)	0.01 (0.09)	0.08 (0.29)	0.10 (0.23)
Eigenvector	0.24 (0.24)	0.34 (0.27)	0.26 (0.22)	0.38 (0.40)	0.25 (0.27)	0.30 (0.24)	0.33 (0.26)	0.14 (0.21)	0.31 (0.37)	0.26 (0.26)
Travel time as weight										
Strength	0.69 (0.31)	0.48 (0.26)	0.57 (0.27)	0.53 (0.36)	0.69 (0.31)	0.84 (0.21)	0.62 (0.35)	0.75 (0.18)	0.71 (0.30)	0.62 (0.28)
Closeness	0.74 (0.29)	0.48 (0.29)	0.56 (0.29)	0.61 (0.36)	0.71 (0.29)	0.86 (0.20)	0.70 (0.33)	0.77 (0.17)	0.73 (0.30)	0.64 (0.28)
Betweenness	0.02 (0.15)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.04 (0.20)	0.03 (0.16)	0.01 (0.09)	0.50 (0.00)	0.12 (0.23)
Eigenvector	0.68 (0.32)	0.45 (0.28)	0.54 (0.28)	0.50 (0.38)	0.67 (0.31)	0.84 (0.21)	0.61 (0.36)	0.73 (0.19)	0.71 (0.29)	0.60 (0.30)
Travel time and population based gravity model weight										
Strength	0.08 (0.16)	0.15 (0.21)	0.11 (0.17)	0.22 (0.31)	0.05 (0.15)	0.07 (0.15)	0.06 (0.17)	0.07 (0.10)	0.31 (0.30)	0.09 (0.17)
Closeness	0.16 (0.21)	0.33 (0.26)	0.19 (0.23)	0.43 (0.29)	0.08 (0.15)	0.13 (0.18)	0.23 (0.26)	0.16 (0.13)	0.39 (0.29)	0.18 (0.20)
Betweenness	0.02 (0.15)	0.05 (0.22)	0.02 (0.13)	0.11 (0.33)	0.02 (0.13)	0.02 (0.15)	0.03 (0.16)	0.01 (0.09)	0.08 (0.29)	0.02 (0.14)
Eigenvector	0.09 (0.18)	0.20 (0.23)	0.12 (0.17)	0.31 (0.33)	0.05 (0.16)	0.08 (0.16)	0.09 (0.18)	0.07 (0.11)	0.35 (0.31)	0.10 (0.18)
Travel time and forest loss based gravity model weight										
Strength	0.15 (0.25)	0.34 (0.28)	0.34 (0.27)	0.35 (0.38)	0.17 (0.24)	0.23 (0.26)	0.32 (0.33)	0.11 (0.16)	0.46 (0.33)	0.22 (0.26)
Closeness	0.25 (0.29)	0.40 (0.26)	0.39 (0.27)	0.46 (0.35)	0.28 (0.30)	0.40 (0.25)	0.44 (0.36)	0.27 (0.23)	0.55 (0.34)	0.33 (0.28)
Betweenness	0.05 (0.22)	0.06 (0.22)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.06 (0.25)	0.05 (0.23)	0.01 (0.10)	0.50 (0.00)	0.12 (0.23)
Eigenvector	0.15 (0.25)	0.34 (0.28)	0.33 (0.27)	0.35 (0.40)	0.16 (0.24)	0.20 (0.26)	0.30 (0.32)	0.10 (0.16)	0.45 (0.34)	0.21 (0.26)

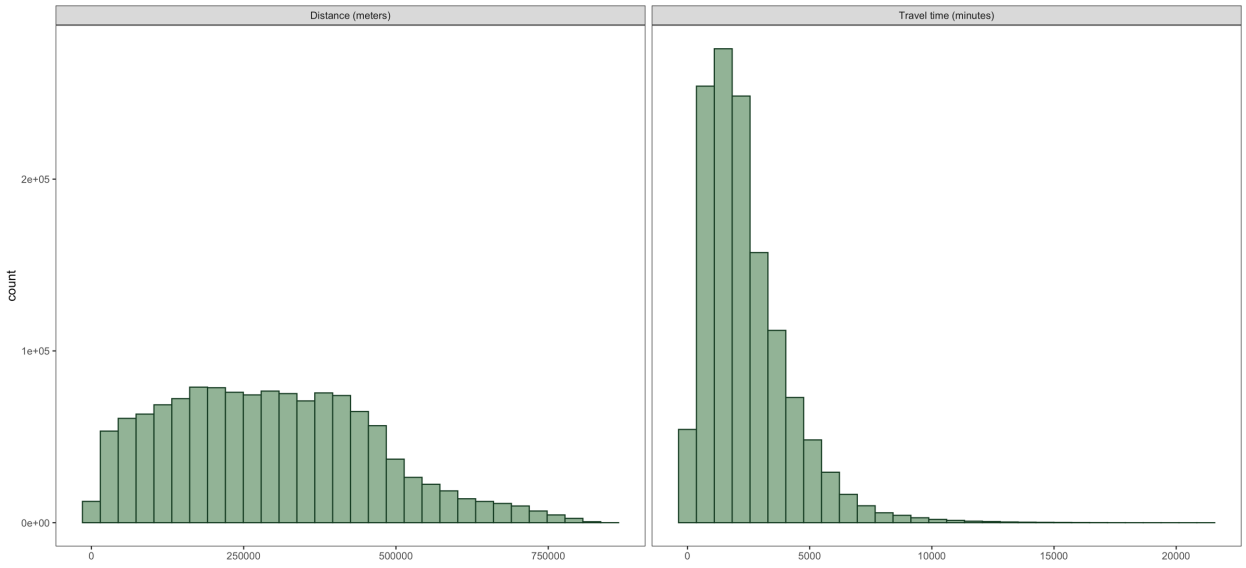
Supplementary Table 2.2. Ordinary least squares Linear Regression model for interaction between centrality and PAMAFRO coverage categories.

	Coeff.	S.E.	t. val.	p-value
High Centrality (Ref: Low Centrality)	1.28	35.34	0.04	0.97
PAMAFRO T2 (Ref: T1)	165.96	48.64	3.41	<0.01*
PAMAFRO T3 (Ref: T1)	112.94	50.87	2.22	0.03 *
High Centrality * PAMAFRO T2	142.59	50.69	2.81	<0.01*
High Centrality * PAMAFRO T3	96.05	46.7	2.06	0.04 *

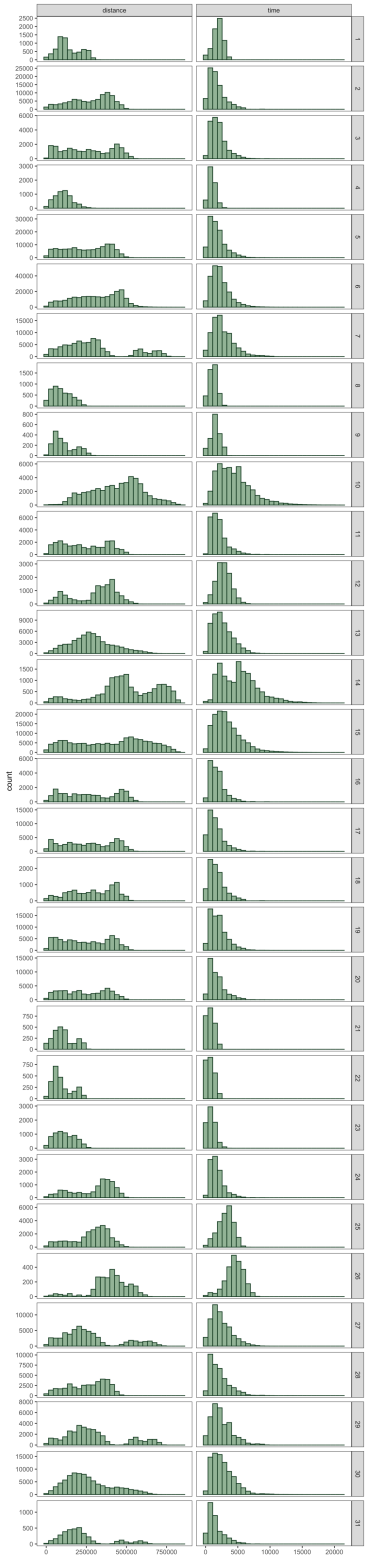
Accounting for watersheds as fixed effect



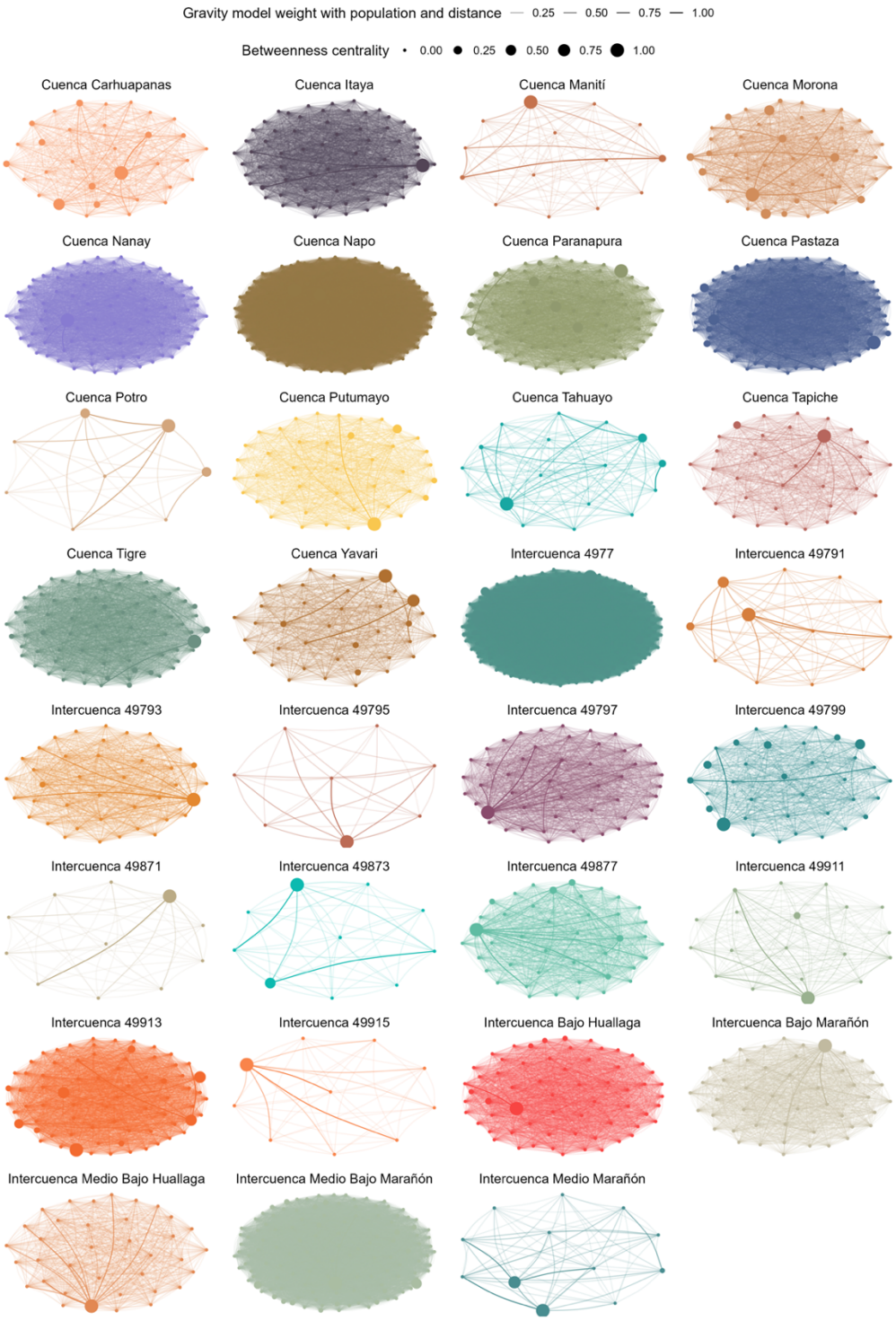
Supplementary Figure 2.1. Data flowchart of the analytical dataset.



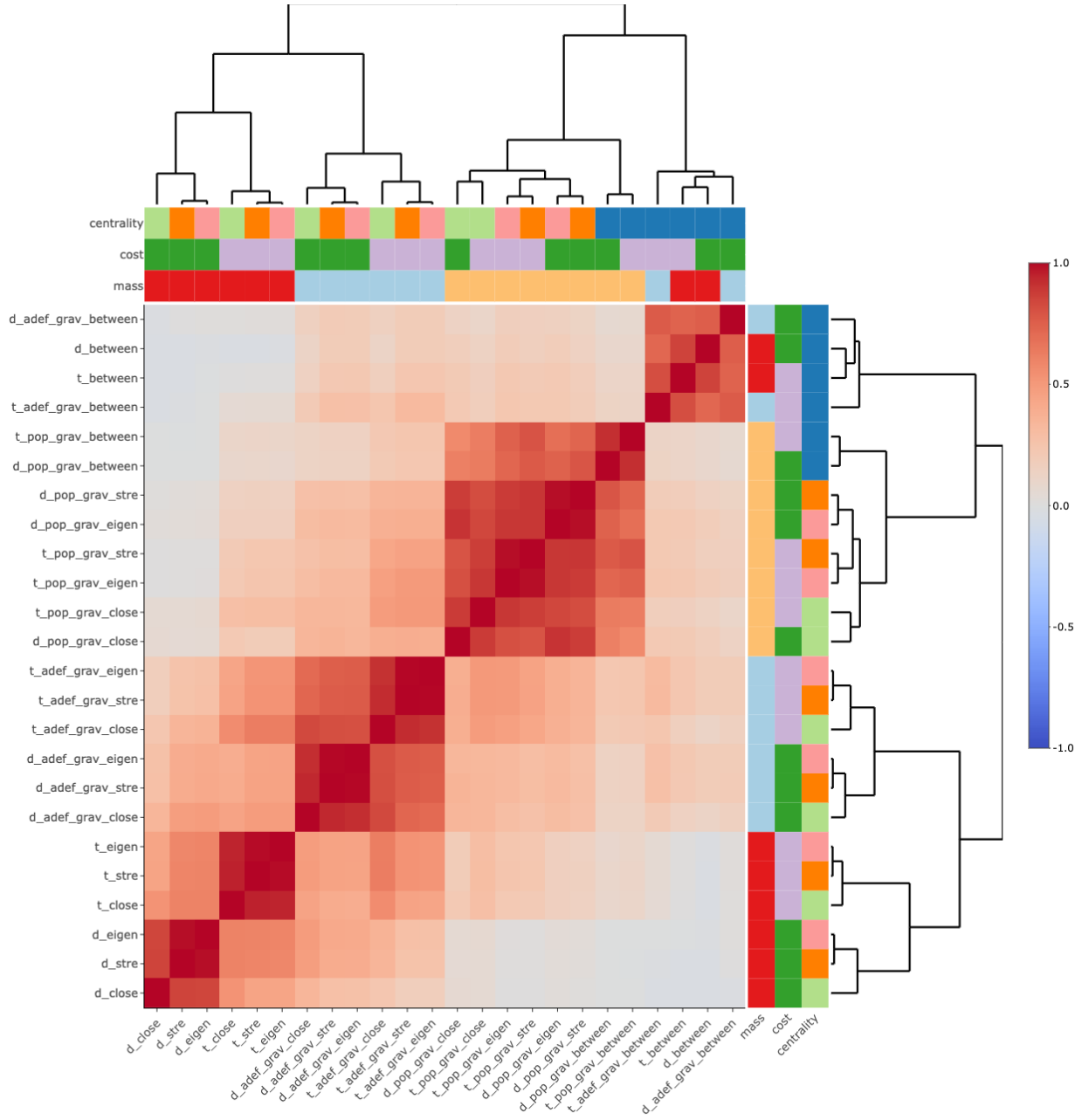
Supplementary Figure 2.2. Overall distribution of distance and travel time between dyad villages in the Loreto department in the Peruvian Amazon.



Supplementary Figure 2.3. Distribution of distance and travel time between dyad villages in each watershed in the Loreto department in the Peruvian Amazon.

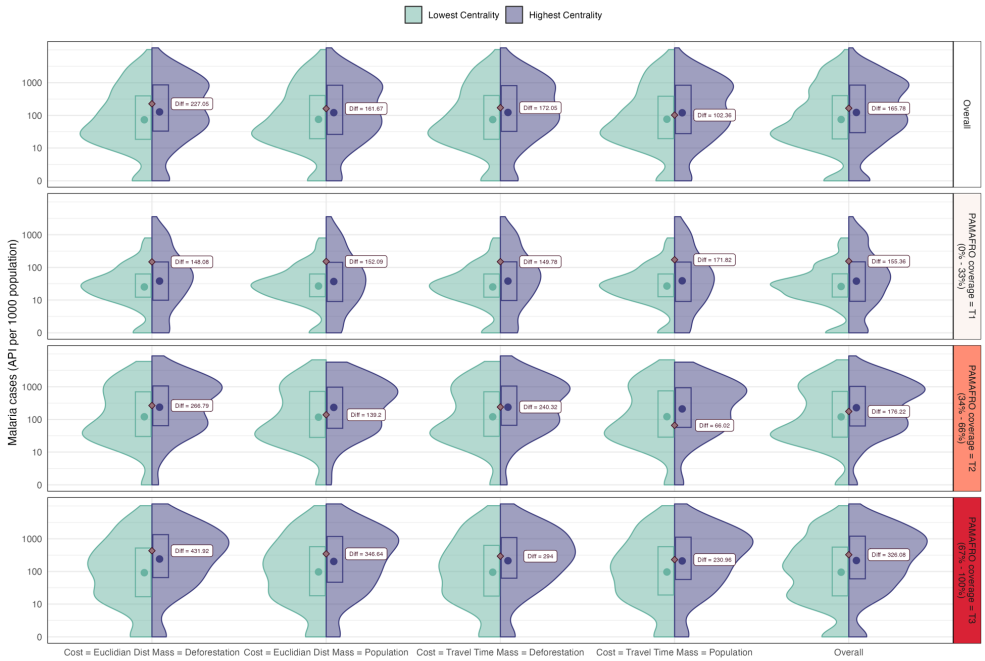


Supplementary Figure 2.4. Consensus graph of the network of villages by watersheds in the Loreto department in the Peruvian Amazon.

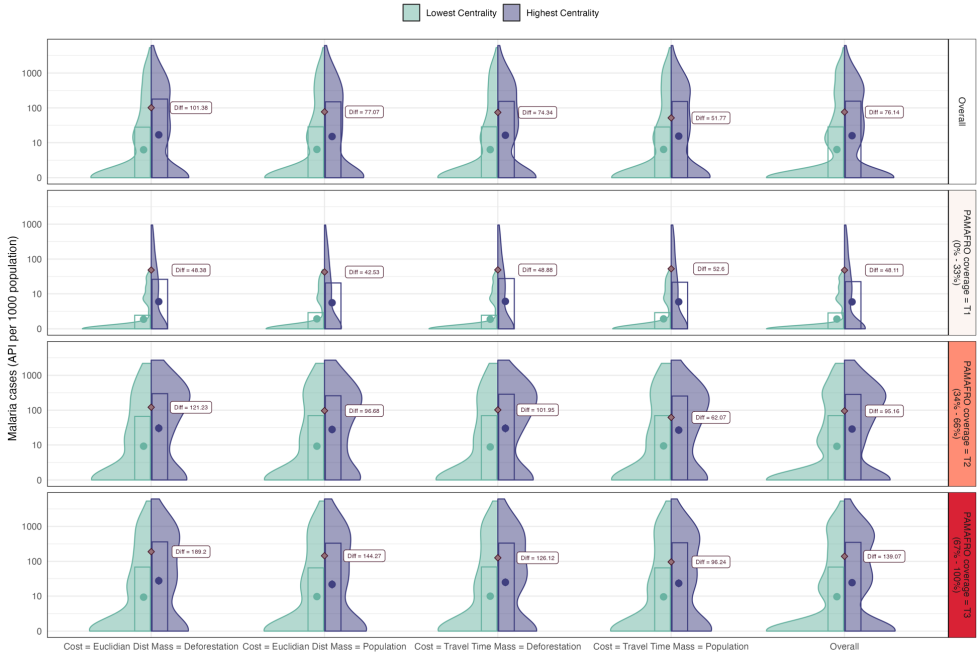


Supplementary Figure 2.5. Correlation of centrality metrics of villages in the Loreto department in the Peruvian Amazon. Margin plots shows the mass (population [pop; light orange], deforested area [adef; light blue], and none [red]), cost (distance [d; green], travel time [t; purple]), and centrality type (betweenness [between; blue], strength [stre; orange], eigen [eigen; pink], closeness [close; light green]) used for the calculations. Dendrograms to cluster centrality metrics are based on a hierarchical cluster analysis using a complete linkage method.

A)

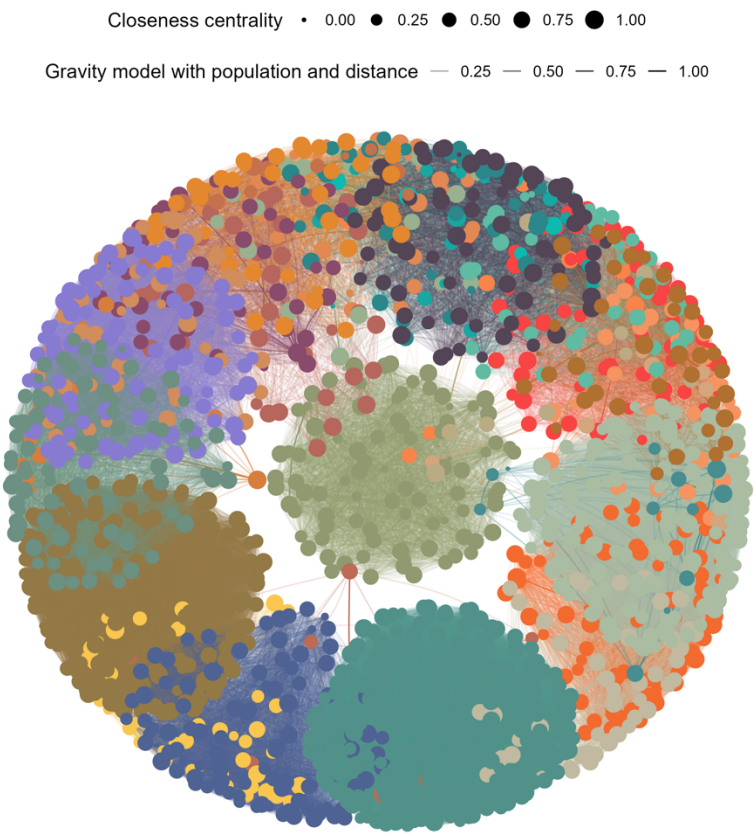


B)

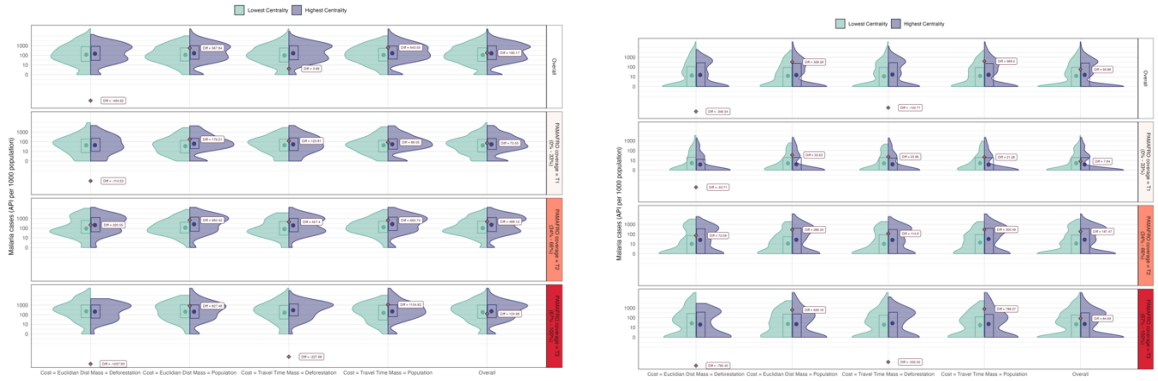


Supplementary Figure 2.6. Distribution of Total Annual Parasite Index (API) per 1,000 individuals (2011–2018) across high and low centrality villages stratified by levels of PAMAFRO intervention coverage and overall distributions in the Loreto department in the Peruvian Amazon. A) For *P. vivax* and B) for *P. falciparum.* Mean difference between groups are represented as diamonds in each panel. T1 = Tier 1 (Low coverage; Low baseline endemicity), T2 = Tier 2 (Moderate coverage; Moderate baseline endemicity), and T3 = Tier 3 (High coverage; High baseline endemicity).

A)

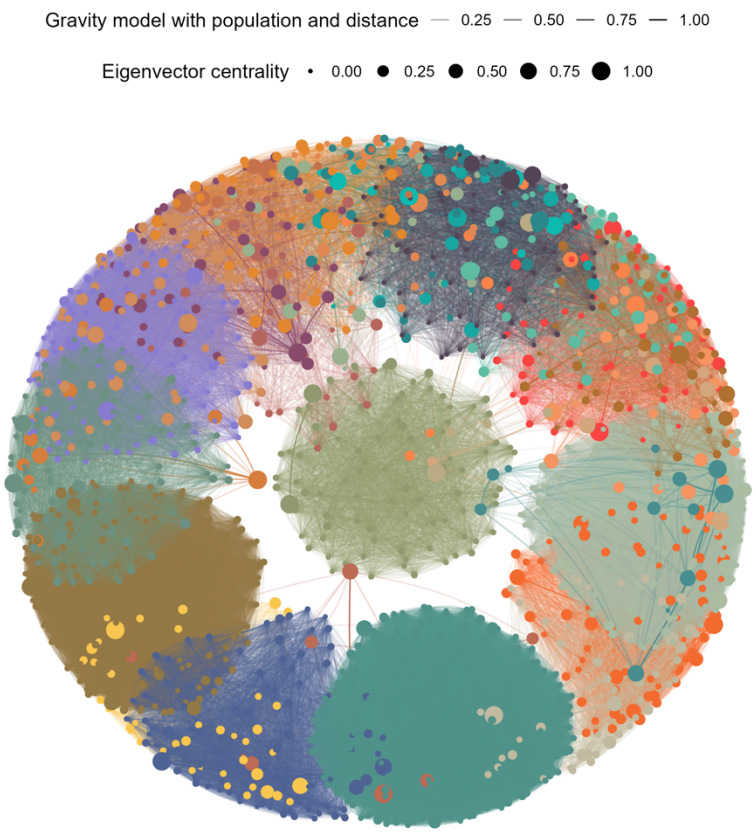


B) ***C)***

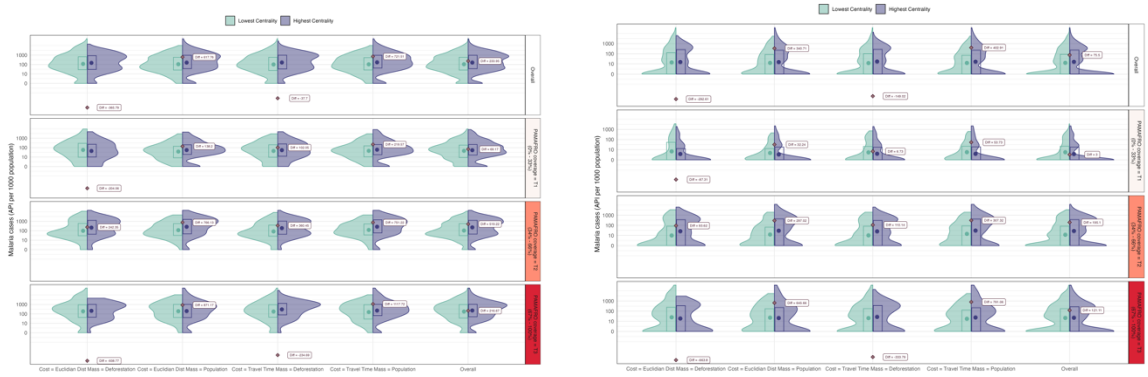


Supplementary Figure 2.7. A) Consensus graph of the network of villages in the Loreto department in the Peruvian Amazon using closeness centrality and Distribution of Total Annual Parasite Index (API) per 1,000 individuals (2011–2018) across high and low centrality villages stratified by levels of PAMAFRO intervention coverage and overall distributions in the Loreto department in the Peruvian Amazon. B) For *P. vivax* and C) for *P. falciparum.* Mean difference between groups are represented as diamonds in each panel. T1 = Tier 1 (Low coverage; Low baseline endemicity), T2 = Tier 2 (Moderate coverage; Moderate baseline endemicity), and T3 = Tier 3 (High coverage; High baseline endemicity).

A)

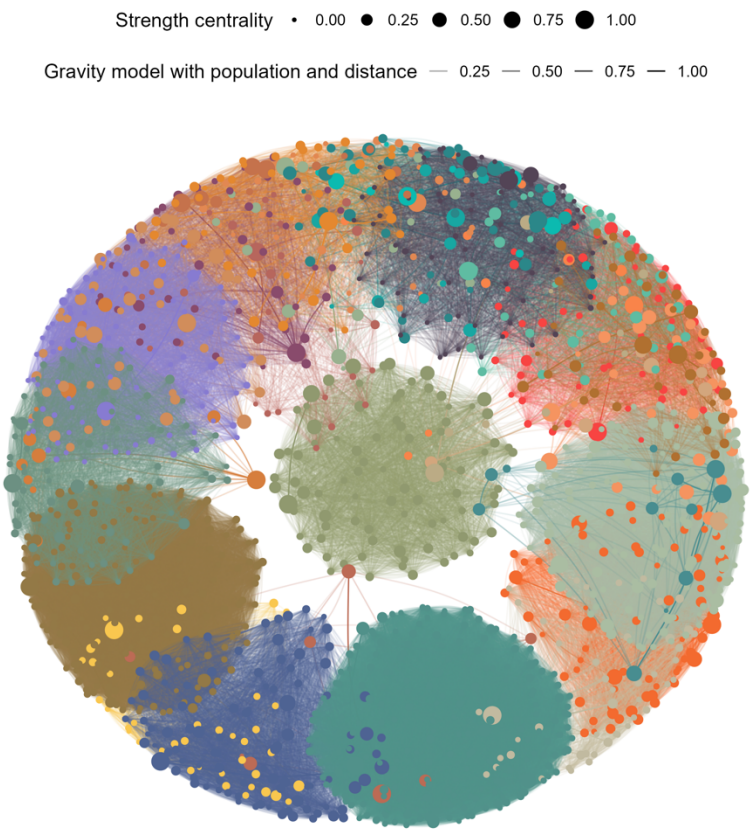


B) ***C)***

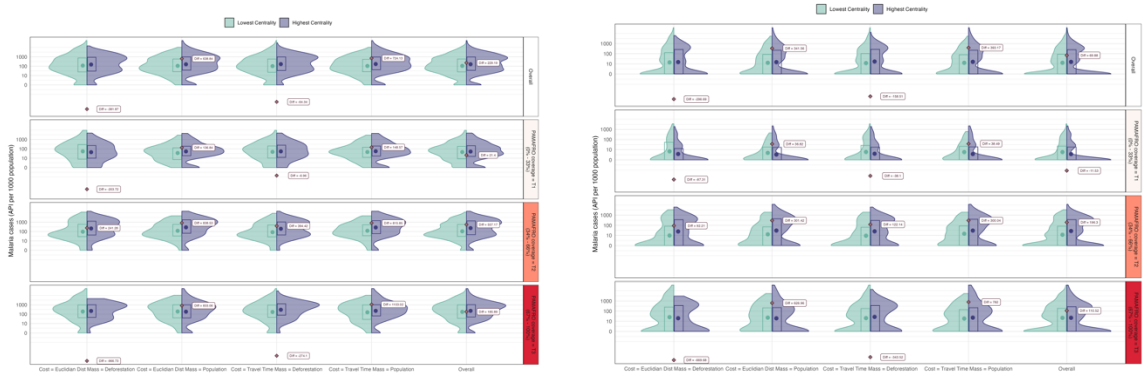


Supplementary Figure 2.8. A) Consensus graph of the network of villages in the Loreto department in the Peruvian Amazon using eigen centrality and Distribution of Total Annual Parasite Index (API) per 1,000 individuals (2011–2018) across high and low centrality villages stratified by levels of PAMAFRO intervention coverage and overall distributions in the Loreto department in the Peruvian Amazon. B) For *P. vivax* and C) for *P. falciparum.* Mean difference between groups are represented as diamonds in each panel. T1 = Tier 1 (Low coverage; Low baseline endemicity), T2 = Tier 2 (Moderate coverage; Moderate baseline endemicity), and T3 = Tier 3 (High coverage; High baseline endemicity).

A)



B) ***C)***



Supplementary Figure 2.9. A) Consensus graph of the network of villages in the Loreto department in the Peruvian Amazon using strength centrality and Distribution of Total Annual Parasite Index (API) per 1,000 individuals (2011–2018) across high and low centrality villages stratified by levels of PAMAFRO intervention coverage and overall distributions in the Loreto department in the Peruvian Amazon. B) For *P. vivax* and C) for *P. falciparum.* Mean difference between groups are represented as diamonds in each panel. T1 = Tier 1 (Low coverage; Low baseline endemicity), T2 = Tier 2 (Moderate coverage; Moderate baseline endemicity), and T3 = Tier 3 (High coverage; High baseline endemicity).

2.10. Acknowledgements

Chapter 2, in full, has been submitted for publication in Nature Communications. Gabriel Carrasco-Escobar, Diego Villa, Antony Barja, Rachel Lowe, Mercedes Pascual, Alejandro Llanos-Cuentas, Tarik Benmarhnia. The paper title is “The role of connectivity on malaria dynamics across areas with contrasting control coverage in the Peruvian Amazon”. The dissertation author was the primary researcher and author of this paper.

Chapter 3: Effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon using parametric g-formula

3.1. Abstract

In the Amazon Region of Peru, occupational activities are important drivers of human mobility and may increase the individual risk of being infected while contributing to increasing malaria community-level transmission. Even though out-of-village working activities and other mobility patterns have been identified as determinants of malaria transmission, no studies have quantified the effect of out-of-village working activities on recent malaria exposure and proposed plausible intervention scenarios. Using two population-based cross-sectional studies in the Loreto Department in Peru, and the parametric g-formula method, we simulated various hypothetical scenarios intervening in out-of-village working activities to reflect their potential health benefits. This study estimated that the standardized mean outcome (malaria seroprevalence) in the unexposed population (no out-of-village workers) was 44.6% (95% CI: 41.7 % - 47.5%) and 66.7% (95% CI: 61.6% - 71.8%) in the exposed population resulting in a risk difference of 22.1% (95% CI: 16.3% - 27.9%). However, heterogeneous patterns in the effects of interest were observed between peri-urban and rural areas (Cochran's Q test = 15.5, p-value <0.001). Heterogeneous patterns were also observed in scenarios of increased prevalence of out-of-village working activities and restriction scenarios by gender (male vs. female) and age (18 and under vs. 19 and older) that inform possible occupational interventions targeting population subgroups. The findings of this study support the hypothesis that targeting out-of-village workers will considerably benefit current malaria elimination strategies in the Amazon Region. Particularly, males and adult populations that carried out out-of-village working

activities in rural areas contribute the most to the malaria seropositivity (recent exposure to the parasite) in the Peruvian Amazon.

3.2. Introduction

The Amazon rainforest located in the World Health Organization (WHO) Region of the Americas remains a malaria hotspot. Within the 19 countries in this Region, more than 600,000 (presumed and confirmed) incident cases were estimated in 2020 [144] and 9 countries shared the Amazonian territory and most of the malaria cases: Bolivia, Brazil, Colombia, Ecuador, French Guiana (France), Guyana, Suriname, Venezuela, and Peru [144]. Malaria transmission in this area is dominated by *Plasmodium vivax* (75%) followed by *P. falciparum* and mixed (25%) infections. Most of these cases are located within 14 subnational units only [144] which include the Loreto Region located in the Peruvian Amazon. Historically, 93.1% of cases in Peru were reported in this Region [145] that are mainly transmitted by *Nyssorinchus* (*Anopheles) darlingi* [18,60].

In the last two decades, many interventions aiming at reducing the incidence of malaria in Peru have been implemented. For example, the PAMAFRO project (2005–2010) [146] focused on training community health workers for early diagnosis, monitoring, and treatment of malaria, the use of long-lasting insecticide-treated nets (LLINs), and community education in malaria prevention measures. Shortly after the interruption of the PAMAFRO project, the "Plan Malaria Cero" (PMC; 2017-2021) [147] was implemented with the aim to eliminate malaria transmission in three stages over a 25-year timeframe. More recently, a new "Plan Hacia la Eliminación de la Malaria en el Perú" (2022 – 2030) [148] aims to provide the legal, economic, and political support to achieve malaria elimination in Peru by applying a set of evidence-based interventions.

However, while most of these activities contributed to a reduction of malaria incident cases [23], many dimensions regarding malaria transmission, such as human mobility, are still not considered and constitute missed opportunities for alternative interventions to ultimately eliminate malaria. The WHO guidelines for elimination and prevention of reintroduction strategies recently highlighted the important role of human mobility as a challenge for sustaining malaria elimination efforts [149,150]. Human mobility was described as an associated factor for both malaria exposure [106] and infection [13] in the Amazon Region.

In the Amazon Region of Peru, an important driver of human mobility is related to occupational activities [74,98]. Indeed, many workers engage in out-of-village working activities in order to meet job opportunities and thus may increase their risk of being infected but also may contribute to increasing community-level transmission [13,74]. Previous studies have identified occupational mobility as a determinant for malaria risk and used forest goers [151–153] or out-of-village activities [13,141] as measures of such exposure but did not rely on causal modelling nor objectively measured exposures and malaria outcomes. Furthermore, no study simulated plausible scenarios to assess the potential benefits of hypothetical interventions.

The contribution of occupational determinants of infectious diseases has received more attention recently in the context of the COVID-19 pandemic [154] but evidence about which strategies targeting occupational mobility may be most effective at reducing malaria risk is lacking. Therefore, simulating the potential benefits of various intervention scenarios could be particularly helpful to design future occupational interventions to complement already implemented community-based actions to ultimately reach malaria elimination. Some modern

causal inference methods, including the parametric g-formula, have been proposed to flexibly estimate the effect of different exposure regimes. Parametric g-formula methods are a generalization of standardization methods that can simulate different hypothetical interventions on the exposure of interest [155]. While many studies have recently relied on such an approach, including in occupational settings [156–158], to simulate hypothetical interventions, such methods have been applied to a limited extent in the context of malaria epidemiology [159,160].

In addition, many population characteristics may modulate the effect of out-of-village activities on malaria risk and may inform targeted interventions. First, individual-level characteristics such as gender and age may constitute important effect modifiers. Yet, cultural, geographical, and social characteristics may also greatly differ between rural and peri-urban areas in the Peruvian Amazon. Previous studies found contrasting differences in the proportion of inhabitants that participate in out-of-village activities between rural and peri-urban areas in Iquitos [106]. Also, contrasting patterns were reported in malaria infection rate [13,141], seropositivity (exposure to previous infection) [106], and parasite genetic population structure [161,162] between these areas. In regions with rural-to-urban gradients, ecological factors increase disparities in malaria susceptibility [63] driven by marked variations in *Ny. darlingi* abundance and biting behavior across the forest, chacra (crop fields) (perturbed secondary forest), and urban settings [113,163].

Thus, this study aims to estimate the effect of out-of-village mobility on malaria exposure in contrasting geographic areas to better inform occupational interventions related to malaria elimination strategies in the Amazon Region. Using two population-based cross-sectional studies

in the Loreto Department in Peru, and g-formula methods, we simulated various hypothetical scenarios intervening in out-of-village working activities and various population subgroups to reflect the potential health benefits of future interventions.

3.3. Methods

3.3.1. Ethics

This study analyzed data from two studies that were approved by the Ethics Review Board of the Regional Health Directorate of Loreto and Universidad Peruana Cayetano Heredia in Lima: the Circles of Research on Arboviruses and Malaria (CAM) study (SIDISI 101645/2017) and the Amazonia International Center of Excellence in Malaria Research (ICEMR) study (SIDISI 101518/2018). Participants in both studies were enrolled upon signing an informed consent or informed assent in case of participants under 18 years old. All the methods were carried out in accordance with the approved guidelines.

3.3.2. Study design

We conducted etiological and simulation studies to quantify the role of out-of-village working activities on recent malaria exposure in two population-based cross-sectional studies carried out in the Loreto Department, Peru. The designs of both studies were described elsewhere [164]. Briefly, both studies were conducted by the same research team in different months in 2018. A structured questionnaire, georeferencing of households, and blood samples were collected in 10 villages in two districts of Loreto: Iquitos – mostly urban – in April 2018, and Mazán – mostly rural – in July 2018. Here, the –previously reported [164] – seropositivity status of the

participants (based on a random forest classifier) was used in combination with a parametric g-formula (see details below) to compute the average causal effect of out-of-village mobility on malaria exposure. In addition, we simulated multiple scenarios of mobility restrictions (by proportion of travelers, gender, and age) to estimate the impact of such restriction policies in reducing malaria exposure in the Peruvian Amazon.

3.3.3. Study site and population

High-risk malaria villages were selected in peri-urban and rural areas based on Ministry of Health (MoH) historical data (Figure 3.1A). Three villages were selected in the peri-urban area: Rumococha (RM), Santo Tomás (ST), and Quistococha (QC). These villages are located on the outskirts of Iquitos district, 10 km from Iquitos City (capital of Loreto; lat: 03°44.591′S, long: 73°19.615′W), accessible by road and highly deforested. Seven villages were selected in the rural area: Gamitanacocha (GC), Libertad (LB), Primero de Enero (PE), Puerto Alegre (PA), Salvador (SL), Lago Yuracyacu (LY), and Urco Miraño (UM). These villages are located in the Mazán district, accessible only by boat (~2-7 h from Iquitos city) and characterized by dense primary and secondary forest cover. All participants 6 months or older at the date of survey were invited to the study if they lived in the selected village and gave consent to donate a blood sample by venipuncture for malaria diagnosis.

3.3.4. Data collection and variable definitions

A full census of the study populations was conducted in April-July 2018. Individual and household data on socio-demographics (age, gender, education, occupation), self-reported previous history of clinical malaria, and structural characteristics of the household were

collected. All households and participants were encoded and geo-referenced using a Global Positioning System (GPS) handheld device (Garmin's GPSMAP 60CSx, Garmin International Inc., USA).

A blood sample of 6 mL for adults or 3 mL for children of whole blood was collected by venipuncture in tubes with EDTA (BD Vacutainer, BD Franklin Lakes, USA) as a preservative. Venipuncture blood samples were separated by centrifugation (3500 rpm) into plasma and packed red blood cells (PRBC) for serological analysis.

The primary exposure – out-of-village working activities – and covariates were collected in structured questionaries. All villagers self-reported whether they traveled in the previous month (travel history), sex, age, and occupation. All occupational activities were grouped into a binary variable according to the location where the activities were carried out (inside or outside their home village). Previous studies identified that out-of-village working activities in these areas include logging, hunting, fishing, trading, and farming [165,13,98].

The primary outcome –malaria serological exposure– was defined according to a serological assay that target *Plasmodium* species-specific levels (recent infection up to 9 months in the past) [166]. IgG antibody responses to 8 serological exposure markers (SEM) to *P. vivax* were measured using a Luminex® platform, as described elsewhere [167]. The 8 SEM panel has been previously validated [166] and consisted of the following proteins: PVX_099980 (19 kDa C-terminal region of merozoite surface protein 1, PvMSP119), PVX_096995 (tryptophan-rich antigen, Pv-fam-a, PvTRAg_2), PVX_112670 (PvTRAg_28), PVX_097625 (merozoite surface

protein 8, putative, PvMSP8), PVX_097720 (merozoite surface protein 3, PvMSP3.10), PVX_087885 (rhoptry-associated membrane antigen, putative, PvRAMA), PVX_094255 (reticulocyte binding protein 2b, PvRBP2b) and KMZ83376.1d (erythrocyte-binding protein II, PvEBPII). To normalize and diminish inter-plate variation, a standard curve was prepared using a plasma pool of hyper-immune adults from Papua New Guinea. Relative Antibody Units (RAU) or dilutions were obtained by extrapolating the Median Fluorescence Intensity (MFI) in a standard curve by a 5 parameters logistic model. Seropositivity to each marker was defined by using a Random Forests based classification algorithm previously validated in low *P. vivax* transmission contexts [166]. Further description of the serological makers and the measured structure of the transmission in the area could be found elsewhere [164].

3.3.5. Estimating the average causal effect of out-of-village working activities on malaria

To estimate the average causal effect of out-of-village working activities on malaria a parametric g-computation described previously [155,158,168–171] was used. G-formula (also known as g-computation) can be seen as a generalization of standardization methods applied to multiple settings and first described in 1986 [172]. In the g-formula, under identification assumptions such as exchangeability, consistency and positivity conditional on the variables in L (potential confounders), the standardized mean outcome is the weighted average of the conditional means using as weights the prevalence of each stratum l of the vector of confounders L in the study population computed as follows:

$$\sum_{l} E[Y|A = a, L = l] \times Pr[L = l]$$

where $E[Y|A = a, L = l]$ are the conditional means in each of the strata l and $Pr[L = l]$ is the prevalence of l. Such quantities are estimated parametrically. The following 4-step process

was adopted. First, expansion of the original dataset; a new set of analytic datasets was created by repeating the original dataset in three blocks. The first block was identical to the original dataset, the second block was modified and set the values of A (of out-of-village working activities) to unexposed ($A=0$), the third block was modified and set the values of A to exposed ($A=1$). In the second and third blocks, the values of the outcome (Y – malaria exposure) were removed and set as missing. Second, a regression model (a modified Poisson regression [173] to consider the highly prevalent outcome) was fitted for the outcome (i.e. malaria) given exposure A (out-of-village working activities) and confounders L (including villages as fixed effects). The variables used for the model estimation were age, sex, education, and fever history. The final model included interactions between the main exposure and age and sex. It is worth mentioning that only data in the first block contributed to the estimation (as Y was absent from the created blocks). Third, the parameters estimated using data from the first block were used to predict the outcome values for all observations in the second and third blocks, which standardizes based on the empirical distribution of confounders. The average of all predicted values in the second and third block is precisely the standardized mean outcome in the unexposed and exposed, respectively. Finally, risk differences and ratios can be estimated by comparing such estimated counterfactual quantities. To obtain 95% Confidence Intervals (CI), a Monte Carlo resample with 999 replicates was drawn with replacement from the original data. These analyses were further explored by stratifying by age, proportion of travelers, and gender as well as location in peri-urban or rural settings.

3.3.6. Simulation of restriction scenarios

We then conducted a series of simulations to (synthetically) modify the prevalence of the main exposure (out-of-village working activities) while keeping the confounding structure, to explore scenarios where the main exposure would vary and compared to the natural course (i.e., the initial/observed setting or said differently, in the absence of any interventions) to inform future interventions. We tested scenarios of the prevalence of out-of-village working activities ranging from 0 to 1 through incremental steps of 0.1. We stratified our analyses by peri-urban or rural settings. In addition, 4 scenarios were tested based on full (FE) and null (NE) exposure in relation to gender (male vs. female) and age (18 and under vs. 19 and older) to inform possible occupational interventions targetting population subgroups.

3.4. Results

3.4.1. Baseline characteristics

A total of 785 individuals from 421 households were enrolled in the Iquitos district (peri-urban setting) and 1005 individuals from 419 households in the Mazán district (rural setting). The village sample size range between 250 and 273 individuals in peri-urban settings and between 47 and 270 individuals in rural areas. The average age of the population was 27.7 (SD = 22.2) years in the rural area and 30.6 (SD = 21.8) years in the peri-urban area. Important differences were observed in the proportion of females (59% vs 51%), secondary or superior education (46% vs 27%), work inside the village (88% vs 59%), and travel in the last month (2% vs 33%) between peri-urban and rural settings.

3.4.2. Malaria seroprevalence rate

An overall seroprevalence rate of 49% was observed in the study population. However important differences were observed across settings (Table 3.1). A higher seroprevalence was observed in rural areas (57%) in comparison to peri-urban areas (39%). Importantly, a seroprevalence rate higher than 40% was observed in 6 (GC, LB, PE, SL, PA, and UM) out of the 7 rural villages and only in 1 (ST) out of the 3 peri-urban villages. The highest seroprevalences were observed in GC (87%), LB (75%), and PE (69%), all located in the rural district of Mazán. On average, the age of seropositive participants is higher (37 years old) in comparison to seronegative participants (21 years old). Slight differences in the seropositivity status were observed in relation to gender and education, however, contrasting patterns were observed in relation to outside (77%) in comparison to inside (38%) workers and recent travelers (66%) in comparison to no travelers (45%).

The spatial distribution of the seropositivity rates is shown in Figure 3.1A and Supplementary Figure 3.1. A clustered pattern at the household level was observed in both study settings. Out of the 421 households surveyed in the peri-urban area, a seropositive individual was detected in 290 (68%) households, ranging from 55% to 82% at the village level. Conversely, in the rural setting, a seropositive individual was detected in 396 (94%) out of 419 households. Remarkably, at least one seropositive participant was detected in all households in GC and PE.

The seroprevalence rates were further explored across age categories and different socioeconomic variables in Figure 3.1B. In addition to the overall higher malaria exposure (seroprevalence) in rural than peri-urban areas, the age breakdown showed contrasting patterns

between these areas. A smoother increase in the age-seroprevalence trend was observed in peri-urban areas. In contrast, an abrupt disruption at age 15 was observed in rural areas. It is important to notice that the age composition is different between rural/peri-urban areas (distinguished by the width of the bars), much younger in rural areas. Most noticeable differences in the age-structure and age-seroprevalence trend were observed between type of activities (inside/outside village) and recent travelers. Out-of-village working activities were carried out by the older population (>30 years old) that showed high seroprevalence rates (>60%). A similar pattern was observed for recent travelers, most of them were adults (> 30 years old) with very high seroprevalence rates (>80%). The frequency distribution of out-of-village working activities and seropositivity status is shown in Supplementary Figure 3.2 and spatial distribution of household work out-of-village rate spatial distribution in the villages in the study area is shown in Supplementary Figure 3.3.

3.4.3. Average causal effect estimation

Using a parametric g-computation, this study estimated that the standardized mean outcome (malaria seroprevalence) in the unexposed population (i.e. if all participants do not carry out out-of-village working activities) is 44.6% (95% CI: 41.7 % - 47.5%) and 66.7% (95% CI: 61.6% - 71.8%) in the exposed population (i.e. if all participants carry out out-of-village working activities) (Figure 3.2). The role of out-of-village working activities on recent malaria exposure (seroprevalence) was estimated as the difference between these quantities (standardized mean outcome) in the exposed and unexposed. This results in an important and precise average causal effect with a risk difference of 22.1% (95% CI: 16.3% - 27.9%).

However, contrasting patterns in standardized mean outcomes and the average causal effects were observed between peri-urban and rural areas. In peri-urban settings, similar standardized mean outcomes were estimated among exposed (42.5%; 95% CI: 29.7% - 55.4%) and unexposed (38.6%; 95% CI: 35.0% - 42.1%) resulting in no detected differences in malaria risk associated with out-of-village working activities. In the rural areas, the standardized mean outcomes among exposed was 78.4% (95% CI: 72.7% - 84.1%) and 47.7% (95% CI: 43.2% - 52.1%) among unexposed, resulting in an average causal effect of 30.7% (95% CI: 23.8% - 37.6%). Significant heterogeneity in the average causal effect was observed between peri-urban and rural areas (Cochran's Q test = 15.5, p-value <0.001).

3.4.4. Restriction scenarios

Multiple scenarios were tested to inform policy making by simulating the prevalence of out-of-village working activities. Overall, the observed prevalence of out-of-village working activities was 28.6% (Table 3.1) and the estimated standardized mean outcome (malaria seroprevalence) was 48.8% (95% CI: 46.4% - 51.1%) (Figure 3.2). After manipulating (by simulation) the prevalence of the exposure (herein referred to as simulated exposure – SE) and computing the corresponding standardized mean outcome, a dose-response curve was constructed for overall, peri-urban, and rural areas (Figure 3.3). The average causal effect is –in consequence– the difference between the standardized mean outcome at both extremes of these dose-response curves (0% exposed vs. 100% exposed). Further explorations were conducted by comparing the SE against no (0%) exposure (NE) and the natural course (NC) in each geographic area (Figure 3.3). The main role of out-of-village working activities on malaria

seroprevalence in rural in comparison to peri-urban areas was further depicted by comparing the dose-response curves between these areas.

Further scenarios were explored based on gender and age travel restriction policies. The standardized mean outcome when simulating a full exposure (FE – 100% prevalence of out-of-village working activities) was 57.3% (95% CI: 53.6% - 61.0%) in males and 58.2% (95% CI: 54.8% - 61.6%) in females (Figure 3.4A). In contrast, the standardized mean outcome when simulating the NE was 45.8% (95% CI: 43.0% - 48.6%) and 47.6% (95% CI: 45.0% - 50.2%) in males and females, respectively. Overall, a slightly greater impact was observed when restriction policies targeted males. The average causal effect in males was 11.5% (95% CI: 8.0% - 15.0%) and 10.5% (95% CI: 7.8% - 13.3%) in females. Importantly, this type of policy is most effective in rural areas (Figure 3.4A). The average causal effect is 3- and 15-folds higher in rural than peri-urban areas, in females and males respectively.

A contrasting pattern was observed when simulating travel restrictions based on legal adult age (18 years old) in Peru. The standardized mean outcome when simulating the FE was greater in adults –18 years old and older– (63.8%; 95% CI: 59.6% - 68.0%) than in children and adolescents –17 years old and under– (51.7%; 95% CI: 48.7% - 54.7%) and (Figure 3.4B). In addition, a pronounced reduction in the standardized mean outcome was observed in adults (44.9%; 95% CI: 42.0% - 47.9%) in comparison to children and adolescents (48.5%; 95% CI: 46.0% - 50.9%) when simulating the NE. Overall, a greater impact was observed when restriction policies were targeted at adults than children and adolescents. The average causal effect in adults was 18.8% (95% CI: 14.2% - 23.5%) and 3.2% (95% CI: 1.5% - 5.0%) in

children and adolescents. As previously observed, this policy scenario (targeting mobility restrictions to adults) is only effective in rural than peri-urban areas (Figure 3.4B). In rural areas, we identified an average causal effect of 27.9% (95% CI: 21.5% - 34.3%) while no effect in peri-urban areas.

3.5. Discussion

Despite numerous studies highlighting the links between human mobility and malaria in Amazonian contexts, the quantification of occupational-driven mobility was lacking. Using two population-based studies we determined the average causal effect of out-of-village working activities on malaria seropositivity (recent exposure to the malaria parasite). This study highlighted the critical role of human population mobility in sustaining malaria transmission in the Peruvian Amazon. By simulating the prevalence of out-of-village working activities to reflect different policy scenarios, this study showed the importance of targeting key subpopulations when designing such occupational interventions. Particularly, targeting males and adult (18 years old and older) populations causes the greatest effect on malaria seropositivity. Finally, in all these scenarios, the effect is highly pronounced in rural in comparison to peri-urban areas. The findings of this study are substantial to tailor current and future malaria elimination programs in the Amazon Region.

The role of human population mobility is of particular importance under elimination and prevention of reintroduction frameworks [149,150,174]. In areas where malaria transmission is heterogeneous –such as the Peruvian Amazon–, human mobility increases the importation risk (formerly known as vulnerability) [63]. Multiple mechanisms originate different mobility patterns

[61,65,66,68] as described in Africa [83,91], Southeast Asia [94], and more recently in Latin America [74,100]. Out-of-village working activities are central in the Peruvian Amazon since it is the most frequent reason for human mobility [98]. However, as previously described [15,61,175,176], subnational and local approaches should be considered since human mobility and malaria dynamics are tightly related micro-geographical and local contexts. In this study, we estimated the contrasting effect of a set of policy scenarios in rural and peri-urban areas. Importantly, both settings are located in contiguous districts (administrative level 3). This emphasizes the importance to adopt flexible malaria elimination approaches since a variety of scenarios could be found at neighbor subnational levels.

The goal of this study was to simulate the benefits of a new set of interventions by focusing solely on hypothetical interventions linked to out-of-village working activities. We recommend that future directions in malaria elimination research and policy would quantify other historical interventions based on pharmacological, environmental, and social/lifestyle factors to define a cost-effective set of interventions to achieve local malaria elimination goals. The hypothetical set of interventions tested in this study does not intend to suggest limiting the mobility of habitants in the Amazon Region, rather, intends to highlight their key role in sustaining malaria transmission. In consequence, our findings emphasized the urge to design tailored interventions for subpopulations that contribute the most to the malaria exposure such as males and adults in rural areas. Based on previous experiences [146], community health workers (CHW) may play a key role in deploying such kind of targeted strategies.

This study concludes that out-of-village working activities potentially encompass a wide range of mobility patterns that, if correctly identified, may help to enhance targeted interventions such as screening or surveillance strategies. Furthermore, recent studies leverage detailed GPS data to identify where people go and spend their time, allowing them to obtain accurate measurements of what they are exposed to within their activity spaces [177]. Studies in Southeast Asia [94] and Latin America [74] showed the interaction between travel/commuting patterns and land coverage as a main driver of malaria endemicity. Importantly, *Ny. darlingi* (dominant malaria vector) demonstrates an increased exophagic –outdoors– biting behavior [8,163,178] and a breeding site preference in the forest fringes [179–181] in rural Amazon. Taken together, if both environmental and health policies are combined, it is hypothesized that amplified impacts in both fields can be achieved [136].

We acknowledge the following limitations in this study. First, as a cross-sectional, this study is not designed to infer malaria transmission intensity. In this case, the seroprevalence reflects recent exposure to malaria infection (up to 9 months in the past) [164,166], however no active infection data was used. Despite other studies demonstrating that malaria seroprevalence is a good proxy for malaria transmission intensity [106,182,183], further longitudinal studies, to deal with potential regressions to the mean issues, are suggested to determine the causal effect on malaria transmission intensity. Second, personal protective measures (i.e., seasonal mobility, the use of bed nets or pharmacological prophylaxis) may play a key role in effect modification. Furthermore, besides the inclusion of these behaviors, considering the timing during transit or return of out-of-village activities may be important to consider for future studies. Finally, a potential threat to causal identifiability in this study may arise from the fact the outcome

(secondary infection) in one individual may be dependent on the outcome (primary infection) in other individuals (in other words, it would violate the stable unit treatment value assumption) [184]. Given the design of this study, the main assumption relies on that the main outcome (recent malaria exposure) is independent across study participants (i.e., no interference). Further longitudinal studies including GPS data may explicitly determine the interactions (matrices) between primary and secondary infections to estimate causal effects that are conditional on contact with an exposed individual [185,186].

3.6. Conclusion

The findings of this study support the hypothesis that targeting out-of-village workers will considerably benefit current malaria elimination strategies in the Amazon Region. Particularly, males and adult populations that carried out out-of-village working activities in rural areas contribute the most to the malaria seropositivity (recent exposure to the parasite) in the Peruvian Amazon. This study contributed to designing a new set of interventions that will potentially prevent one-third of recent malaria exposures. An optimal set of interventions to achieve malaria elimination goals in Amazonia should be driven by further exploring the causal effects of innovative and traditional policies.

3.7. Tables

Table 3.1. Baseline characteristics of the study population and their malaria seropositive status.

	Negative (N=917)	Positive (N=873)	Overall (N=1790)
Area			
Peri-urban	481 (52.5%)	304 (34.8%)	785 (43.9%)
Rural	436 (47.5%)	569 (65.2%)	1005 (56.1%)
Age			
[0,5]	147 (16.0%)	31 (3.6%)	178 (9.9%)
(5,15]	405 (44.2%)	163 (18.7%)	568 (31.7%)
(15,30]	148 (16.1%)	151 (17.3%)	299 (16.7%)
(30,50]	126 (13.7%)	262 (30.0%)	388 (21.7%)
(50,117]	91 (9.9%)	266 (30.5%)	357 (19.9%)
Gender			
Female	536 (58.5%)	437 (50.1%)	973 (54.4%)
Male	381 (41.5%)	436 (49.9%)	817 (45.6%)
Education			
Primary or less	582 (63.5%)	601 (68.8%)	1183 (66.1%)
Secondary or Superior	335 (33.9%)	272 (31.2%)	607 (33.9%)
Work type			
Inside village	798 (87.0%)	480 (55.0%)	1278 (71.4%)
Outside village	119 (13.0%)	393 (45.0%)	512 (28.6%)
Travel in the last month *			
No	794 (86.6%)	642 (73.5%)	1436 (80.2%)
Yes	120 (13.1%)	229 (26.2%)	349 (19.5%)
Microscopy result*			
Negative	894 (97.5%)	835 (95.6%)	1729 (96.6%)
Positive	12 (1.3%)	26 (3.0%)	38 (2.1%)
Fever			
No	900 (98.1%)	864 (99.0%)	1764 (98.5%)
Yes	17 (1.9%)	9 (1.0%)	26 (1.5%)

* variable with missing data

3.8. Figures

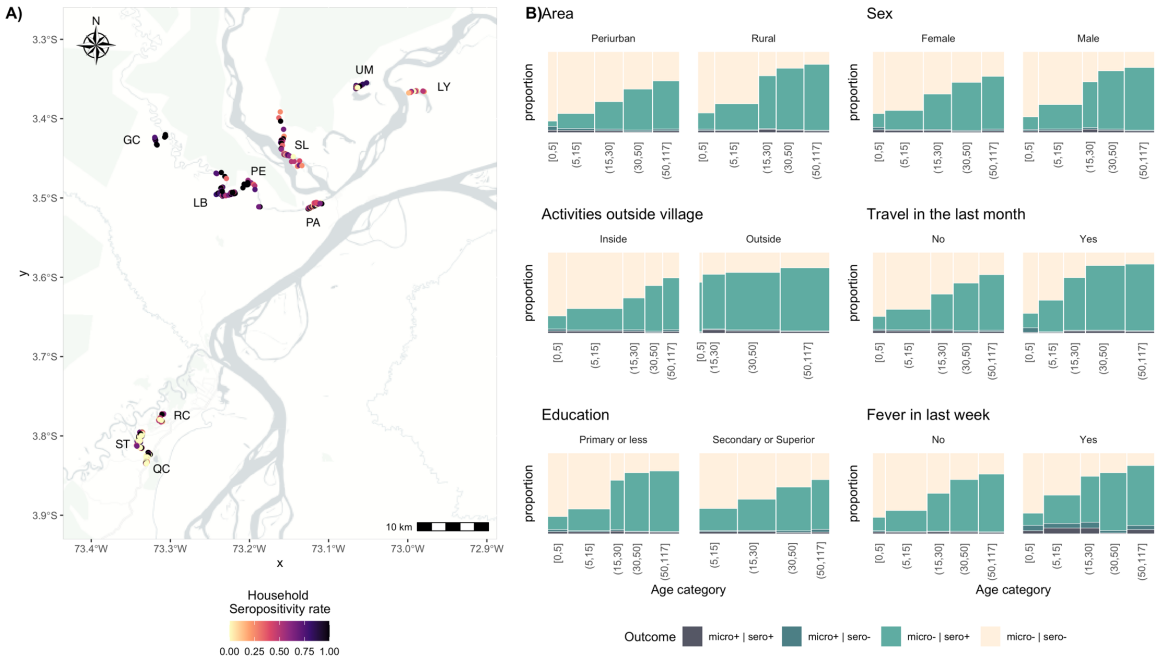


Figure 3.1. Study area and socio-demographic distribution of seropositivity in the Loreto department in the Peruvian Amazon. A) Seropositivity rate at household level in the villages of Iquitos district (bottom left): Rumococha (RC), Santo Tomas (ST), Quistococha (QC), and Mazán district (top right): Gamitanacocha (GC), Libertad (LB), Primero de Enero (PE), Puerto Alegre (PA), Salvador (SL), Lago Yuracyacu (LY), and Urco Miraño (UM). B) Distribution of serology (sero) and microscopy (micro) rates across age categories and sociodemographic variables. Maps were produced using R v.4.1 (R Development Core Team, R Foundation for Statistical Computing, Australia) based on public geographic data extracted from OpenStreetMap contributors (www.openstreetmap.org) under Open Data Commons Open Database License (ODbL) 1.0 (http://openstreetmap.org/copyright).

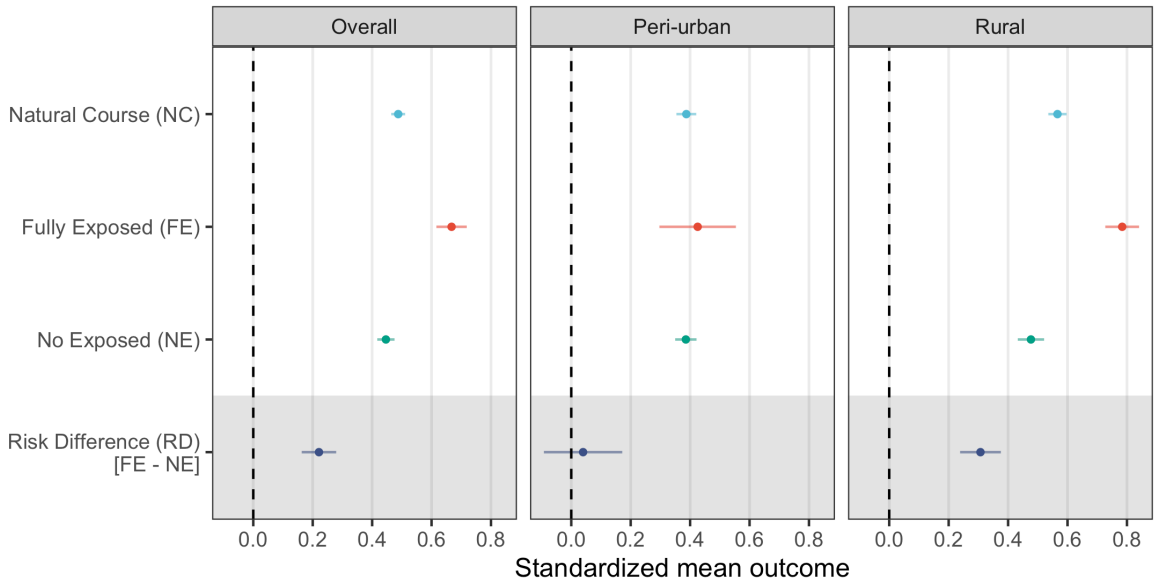


Figure 3.2. Summary of g-computation estimates by geographic area in the Loreto department in the Peruvian Amazon. Standardized mean outcome estimations for the natural course (NC) and simulated scenarios of full exposure (FE), no exposure (NE) and their risk differences (RD, grey area).

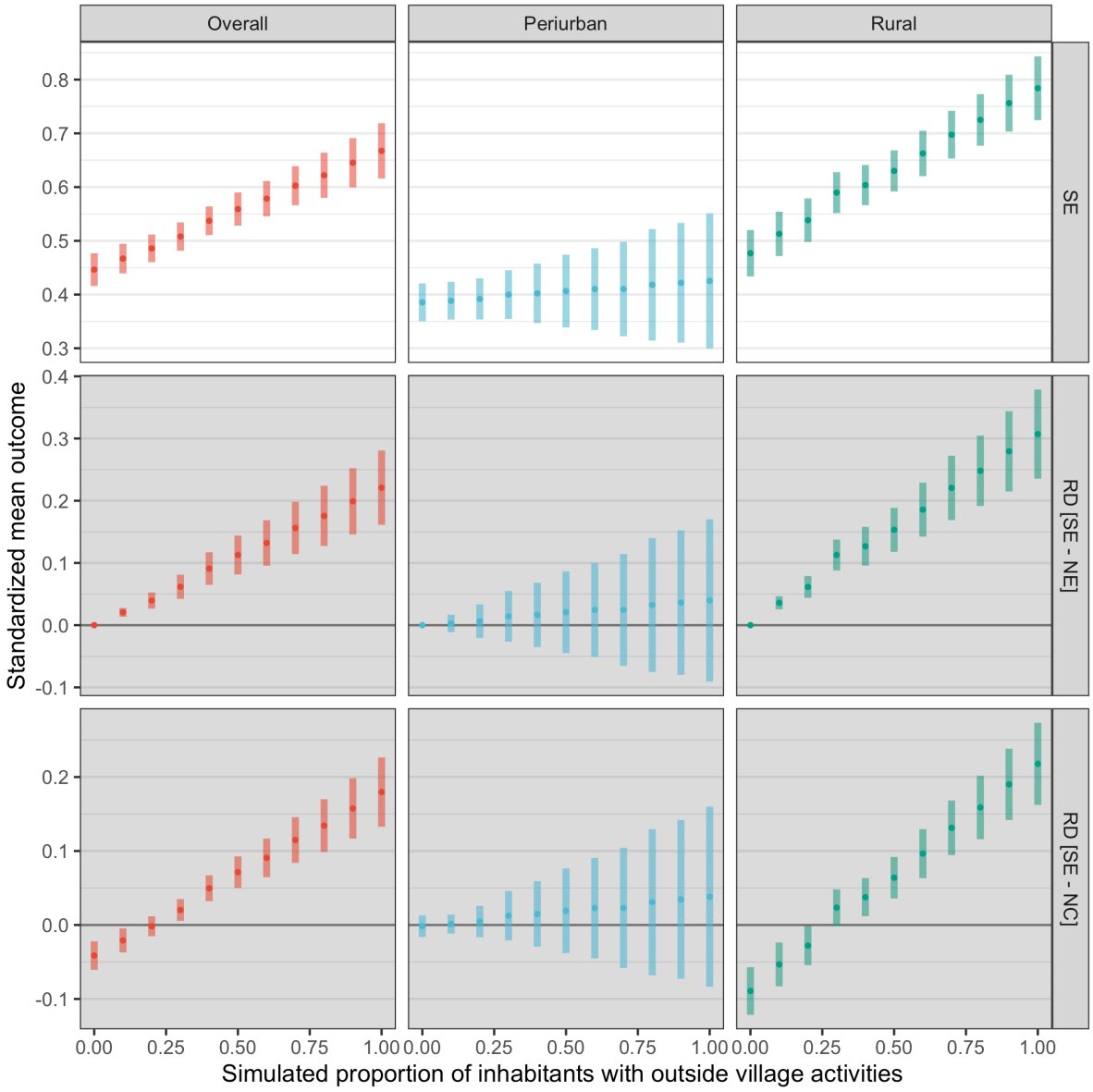


Figure 3.3. Standardized mean outcome estimated under simulated scenarios by proportion of travelers and geographic area in the Loreto department in the Peruvian Amazon. Standardized mean outcome estimations for the simulated exposure (SE; level of out-of-village working activities) and its risk difference (RD, grey panels) in comparison to no exposure (NE) and the natural course (NC).

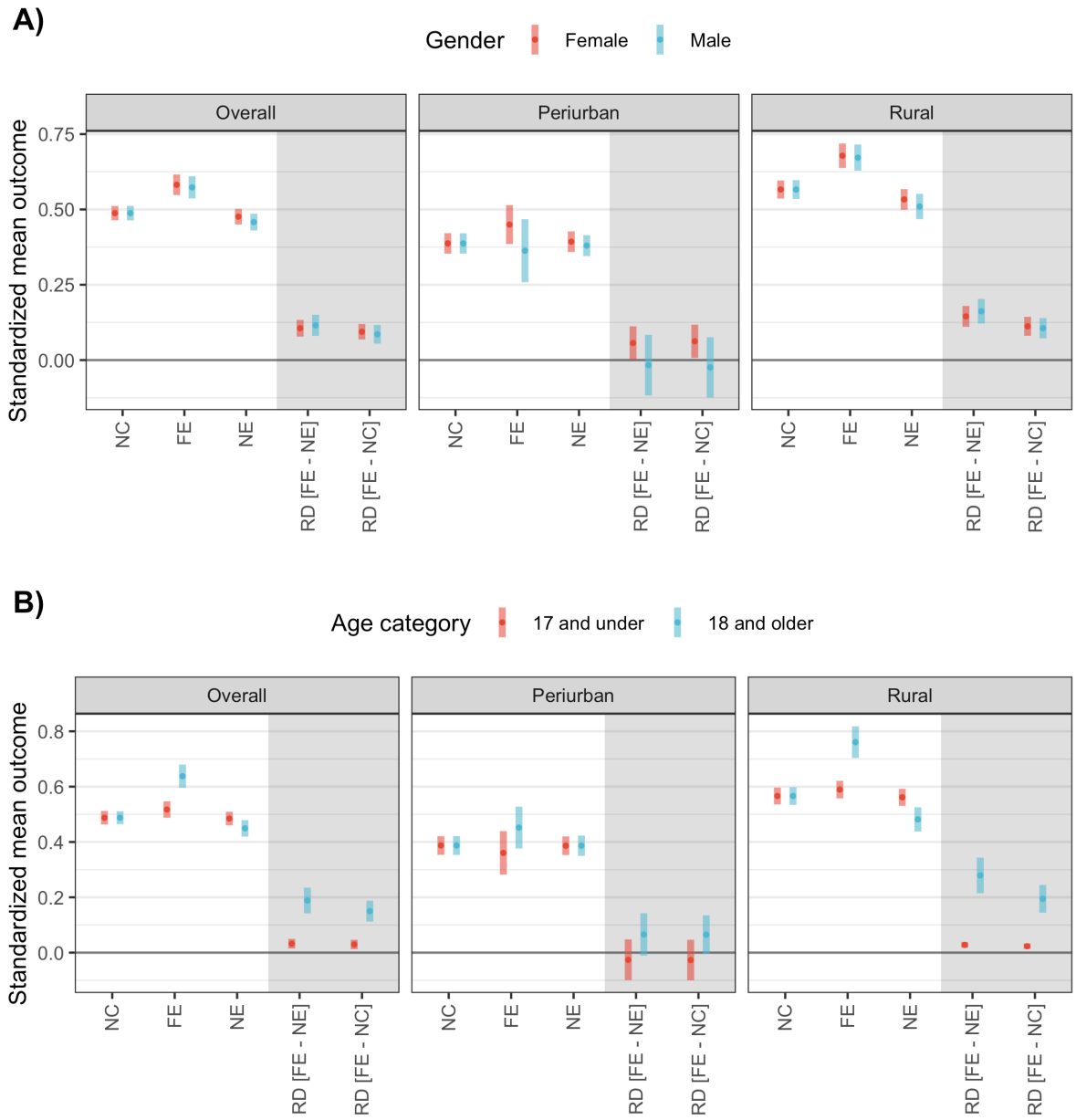
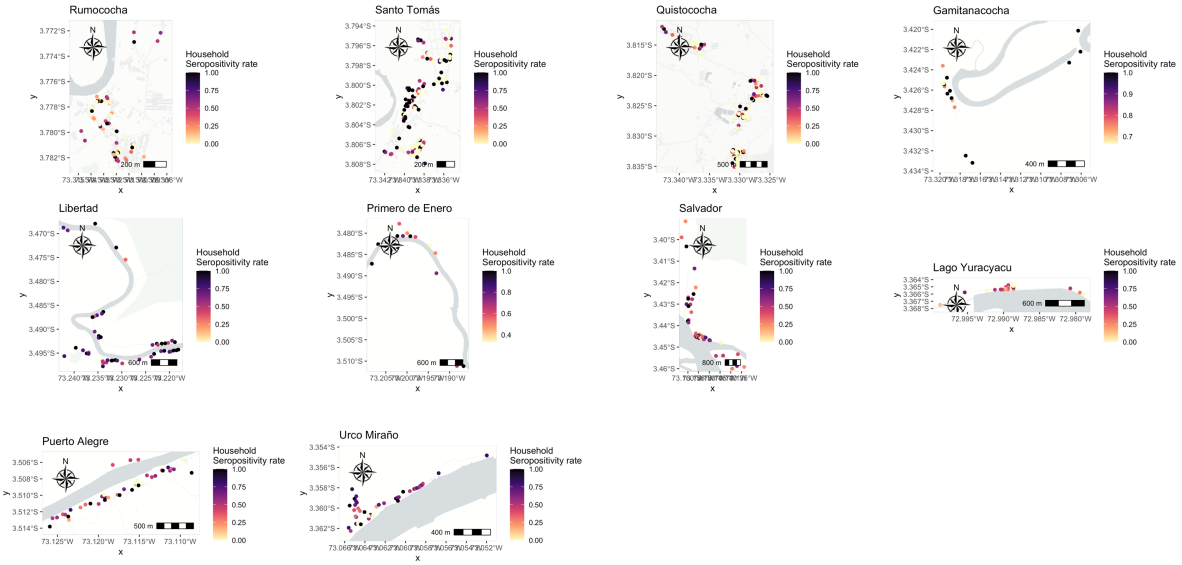
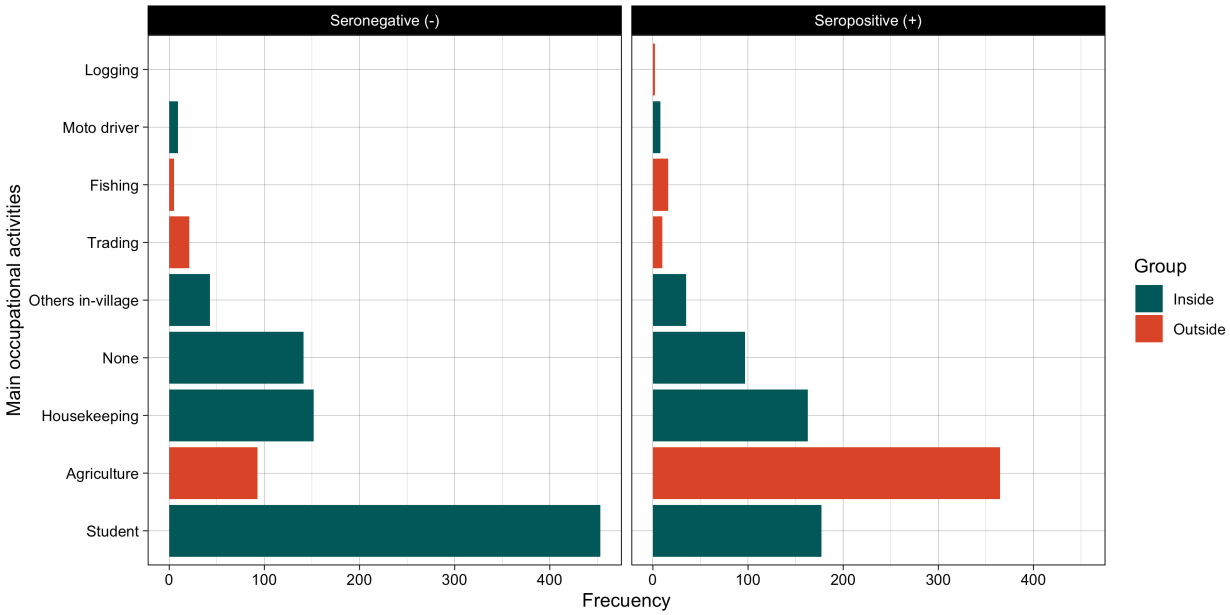


Figure 3.4. Standardized mean outcome estimated under simulated scenarios by geographic area and A) gender and B) age in the Loreto department in the Peruvian Amazon. Standardized mean outcome estimations for the natural course (NC) and simulated scenarios of full exposure (FE), no exposure (NE) and their risk differences (RD, grey area).

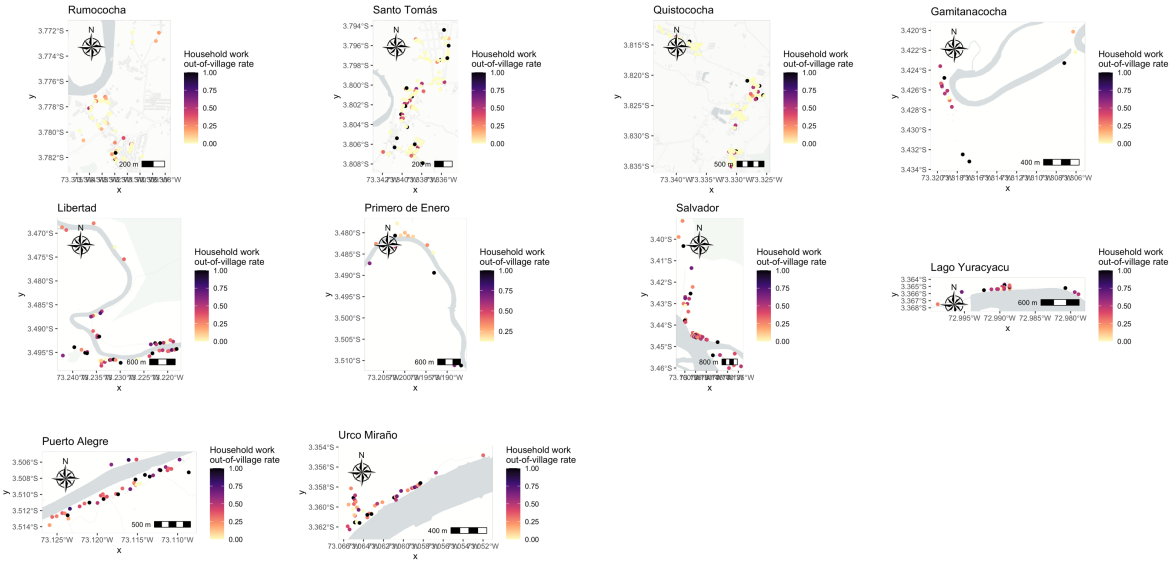
3.9. Supplementary information



Supplementary Figure 3.1. Household seropositivity rate spatial distribution in villages in the study area in the Loreto department in the Peruvian Amazon. Maps were produced using R v.4.1 (R Development Core Team, R Foundation for Statistical Computing, Australia) based on public geographic data extracted from OpenStreetMap contributors (www.openstreetmap.org) under Open Data Commons Open Database License (ODbL) 1.0 (http://openstreetmap.org/copyright).



Supplementary Figure 3.2. Frequency of main occupational activity grouped by out-of-village and in-village activities and seropositivity status.



Supplementary Figure 3.3. Household work out-of-village rate spatial distribution in villages in the study area in the Loreto department in the Peruvian Amazon. Maps were produced using R v.4.1 (R Development Core Team, R Foundation for Statistical Computing, Australia) based on public geographic data extracted from OpenStreetMap contributors (www.openstreetmap.org) under Open Data Commons Open Database License (ODbL) 1.0 (http://openstreetmap.org/copyright).

3.10. Acknowledgements

Chapter 3, in full, is a reprint of the material as it appears in Scientific Reports 2022. Gabriel Carrasco-Escobar, Jason Rosado, Oscar Nolasco, Michael T. White, Ivo Mueller, Marcia C. Castro, Hugo Rodriguez-Ferruci, Dionicia Gamboa, Alejandro Llanos-Cuentas, Joseph M. Vinetz, Tarik Benmarhnia. The paper title is “Effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon using parametric g-formula”. The dissertation author was the primary researcher and author of this paper.

Chapter 4: Quantifying the effect of Human Population Mobility on Malaria risk in the Peruvian Amazon

4.1. Abstract

The impact of Human population movement (HPM) on the epidemiology of vector-borne diseases, such as malaria, has been described. However, there are limited data on the use of new technologies for the study of HPM in endemic areas with difficult access such as the Amazon. In this study conducted in rural Peruvian Amazon, we used self-reported travel surveys and GPS trackers coupled with a Bayesian spatial model to quantify the role of HPM on the malaria risk. By using a densely sampled population cohort, this study highlighted the elevated malaria transmission in a riverine community of the Peruvian Amazon. We also found that the high connectivity between Amazon communities for reasons such as work, trading or family plausibly sustain such transmission levels. Finally, by using multiple human mobility metrics including GPS-trackers, and adapted causal inference methods we identified for the first time the effect of human mobility patterns on malaria risk in rural Peruvian Amazon. This study provides evidence of the causal effect of HPM on malaria that may help to adapt current malaria control programs in the Amazon.

4.2. Introduction

During 2019, 229 million malaria cases worldwide and 409 deaths occurred in 87 endemic countries which surpassed projections made few years earlier [187]. In Latin America, countries that share the Amazon region account for about 86% of all of the cases in the continent, despite control programs implemented for several years [187,188]. In Peru, >90% of malaria cases

are concentrated in the Loreto region, a mainly rural area with no electricity supply, no piped water, mostly accessible only by river, poor housing conditions and high mobility of people between villages [187,189]. Well-identified individual and household level risk factors for malaria risk such as the misuse of personal protection measures (e.g., bed nets), knowledge of malaria, occupation household infrastructure, overcrowding, indoor animals, and proximity to mosquito breeding sites have been widely described [190–194]. In addition, a number of studies showed the interplay of these individual and household factors with large-scale processes such as climate, deforestation, control programs and cultural aspects [195–198]. However, most of these previous epidemiological studies relied on static exposures assuming individuals are not moving across different areas. Human population movement (HPM) has been hypothesized as a potential driver of malaria transmission [199,200] but empirical evidence is lacking in the Amazon region.

The HPM between different villages impacts the spread of multiple infectious diseases and it has been observed that HPM is responsible for the spread of vector-borne diseases on scales that exceed the areas covered by their main arthropod vectors [201,202]. In previous studies considering HPM, standardized (self-reporting) travel questionnaires have been used for the study of HPM [203–205], however, the lack of space and time granularity collected with these questionnaires often masked the effect of complex travel patterns [206]. In the context of rural and riverine areas such as the Peruvian Amazon, labor activities (e.g., fishermen, loggers, and trading) involve long-distance travel through the rivers that connect the whole region, multimodal transportation, and multiple intermediate destinations that often include villages that are crossing borders to other countries with heterogeneous endemicity level [207,208]. The particular HPM characteristics in this area require the use of innovative geographic positioning

technologies that help the characterization of HPM at a fine scale and with greater precision [204,209].

Previous studies described multiple options to collect data for the study of HPM and its role in disease transmission, such as the use of GPS trackers, cellphone records, and participatory mapping (GeoODK) [199,210,211]. These approaches allow the collection of data for spatiotemporal analysis at different levels (neighborhood to international temporal scale, or daily to seasonal temporal scale). Another approach is based on using data from mobile phones to capture mobility patterns. Data collected with mobile phones generates a large number of records to capture people's mobility that can be then used to quantify HPM patterns. The combination of mobile records with epidemiological surveillance and genetic data improves the estimation of the flows of parasite between localities due to HPM [212]. However, due to privacy concerns, it is not possible to obtain sociodemographic data of the mobile users and in rural areas, such as the Amazon region, the communication network is scarce and limits the collection of data through these devices [210,213]. Another option to capture population mobility is based on GPS data. GPS data is increasingly being applied to a variety of statistical methods to quantitatively characterize mobility patterns, such as time-weighted spatial averaging (TWSA) approaches (e.g., utilization distribution, kernel density, density ranking) and models based on activity spaces (e.g., daily path area, minimum convex polygon, standard deviation ellipse) at a fine spatio-temporal resolution. The integration of sociodemographic data from travel surveys with the accuracy of GPS data trackers would improve the assessment of HPM and quantify its role on the malaria epidemiology in the Peruvian Amazon. In the previous study [214], we considered only 20 participants with GPS and movement ecology methods to describe the mobility between infected

and non-infected participants. For the current paper, we extend the analysis to the whole cohort, combining and comparing two methodologies for the study of HPM (GPS and surveys) to investigate the etiological effect of human mobility on malaria.

Another important challenge for the study of the effect of HPM on malaria risk is the complex confounding structure which require a clear causal inference framework and adapted methods. In the presence of complex confounding structures, methods based on propensity score estimation such as Inverse Probability of Treatment Weighting (IPTW) [215] have been proposed to deal with high-dimensional settings while allowing to check and optimize covariate balance to ensure that exposed and unexposed individuals are as similar as possible thus emulating a target trial [216]. However, such approaches are still underused in the malaria epidemiology literature [217–219] and to the best of our knowledge never applied in the context of HPM.

To bridge this gap in the literature, our study sought to characterize and quantify the mobility patterns and their effects on malaria risk using a densely sampled study in rural Amazonia as a case study. This study leverages a rich human movement dataset (previously reported [199]) with a causal inference framework to estimate the etiological effect of multiple mobility patterns on malaria risk that will provide important insights for malaria control and elimination in rural areas such as the Amazon rainforest region.

4.3. Methods

4.3.1. Ethics

All participants were included in the study (data collection and blood sample) after the signature of an informed consent form. The study was approved by the ethics review board of Universidad Peruana Cayetano Heredia (UPCH) in Lima (SIDISI: 100469).

4.3.2. Study design

A population cohort study was conducted to assess the contribution of human population movement to the malaria epidemiology in the rural Peruvian Amazon. All inhabitants aged 18 years or older were invited to participate and upon signed consent were followed for two months with weekly measurements. Sociodemographic and epidemiological data were collected at census (week 0), parasitological surveys (weeks 1 and 8), and follow-up surveys (weeks 2 to 7). In addition, a sub-cohort was conducted with 20 selected inhabitants of the main cohort who were given a 3D-printed GPS tracker developed for the study, described elsewhere [199] during the last 4 weeks of the study (weeks 5 to 8). Malaria infection status was diagnosed weekly by molecular testing. Finger prick blood samples were taken to all participants during the parasitological surveys regardless their symptoms or travel status. During the follow-up surveys samples were collected if at least one of the following three conditions occurred: (i) the participant reported a travel outside the village the last week; (ii) the GPS tracker recorded travel outside the village (for the sub-cohort participants) and (iii) the participant presented clinical symptoms compatible with malaria.

4.3.3. Study site and population

This study was conducted in Gamitanacocha (3.426°S, 73.318°W), in the Mazan district, Maynas province in the Loreto Region (Figure 4.1) with a total population of 92 inhabitants. Gamitanacocha is a community located north of Iquitos, capital of the Loreto. The community is only accessible using boat transportation, which takes ~6 hours from Iquitos. Gamitanacocha is a community surrounded by dense primary and secondary tropical forest with tropical weather and two marked seasons: rainy season from November to May and a dry season from June to October. This climate is optimal for the development of the primary vector of malaria in the region, *Nyssorhynchus (Anopheles) darlingi* [220], making it one of the villages with the highest risk of malaria [205]. All villagers aged 18 years or older (N = 50) were included in this study. The GPS trackers developed for this proof-of-concept study were distributed by purposive sampling of participants, taking into account whether the participant had self-reported a trip outside the community in the last month [199].

4.3.4. Data collection

A census of all inhabitants aged 18 years or older in all households was conducted and included sociodemographic data (age, sex, occupation, migratory status, birthplace, time in community, pregnant, chronic disease, educational level, household structure). A parasitological survey (weeks 1 and 8) was conducted to collect epidemiological and mobility history in the last month for the entire population. In addition, the follow-up survey (weeks 2 to 7) was used to collect epidemiological and mobility data for the last week, including whether the GPS tracker recorded participants departure from the community. These surveys included data such as place of sleeping during travel, reason for travel, work conducted during travel, travel destination and

travel duration. It is important to mention that the difference between the main occupation (census) and the work performed during the travel (follow-up survey) was considered because the study population had the characteristic of having occasional or seasonal jobs that did not necessarily coincide with their main occupation. Capillary blood samples were taken on filter paper from all participants to determine the basal and final infection status (parasitological surveys) and the infection status per week of each participant (follow-up survey). Infection status was determined by microscopy and PCR for *Plasmodium* at the species level [205], because infections can be submicroscopic. In addition, a second blood sample was taken 4 days later to avoid false negatives, because the parasitemia may be undetectable at the beginning of the disease. All this information was collected in both the main cohort and the sub-cohort.

4.3.5. Laboratory procedure

In the case of detection by PCR, Genomic DNA was extracted from dry blood spots of ~6 mm2 sections using E.Z.N.A.® Blood DNA Kit (Omega Bio-tek®, USA), according to the manufacturer guidelines with slight modifications – addition of TEN (20 mM Tris-HCl, pH 8.0; 2 mM EDTA, pH 8.0; 0.2 M NaCl) buffer, supplemented with SDS 10% w/v. Subsequent amplification was done by a real-time quantitative PCR (qPCR) method targeting the 18SSU rRNA gene region of the *Plasmodium* species-specific region. Oligonucleotides 5-TAACGAACGA- GATCTTAA-3 and 5-GTTCCTCTAAGAAGCTTT-3 were used as primers as reported by Mangold et al. [205] and a modified protocol was used including PerfeCta SYBR Green Fast Mix (Quanta Biosciences, MD, USA). Ambiguous diagnostic results were confirmed by using a nested ssPCR method [221].

4.3.6. Statistical analysis

The sociodemographic data of all participants and the characteristics of their households were summarized in proportions for categorical variables, and in median and interquartile range for continuous variables with skewed distribution.

4.3.6.1. Mobility patterns

For the main cohort, mobility patterns derived from the parasitological and follow-up surveys were described, such as travel frequency, destination and reason for travel, work performed during the travel, number of destinations, and travel time. For the sub-cohort, GPS trackers were used to record the total distance covered and total time covered (per participant) to observe their relationship with *Plasmodium spp.* infection. The features of the GPS tracker (hardware architecture, code and performance characteristics) were described elsewhere [199].

The GPS records after being integrated with the data of the surveys, were characterized by a non-parametric Bayesian framework using the *Bayesmove* package in R programming software v.4.0.3 [222] , to accurately characterize the mobility patterns of the sub-cohort [223]. The model consists of a two-step framework, which divides individual tracks into segments (Supplementary Figure 4.1) and then groups these segments into possible movement patterns called " latent behavioral states" which, depending on whether their characteristics (in terms of step length-SL and turning angle-TA) have biological plausibility, are finally selected. The detailed procedure can be found in Supplementary methods 1. To obtain a mobility pattern per individual using these generated behavioral states, a summary of these results was performed. First, the proportion of each behavioral state generated was calculated, and second, the optimal

value of the proportion of a behavioral state that predicts *Plasmodium* infection was calculated as a breakpoint to define which behavioral state will be assigned as the mobility pattern. This procedure was performed using a ROC curve.

In addition, we compared the performance of GPS trackers and surveys to identify whether participants leave the village, destinations and number of travels. Community departure identified by the GPS trackers was based on the detection of movement out of a 500 m radius with respect to the village centroid. In the case of GPS tracker travel destinations, they were identified by the presence of GPS records within a buffer (~500 m) of the Mazan riverine communities based on the results of the Bayesmove spatial model.

4.3.6.2. Mobility patterns as risk factors for malaria infection

Multiple mobility pattern metrics were explored, such as traveled more than 4 times (TN), traveled to Mazan (TD), traveled for work (TR), use a non-motorized boat for travel (TT), sleep outside during the trip (TSP), traveled to Mazan for more than 24 hours (M24), traveled to Mazan for reasons other than work (MnW), which were constructed from the weekly surveys, and the displacement pattern obtained by the Bayesmove model (MovT). and commuting pattern by Bayesmove model (MovT). Since the infection status was tested each week, we selected participants free of infection at week 1 (N = 30). From these participants, a positive infection status was determined if the participants developed infection during the study period (the first positive result obtained) and, in addition, we computed the time (in weeks) from enrollment to the first malaria infection.

To determine the effect of multiple mobility patterns while dealing with multiple a priori identified confounders, we used inverse probability treatment weighting (IPTW) to minimize the differences in baseline characteristics between the groups for each exposure (mobility pattern). The weights were created from the inverse of the propensity score (PS) if the individual was exposed, and for the unexposed it was the inverse of 1 minus the propensity score [224]. The variables used to create the propensity scores for each mobility pattern were a priori selected using a directed acyclic graph (DAG) [225] (Supplementary Figure 4.2) considering all possible causal paths between each exposure (mobility pattern) and malaria infection. We first optimized the propensity score models (considering multiple functional forms and interactions using identified covariates) by comparing the AIC between models. Once we obtained an optimized propensity score (for each comparison of interest), we then created weights as described above [226]. We then checked the covariate balance between exposure groups using standardized mean differences and love plots. We developed 3 different models to determine the risk malaria infection associated with HPM exposures through different incidence-based approaches: (1) Log-binomial model to estimate the Incidence Proportion Ratio (IPR), (2) Poisson model controlling by person-week of follow-up as an offset to estimate the Incidence Rate Ratio (IRR), and (3) Cox Proportional-hazards model to estimate the Hazard Ratio (HR). All analysis and figures were conducted in R programming software v.4.0.3 [222].

4.4. Results

4.4.1. Baseline characteristics and infection status

From the 30 participants in the cohort (excluding positives at baseline), the median age was 32 years (IQR = 25-48), the proportion of males and females were the same, and the most

frequent occupation was farmer (73%). More than 70% of participants had at least primary education and were literate, only 2 people had a diagnosis of chronic disease, and 43% of the total population was born in the community. The main reason for migration to Gamitanacocha was for family and economic reasons (30% and 23%, respectively). The distribution of socioeconomic variables by at least one episode of malaria infection during the study is presented in Table 4.1. The incidence proportion of malaria by PCR in the study was 40%, and the incidence rate was 1 per 10 person-week with a median survival time of 2 weeks, and the incidence proportion of malaria by microscopy was 17 %. Figure 4.2 shows the distribution of malaria cases by species in the total number of inhabitants of the village during the 8 weeks of the study.

Regarding households' characteristics, 86% of the houses use the river as a source of water and 14% use rainwater. Most (86%) households did not have a bathroom, 67% of the households had electricity, and only 14% were fumigated. The floor and walls were made of wood in more than 90% of the households, and the roofs were made of thatch in more than 50%. The characteristics of the total population (N=50) of the village are shown in Table 4.1 in Supplementary Table 4.1.

4.4.2. Mobility patterns

In the main cohort (n=30), a total of 132 trips were reported by participants during the study (using surveys). Work was the most frequent reason to travel out of the village (42%). Mazan was reported as the most frequent travel destination (36%) and logger the most frequent occupation (Table 4.2). Table 4.3 shows the mobility patterns and person-weeks for each

category for participants (free of infection at baseline). The 77% of participants traveled to Mazan at least once, 83% traveled for work at least once, 73% traveled to Mazan not for work at least once, and 53% traveled to Mazan for more than 24 hours at least once. The travel characteristics of the total population of the village are presented in Supplementary Table 4.2.

For the sub-cohort (n=20), mobility patterns were described in relation to time and distance covered by the participants (using GPS data). Malaria cases increased in those who stayed more time outside the village and traveled more distance, both for the infections summarized per participant (Figure 4.3A) and per participant-week of follow-up (Figure 4.3B). From the Bayesmove model we obtained probabilities for each GPS record of belonging to a latent behavioral state, showing that the first and second behavioral states grouped the largest number of GPS records. We analyzed the characteristics of the two behavioral states with the largest amount of data and identified that the first pattern consisted of long step length movements without a defined turning angle, and another pattern of medium to short step length, with varying turning angles (Supplementary Figure 4.3). From these results, the first movement pattern identified was named “displacement pattern” and the second, “community pattern”. With these mobility patterns, the villages in which participants developed a community pattern, that is, their travel destinations, were identified. The Figure 4.3 shows selected trajectories followed by the participants, with their respective behavioral states. It can also be observed that the trajectories marked by the displacement patterns were variable, from leaving up to a radius of 3 km relative to Gamitanacocha, to beyond 20 km. Graphics of all participants are in Supplementary Figure 4.4 The results of the comparison of GPS trackers and surveys to detect community departures, number of trips and destinations are shown in Supplementary Table 4.3.

4.4.3. Effect of mobility patterns on risk of infection by *Plasmodium spp.*

Figure 4.4 shows the results of IPR, IRR, and HR (in balanced samples) to estimate the effect of multiple mobility patterns on the risk of malaria infection. A travel to Mazan at least once (IPR: 6.61 [1.87-40.32], IRR: 7.33 [2.07-44.66]), travel to Mazan for other than work (IPR: 3.22 [1.18-10.65], IRR: 3.43 [1.26-11.33]), and travel to work (IPR: 4.35 [1.19-29.76], IRR: 4.79 [1.31-32.76]) were the mobility metrics that increased risk of acquiring a malaria infection in 2 of the 3 models considered. Travel to Mazan for more than 24 hours (IRR: 2.81 [1.06 - 7.97]) was the mobility metrics shown to be a risk for *Plasmodium* infection in one of the 3 models. The point estimation of the 3 models had similar results although for the interval estimates the Cox model less precise. Kaplan-Meier plots (Supplementary Figure 4.5) show that, although not statistically significant, it takes less time for the exposed group (for each mobility pattern) to reach 50% infection compared to the non-exposed.

4.5. Discussion

The estimation of the etiological effect of human population mobility on malaria risk has been elusive in rural areas in Latin America where moderate-to-high malaria transmission occurs. By using a densely sampled population cohort, this study highlighted the elevated malaria transmission in a riverine community of the Peruvian Amazon. We also found that the high connectivity between Amazon communities for reasons such as work, trading or family plausibly sustain such transmission levels. Finally, by using multiple human mobility metrics including GPS-trackers and adapted causal inference methods we identified for the first time the effect of human mobility patterns on malaria risk in rural Peruvian Amazon.

This area was previously characterized as a moderate transmission setting [227], however, this is the first report of hyper-endemicity levels comparable to some hyper-endemic areas of SSA such as Guinea (43.9%), Mozambique (38.3%), and ivory coast (41.5%) [228]. Importantly, it has been previously reported a high level of asymptomatic malaria cases (~88%) in the villages in the Mazan basin, that potentially underestimates actual transmission in the area [221]. In this study, a high detection of infections was accounted by sampling the population twice a week, highlighting the importance of prospective surveillance of malaria. This is particularly important in populations with low parasite densities or close to the limit of detection of diagnostic methods due to the fitness of the immune response in consequence to constant infections [229]. The results obtained in this study are comparable with the 54.3% previously obtained in the district of Mazan [220], 40.9% and 31.8% in the village of Libertad [205,230], Primero de Enero village with 44,4% [205], and in villages that belong to other river basins, such as the Napo river basin, where the Urco Miraño village had a prevalence of 30.7% [205,230]. Further studies are needed to determine the generalizability of the results of this study to the entire Loreto Region, mainly due to the high heterogeneity in malaria endemicity and HPM patterns.

This study highlights to importance of considering population mobility when designing strategies to reduce malaria risk. For example, environmental interventions such as vegetation clearance or modification of river boundaries may target areas characterized with higher population mobility patterns. This high transmission observed is possibly fueled by the high connectivity between communities in the Amazon due to highly mobile populations. In this study, all participants have left the community at least once during the study period (2 months).

We determined the mobility of the participants through self-reporting travel questionnaires and the use of GPS trackers. By using questionnaires, we determined that travel reasons of the inhabitants of Gamitanacocha were mainly due to work, followed by trading, and family. Collecting mobility data from both questionnaires and GPS are complementary and here we showed how such data can be combined to study the effect of HPM patterns on malaria risk. We therefore encourage future studies to rely on diverse sources of data collection. Furthermore, collecting survey data regarding travel reasons has been shown to limit the "residential" effect fallacy [231]. The results are consistent with those reported in Brazil [204,208], where labor was the third cause of travel; however, these reports were collected from surveys in rural and urban areas, where workplaces may be in the same locality. In our study, logging was the most frequent work. Despite the community is surrounded by primary and secondary tropical forest, which allows the development of this activity, long distances were recorded to logging areas where high human-biting behaviors of mosquitoes were recorded [221]. Another important factor is the seasonality of both, vector and work-related travel. This study was conducted during the rainy season, when the malaria vector has favorable climatic conditions to develop. In addition, the activities carried out by people are also conditioned by seasonality since there are activities that are carried out on a daily (trading, fishing), periodically (tourism, mining, logging) or long term (migration, colonization) basis [232]. This connectivity between communities has also been inferred through genetic analysis of parasite populations found in the communities of the same watershed. The level of genetic similarity between parasites sampled in the Mazan watershed indicated an intense flow of parasites between communities, highly compatible with human populations with frequent mobility between them [200].

In this study we analyzed GPS data –to objectively characterize the HPM among villagers– with a Bayesian behavioral state approach. This analysis revealed that there are at least two mobility behaviors in the study participants. We were able to differentiate travel trajectories through the river network from travel in communities or work zones within the forest or forest fringes. It is important to highlight that previous studies reported that these displacements in different communities or forest areas put people at great risk (2-fold higher) of get a malaria infection, because the vector prefers peri-domestic areas and the forest fringe [233]. This suggests that interventions to prevent malaria should take particular attention not only to local transmission but HPM as a key mechanism in moderate-to-high transmission settings. Additionally, we were able to identify the communities most frequented by participants, which for 2018, the Regional Health Direction of Loreto reported Annual Parasitic Indices (API) of 403.4 and 537.5 for Libertad and Visto Bueno villages, 138.8 for Santa Cruz village, 51.9 for Puerto Alegre and 11.6 for Mazan village. Although Mazan does not present high API levels, the malaria risk obtained by the models may be due to these villages, for their location in the basin, are transitory stops on the way to the final destination, which is Mazan. Rich GPS tracking data were leveraged by movement ecology methods such as utilization distribution or kernel density, and other post-hoc analyses such as step selection functions (SSF), calculation of activity budgets, or behavior-specific measures of landscape resistance [199,234,235] in the context of non-human animals; however, a body of studies examining these methods for infectious diseases epidemiology are still scarce. Moreover, the fine-scale characterization of HPM obtained in this study is consistent with that described elsewhere [236,237], which gives it an advantage over more commonly used methodologies such as telephone records or self-report surveys. We should also highlight that, in this study, low-cost open-source GPS devices were used [214], which overcomes

the logistical limitation that is usually present when researchers want to implement HPM studies at this level of resolution. It is also important to mention that, in this study, the surveys had highly comparable results with the GPS trackers for the detection of community departures by participants. This is possibly since the surveys –like the malaria tests– were conducted on a weekly basis during the whole study, decreasing possible recall bias due to participant omission.

We determined that travel to Mazan city (~3 Km from study area) increases the risk of malaria infection, compatible with previous reports that also highlight that the vector is present in the peri-domestic zones of the community and in greater frequency in all villages in the watershed of the Mazan river [230,233]. It is also important to note that Mazan is the main city in the entire watershed, so all the other communities interact there for different reasons. Several previous studies have reported the association between traveling for work and risk of malaria, although due to the characteristics of the type of work carried out in this study (mainly logging) it resulted in a much higher risk compared to the other studies [205,207,208]. From the 3 types of models developed, the estimates obtained from calculating the IPR and IRR were consistent in the point estimation and precision, in contrast to the HR, which obtained wider CIs, although with the same direction of effect and similar point estimate.

Some limitations in this study must be acknowledged. First, the mobility data from the overall cohort were collected over an 8-week period and the GPS data over a 4-week period. This condition would not capture seasonal changes in both mobility patterns and vector development in the area [232]. Future, longitudinal studies are suggested to account for potential seasonal differences in malaria transmission and mobility patterns. Second, the GPS tracker had

moments of low battery and had to stop recording data, but the spatial model developed in the study homogenized the recording times in a step prior to the analysis, and the records that could not be smoothed were excluded from the analysis without altering the results. Third, there were participants who were not sampled in all weeks of the study, which could affect the time at which they truly became infected. The weekly sampling strategy was intended to minimize this missing data problem. Fourth, although there were variables that could not be collected, these variables were identified as latent variables through the DAGs and their position as confounders in the causal path were examined. No latent variables were required in the minimum adjustment sets for all DAGs evaluated. Fifth, the study was designed to detect malaria infections during travel and not locally in the community. This could affect the differentiation of whether the participant became infected during travel or in the same community. In addition, we were not able to consider the role non-travelling asymptomatic individuals in the study area regarding malaria risk. Sixth, personal protection measures (e.g., using of repellent) may act as important effect modifiers. The survey included questions about such individual behaviors' measures, but most participants did not answer to these questions. It would be important to investigate the potential protective role of such behaviors measures regarding the effect of mobility and malaria risk. Finally, the small sample size of this study, in addition to the exclusion of participants infected at week 1, may affect the precision of our estimates; however, the different metrics analyzed were consistent in direction and magnitude despite the precision estimates.

4.6. Conclusion

This study shows the potential of hyper-endemicity of malaria in riverine communities of the Peruvian Amazona that were supported by high connectivity between villages in the same

watershed. We objectively characterized the human population movement through the river network and forested areas using GPS tracking data and highlight the mobility patterns that strongly increase the risk of malaria in the area. Finally, this study provides evidence of the etiological effect of human mobility on the risk of malaria that may help tailoring current malaria control strategies in areas with moderate-to-high malaria transmission such as the Amazon rainforest.

4.7. Tables

Table 4.1. Baseline characteristics of the study population infection-free at the beginning of the study.

		P. vivax		*P. falciparum*	
		Negative	Positive	Negative	Positive
	N = 30	n = 18 (%)	n = 12 (%)	n = 26 (%)	n = 4 (%)
Age (years)					
18–50	23 (77)	13 (72)	10 (83)	20 (77)	3 (75)
51–65	5 (17)	4 (22)	1 (8.3)	5 (19)	0 (0)
> 65	2 (6.7)	1 (5.6)	1 (8.3)	1 (3.8)	1 (25)
Sex					
Male	15(50)	8(44)	7(58)	12(43)	3(75)
Female	15(50)	10(56)	5(42)	14(54)	1(25)
Occupation					
Farmer	22(73)	14(78)	8(67)	19(73)	3(75)
Housewife	6(20)	3(17)	3(25)	5(19)	1(25)
Teacher	1(3.3)	1(5.6)	0	1(3.8)	0(0)
Health promoter	1(3.3)	0 (0)	1(8.3)	1(3.8)	0(0)
Education level					
None	2(6.7)	2(11)	0(0)	2(7.7)	0(0)
Primary school	21(70)	12(67)	9(75)	19(73)	2(50)
Secondary school	5(17)	2(11)	3(25)	3(12)	2(50)
Higher education	2(6.7)	2(11)	0(0)	2(7.7)	0(0)
Literate					
Yes	26(87)	15(83)	11(92)	22(85)	4(100)
No	4(13)	3(17)	1(8.3)	4(15)	0(0)
Other disease					
None	28(93)	16(89)	12(100)	24(92)	4(100)
Anemia	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
Diabetes	0(0)	0(0)	0(0)	0(0)	0(0)
Rheumatism	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
Pregnant					
Yes	2(13)	0(0)	2(40)	2(14)	0(0)
No	13(87)	10(100)	3(60)	12(86)	1(100)
Born in community					
Yes	13(43)	8(44)	5(42)	12(46)	1(25)
No	17(57)	10(56)	7(58)	14(54)	3(75)
Time in community (months)*					
< 12	2 (6.7)	1 (5.6)	1 (8,3)	2 (7.7)	0 (0)
12–24	1 (3.3)	1 (5.6)	0 (0)	1 (3.8)	0 (0)
24–48	4 (13)	2 (11)	2 (17)	4 (15)	0 (0)
48–60	1 (3.3)	1 (5.6)	0 (0)	1 (3.8)	0 (0)
> 60	22 (73)	13 (72)	9 (75)	18 (69)	4 (100)
Reason for migration					
None	13 (43)	8(44)	5(42)	12(46)	1(25)
Economic	7(23)	3(17)	4(33)	5(19)	2(50)
Family	9(30)	6(33)	3(25)	8(31)	1(25)
Other	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
Birthplace					
Gamitanacocha	13(43)	8(44)	5(42)	12(46)	1(25)
Marañon	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
Loreto	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
Maynas	12(40)	6(33)	6(50)	9(35)	3(75)
Putumayo	1(3.3)	0(0)	1(8.3)	1(3.8)	0(0)
Ramon Castilla	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)
San Martin	0(0)	0(0)	0(0)	0(0)	0(0)
Yurimaguas	1(3.3)	1(5.6)	0(0)	1(3.8)	0(0)

Table 4.2. Total number of travels of the participants during the whole study.

		P. vivax		*P. falciparum*	
		Negative	**Positive**	**Negative**	**Positive**
	N = 132	**n = 113(%)**	**n = 19 (%)**	**n = 127 (%)**	**n = 5 (%)**
Travel work					
No work on the travel	91(69)	77(68)	14974)	88(69)	3(69)
Farmer	8(6.1)	6(5.3)	2(11)	7(5.5)	1(20)
Logger	14(11)	14(12)	0(0)	14(11)	0(0)
Fisher	8(6.1)	7(6.2)	1(5.3)	8(6.3)	0(0)
Laborer	3(2.3)	1(0.9)	2(11)	3(2.4)	0(0)
Other	8(6.1)	8(7.1)	0(0)	7(5.5)	1(20)
Travel destination					
Libertad	7(5.3)	7(6.2)	0(0)	7(5.5)	0(0)
Mazan	47(36)	38(34)	9(47)	45(35)	2(40)
Visto Bueno	20(15)	19(17)	1(5.3)	19(15)	1(20)
Other	58(44)	49(43)	9(47)	56(44)	2(40)
Travel reason					
Trade	26(20)	22(19)	4(21)	25(20)	1(20)
Studies	2(1.5)	2(1.8)	0(0)	2(1.6)	0(0)
Family	21(16)	16(14)	5(26)	20(16)	1(20)
Health	6(4.5)	4(3.5)	2(11)	6(4.7)	0(0)
Work	56(42)	49(43)	7(37)	54(43)	2(40)
Other	21(16)	20(18)	1(5.3)	20(16)	1(20)
Travel transportation					
Canoe	13(9.4)	11(9.7)	2(11)	12(9.4)	1(20)
Motorized boat	109(83)	92(81)	17(89)	105(83)	4(80)
Other	10(7.6)	10(8.8)	0(0)	10(7.9)	0(0)
Travel sleep place					
No sleep during travel	73 (55)	65(58)	8(42)	69(54)	4(80)
House	49(37)	38(34)	11(58)	48(38)	1(20)
Outside	6(4.5)	6(5.3)	0(0)	6(4.7)	0(0)
Other	4(3)	4(3.5)	0(0)	4(3.1)	0(0)
Travel duration (hours)*	40 (9–107)	49 (8–124)	28 (27–51)	49 (9–123)	27 (18–39)

*Median (IQR)

Table 4.3. Mobility patterns of the participants for the whole study period.

	Positive N = 12 (%)	Population N = 30 (%)	Pseudo-population	Person-week (N = 120)
Traveled to Mazan not for work				
No	2 (17)	8(27)	33	35
Yes	10(83)	22(73)	30	85
Traveled to Mazan > 24 h				
No	4(33)	14(47)	36	61
Yes	8(67)	16(53)	21	59
Number of travels (> = 4)				
No	7(58)	19(63)	30	72
Yes	5(42)	11(37)	25	48
Slept outside during travel				
No	8(67)	22(73)	30	90
Yes	4(33)	8(27)	24	30
Traveled in motorized boat				
No	7(58)	17(57)	31	63
Yes	5(42)	13(43)	30	57
Traveled to Mazan				
No	1(8.3)	7(23)	31	33
Yes	11(91.7)	23(77)	30	87
Traveled to work				
No	1(8.3)	5(17)	32	22
Yes	11(91.7)	25(83)	55	98
Displacement pattern*				
No	3(38)	8(53)	15	36
Yes	3(43)	7(47)	15	32

*Percent for N=15.

4.8. Figures

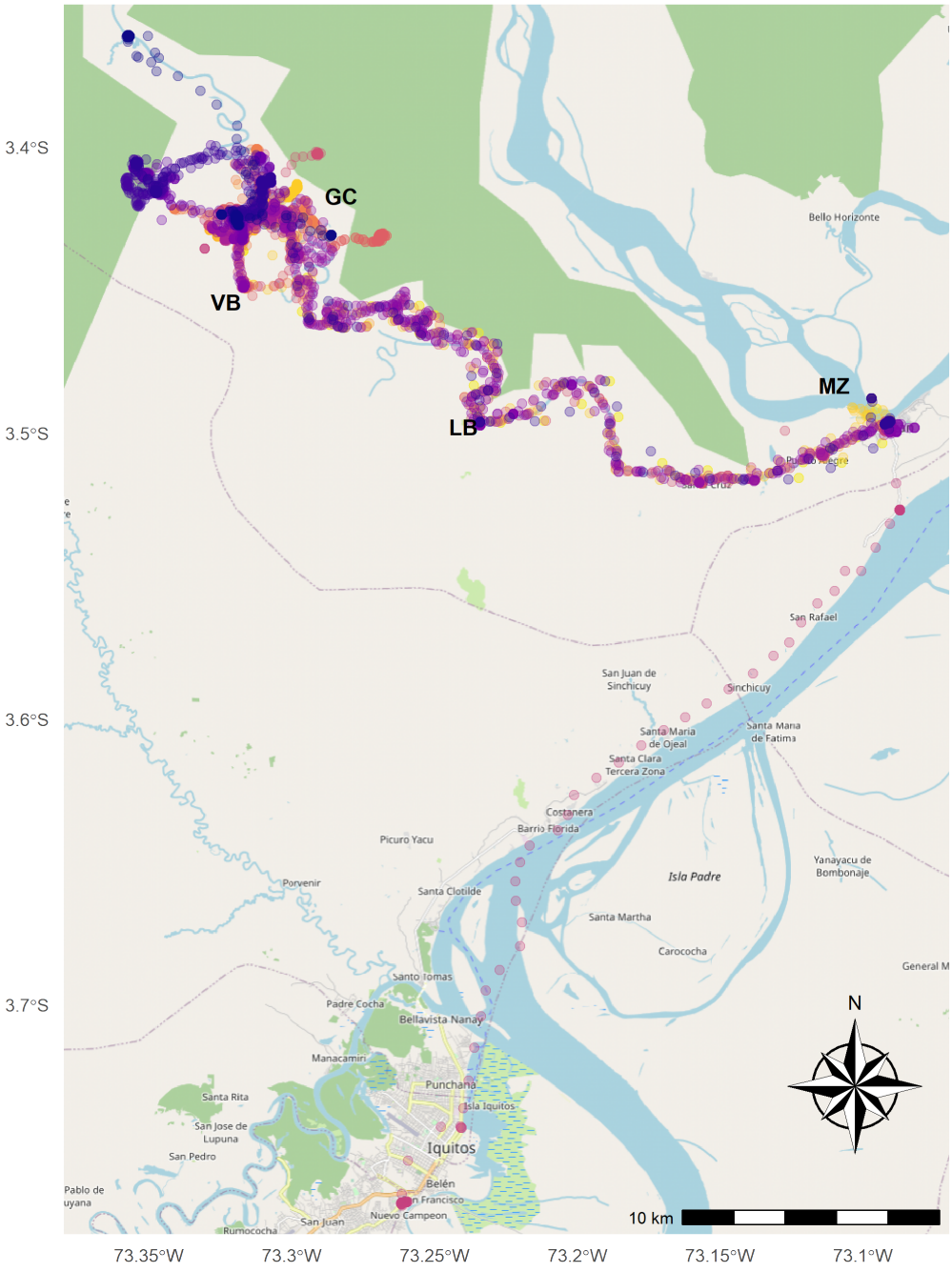


Figure 4.1. Location of Gamitanacocha in the Peruvian Amazon and travel records. Located in the district of Mazan, province of Maynas, region Loreto, Peru. The GPS trackers recorded the trajectories developed by the participants during 4 weeks. (GC: Gamitanacocha, MZ: Mazan, VB: Visto Bueno, LB: Libertad).

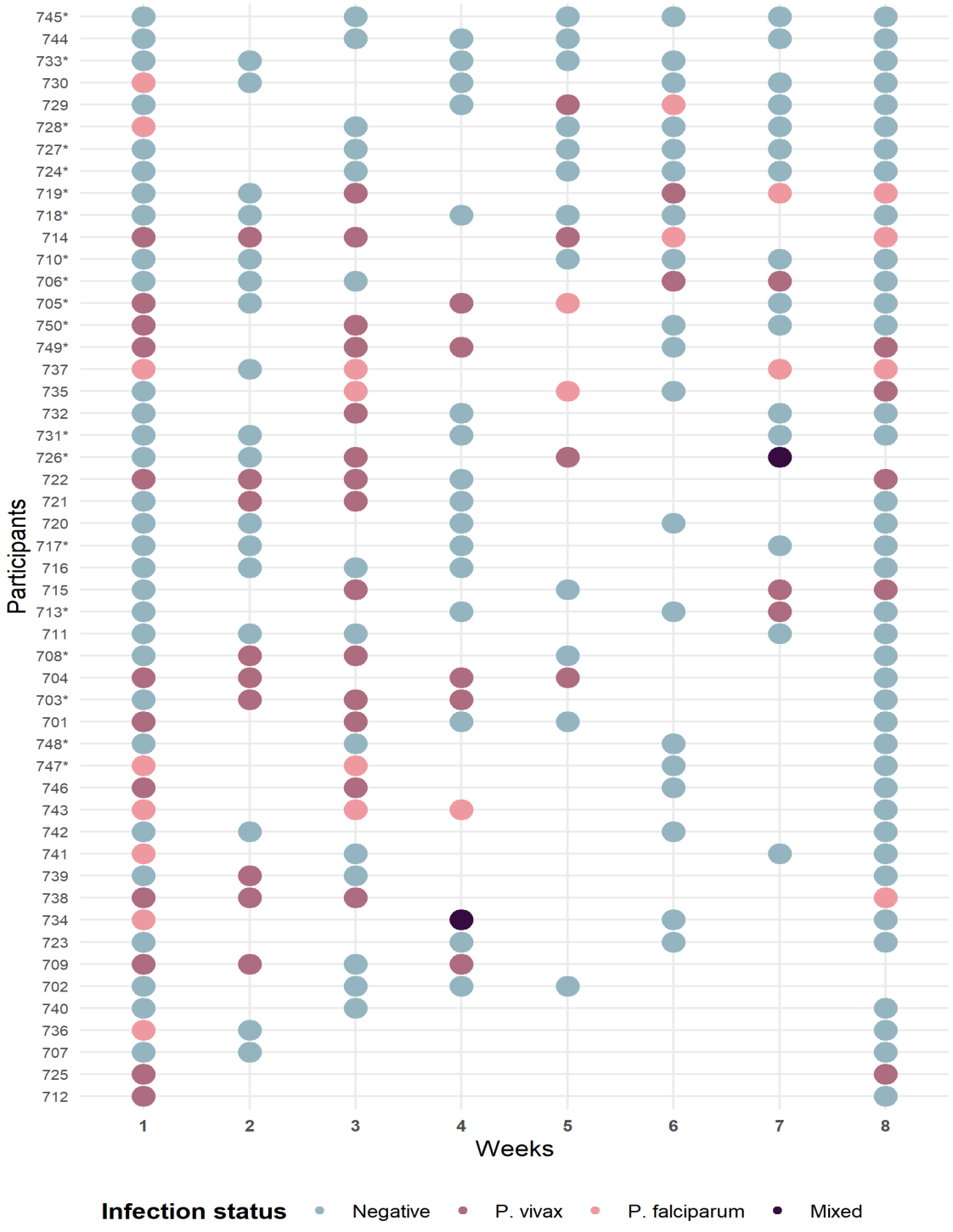


Figure 4.2. Cases per species detected by PCR weekly of the 50 study participants. Participants who were included in the sub cohort are indicated with (*). Empty spaces indicate that the participant was not sampled in that week because did not meet the established criteria for being sampled.

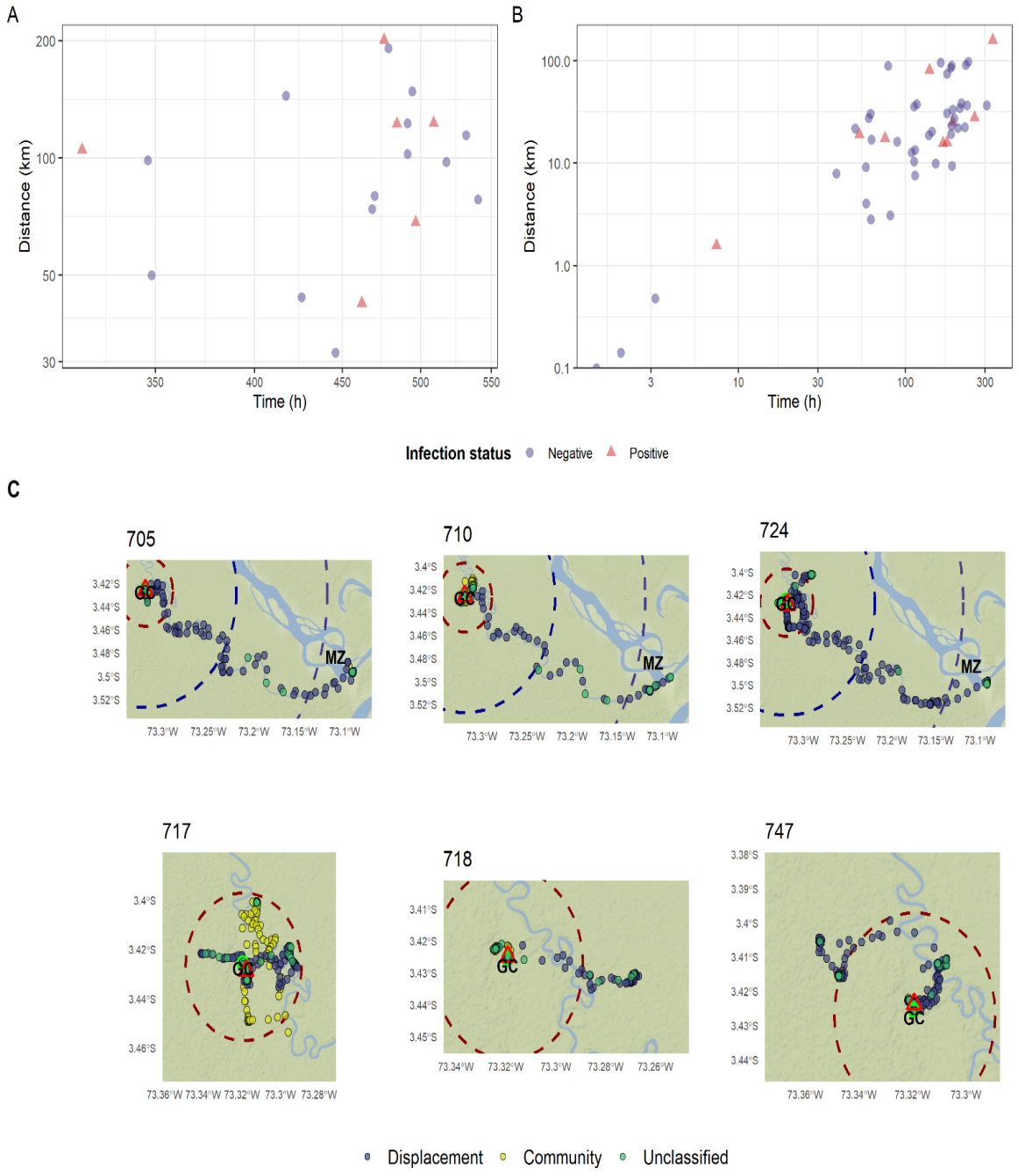


Figure 4.3. Travel patterns of study participants in Gamitanacocha in the Peruvian Amazon A. Cumulative distance and time traveled by subcohort participants during the whole study by infection status. B. Distance and time traveled weekly by participants in the subcohort during the whole study by infection status. C. Trajectories of selected participants outside the village taking into account the type of mobility pattern performed and the distance from Gamitanacocha (Red buffer: 3 km, Blue buffer: 10 km, Violet buffer: 20 km).

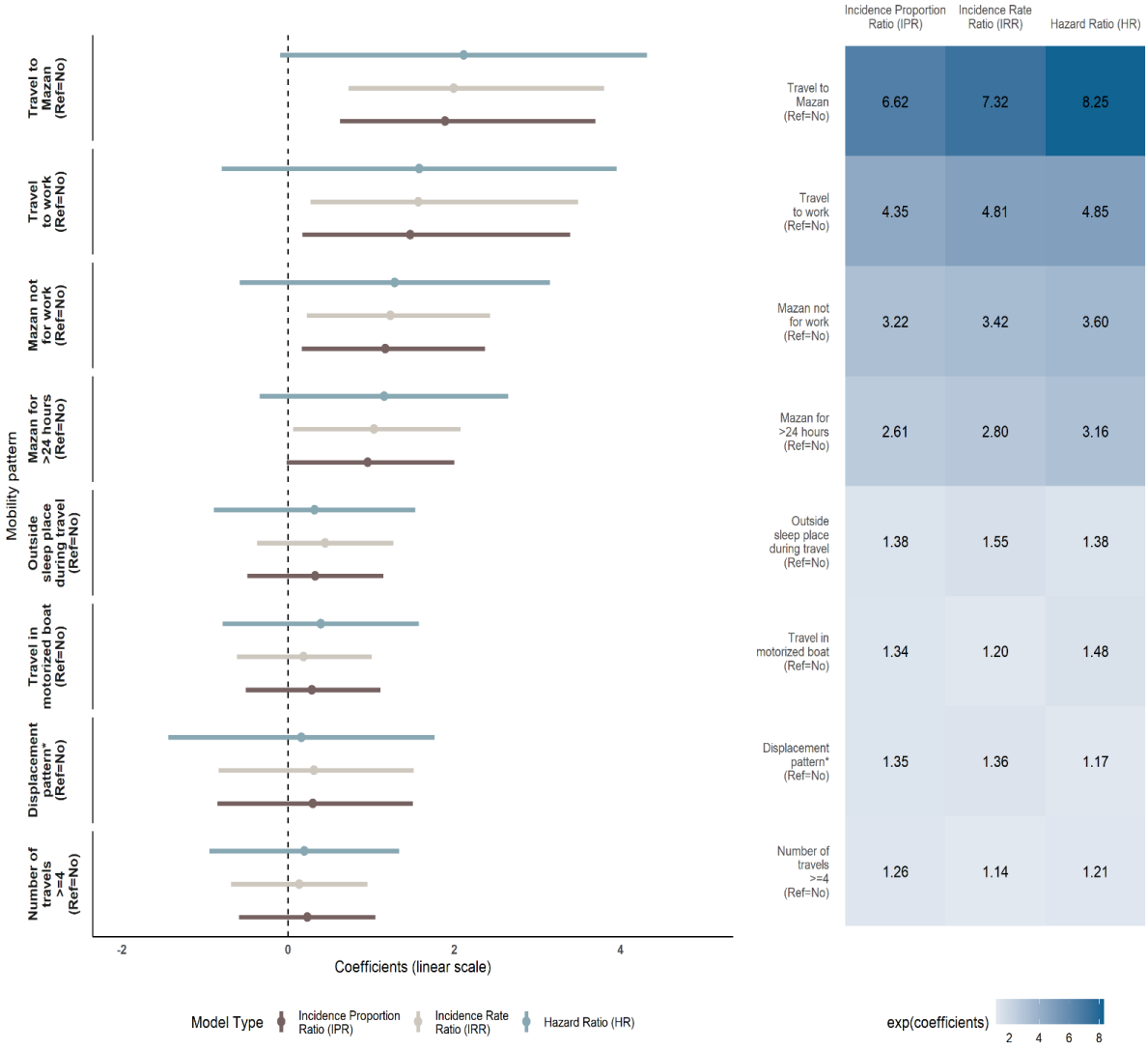


Figure 4.4. Forest plot of the models for each exposure applying the IP weighting method for each type of model developed.

4.9. Supplementary information

Supplementary methods 4.1. Mobility characterization using the Bayesmove package

GPS data processing

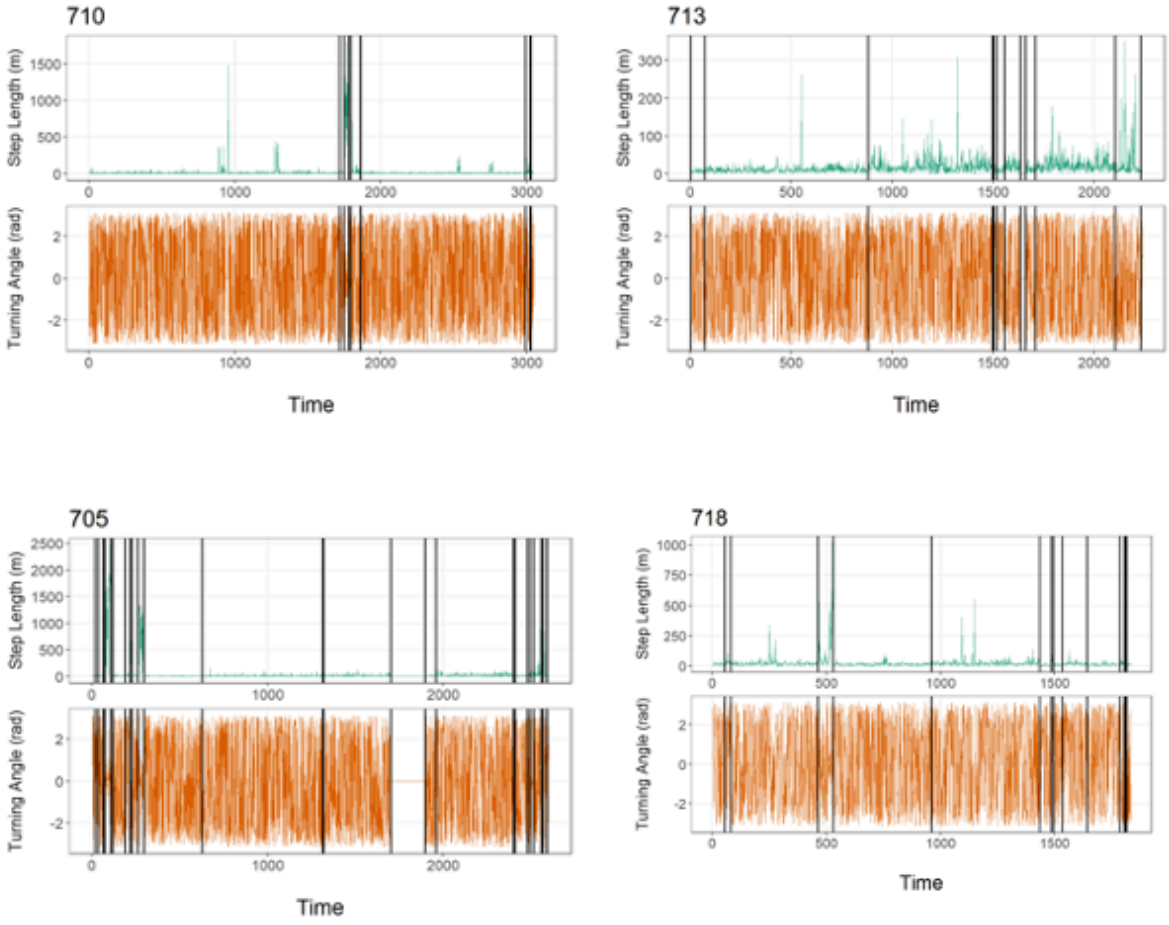
The model consists of a two-step framework. The step length (SL), turning angle (TA) and time difference (DT) of the records were calculated based on the GPS tracks. The records with the most frequent DT value (five minutes) were selected. From these data, the SL and TA values were discretized, using quantiles for SL and by □□4 from -2 □ to 2□ for TA. These new discrete variables were used for the first step of the model.

Segment the tracks

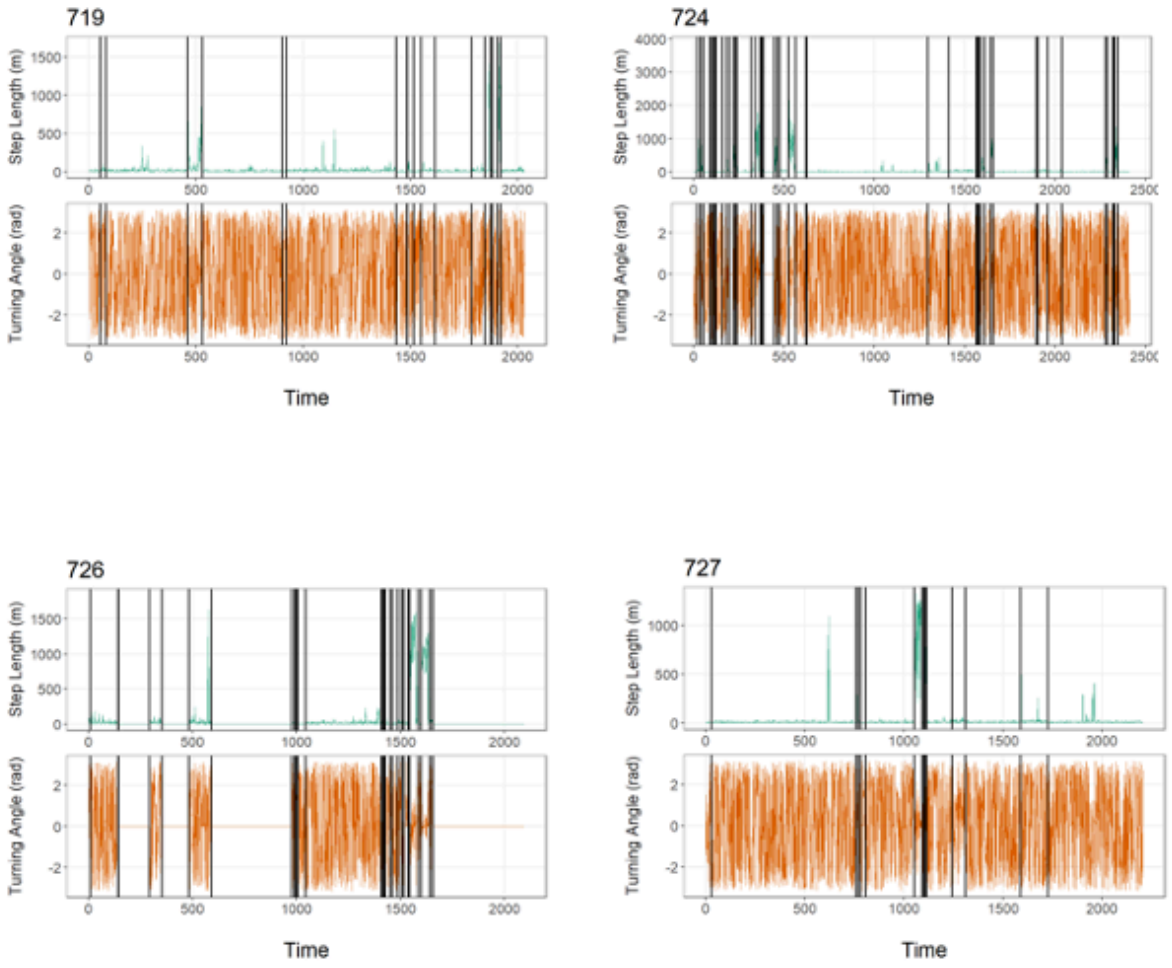
This first step consists in a Bayesian segmentation model, which uses a reversible-jump Markov chain Monte Carlo (RJMCMC) algorithm to estimate the breakpoints where values of SL and TA substantially change. These breakpoints were verified visually using a function from the Bayesmove package to generate graphs with the breakpoints generated, in addition to verifying the convergence of the model per individual. These breakpoints generated segments in the data that represent different changes in movement patterns of the participants based on the variables (SL and TA).

Cluster track segments

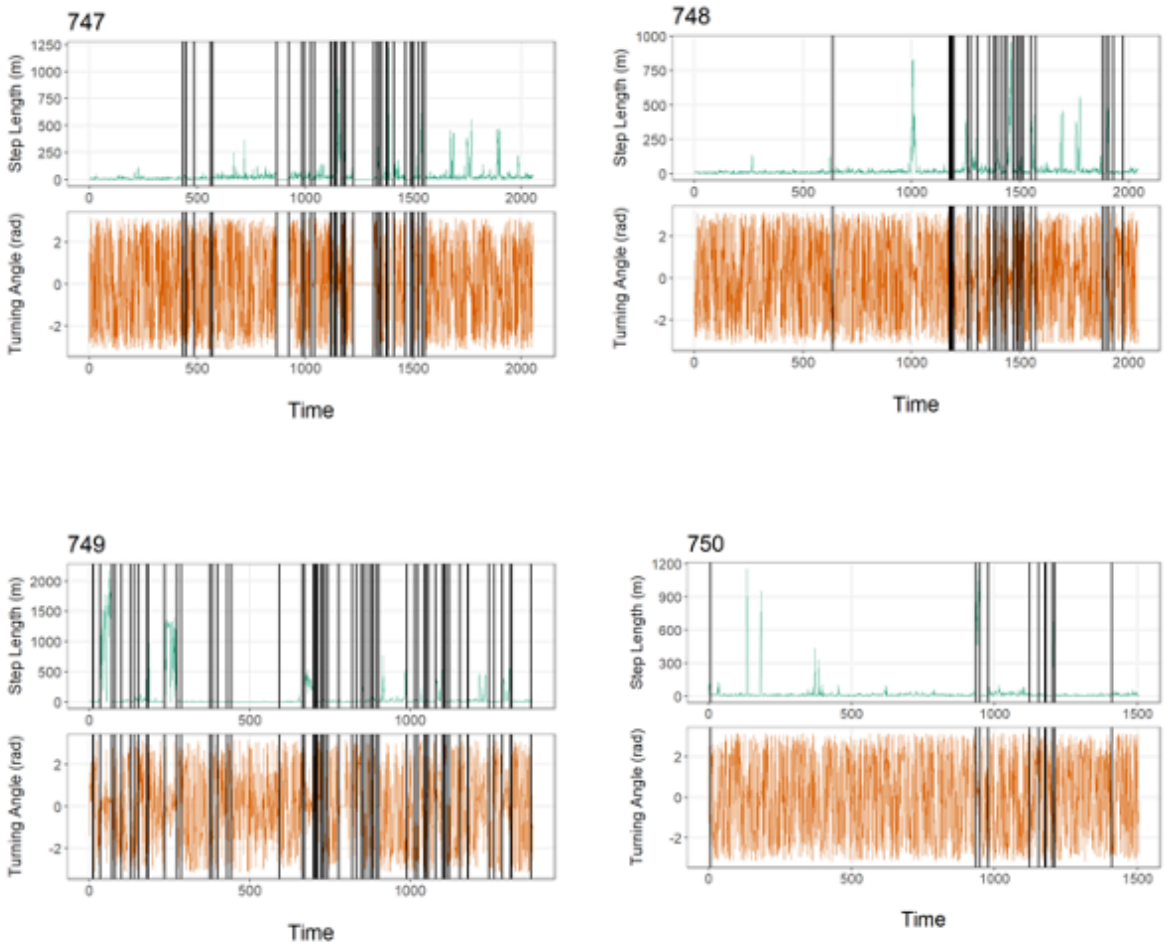
Once observations from all participants have been assigned to the segments obtained in the previous step, a Latent Dirichlet Allocation (LDA) model was performed to estimate the latent behavioral states. For the selection of the final behavioral states, we considered the proportions of each latent behavioral state with respect to all the records and the biological sense, that is, that according to the characteristics of the behavioral state, it can be explained as an activity of the individual.



Supplementary Figure 4.1. Graphs showing the breakpoints generated in the first step of the spatial model (Bayesmove). the vertical black lines indicate where the model detected a significant change in SL and TA values for each participant.

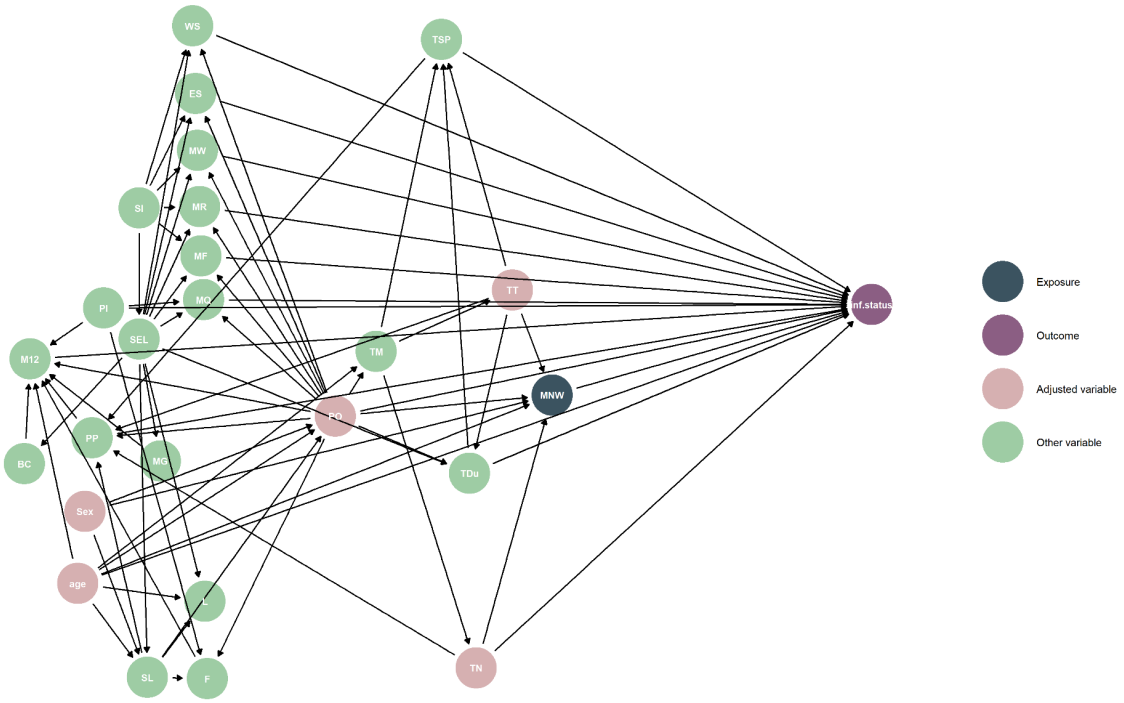


Supplementary Figure 4.1. (Continued) Graphs showing the breakpoints generated in the first step of the spatial model (Bayesmove). the vertical black lines indicate where the model detected a significant change in SL and TA values for each participant.



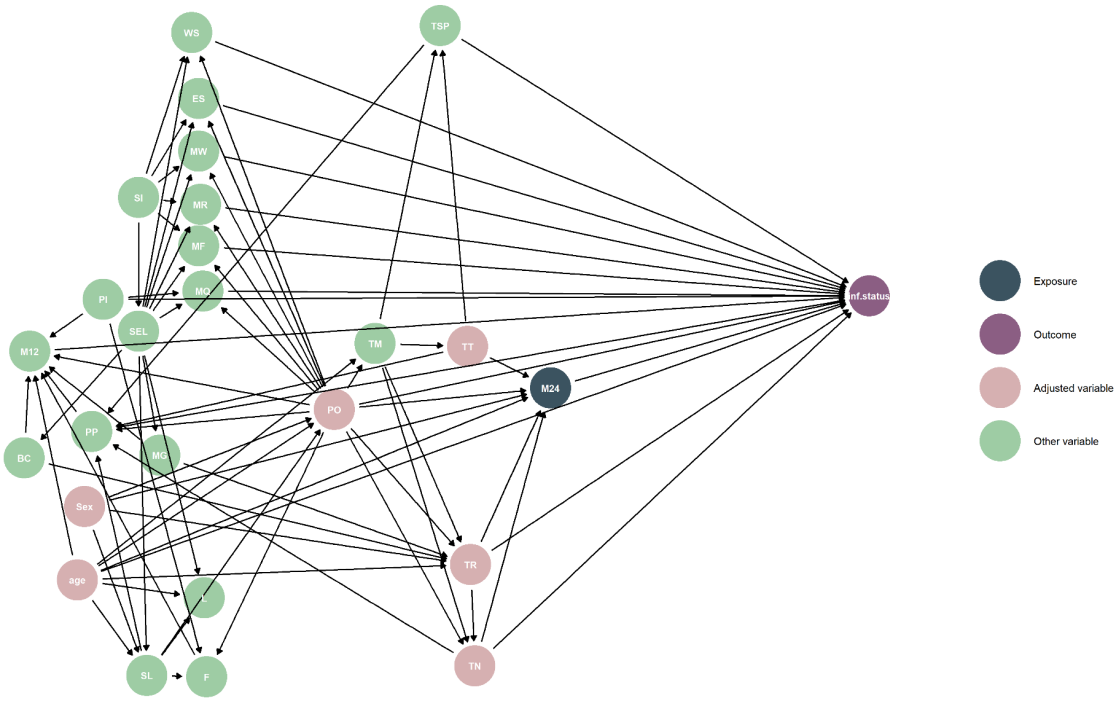
Supplementary Figure 4.1. (Continued) Graphs showing the breakpoints generated in the first step of the spatial model (Bayesmove). the vertical black lines indicate where the model detected a significant change in SL and TA values for each participant.

A



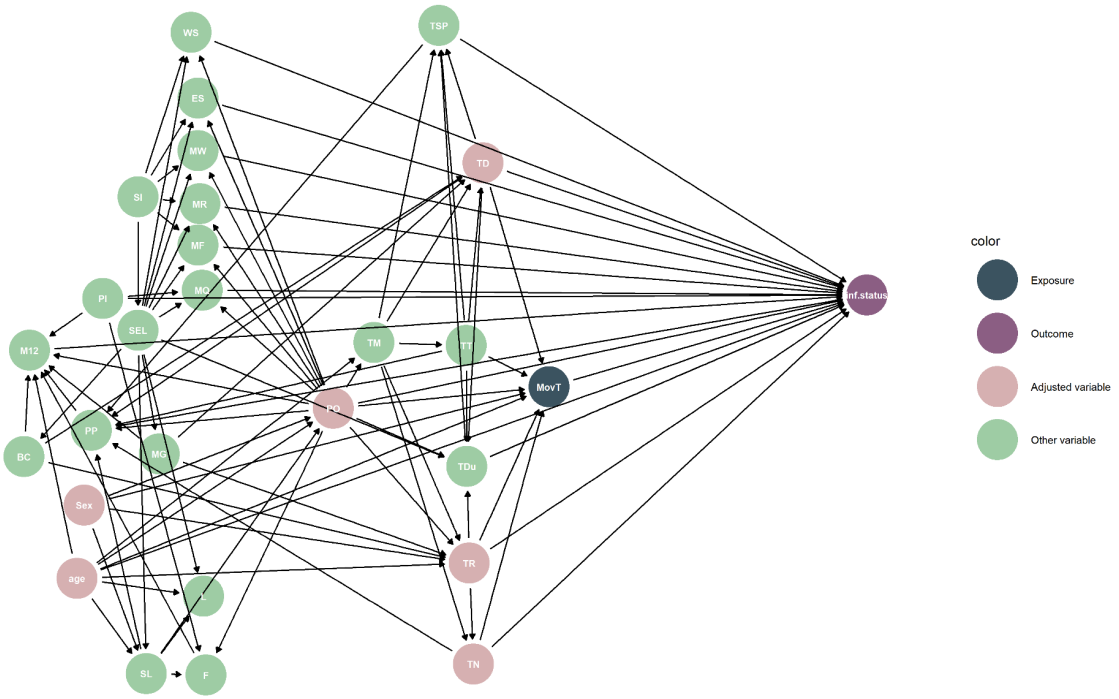
Supplementary Figure 4.2. DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

B



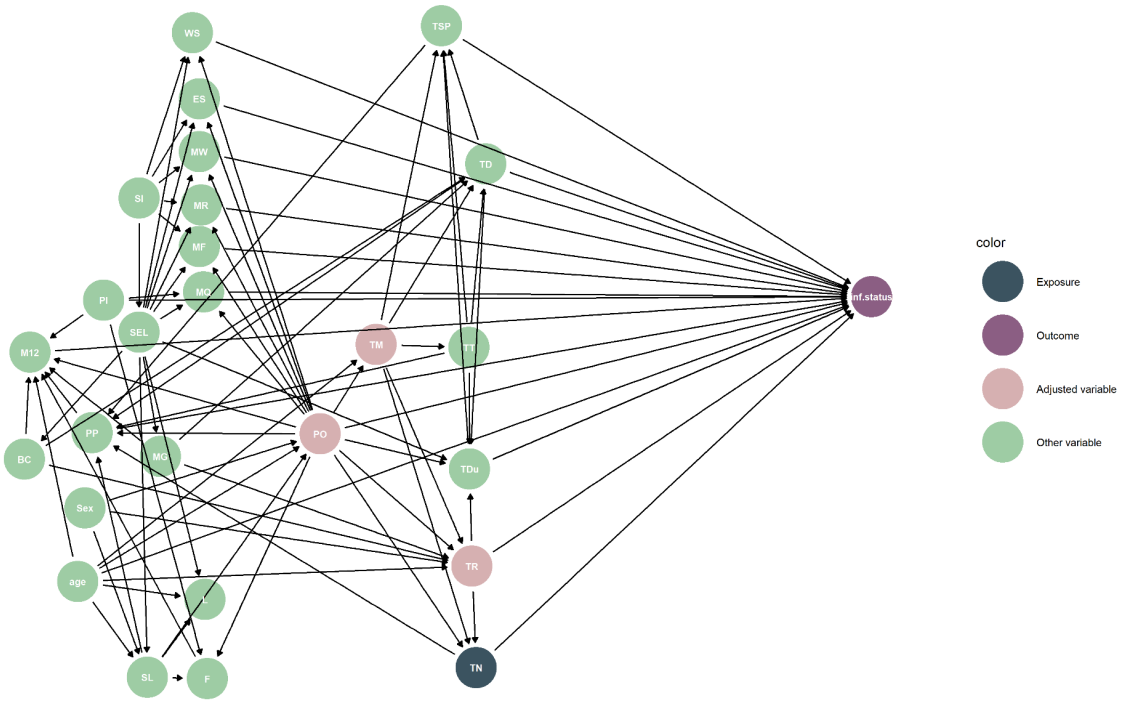
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

C



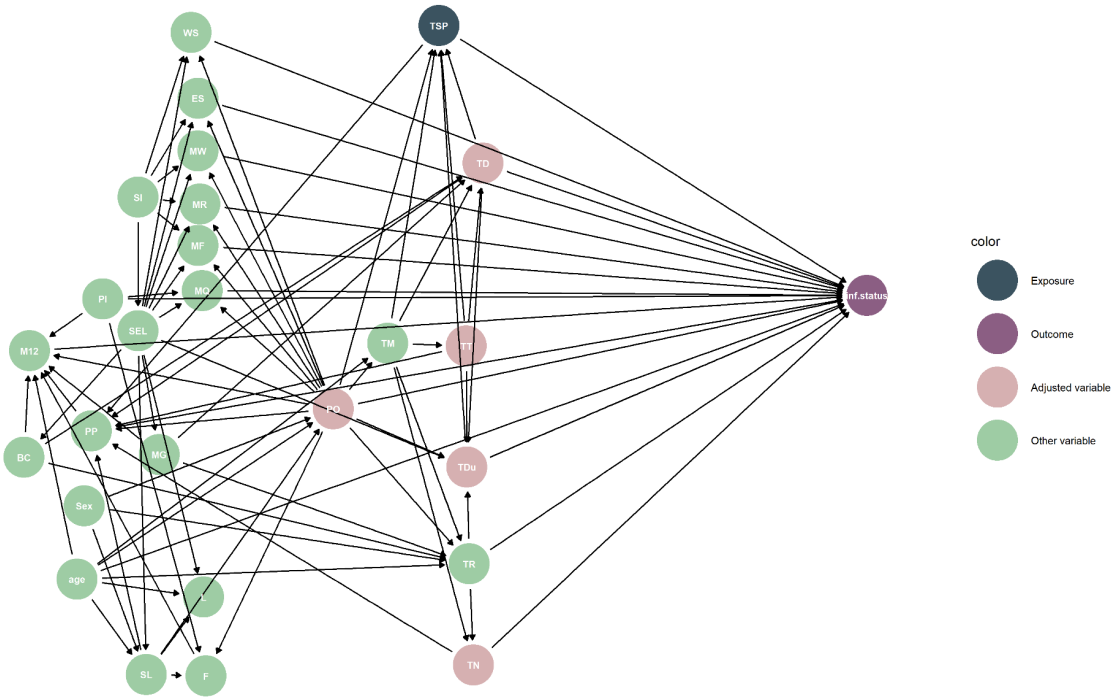
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

D



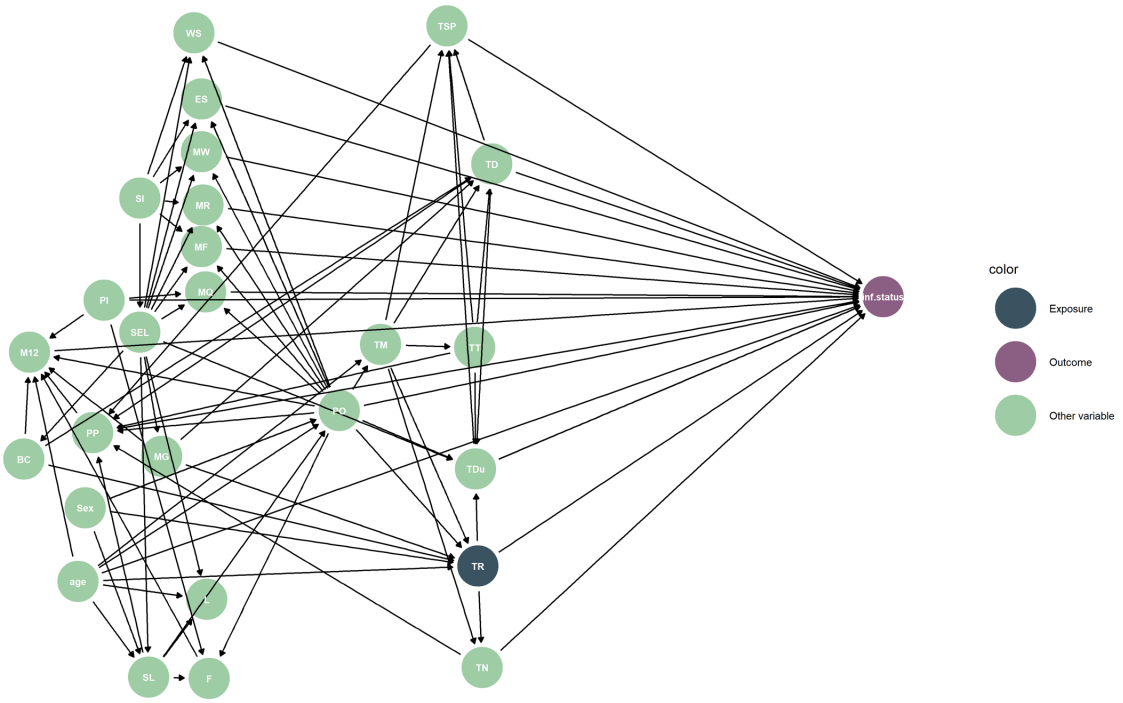
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

E



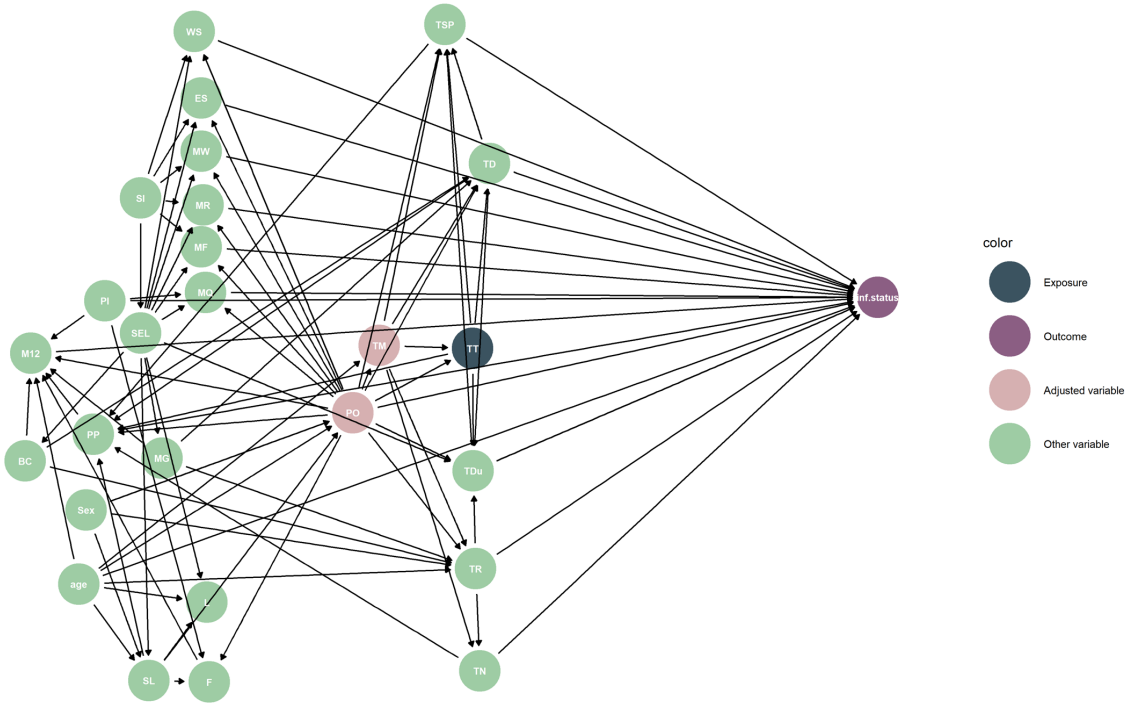
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

F



Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

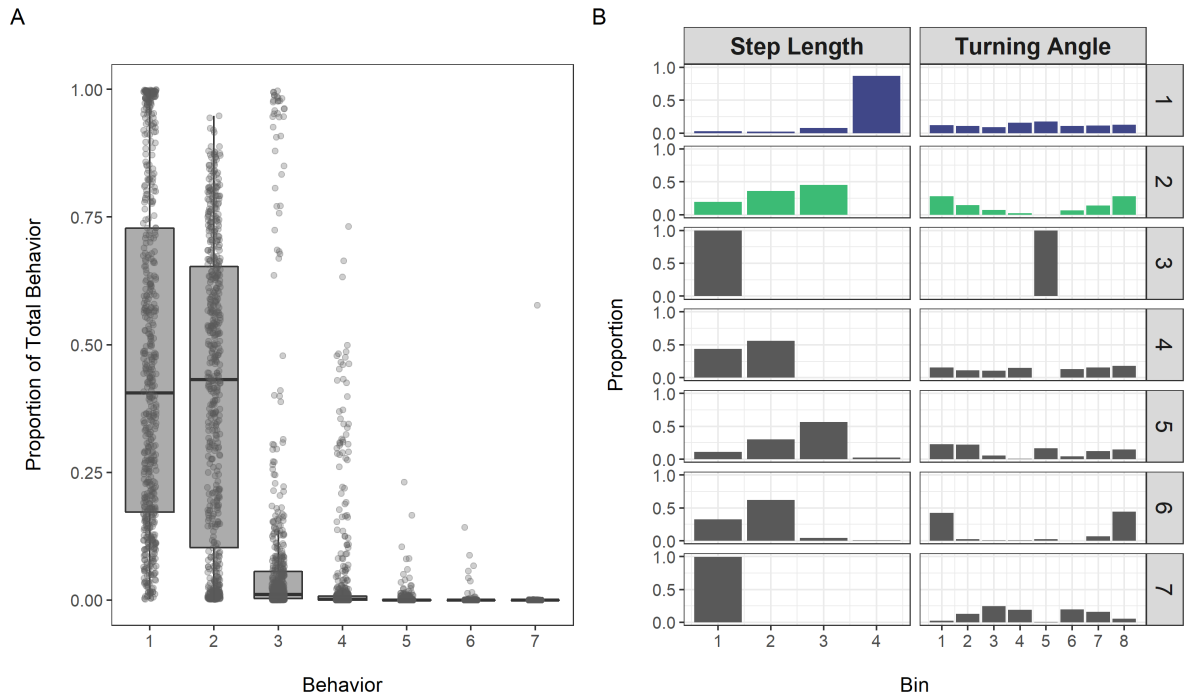
G



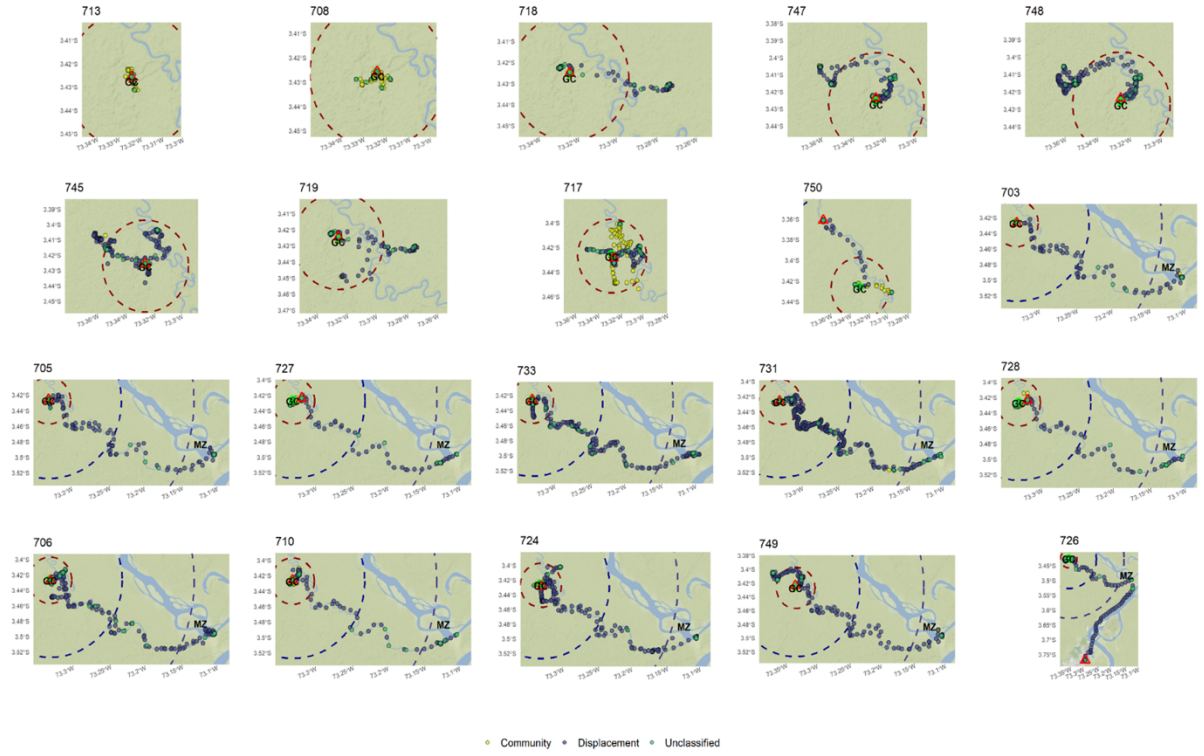
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).

Definition of terms – DAGs	
BC	born in community
M12	malaria in the last 12 months
PP	personal protection
PI	preventive intervention
SL	study level
SEL	socioeconomic level
SI	structure intervention
MQ	use flynet
MF	floor material
MR	roof material
MW	wall material
ES	electricity source
WS	water source
F	fumigate
L	literate
MG	migration
PO	principal occupation
TM	
TN	travel number
TR	travel reason
TDu	travel duration
TT	travel transportation
TD	travel destination
TSP	travel sleep place
MovT	Movement type
Inf.status	Infection status
M24	Mazan at least 24 hours
MNW	Mazan not for work

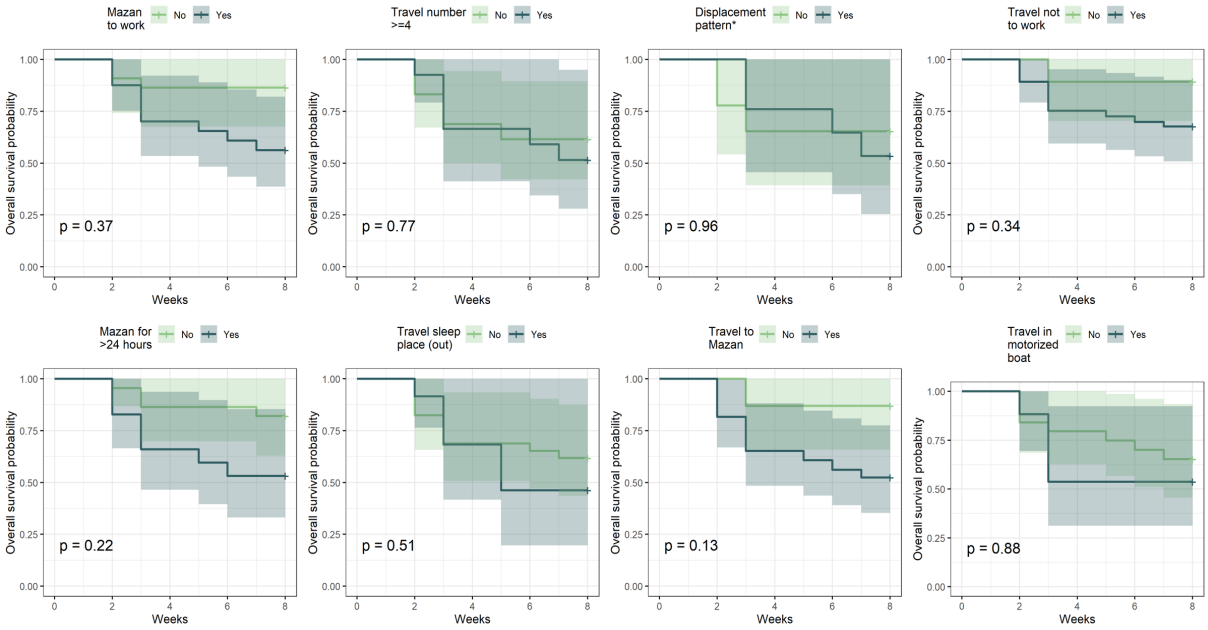
Supplementary Figure 4.2. (Continued) DAGs for the models adjusted for each exposure. **A**. Travel to Mazan not for work (MW), **B.** Travel to Mazan for at least 24 hours (M24), **C**. Mobility pattern (MovT), **D.** Travel Number >4 (TN), **E.** Travel sleep place -out (TSP), **F.** Travel to work (TR), **G.** Travel to Mazan (TD).



Supplementary Figure 4.3. A. Bar chart showing the probabilities that a GPS record belongs to a particular latent behavioral state. B. Step length and turning angle characteristics of the chosen behavioral states (in color).



Supplementary Figure 4.4. Behavioral state per participant and trajectory. The dotted lines indicate the radius of mobilization traveled taking Gamitanacocha as reference (Red: 3 km, Blue: 10 km, Violet: 20 km).



Supplementary Figure 4.5. Kaplan Meier plot of mobility patterns for the main cohort.

Supplementary Table 4.1. Baseline characteristics of study population

		P. vivax		*P. falciparum*	
		Negative	**Positive**	**Negative**	**Positive**
	N = 50	**n = 25 (%)**	**n = 25 (%)**	**n = 35 (%)**	**n = 15 (%)**
Age (years)*	32(23)	31 (29)	32(14)	32(18)	32(31)
Sex					
Male	28(56)	13(52)	15(60)	18(51)	10(67)
Female	22(44)	12(48)	10(40)	17(49)	5(33)
Occupation					
Farmer	37(74)	20(80)	17(68)	25(71)	12(80)
Housewife	10(20)	4(16)	6(24)	7(20)	3(20)
Teacher	2(4)	1(4)	1(4)	2(5.7)	0(0)
Health promoter	1(2)	0 (0)	1(4)	1(2.9)	0(0)
Education level					
None	3(6)	2 (8)	1(4)	3(8.6)	0(0)
Primary school	35(70)	18(72)	17(68)	24(69)	11(73)
Secondary school	9(18)	3(12)	6(24)	5(14)	4(27)
Higher education	3(6)	2(8)	1(4)	3(8.6)	0(0)
Literate					
Yes	40(80)	19(76)	21(84)	28(80)	12(80)
No	10(20)	6(24)	4(16)	7(20)	3(20)
Other disease					
None	47(94)	22(88)	25(100)	33(94)	14(93)
Anemia	1(2)	1(4)	0(0)	1(2.9)	0(0)
Diabetes	1(2)	1(4)	0(0)	0(0)	1(6.7)
Rheumatism	1(2)	1(4)	0(0)	1(2.9)	0(0)
Pregnant					
Yes	3(14)	0(0)	3 (30)	2(12)	1(20)
No	19(86)	12(100)	7 (70)	20(57)	12(80)
Born in community					
Yes	18(36)	9(36)	9(36)	15(43)	3(20)
No	32(64)	16(64)	16(64)	20(57)	12(80)
Time in community (months)*	168(222)	156 (228)	216(192)	180(294)	120(180)
Migration reason					
None	18(36)	9(36)	9(36)	15(43)	3(20)
Economic	12(24)	6(24)	6(24)	7(20)	5(33)
Family	18(36)	8(32)	10(40)	12(34)	6(40)
Others	2(4)	2(8)	0(0)	1(2.9)	1(1.67)
Birthplace					
Gamitanacocha	18(36)	9(36)	9(36)	15(43)	3(20)
Marañon	1(2)	1(4)	0(0)	1(2.9)	0(0)
Loreto	1(2)	1(4)	0(0)	1(2.9)	0(0)
Maynas	23(46)	11(44)	12(48)	12(34)	11(73)
Putumayo	1(2)	0(0)	1(4)	1(2.9)	0(0)
Ramon Castilla	2(4)	1(4)	1(4)	2(2.5)	0(0)
San Martin	1(2)	0(0)	1(4)	1(2.9)	0(0)
Yurimaguas	3(6)	2(8)	1(4)	2(5.7)	1(6.7)

*Median (IQR)

Supplementary Table 4.2. Total travels of the total number of participants over 8 weeks

		P. vivax		*P. falciparum*	
		Negative	**Positive**	**Negative**	**Positive**
	N = 210	**n = 162 (%)**	**n = 48 (%)**	**n = 190 (%)**	**n = 20 (%)**
Travel work					
No work on the travel	147(70)	110(68)	37(77)	133(70)	14(70)
Farmer	14(6,7)	10(6,2)	4(8,3)	12(6,3)	2(10)
Logger	23(11)	21(13)	2(4,2)	21(11)	2(10)
Fisher	11(5,2)	9(5,6)	2(4,2)	10(5,3)	1(5)
Laborer	6(2,9)	3(1,9)	3(6,2)	6(3,2)	0(0)
Other	9(4,3)	9(5,6)	0(0)	8(4,2)	1(5)
Travel destination					
Libertad	10(4,8)	9(5,6)	1(2,1)	10(5,3)	0(0)
Mazan	75(36)	61(38)	14(29)	65(34)	10(50)
Visto Bueno	28(13)	69(43)	28(58)	25(13)	3(15)
Other	97(46)	23(14)	5(10)	90(47)	7(35)
Travel reason					
Trade	46(22)	37(23)	9(19)	39(21)	7(35)
Studies	2(1)	2(1,2)	0(0)	2(1,1)	0(0)
Family	32(15)	23(14)	9(19)	29(15)	3(15)
Health	10(4,8)	7(4,3)	3(6,2)	9(4,7)	1(5)
Work	92(44)	70(43)	22(46)	86(45)	6(30)
Other	28(13)	23(14)	5(10)	25(13)	3(15)
Travel transportation					
Canoe	19(9)	13(8)	6(12)	17(8,9)	2(10)
Motorized boat	176(84)	135(83)	41(85)	159(84)	17(85)
Other	15(7,1)	14(8,6)	1(2,1)	14(7,4)	1(5)
Travel sleep place					
No sleep during travel	107(51)	82(51)	25(52)	97(51)	10(50)
House	84(40)	62(38)	22(46)	77(41)	7(35)
Outside	13(6,2)	12(7,4)	1(2,1)	11(5,8)	2(10)
Other	6(2,4)	6(3,1)	0(0)	6(2,6)	0(0)
Travel duration (hours)*	49(112)	51(122)	28(44)	49(112)	52(98)

Statistical test performed: Wilcoxon rank-sum test, Fisher's exact test

*Median (IQR)

Supplementary Table 4.3. Performance comparison between GPS trackers and surveys

		GPS trackers	Self-reported surveys
Number of trips		46	48
Travel destination			
	Number		
	1–3	16	37
	4–6	-	4
	7–10	-	6
	> 10	-	1
	Places		
	Mazan	11	19
	Visto Bueno	4	8
	Libertad	2	1
	Others	5	20

4.10. Acknowledgements

Chapter 4, in full, is a reprint of the material as it appears in Royal Society Open Science 2022. Gabriel Carrasco-Escobar, Jose Matta-Chuquisapon, Edgar Manrique, Jorge Ruiz-Cabrejos, Jose Luis Barboza, Daniel Wong, German Henostroza, Alejandro Llanos-Cuentas, Tarik Benmarhnia. The paper title is “Quantifying the effect of human population mobility on malaria risk in the Peruvian Amazon”. The dissertation author was the primary researcher and author of this paper.

Chapter 5: Overall Discussion

5.1. Summary of dissertation research

Results from previous ecological and epidemiological studies provided preliminary evidence for the relationship between urban expansion and increased human mobility on malaria epidemiology [61,130,136]. Independent and combined effects of these factors impact malaria epidemiology particularly in areas with unstable transmission and in early stages of development such as the Amazon rainforest. However, no studies have been conducted to comprehensively understand these multiple (and multilevel) pathways that condition malaria endemicity in many Amazonian countries. The current pathway towards malaria elimination is severely threatened by growing environmental and human activity changes, yet there is still scarce evidence of how changing land cover affects HPM dynamics and eventually malaria transmission, especially in the Amazon region.

The purpose of this dissertation was to provide evidence of the epidemiological impacts of human mobility and land cover change in an understudied region and population in the Peruvian Amazon. This research expands on previous work by 1) using novel data sources to represent human mobility at a range of spatio-temporal scales from village-level to individual-level trip-level records, 2) applying novel analysis methodologies such as network analysis, time-weighted spatial averaging approaches (TWSA) for GPS-tracking data, and causal inference methods, and 3) identifying target populations and proposing policy strategies for malaria control in Amazonian settings with high HPM and LULC changes. In the first aim of the dissertation, a comprehensive set of village-level connectivity metrics were computed using network analysis and their relationships with malaria transmission levels were tested. In the second aim, the effect

of out-of-village working activities on recent malaria exposure was determined by using g-computation methods on two population-based studies in rural and peri-urban areas of the Peruvian Amazon. Lastly, the third aim quantified the effect of human mobility on malaria risk using GPS data and fine-scale mobility metrics.

Aim 1 of this dissertation (Section 2) represents the first study in the Peruvian Amazon that has applied novel network analysis on granular PCD routine data to understand the role of connectivity on malaria endemicity. This expanded on previous studies published in the Brazilian Amazon [15] by testing an ample set of connectivity metrics on networks constructed based on watershed distribution. Betweenness centrality outperformed other metrics based on its consistency and stability, and no important differences were detected when using multiple versions of weights (i.e., population, deforested area, Euclidian distance, or travel time) for its construction. In this region, this study found that localities and cities with high connectivity consistently have higher malaria endemicity. Subsequent explorations revealed that these differences in malaria endemicity due to connectivity is exacerbated in regions with the highest baseline malaria transmission rates. These results contradict prior research [61–63,82,130] that suggested that HPM was more significant in low-transmission and close-to-elimination settings than in moderate- and high-transmission settings. This study's findings support a transition in existing malaria control methods from targeting specific villages based on malaria metrics to strategies targeting interconnected “neighborhoods” that include influential connected villages that share a flow of parasites and hosts.

Aim 2 of this dissertation (Section 3) applied a g-computation methodology [155,158,168–171] to estimate the average causal effect of out-of-village working activities on malaria exposure. In g-computation, the standardized mean outcome is the weighted average of the conditional means using the prevalence of each stratum l of the vector of confounders L in the study population as weights. With this framework, this study was able to simulate multiple scenarios of mobility restrictions (by proportion of travelers, gender, and age) to quantify the impact of such restriction policies on malaria exposure reduction in the Peruvian Amazon. This study stressed the importance of human population mobility in supporting malaria transmission in the Peruvian Amazon. It also demonstrated the relevance of targeting key subpopulations when creating interventions based on employment activities. Targeting males and adults (18 years and older) groups had the greatest influence on malaria seropositivity. Finally, across all these scenarios, the effect of working activities on malaria exposure was far stronger in rural areas than in peri-urban areas. The findings of this study are important for tailoring existing and future malaria elimination strategies in the Amazon region and identifying sub-populations with important potential for intensified epidemiological surveillance.

Finally, Aim 3 of this dissertation (Section 4) leveraged GPS-tracking data from a proof-of-concept study in a village with high malaria incidence in the Loreto Region of Peru. Multiple detailed mobility metrics were computed using GPS data records. In addition, a novel non-parametric Bayesian framework was applied, which addresses several limitations of existing segmentation methods and state-space models in movement analysis. Using this detailed information, this study showed that the high interaction between Amazon villages for reasons

such as labor, commerce, or recreation may sustain these high endemicity levels through increased exposure to the malaria parasite and in consequence the importation risk.

The implications of this work are threefold. First, the results can be used to better understand the potential of human mobility in sustaining malaria endemicity in rural villages in the Peruvian Amazon. Second, the findings can be used to inform malaria control strategies and policies in the Amazon region that incorporate connectivity structure to prioritize connected areas instead of single villages and intensify malaria screening in sub-populations defined by their mobility patterns. Finally, the methodology implemented in this dissertation can be further adapted to other infectious diseases to test the importance of human mobility on their transmission, particularly in rural Amazonian regions with potential to expand to other contexts.

5.2. Methodological contributions

The methodological approaches applied in this dissertation have not been traditionally used in malaria epidemiology in the Amazonian region. The estimation of the etiological effect of human population mobility on malaria risk has been elusive in rural areas in Latin America where moderate-to-high malaria transmission occurs. The network analysis used in Aim 1 was applied previously to international migration [84], as well as to other diseases in the Brazilian Amazon [15,97,238]. The novelty of this study relies on three characteristics. First, the analysis of a broad set of centrality metrics and multiple versions of weights based on population, deforested area, distance, and travel time. Second, the coverage of an extensive territory with more than 1,600 villages across the Loreto region. And finally, the focus on rural areas and the use of watershed composition for the creation of the graph representation.

In Aim 2, this study made use of causal inference methods (parametric g-formula) to simulate various hypothetical scenarios targeting out-of-village working activities and determine their potential health benefits in terms of malaria exposure reduction. The use of such causal inference methods in the literature of malaria epidemiology is still in its infancy [160,239]. This research is intended to expand the set of tools applied for malaria epidemiology, particularly in remote areas, to better simulate interventions and guide control and elimination strategies. Finally, this study presented and tested a novel set of policy interventions that included scenarios involving mobility restrictions (by proportion of travelers, gender, and age) in peri-urban and rural settings.

In Aim 3, this dissertation explored the use on novel time-weighted spatial averaging approaches (TWSA). In particular, the granular GPS tracking data collected in this study was leveraged by the use of movement ecology methods such as the mixed-membership method for movement (M4) that provides segment-level behavioral state estimation. This approach as well as other movement ecology methods such as utilization distribution or kernel density, and other post-hoc analyses such as step selection functions (SSF)[199,234,235] were only traditionally used for the analysis of non-human animals. There is still a dearth of research evaluating these techniques for infectious disease epidemiology that involve human mobility.

Finally, capturing signals of human mobility and its impact on infectious diseases remains elusive and difficult to apply to most malaria control initiatives. Overall, this dissertation presented a multilevel methodology that can be extended to other infectious diseases and regions.

This approach comprises 1) connectivity metrics at the village level to understand village level flow, 2) standardized travel questionnaires to understand individual importation risk and g-computation simulation to design targeted populations, and 3) GPS-tracking to characterize detailed mobility patterns to identify sub-populations for improved malaria screening upon departure or return.

5.3. Policy implications

Overall, the findings in this dissertation suggest that human mobility is an important driver of malaria in the Amazon region and should be incorporated into current malaria control strategies. In the Peruvian Amazon in particular, control strategies may shift from prioritizing singular locations based on local malaria metrics to strategies based on connectivity neighborhoods that include influential connected villages under active landscape modification. As villages and cities continue to expand, human mobility intensifies. The magnitude of connectivity in these areas is conditioned by the size of the cities as a proxy of the number of services and commerce in place. This results in increased connectivity, leading to a complex network of cities/villages that influence the transmission of the malaria parasite and which should also be considered in regional and urban planning.

Current literature and malaria elimination programs have placed a greater importance on HPM in low-transmission and close-to-elimination settings than in moderate- and high-transmission settings [61–63,82,130]. However, the findings of this dissertation challenge this previous knowledge by showing that higher malaria endemicity was observed in high centrality villages (Aim 1). In addition, policies to restrict human mobility have a greater impact in rural villages

with a higher malaria incidence (Aim 2). Finally, in villages with high local malaria transmission, particular human mobility patterns contribute to malaria importation and sustaining of malaria transmission (Aim 3). Taken together, a re-direction of current malaria control strategies to target high centrality villages and implement mobility restrictions in the Peruvian Amazon is strongly suggested, in addition to including travel information in current routine data collection so that these strategies can be modified iteratively.

Lastly, evidence and methods described in this dissertation may help to prioritize sub-populations for targeting interventions. Evidence from Vietnam [151], Cambodia [152], and Indonesia [240] has shown that improved strategies targeting actively commuting subpopulations may greatly impact malaria transmission in origin and destination localities. Interventions include early diagnosis and screening, provision of prophylaxis, improved collection of travel-related data, and case investigation to prevent secondary cases. The findings in this dissertation stress the importance of intervening in these high-mobility populations, however, further studies are needed to define the most appropriate intervention strategies. Considering the lack of accessibility to health facilities in most villages in the Peruvian Amazon [29], an important factor in implementing most strategies that target mobile populations will be community health workers (CHW), who were also a key component of previous malaria control campaigns in Peru [146].

5.4. Future directions

This dissertation illustrates the application of novel methods to study the effect of human mobility and land cover change on the epidemiology of malaria in the Peruvian Amazon. These findings may inform the direction of future research to better understand this relationship and

design targeted interventions. The methodologies presented in this dissertation could be extended to refine the conceptualization of human mobility variables. Connectivity metrics covered in Aim 1 could be further explored to include more flexible and dynamic network structures [77,80,241,242]. In rural context, important change in the mobility patterns (and in consequence the connectivity between villages) occurs seasonally. For instance, labor season for occupations such as logging, fishing, and agriculture varies throughout the year. Albeit less marked than in labor seasons, hydrological conditions affect the links between villages. In the dry season, many fluvial routes are interrupted due to the low level of the river. In contrast, in the wet season, new shortcuts emerged that favor the links between villages.

Travel questionaries applied in Aim 2 could be revised and synthesized to determine a minimum set of variables that capture enough travel information to inform models. In addition, group (i.e. family) travel patterns should be accounted since the potential for malaria importation and subsequently secondary infections (local infections acquired from an imported parasite) is larger as multiple hosts share similar travel patterns.

The adherence to the use of the GPS tracking devices tested in Aim 3 requires intense logistics and supervision that could be hard to maintain at a large geographical scale. Studies that evaluate attaching the GPS tracking devices to the commuting vehicle instead of humans should be conducted to assess their added value with lower logistic costs. In addition, community health workers (CHW) could be integrated in the malaria importation surveillance based on the CHW network deployed in the PAMAFRO project. Finally, movement ecology is an active field under

development; therefore, mobility patterns computed in Aim 3 might be further revised to incorporate human behaviors that shape travel patterns.

Following our recommendation to shift malaria strategies towards those that prioritize connectivity neighborhoods, further studies may explore the operational definition of this suggested area. This definition will be influenced by the frequency and extent of commuting but also by the malaria endemicity in the destination locations. Both might be improved with robust data collection methods able to capture high spatial and temporal resolution. In this regard, sentinel villages could be selected for detailed data collection (travel questionnaires and GPS tracking) that could inform a general connectivity model.

Lastly, recent studies in Southeast Asia have started to evaluate the use of improved malaria control strategies informed by human mobility metrics vs. standard strategies that do not consider HPM [152,243]. Future work should consider this type of community randomized trial in the Amazon region, which has unique behavioral, socio-demographic, ecological, and climatic conditions, in order to better understand the benefits provided by human mobility informed systems vs. current malaria control strategies.

5.5. Concluding remarks

In the context of climate change and an ever-growing urban population, the evidence of this study is of particular interest for a global public health audience. Addressing the aims of this dissertation provided valuable knowledge of the relationship between urban expansion, human mobility, and malaria transmission using a combination of multiple methodological approaches

from global earth observation at a regional scale to detailed individual mobility data. The findings of this study refute prior assumptions that human mobility is only of interest in low transmission settings. These findings also offer valuable information for redesigning current malaria control strategies in rural areas, where populations are particularly vulnerable to the devastating effects of malaria.

References

1. World Health Organization. *World malaria report 2022*. (World Health Organization, 2022).

2. Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. & Hay, S. I. The global distribution of clinical episodes of Plasmodium falciparum malaria. *Nature* **434**, 214–217 (2005).

3. Su, X. Human Malaria Parasite: Are We Ready for a New Species? *J. Infect. Dis.* **201**, 1453–1454 (2010).

4. Organization, W. H. World malaria report 2019. (2019).

5. Harbach, R. E. The classification of genus Anopheles (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. *Bull. Entomol. Res.* **94**, 537–553 (2004).

6. Sinka, M. E. *et al.* The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasit. Vectors* **3**, 72 (2010).

7. Vittor, A. Y. *et al.* Linking deforestation to malaria in the Amazon: characterization of the breeding habitat of the principal malaria vector, Anopheles darlingi. *Am. J. Trop. Med. Hyg.* **81**, 5–12 (2009).

8. Saavedra, M. P. *et al.* Higher risk of malaria transmission outdoors than indoors by Nyssorhynchus darlingi in riverine communities in the Peruvian Amazon. *Parasit. Vectors* **12**, 374 (2019).

9. Moreno, M. *et al.* Implications for changes in Anopheles darlingi biting behaviour in three communities in the peri-Iquitos region of Amazonian Peru. *Malar. J.* **14**, 290 (2015).

10. Moreno, M. *et al.* Intensive trapping of blood-fed Anopheles darlingi in Amazonian Peru reveals unexpectedly high proportions of avian blood-meals. *PLoS Negl. Trop. Dis.* **11**, e0005337 (2017).

11. de Barros, F. S. M. & Honório, N. A. Man biting rate seasonal variation of malaria vectors in Roraima, Brazil. *Mem. Inst. Oswaldo Cruz* **102**, 299–302 (2007).

12. Achee, N. L., Grieco, J. P., Andre, R. G., Rejmankova, E. & Roberts, D. R. A mark-release-recapture study using a novel portable hut design to define the flight behavior of Anopheles darlingi in Belize, Central America. *J. Am. Mosq. Control Assoc.* **21**, 366–379 (2005).

13. Carrasco-Escobar, G. *et al.* Micro-epidemiology and spatial heterogeneity of P . vivax parasitaemia in riverine communities of the Peruvian Amazon: A multilevel analysis. *Sci. Rep.* **7**, 8082 (2017).

14. Peeters Grietens, K. *et al.* Characterizing Types of Human Mobility to Inform Differential and Targeted Malaria Elimination Strategies in Northeast Cambodia. *Sci. Rep.* **5**, 16837 (2015).

15. Gomes, M. F. C., Codeço, C. T., Bastos, L. S. & Lana, R. M. Measuring the contribution of human mobility to malaria persistence. *Malar. J.* **19**, 404 (2020).

16. Espinoza, J. L. Malaria Resurgence in the Americas: An Underestimated Threat. *Pathogens* **8**, (2019).

17. Recht, J. *et al.* Malaria in Brazil, Colombia, Peru and Venezuela: current challenges in malaria control and elimination. *Malar. J.* **16**, 273 (2017).

18. Conn, J. E., Grillet, M. E., Correa, M. & Sallum, M. A. M. Malaria Transmission in South America—Present Status and Prospects for Elimination. *Malar. Elimin. - Leap Forw.* (2018) doi:10.5772/intechopen.76964.

19. Hiwat, H. & Bretas, G. Ecology of Anopheles darlingi Root with respect to vector importance: a review. *Parasit. Vectors* **4**, 177 (2011).

20. Prussing, C. *et al.* Malaria vector species in Amazonian Peru co-occur in larval habitats but have distinct larval microbial communities. *PLoS Negl. Trop. Dis.* **13**, e0007412 (2019).

21. Rozendaal, J. A. Relations between Anopheles darlingi breeding habitats, rainfall, river level and malaria transmission rates in the rain forest of Suriname. *Med. Vet. Entomol.* **6**, 16–22 (1992).

22. Charlwood, J. D. Biological variation in Anopheles darlingi Root. *Mem. Inst. Oswaldo Cruz* **91**, 391–398 (1996).

23. Rosas-Aguirre, A. *et al.* Epidemiology of Plasmodium vivax Malaria in Peru. *Am. J. Trop. Med. Hyg.* (2016) doi:10.4269/ajtmh.16-0268.

24. Ministerio de Salud del Peru. Sala de Situación de Salud: Malaria 2019. (2019).

25. Griffing, S. M., Gamboa, D. & Udhayakumar, V. The history of 20th century malaria control in Peru. *Malar. J.* **12**, 303 (2013).

26. Soto-Calle, V. *et al.* Spatio-temporal analysis of malaria incidence in the Peruvian Amazon Region between 2002 and 2013. *Sci. Rep.* **7**, 40350 (2017).

27. Alves, F. P. *et al.* High prevalence of asymptomatic Plasmodium vivax and Plasmodium falciparum infections in native Amazonian populations. *Am. J. Trop. Med. Hyg.* **66**, 641–648 (2002).

28. Cotter, C. *et al.* The changing epidemiology of malaria elimination: new strategies for new challenges. *The Lancet* **382**, 900–911 (2013).

29. Carrasco-Escobar, G., Manrique, E., Tello-Lizarraga, K. & Miranda, J. J. Travel Time to Health Facilities as a Marker of Geographical Accessibility Across Heterogeneous Land Coverage in Peru. *Front. Public Health* **8**, (2020).

30. Solano-Villarreal, E. *et al.* Malaria risk assessment and mapping using satellite imagery and boosted regression trees in the Peruvian Amazon. *Sci. Rep.* **9**, 1–12 (2019).

31. Parker, B. S. *et al.* Hyperendemic malaria transmission in areas of occupation-related travel in the Peruvian Amazon. *Malar. J.* **12**, 178 (2013).

32. Aramburú Guarda, J., Ramal Asayag, C. & Witzig, R. Malaria reemergence in the Peruvian Amazon region. *Emerg. Infect. Dis.* **5**, 209–215 (1999).

33. Chowell, G., Munayco, C. V., Escalante, A. A. & McKenzie, F. E. The spatial and temporal patterns of falciparum and vivax malaria in Perú: 1994–2006. *Malar. J.* **8**, 142 (2009).

34. Ramal, C., Vásquez, J., Magallanes, J. & Carey, C. Variables climáticas y transmisión de malaria en Loreto, Perú: 1995-2007. *Rev. Peru. Med. Exp. Salud Publica* **26**, 9–14 (2009).

35. Gagnon, A. S., Smoyer-Tomic, K. E. & Bush, A. B. G. The El Niño southern oscillation and malaria epidemics in South America. *Int. J. Biometeorol.* **46**, 81–89 (2002).

36. Guthmann, J. P., Llanos-Cuentas, A., Palacios, A. & Hall, A. J. Environmental factors as determinants of malaria risk. A descriptive study on the northern coast of Peru. *Trop. Med. Int. Health TM IH* **7**, 518–525 (2002).

37. Ventosilla, P., Huarcaya, E., Gutierrez, P. & Chauca, J. A statistical model for assessing the relationship between meteorological variables and the incidence of Plasmodium falciparum and Plasmodium vivax in a Peruvian endemic area. *Int. J. Environ. Health* (2008).

38. Phillips, D. R. Urbanization and human health. *Parasitology* **106 Suppl**, S93-107 (1993).

39. Harpham, T. Urbanisation and health in transition. *The Lancet* **349**, S11–S13 (1997).

40. Hay, S. I., Guerra, C. A., Tatem, A. J., Atkinson, P. M. & Snow, R. W. Urbanization, malaria transmission and disease burden in Africa. *Nat. Rev. Microbiol.* **3**, 81–90 (2005).

41. Tatem, A. J., Gething, P. W., Smith, D. L. & Hay, S. I. Urbanization and the global malaria recession. *Malar. J.* **12**, 1–11 (2013).

42. Santos-Vega, M., Bouma, M. J., Kohli, V. & Pascual, M. Population density, climate variables and poverty synergistically structure spatial risk in urban malaria in India. *PLoS Negl. Trop. Dis.* **10**, e0005155 (2016).

43. Alimi, T. O. *et al.* Predicting potential ranges of primary malaria vectors and malaria in northern South America based on projected changes in climate, land cover and human population. *Parasit. Vectors* **8**, 431 (2015).

44. MacDonald, A. J. & Mordecai, E. A. Amazon deforestation drives malaria transmission, and malaria burden reduces forest clearing. *Proc. Natl. Acad. Sci.* **116**, 22212–22218 (2019).

45. Pimenta, P. F. *et al.* An overview of malaria transmission from the perspective of Amazon Anopheles vectors. *Mem. Inst. Oswaldo Cruz* **110**, 23–47 (2015).

46. Patz JA, Epstein PR, Burke TA, & Balbus JM. GLobal climate change and emerging infectious diseases. *JAMA* **275**, 217–223 (1996).

47. Parham, P. E. & Michael, E. Modeling the effects of weather and climate change on malaria transmission. *Environ. Health Perspect.* **118**, 620–626 (2010).

48. Castro, M. C. *et al.* Development, environmental degradation, and disease spread in the Brazilian Amazon. *PLOS Biol.* **17**, e3000526 (2019).

49. Feachem, R. G. A. *et al.* Malaria eradication within a generation: ambitious, achievable, and necessary. *The Lancet* **394**, 1056–1112 (2019).

50. Vittor, A. Y. *et al.* The effect of deforestation on the human-biting rate of Anopheles darlingi, the primary vector of Falciparum malaria in the Peruvian Amazon. *Am. J. Trop. Med. Hyg.* **74**, 3–11 (2006).

51. de Castro, M. C., Monte-Mór, R. L., Sawyer, D. O. & Singer, B. H. Malaria risk on the Amazon frontier. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2452–2457 (2006).

52. de Oliveira Padilha, M. A. *et al.* Comparison of malaria incidence rates and socioeconomic-environmental factors between the states of Acre and Rondônia: a spatio-temporal modelling study. *Malar. J.* **18**, 306 (2019).

53. Myers, S. S. *et al.* Human health impacts of ecosystem alteration. *Proc. Natl. Acad. Sci.* **110**, 18753–18760 (2013).

54. Gottdenker, N. L., Streicker, D. G., Faust, C. L. & Carroll, C. R. Anthropogenic land use change and infectious diseases: a review of the evidence. *EcoHealth* **11**, 619–632 (2014).

55. Gibb, R. *et al.* Zoonotic host diversity increases in human-dominated ecosystems. *Nature* **584**, 398–402 (2020).

56. Stefani, A. *et al.* Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. *Malar. J.* **12**, 192 (2013).

57. Baeza, A., Santos-Vega, M., Dobson, A. P. & Pascual, M. The rise and fall of malaria under land-use change in frontier regions. *Nat. Ecol. Evol.* **1**, 108–108 (2017).

58. Chaves, L. S. M., Conn, J. E., López, R. V. M. & Sallum, M. A. M. Abundance of impacted forest patches less than 5 km 2 is a key driver of the incidence of malaria in Amazonian Brazil. *Sci. Rep.* **8**, 1–11 (2018).

59. da Silva, N. S. *et al.* Epidemiology and control of frontier malaria in Brazil: lessons from community-based studies in rural Amazonia. *Trans. R. Soc. Trop. Med. Hyg.* **104**, 343–350 (2010).

60. Conn, J. E. *et al.* Emergence of a new neotropical malaria vector facilitated by human migration and changes in land use. *Am. J. Trop. Med. Hyg.* **66**, 18–22 (2002).

61. Pindolia, D. K. *et al.* Human movement data for malaria control and elimination strategic planning. *Malar. J.* **11**, 205 (2012).

62. Mendis, K. *et al.* From malaria control to eradication: The WHO perspective. *Trop. Med. Int. Health TM IH* **14**, 802–809 (2009).

63. WHO, Malaria Policy Advisory Committee. *Meeting report of the WHO Evidence Review Group on the assessment of malariogenic potential to inform elimination strategies and plans to prevent re-establishment of malaria*. https://www.who.int/malaria/mpac/mpac-april2019-session7-report-erg-malariogenic-potential.pdf?ua=1 (2018).

64. Organization, W. H. *A framework for malaria elimination*. (World Health Organization, 2017).

65. Sturrock, H. J. W., Roberts, K. W., Wegbreit, J., Ohrt, C. & Gosling, R. D. Tackling Imported Malaria: An Elimination Endgame. *Am. J. Trop. Med. Hyg.* **93**, 139–144 (2015).

66. Le Menach, A. *et al.* Travel risk, malaria importation and malaria transmission in Zanzibar. *Sci. Rep.* **1**, (2011).

67. Wesolowski, A. *et al.* Quantifying the Impact of Human Mobility on Malaria. *Science* **338**, 267–270 (2012).

68. Ruktanonchai, N. W. *et al.* Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data. *PLOS Comput. Biol.* **12**, e1004846 (2016).

69. Ihantamalala, F. A. *et al.* Estimating sources and sinks of malaria parasites in Madagascar. *Nat. Commun.* **9**, 3897 (2018).

70. Buckee, C. O., Wesolowski, A., Eagle, N. N., Hansen, E. & Snow, R. W. Mobile phones and malaria: Modeling human and parasite travel. *Travel Med. Infect. Dis.* **11**, 15–22 (2013).

71. Tatem, A. J. *et al.* Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malar. J.* **13**, 52 (2014).

72. Deville, P. *et al.* Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci.* **111**, 15888–15893 (2014).

73. Wesolowski, A., Buckee, C. O., Engø-Monsen, K. & Metcalf, C. J. E. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. *J. Infect. Dis.* **214**, S414–S420 (2016).

74. Carrasco-Escobar, G. *et al.* Open-Source 3D Printable GPS Tracker to Characterize the Role of Human Population Movement on Malaria Epidemiology in River Networks: A Proof-of-Concept Study in the Peruvian Amazon. *Front. Public Health* **8**, (2020).

75. Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network Analysis in the Social Sciences. *Science* **323**, 892–895 (2009).

76. Newman, M., Barabási, A.-L. & Watts, D. J. *The Structure and Dynamics of Networks*. (Princeton University Press, 2006).

77. Barabási, A.-L. Scale-Free Networks: A Decade and Beyond. *Science* **325**, 412–413 (2009).

78. Poulin, R. Network analysis shining light on parasite ecology and diversity. *Trends Parasitol.* **26**, 492–498 (2010).

79. Freeman, L. C. A Set of Measures of Centrality Based on Betweenness. *Sociometry* **40**, 35–41 (1977).

80. Barrat, A., Barthélemy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3747–3752 (2004).

81. Buckee, C. O., Bull, P. C. & Gupta, S. Inferring malaria parasite population structure from serological networks. *Proc. R. Soc. B Biol. Sci.* **276**, 477–485 (2009).

82. Tatem, A. J. & Smith, D. L. International population movements and regional Plasmodium falciparum malaria elimination strategies. *Proc. Natl. Acad. Sci.* **107**, 12222–12227 (2010).

83. Pindolia, D. K. *et al.* The demographics of human and malaria movement and migration patterns in East Africa. *Malar. J.* **12**, 397 (2013).

84. Huang, Z. & Tatem, A. J. Global malaria connectivity through air travel. *Malar. J.* **12**, 269 (2013).

85. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).

86. Sorichetta, A. *et al.* Mapping internal connectivity through human migration in malaria endemic countries. *Sci. Data* **3**, 160066 (2016).

87. Marshall, J. M. *et al.* Mathematical models of human mobility of relevance to malaria transmission in Africa. *Sci. Rep.* **8**, 7713 (2018).

88. Prothero, R. M. Disease and mobility: a neglected factor in epidemiology. *Int. J. Epidemiol.* **6**, 259–267 (1977).

89. González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).

90. Hasan, S., Schneider, C. M., Ukkusuri, S. V. & González, M. C. Spatiotemporal Patterns of Urban Human Mobility. *J. Stat. Phys.* **151**, 304–318 (2013).

91. Wesolowski, A. *et al.* Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Sci. Rep.* **4**, 5678 (2014).

92. Perchoux, C., Chaix, B., Cummins, S. & Kestens, Y. Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. *Health Place* **21**, 86–93 (2013).

93. Fornace, K. M. *et al.* Use of mobile technology-based participatory mapping approaches to geolocate health facility attendees for disease surveillance in low resource settings. *Int. J. Health Geogr.* **17**, (2018).

94. Fornace, K. M. *et al.* Local human movement patterns and land use impact exposure to zoonotic malaria in Malaysian Borneo. *eLife* **8**, e47602 (2019).

95. Kar, N. P., Kumar, A., Singh, O. P., Carlton, J. M. & Nanda, N. A review of malaria transmission dynamics in forest ecosystems. *Parasit. Vectors* **7**, 265 (2014).

96. Stoddard, S. T. *et al.* The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl. Trop. Dis.* **3**, e481 (2009).

97. Angelo, J. R. *et al.* The role of spatial mobility in malaria transmission in the Brazilian Amazon: The case of Porto Velho municipality, Rondônia, Brazil (2010-2012). *PLOS ONE* **12**, e0172330 (2017).

98. Carrasco-Escobar, G. *et al.* Use of open mobile mapping tool to assess human mobility traceability in rural offline populations with contrasting malaria dynamics. *PeerJ* **7**, e6298 (2019).

99. Sevilla-Casas, E. Human mobility and malaria risk in the Naya river basin of Colombia. *Soc. Sci. Med. 1982* **37**, 1155–1167 (1993).

100. Johansen, I. C., Rodrigues, P. T. & Ferreira, M. U. Human mobility and urban malaria risk in the main transmission hotspot of Amazonian Brazil. *PLOS ONE* **15**, e0242357 (2020).

101. Paz-Soldan, V. A. *et al.* Strengths and Weaknesses of Global Positioning System (GPS) Data-Loggers and Semi-structured Interviews for Capturing Fine-scale Human Mobility: Findings from Iquitos, Peru. *PLoS Negl. Trop. Dis.* **8**, e2888 (2014).

102. Vazquez-Prokopec, G. M. *et al.* Using GPS Technology to Quantify Human Mobility, Dynamic Contacts and Infectious Disease Dynamics in a Resource-Poor Urban Environment. *PLoS ONE* **8**, (2013).

103. Erhart, A. *et al.* FOREST MALARIA IN VIETNAM: A CHALLENGE FOR CONTROL. *Am. J. Trop. Med. Hyg.* **70**, 110–118 (2004).

104. Chuquiyauri, R. *et al.* Socio-demographics and the development of malaria elimination strategies in the low transmission setting. *Acta Trop.* **121**, 292–302 (2012).

105. Chuquiyauri, R. *et al.* Microgeographical differences of Plasmodium vivax relapse and re-infection in the Peruvian Amazon. *Am. J. Trop. Med. Hyg.* **89**, 326–338 (2013).

106. Rosas-Aguirre, A. *et al.* Hotspots of Malaria Transmission in the Peruvian Amazon: Rapid Assessment through a Parasitological and Serological Survey. *PloS One* **10**, e0137458 (2015).

107. Smith, L., Foley, L. & Panter, J. Activity spaces in studies of the environment and physical activity: A review and synthesis of implications for causality. *Health Place* **58**, 102113 (2019).

108. Patterson, Z. & Farber, S. Potential Path Areas and Activity Spaces in Application: A Review. *Transp. Rev.* **35**, 679–700 (2015).

109. Ministerio de Salud del Peru. Sala de Situación de Salud: Malaria 2012. (2012).

110. Branch, O. *et al.* Clustered local transmission and asymptomatic Plasmodium falciparum and Plasmodium vivax malaria infections in a recently emerged, hypoendemic Peruvian Amazon community. *Malar. J.* **4**, 27 (2005).

111. Moreno, M. *et al.* Insights into Plasmodium vivax Asymptomatic Malaria Infections and Direct Skin-Feeding Assays to Assess Onward Malaria Transmission in the Amazon. *Am. J. Trop. Med. Hyg.* **1**, (2022).

112. Carrasco-Escobar, G. *et al.* Time-Varying Effects of Meteorological Variables on Malaria Epidemiology in the Context of Interrupted Control Efforts in the Amazon Rainforest, 2000–2017. *Front. Med.* **8**, 1649 (2021).

113. Lainhart, W. *et al.* Evidence for temporal population replacement and the signature of ecological adaptation in a major Neotropical malaria vector in Amazonian Peru. *Malar. J.* **14**, 375 (2015).

114. United Nations. The World's Cities in 2018-Data Booklet. (2018).

115. Roberts, M., Blankespoor, B., Deuskar, C. & Stewart, B. *Urbanization and development: Is latin america and the caribbean different from the rest of the world?* (The World Bank, 2017).

116. Instituto Nacional de Estadística e Informática. *Perú: Perfil sociodemográfico. Informe nacional. Censos Nacionales 2017: XII de población, VII de vivienda y III de comunidades indígenas*. (Instituto Nacional de Estadística e Informática Lima, 2018).

117. Brierley, C. K., Suarez, N., Arora, G. & Graham, D. Healthcare Access and Health Beliefs of the Indigenous Peoples in Remote Amazonian Peru. *Am. J. Trop. Med. Hyg.* **90**, 180–183 (2014).

118. Ministerio de Salud. Instituto Nacional de Salud. *Manual de procedimientos de laboratorio para el diagnóstico de malaria*. (MINSA, 2003).

119. Ministerio de Salud del Perú. Norma Técnica Para la Atención de la Malaria y Malaria Severa en el Perú; NTS Nro. 054-MINSA/DGSP-V.01, modificada en Febrero 2015. http://bvs.minsa.gob.pe/local/minsa/2949.pdf (2015).

120. Hansen, M. C. *et al.* High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).

121. Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).

122. Verdin, K. L. & Verdin, J. P. A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology* vol. 218 12 (1999).

123. Aybar, C., Wu, Q., Bautista, L., Yali, R. & Barja, A. rgee: An R package for interacting with Google Earth Engine. *J. Open Source Softw.* **5**, 2272 (2020).

124. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).

125. Pedersen, T. tidygraph: A Tidy API for Graph Manipulation. https://tidygraph.data-imaginist.com/ (2023).

126. Pedersen, T. ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. https://github.com/thomasp85/ggraph (2023).

127. Gansner, E. R., Koren, Y. & North, S. Graph Drawing by Stress Majorization. in *Graph Drawing* (ed. Pach, J.) 239–250 (Springer, 2005). doi:10.1007/978-3-540-31843-9_25.

128. Guerra, C. A., Citron, D. T., García, G. A. & Smith, D. L. Characterising malaria connectivity using malaria indicator survey data. *Malar. J.* **18**, 440 (2019).

129. Ruktanonchai, N. W., Ruktanonchai, C. W., Floyd, J. R. & Tatem, A. J. Using Google Location History data to quantify fine-scale human mobility. *Int. J. Health Geogr.* **17**, 28 (2018).

130. Sturrock, H. J. W. *et al.* Targeting Asymptomatic Malaria Infections: Active Surveillance in Control and Elimination. *PLoS Med* **10**, e1001467 (2013).

131. Ruktanonchai, N. W. *et al.* Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data. *PLOS Comput. Biol.* **12**, e1004846 (2016).

132. Manrique, P. *et al.* Microsatellite analysis reveals connectivity among geographically distant transmission zones of Plasmodium vivax in the Peruvian Amazon: A critical barrier to regional malaria elimination. *PLoS Negl. Trop. Dis.* **13**, e0007876 (2019).

133. Carrasco-Escobar, G. *et al.* Quantifying the effect of human population mobility on malaria risk in the Peruvian Amazon. *R. Soc. Open Sci.* **9**, 211611.

134. Carrasco-Escobar, G. *et al.* Effect of out-of-village working activities on recent malaria exposure in the Peruvian Amazon using parametric g-formula. *Sci. Rep.* **12**, 19144 (2022).

135. Pizzitutti, F., Mena, C. F., Feingold, B. & Pan, W. K. Modeling asymptomatic infections and work-related human circulation as drivers of unstable malaria transmission in low-prevalence areas: A study in the Northern Peruvian Amazon. *Acta Trop.* **197**, 104909 (2019).

136. Castro, M. C. *et al.* Development, environmental degradation, and disease spread in the Brazilian Amazon. *PLOS Biol.* **17**, e3000526 (2019).

137. Nilsen, K. *et al.* A review of geospatial methods for population estimation and their use in constructing reproductive, maternal, newborn, child and adolescent health service indicators. *BMC Health Serv. Res.* **21**, 370 (2021).

138. Boo, G. *et al.* High-resolution population estimation using household survey data and building footprints. *Nat. Commun.* **13**, 1330 (2022).

139. Weber, E. M. *et al.* Census-independent population mapping in northern Nigeria. *Remote Sens. Environ.* **204**, 786–798 (2018).

140. Leyk, S. *et al.* The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth Syst. Sci. Data* **11**, 1385–1409 (2019).

141. Carrasco-Escobar, G. *et al.* High prevalence of very-low Plasmodium falciparum and Plasmodium vivax parasitaemia carriers in the Peruvian Amazon: insights into local and occupational mobility-related transmission. *Malar. J.* **16**, 415 (2017).

142. Roshanravan, B. *et al.* Endemic malaria in the Peruvian Amazon region of Iquitos. *Am. J. Trop. Med. Hyg.* **69**, 45–52 (2003).

143. Rovira-Vallbona, E. *et al.* Predominance of asymptomatic and sub-microscopic infections characterizes the Plasmodium gametocyte reservoir in the Peruvian Amazon. *PLoS Negl. Trop. Dis.* **11**, e0005674 (2017).

144. World Health Organization. *World Malaria Report 2021*. (World Health Organization, 2021).

145. Centro Nacional de Epidemiología Prevención y Control de Enfermedades. Salas de situación semanal. https://www.dge.gob.pe/portalnuevo/publicaciones/salas-de-situacion-semanal/.

146. Organismo Andino de Salud-PAMAFRO. *Compartiendo lecciones aprendidas*. (2009).

147. Ministerio de Salud. Documento técnico: "Plan Malaria Cero 2017- 2021". https://www.gob.pe/institucion/minsa/noticias/14025-ministerio-de-salud-aprueba-documento-tecnico-plan-malaria-cero-2017-2021.

148. Ministerio de Salud. Documento Técnico: Plan hacia la eliminación de la Malaria en el Perú 2022-2030. Resolución Ministerial N° 034-2022-MINSA. https://www.gob.pe/institucion/minsa/normas-legales/2723757-034-2022-minsa.

149. World Health Organization. *Malaria Elimination: A Field Manual for Low and Moderate Endemic Countries*. (WHO, 2007).

150. World Health Organization. *A framework for malaria elimination*. (WHO, 2017).

151. Canavati, S. E. *et al.* Targeting high risk forest goers for malaria elimination: a novel approach for investigating forest malaria to inform program intervention in Vietnam. *BMC Infect. Dis.* **20**, 757 (2020).

152. Kunkel, A. *et al.* Choosing interventions to eliminate forest malaria: preliminary results of two operational research studies inside Cambodian forests. *Malar. J.* **20**, 51 (2021).

153. Sanann, N. *et al.* Forest work and its implications for malaria elimination: a qualitative study. *Malar. J.* **18**, 376 (2019).

154. Eisen, E. A., Elser, H. & Picciotto, S. Working: The Role of Occupational Epidemiology. *Am. J. Epidemiol.* **191**, 237–240 (2022).

155. Taubman, S. L., Robins, J. M., Mittleman, M. A. & Hernán, M. A. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int. J. Epidemiol.* **38**, 1599–1611 (2009).

156. Keil, A. P. & Richardson, D. B. Reassessing the Link between Airborne Arsenic Exposure among Anaconda Copper Smelter Workers and Multiple Causes of Death Using the Parametric g-Formula. *Environ. Health Perspect.* **125**, 608–614 (2017).

157. Keil, A. P., Richardson, D. B., Westreich, D. & Steenland, K. Estimating the Impact of Changes to Occupational Standards for Silica Exposure on Lung Cancer Mortality. *Epidemiol. Camb. Mass* **29**, 658–665 (2018).

158. Cole, S. R., Richardson, D. B., Chu, H. & Naimi, A. I. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am. J. Epidemiol.* **177**, 989–996 (2013).

159. Cates, J. E. *et al.* Intermittent Preventive Therapy in Pregnancy and Incidence of Low Birth Weight in Malaria-Endemic Countries. *Am. J. Public Health* **108**, 399–406 (2018).

160. Kilpatrick, K. W. & Hudgens, M. G. G-Formula for Observational Studies with Partial Interference, with Application to Bed Net Use on Malaria. Preprint at https://doi.org/10.48550/arXiv.2102.01155 (2021).

161. Delgado-Ratto, C. *et al.* Population Genetics of Plasmodium vivax in the Peruvian Amazon. *PLOS Negl Trop Dis* **10**, e0004376 (2016).

162. Van den Eede, P. *et al.* Multilocus genotyping reveals high heterogeneity and strong local population structure of the Plasmodium vivax population in the Peruvian Amazon. *Malar. J.* **9**, 151 (2010).

163. Moreno, M. *et al.* Implications for changes in Anopheles darlingi biting behaviour in three communities in the peri-Iquitos region of Amazonian Peru. *Malar. J.* **14**, 290 (2015).

164. Rosado, J. *et al.* Malaria transmission structure in the Peruvian Amazon through antibody signatures to Plasmodium vivax. *PLoS Negl. Trop. Dis.* **16**, e0010415 (2022).

165. Rosas-Aguirre, A. *et al.* Micro-heterogeneity of malaria transmission in the Peruvian Amazon: a baseline assessment underlying a population-based cohort study. *Malar. J.* **16**, 312 (2017).

166. Longley, R. J. *et al.* Development and validation of serological markers for detecting recent Plasmodium vivax infection. *Nat. Med.* **26**, 741–749 (2020).

167. Longley, R. J. *et al.* Asymptomatic Plasmodium vivax infections induce robust IgG responses to multiple blood-stage proteins in a low-transmission region of western Thailand. *Malar. J.* **16**, 178 (2017).

168. Austin, P. C. & Lee, D. S. Estimating the Net Benefit of Improvements in Hospital Performance: G-Computation With Hierarchical Regression Models. *Med. Care* **58**, 651–657 (2020).

169. Danaei, G., Pan, A., Hu, F. B. & Hernán, M. A. Hypothetical midlife interventions in women and risk of type 2 diabetes. *Epidemiol. Camb. Mass* **24**, 122–128 (2013).

170. Edwards, J. K. *et al.* Occupational radon exposure and lung cancer mortality: estimating intervention effects using the parametric G formula. *Epidemiol. Camb. Mass* **25**, 829–834 (2014).

171. Garcia-Aymerich, J., Varraso, R., Danaei, G., Camargo, C. A. & Hernán, M. A. Incidence of adult-onset asthma after hypothetical interventions on body mass index and physical activity: an application of the parametric g-formula. *Am. J. Epidemiol.* **179**, 20–26 (2014).

172. Robins, J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Model.* **7**, 1393–1512 (1986).

173. Zou, G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am. J. Epidemiol.* **159**, 702–706 (2004).

174. World Health Organization. *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. (World Health Organization, 2006).

175. Vazquez-Prokopec, G. M. *et al.* Using GPS Technology to Quantify Human Mobility, Dynamic Contacts and Infectious Disease Dynamics in a Resource-Poor Urban Environment. *PLOS ONE* **8**, e58802 (2013).

176. Tessema, S. *et al.* Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. *Elife* **8**, e43510 (2019).

177. Fuller, D. & Stanley, K. G. The future of activity space and health research. *Health Place* **58**, 102131 (2019).

178. Reinbold-Wasson, D. D. *et al.* Determinants of Anopheles seasonal distribution patterns across a forest to periurban gradient near Iquitos, Peru. *Am. J. Trop. Med. Hyg.* **86**, 459–463 (2012).

179. Vittor, A. Y. *et al.* The effect of deforestation on the human-biting rate of Anopheles darlingi, the primary vector of Falciparum malaria in the Peruvian Amazon. *Am. J. Trop. Med. Hyg.* **74**, 3–11 (2006).

180. Vittor, A. Y. *et al.* Linking deforestation to malaria in the Amazon: characterization of the breeding habitat of the principal malaria vector, Anopheles darlingi. *Am. J. Trop. Med. Hyg.* **81**, 5–12 (2009).

181. Singer, B. H. & de Castro, M. C. Agricultural colonization and malaria on the Amazon frontier. *Ann. N. Y. Acad. Sci.* **954**, 184–222 (2001).

182. Corran, P., Coleman, P., Riley, E. & Drakeley, C. Serology: a robust indicator of malaria transmission intensity? *Trends Parasitol.* **23**, 575–582 (2007).

183. Cook, J. *et al.* Using serological measures to monitor changes in malaria transmission in Vanuatu. *Malar. J.* **9**, 169 (2010).

184. Halloran, M. E. & Struchiner, C. J. Causal inference in infectious diseases. *Epidemiol. Camb. Mass* **6**, 142–151 (1995).

185. Tchetgen Tchetgen, E. J. & VanderWeele, T. J. On causal inference in the presence of interference. *Stat. Methods Med. Res.* **21**, 55–75 (2012).

186. Marshall, B. D. L. & Galea, S. Formalizing the Role of Agent-Based Modeling in Causal Inference and Epidemiology. *Am. J. Epidemiol.* **181**, 92–99 (2015).

187. World malaria report 2020. https://www.who.int/publications-detail-redirect/9789240015791.

188. Actualización Epidemiológica: Malaria - 10 de junio de 2020 - OPS/OMS | Organización Panamericana de la Salud. https://www.paho.org/es/documentos/actualizacion-epidemiologica-malaria-10-junio-2020.

189. Montenegro, C. C. *et al.* Plasmodium falciparum outbreak in native communities of Condorcanqui, Amazonas, Perú. *Malar. J.* **20**, 88 (2021).

190. Nahum, A. *et al.* Malaria Incidence and Prevalence Among Children Living in a Peri-Urban Area on the Coast of Benin, West Africa: A Longitudinal Study. *Am. J. Trop. Med. Hyg.* **83**, 465–473 (2010).

191. Gahutu, J.-B. *et al.* Prevalence and risk factors of malaria among children in southern highland Rwanda. *Malar. J.* **10**, 134 (2011).

192. Ayele, D. G., Zewotir, T. T. & Mwambi, H. G. Prevalence and risk factors of malaria in Ethiopia. *Malar. J.* **11**, 195 (2012).

193. Ferrari, G. *et al.* Identifying risk factors for Plasmodium infection and anaemia in Kinshasa, Democratic Republic of Congo. *Malar. J.* **15**, 362 (2016).

194. Essendi, W. M. *et al.* Epidemiological risk factors for clinical malaria infection in the highlands of Western Kenya. *Malar. J.* **18**, 211 (2019).

195. Ikeda, T. *et al.* Seasonally lagged effects of climatic factors on malaria incidence in South Africa. *Sci. Rep.* **7**, 2458 (2017).

196. Bozcal, E., Eldem, V., Aydemir, S. & Skurnik, M. The relationship between phylogenetic classification, virulence and antibiotic resistance of extraintestinal pathogenic Escherichia coli in İzmir province, Turkey. *PeerJ* **6**, e5470 (2018).

197. MacDonald, A. J. & Mordecai, E. A. Amazon deforestation drives malaria transmission, and malaria burden reduces forest clearing. *Proc. Natl. Acad. Sci.* **116**, 22212–22218 (2019).

198. Essé, C. *et al.* Social and cultural aspects of 'malaria' and its control in central Côte d'Ivoire. *Malar. J.* **7**, 224 (2008).

199. Carrasco-Escobar, G. *et al.* Open-Source 3D Printable GPS Tracker to Characterize the Role of Human Population Movement on Malaria Epidemiology in River Networks: A Proof-of-Concept Study in the Peruvian Amazon. *Front. Public Health* **8**, 526468 (2020).

200. Manrique, P. *et al.* Microsatellite analysis reveals connectivity among geographically distant transmission zones of Plasmodium vivax in the Peruvian Amazon: A critical barrier to regional malaria elimination. *PLoS Negl. Trop. Dis.* **13**, e0007876 (2019).

201. Bradley, J. *et al.* Infection importation: a key challenge to malaria elimination on Bioko Island, Equatorial Guinea. *Malar. J.* **14**, (2015).

202. Wesolowski, A. *et al.* Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).

203. Maheu-Giroux, M., Casapía, M. & Gyorkos, T. W. On the validity of self-reports and indirect reports to ascertain malaria prevalence in settings of hypoendemicity. *Soc. Sci. Med. 1982* **72**, 635–640 (2011).

204. Gomes, M. F. C., Codeço, C. T., Bastos, L. S. & Lana, R. M. Measuring the contribution of human mobility to malaria persistence. *Malar. J.* **19**, (2020).

205. Carrasco-Escobar, G. *et al.* Micro-epidemiology and spatial heterogeneity of P . vivax parasitaemia in riverine communities of the Peruvian Amazon: A multilevel analysis. *Sci. Rep.* **7**, 8082 (2017).

206. Smith, D. L. *et al.* Recasting the theory of mosquito-borne pathogen transmission dynamics and control. *Trans. R. Soc. Trop. Med. Hyg.* **108**, 185–197 (2014).

207. Gunderson, A. K. *et al.* Malaria Transmission and Spillover across the Peru–Ecuador Border: A Spatiotemporal Analysis. *Int. J. Environ. Res. Public. Health* **17**, (2020).

208. Carrasco-Escobar, G. *et al.* High prevalence of very-low Plasmodium falciparum and Plasmodium vivax parasitaemia carriers in the Peruvian Amazon: insights into local and occupational mobility-related transmission. *Malar. J.* **16**, (2017).

209. Marasinghe, D. H. *et al.* Risk of malaria associated with travel to malaria-endemic areas to visit friends and relatives: a population-based case-control study. *CMAJ Open* **8**, E60–E68 (2020).

210. Carrasco-Escobar, G. *et al.* Use of open mobile mapping tool to assess human mobility traceability in rural offline populations with contrasting malaria dynamics. *PeerJ* **7**, e6298 (2019).

211. Tatem, A. J. *et al.* Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malar. J.* **13**, 52 (2014).

212. Chang, H.-H. *et al.* Mapping imported malaria in Bangladesh using parasite genetic and human mobility data. *eLife* **8**,.

213. Wesolowski, A. *et al.* Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Sci. Rep.* **4**, 5678 (2014).

214. Carrasco-Escobar, G. *et al.* Open-Source 3D Printable GPS Tracker to Characterize the Role of Human Population Movement on Malaria Epidemiology in River Networks: A Proof-of-Concept Study in the Peruvian Amazon. *Front. Public Health* **8**, 526468 (2020).

215. Moore, K. L., Neugebauer, R., van der Laan, M. J. & Tager, I. B. Causal inference in epidemiological studies with strong confounding. *Stat. Med.* **31**, 1380–1404 (2012).

216. Hernán, M. A. & Robins, J. M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am. J. Epidemiol.* **183**, 758–764 (2016).

217. Dwomoh, D. *et al.* Impact evaluation of the free maternal healthcare policy on the risk of neonatal and infant deaths in four sub-Saharan African countries: a quasi-experimental design with propensity score Kernel matching and difference in differences analysis. *BMJ Open* **10**, e033356 (2020).

218. Cates, J. E. *et al.* Malaria, malnutrition, and birthweight: A meta-analysis using individual participant data. *PLOS Med.* **14**, e1002373 (2017).

219. Leopold, S. J. *et al.* Investigating causal pathways in severe falciparum malaria: A pooled retrospective analysis of clinical studies. *PLoS Med.* **16**, (2019).

220. Moreno-Gutierrez, D. *et al.* Effectiveness of a Malaria Surveillance Strategy Based on Active Case Detection during High Transmission Season in the Peruvian Amazon. *Int. J. Environ. Res. Public. Health* **15**, (2018).

221. Parker, B. S. *et al.* Hyperendemic malaria transmission in areas of occupation-related travel in the Peruvian Amazon. *Malar. J.* **12**, 178 (2013).

222. R: The R Project for Statistical Computing. https://www.r-project.org/.

223. Cullen, J. *joshcullen/bayesmove*. (2020).

224. Kaura, A. *et al.* Invasive versus non-invasive management of older patients with non-ST elevation myocardial infarction (SENIOR-NSTEMI): a cohort study based on routine clinical data. *The Lancet* **396**, 623–634 (2020).

225. DAGitty v3.0. http://www.dagitty.net/dags.html.

226. Somers, E. C. *et al.* Tocilizumab for Treatment of Mechanically Ventilated Patients With COVID-19. *Clin. Infect. Dis.* (2020) doi:10.1093/cid/ciaa954.

227. Torres, K. *et al.* Malaria Situation in the Peruvian Amazon during the COVID-19 Pandemic. *Am. J. Trop. Med. Hyg.* **103**, 1773–1776 (2020).

228. Papaioannou, I., Utzinger, J. & Vounatsou, P. Malaria-anemia comorbidity prevalence as a measure of malaria-related deaths in sub-Saharan Africa. *Sci. Rep.* **9**, (2019).

229. Roshanravan, B. *et al.* Endemic malaria in the Peruvian Amazon region of Iquitos. *Am. J. Trop. Med. Hyg.* **69**, 45–52 (2003).

230. A, R.-A. *et al.* Integrating Parasitological and Entomological Observations to Understand Malaria Transmission in Riverine Villages in the Peruvian Amazon. *J. Infect. Dis.* **223**, S99–S110 (2021).

231. Chaix, B. *et al.* The 'Residential' Effect Fallacy in Neighborhood and Health Studies: Formal Definition, Empirical Identification, and Correction. *Epidemiol. Camb. Mass* **28**, 789–797 (2017).

232. Martens, P. & Hall, L. Malaria on the move: human population movement and malaria transmission. *Emerg. Infect. Dis.* **6**, 103–109 (2000).

233. Saavedra, M. P. *et al.* Higher risk of malaria transmission outdoors than indoors by Nyssorhynchus darlingi in riverine communities in the Peruvian Amazon. *Parasit. Vectors* **12**, 374 (2019).

234. Fornace, K. M. *et al.* Local human movement patterns and land use impact exposure to zoonotic malaria in Malaysian Borneo. *eLife* **8**, e47602 (2019).

235. Cullen, J. A., Poli, C. L., Fletcher, R. J. & Valle, D. Identifying latent behavioral states in animal movement with M4, a non-parametric Bayesian method. *bioRxiv* 2020.11.05.369702 (2021) doi:10.1101/2020.11.05.369702.

236. Searle, K. M. *et al.* Characterizing and quantifying human movement patterns using GPS data loggers in an area approaching malaria elimination in rural southern Zambia. *R. Soc. Open Sci.* **4**, 170046 (2017).

237. Ueno, T. M. R. L. *et al.* Socio-Epidemiological Features and Spatial Distribution of Malaria in an Area under Mining Activity in the Brazilian Amazon Region. *Int. J. Environ. Res. Public. Health* **18**, 10384 (2021).

238. Lana, R. M., Gomes, M. F. da C., Lima, T. F. M. de, Honório, N. A. & Codeço, C. T. The introduction of dengue follows transportation infrastructure changes in the state of Acre, Brazil: A network-based analysis. *PLoS Negl. Trop. Dis.* **11**, e0006070 (2017).

239. Machekano, R., Dorsey, G. & Hubbard, A. Efficacy studies of malaria treatments in Africa: efficient estimation with missing indicators of failure. *Stat. Methods Med. Res.* **17**, 191–206 (2008).

240. Paramesti, S. I. *et al.* Migration Surveillance as a Maintenance Effort of Malaria Elimination Status (Study in Banyumas Regency, Central Java, Indonesia, 2021). *Open Access Maced. J. Med. Sci.* **10**, 1751–1758 (2022).

241. Albery, G. F., Kirkpatrick, L., Firth, J. A. & Bansal, S. Unifying spatial and social network analysis in disease ecology. *J. Anim. Ecol.* **n/a**,.

242. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).

243. Htike, W. *et al.* Reducing malaria transmission in forest-going mobile and migrant populations in Lao PDR and Cambodia: protocol for stepped-wedge cluster-randomised controlled trial. *BMC Infect. Dis.* **22**, 747 (2022).