cidacs
Centro de Integração de Dados e Conhecimentos para Saúde

Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Gonçalo Moniz

Are traditional cohorts outdated?

Limitations of the traditional approach and the building of large cohorts with administrative/routine health and social data

Mauricio L. Barreto

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

COHORT STUDY (Syn: concurrent, follow-up, incidence, longitudinal, panel, prospective study) The analytic epidemiological study in which subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed—or exposed in different degrees—to a factor or factors hypothesized to influence the occurrence of a given outcome.[1-3,5-9,13,24,42,85,128,149,224,225,800] A common feature of a cohort study is comparison of incidences in groups that differ in exposure levels. The denominators used for analysis may be persons or PERSON-TIME.[239] See also BIAS; COHORT ANALYSIS; REVERSE CAUSATION; VALIDITY.

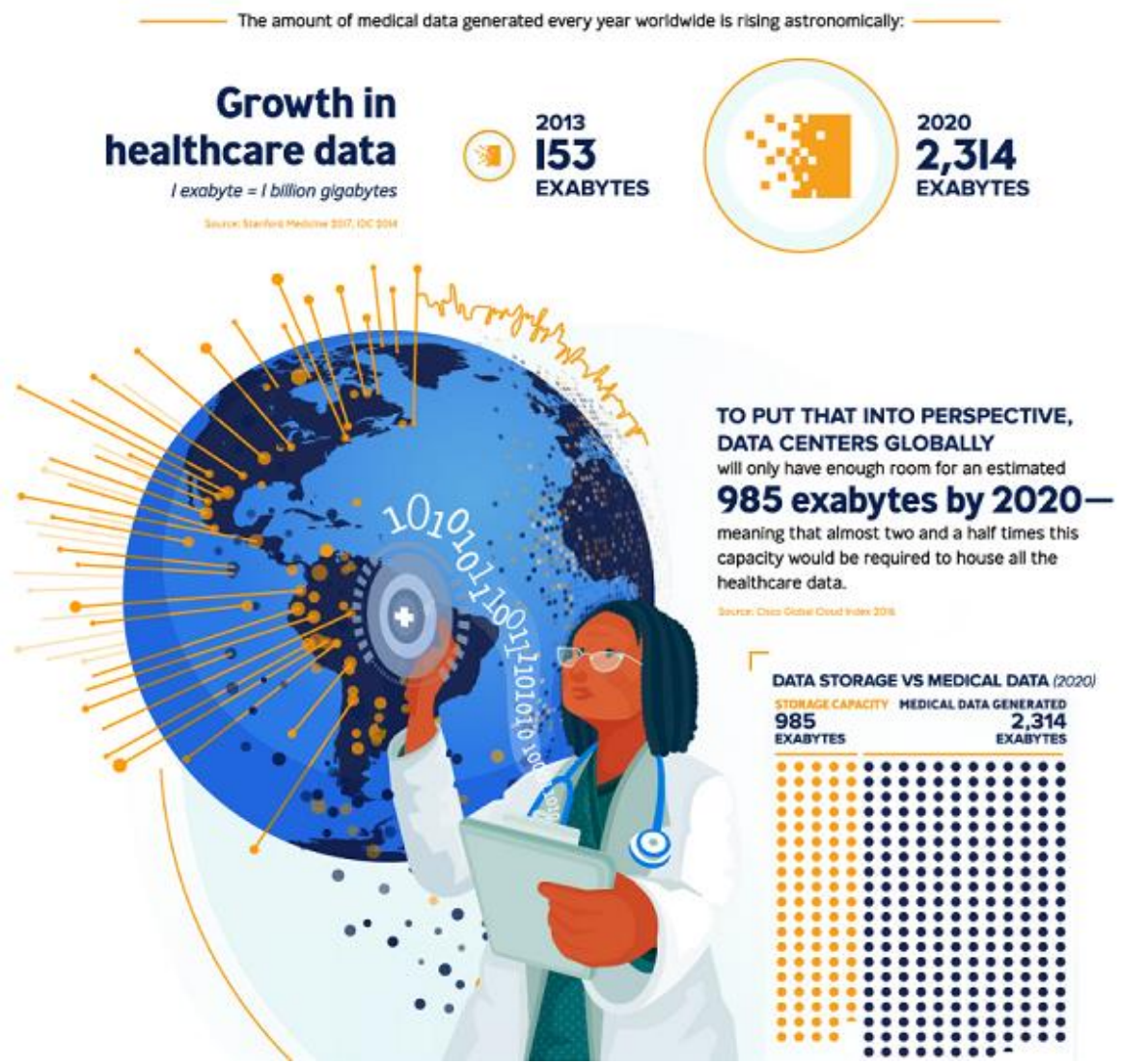
Porta, M. Dictionary of Epidemiology— Sixth edition, 2014

Cohorts **Advantages**

- Clarity of Temporal Sequence
- Allow Calculation of Incidence: Relative risk (risk ratio or rate ratio), Risk difference, Attributable proportion (attributable risk %)
- Facilitate Study of Rare Exposures:
- Allow Examination of Multiple Effects of a Single Exposure
- Avoid Selection Bias at Enrolment

Cohorts **Disadvantages**

- You may have to follow a large number of subjects for a long time.
- They are not good for rare diseases.
- They are not good for diseases with a long latency.
- Differential loss to follow-up can introduce bias.
- They can be very expensive and time-consuming.



How can we transform this growing amount of data into valuable knowledge that serves the common good?

ECO-SYSTEM OF ADMINISTRATIVE HEALTH AND SOCIAL DATA IN BRAZIL

Using administrative data has several benefits:

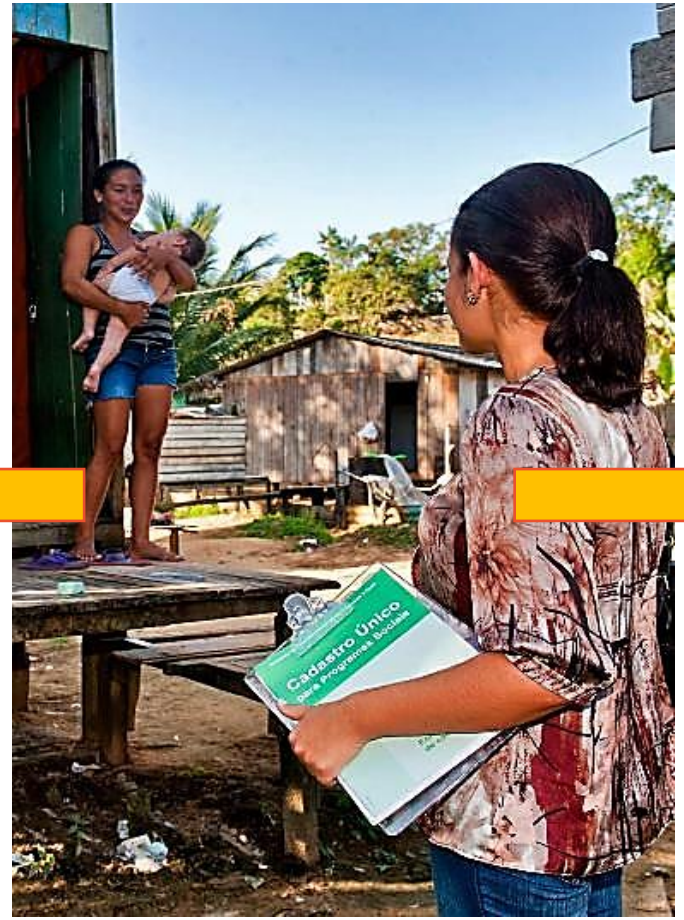
We have data for the entire population and in continuous periods, allowing the construction of a longitudinal structure that makes it possible to follow individuals over time and research critical issues;

Many challenges to connecting these data “silos”

- ✓ Data heterogeneity (accuracy, format);
- ✓ Data fragmentation (multiple databases, multiple owners/ stakeholders);
- ✓ Data availability (protection for commercial or cultural reasons, or related to personal privacy);
- ✓ Data handling (data management, data access, data quality, data querying, data sharing);
- ✓ Data privacy and integrity (prevention of corruption and hacking);
- ✓ Data conceptualisation (ontologies).

Unified Registry for Social Programmes (CadUnico)

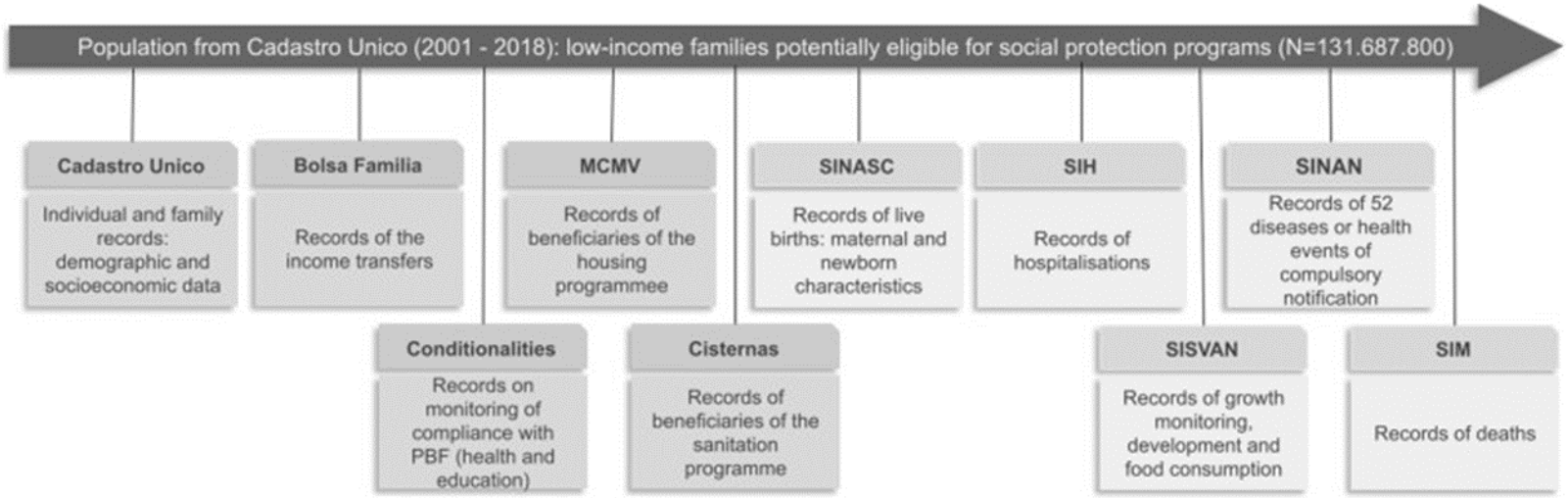
Individual level variables
✓ Age (at time of application)
✓ Sex
✓ Marital status
✓ Relationship to the person responsible for the household
✓ Race/Ethnicity
✓ Literacy status
✓ Level of education
✓ Employment status
✓ Household income (monthly)
✓ Individual income (monthly)
✓ Indigenous, *quilombola*, and other traditional population groups
✓ Experiencing homelessness?
✓ Place of birth



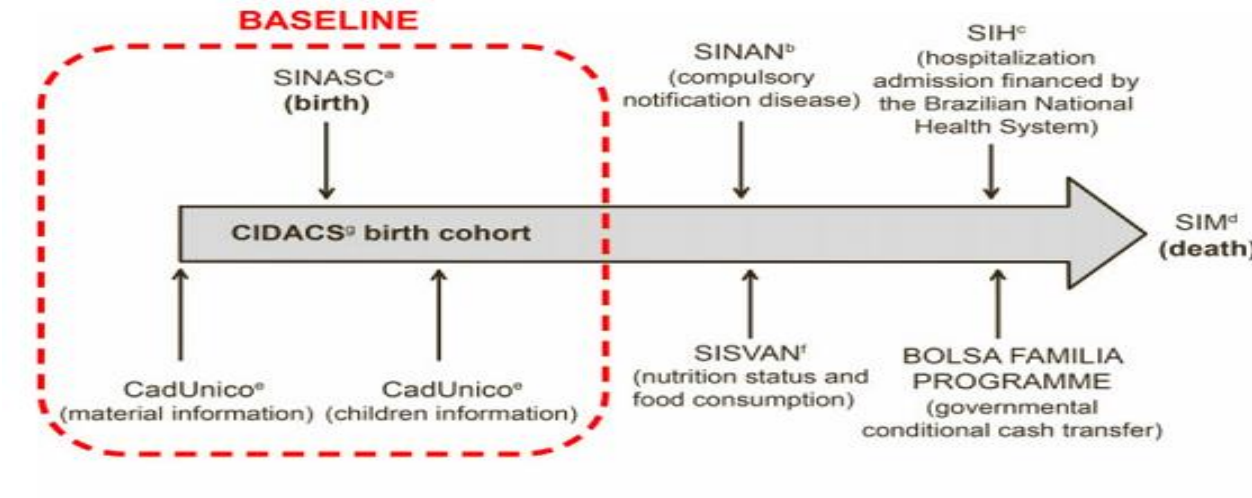
Family level variables
✓ Municipality of family residence
✓ Region of family residence
✓ Location of family residence
✓ Housing and flooring material
✓ Household type
✓ Household water supply
✓ Sewage disposal system
✓ Electricity
✓ Waste collection
✓ Sidewalks around household
✓ Number of individuals in the household
✓ Number of rooms in the household

Cidacs Cohorts

Coorte de 100 milhões de brasileiros



Coorte de Nascimentos



International Journal of Epidemiology
18 June 2021

- ***Editor comments***

- ***Considering that this Cohort is based on administrative data with multiple linkages to other datasets, and there are no waves of data in the usual sense, we feel that this manuscript better suits our Data Resource Profile Category.***

- ***Although below we have asked for clarification under the current subheadings, we would like you to revise this paper using the Data Resource Profile subheadings. Please see the instructions to authors: https://academic.oup.com/ije/pages/General_Instructions#Data%20Resource%20Profiles***

- Firstly, we would like to thank the editorial team for considering our manuscript and for allowing us to resubmit it as a Data Resource Profile.

- **However, respectfully and for the sake of the development of epidemiology as a science, we present the arguments below to convince you to reconsider the decision and keep it as a Cohort Profile.**

- **We argue that our manuscript describes a group of individuals followed during a period with updates and sufficient novelty and quality to fulfil all the fundamental requirements to be classified as a cohort profile article for The International Journal of Epidemiology.**

- **According to the cohort profile criteria: "to be eligible for publication as a cohort profile MUST: have completed baseline data collection *so that some results, in addition to baseline descriptive statistics,* can be presented in the profile".**

- “Describe a study that has either completed prospective follow-up of participants or has funding or clear plans to do so".

- Our Cohort meets these criteria. It is planned to be periodically updated as the databases currently linked are updated. However, more than this, there are plans to link new databases in the Cohort. By contrast, a Data Resource profile is a static data set that "Have completed collection."

- "PREFERENCE is given to cohorts with over 5000 participants at baseline and to cohorts for which follow-up data have been collected specifically for the cohort rather than just from routine data sources".

- Therefore, it is not a mandatory criterion, so it cannot be the only explanation for the paper's rejection. More than that, our Cohort has over 114 million people (and it will certainly grow to over a few million each year), making it possible to explore questions never explored in traditional cohorts, particularly in a LMIC.

- ……..

- Based on the above arguments, we ask you to reconsider our manuscript's status, keeping it as originally submitted – a cohort profile.

International Journal of Epidemiology, 2021, 37–38h
doi: 10.1093/ije/dyaa255
Advance Access Publication Date: 30 December 2020
Cohort Profile

Cohort Profile

Cohort Profile: Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS) Birth Cohort

Enny S Paixao [1,2]* Luciana L Cardim,[1] Ila Rocha Falcao,[1,3] Naiá Ortelan,[1] Natanael de Jesus Silva,[1] Aline dos Santos Rocha,[1,3] Samila Sena,[1] Daniela Almeida,[1] Dandara Oliveira Ramos,[1,4] Flávia Jôse Oliveira Alves,[1,4] Nívea Bispo,[1,5] Sanni Ali,[1] Rosemeire Fiaccone,[1,5] Moreno Rodrigues,[1] Liam Smeeth,[2] Elizabeth B Brickley,[2] Liliana Cabral,[1] Carlos Teles,[1] Maria Conceição N Costa,[1,4] Maria Yury Ichihara,[1,4] Mauricio L Barreto,[1,4] Rita de Cássia Ribeiro Silva[1,3] and Maria Gloria Teixeira[1,4]

[1]Centro de Integração de Dados e Conhecimentos para Saúde, Fiocruz, Salvador, Bahia, Brazil, [2]Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK, [3]Escola de Nutrição, Universidade Federal da Bahia, Salvador, Brazil, [4]Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, Bahia, Brazil, and [5]Departamento de Estatística, Universidade Federal da Bahia, Salvador, Bahia, Brazil

*Corresponding author. Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London,

International Journal of Epidemiology, 2022, e27–e38
https://doi.org/10.1093/ije/dyab213
Advance Access Publication Date: 18 December 2021
Cohort Profile

Cohort Profile

Cohort Profile: The 100 Million Brazilian Cohort

Mauricio L Barreto,[1,2]* Maria Yury Ichihara,[1,2] Julia M Pescarini [1,3] M Sanni Ali,[1,3,4] Gabriela L Borges,[1] Rosemeire L Fiaccone,[1,5] Rita de Cássia Ribeiro-Silva,[1,6] Carlos A Teles,[1] Daniela Almeida,[1] Samila Sena,[1] Roberto P Carreiro,[1] Liliana Cabral,[1] Bethania A Almeida,[1] George CG Barbosa,[1] Robespierre Pita,[1] Marcos E Barreto,[1,7] Andre AF Mendes,[1] Dandara O Ramos,[1,2] Elizabeth B Brickley,[3] Nivea Bispo,[1,5] Daiane B Machado,[1] Enny S Paixao [1,3] Laura C Rodrigues[1,3] and Liam Smeeth[3]

[1]Centre for Data and Knowledge Integration for Health (CIDACS), Fundação Oswaldo Cruz, Salvador, Brazil, [2]Institute of Collective Health, Federal University of Bahia (UFBA), Salvador, Brazil, [3]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK, [4]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Center for Statistics in Medicine, University of Oxford, Oxford, UK, [5]Department of Statistics, Federal University of Bahia, Salvador, Brazil, [6]Department of Nutrition, Federal University of Bahia, Salvador, Brazil and [7]Department of Statistics, London School of Economics and Political Science, London, UK

*Corresponding author. Centre for Data and Knowledge Integration for Health (CIDACS), Parque Tecnológico da Edf. Tecnocentro, R. Mundo, nº 121—Sala 315—Trobogy, Salvador—BA 41745-715, Brazil. E-mail: mauricio.barreto@fiocruz.br

Received 7 April 2021; Editorial decision 2 September 2021, Accepted 17 September 2021

Cidacs: Center for Data and Knowledge Integrationf for Health

Cidacs is a center created to conduct and promote **interdisciplinary research** to produce **knowledge**, develop new scientific **methodologies** and promote professional **training** using linked large-scale databases and high-performance computational resources in a secure environment.

- Founded: December 2016
- TWO Cohorts of **millions** of Brazilian individuals

International Journal of Population Data Science (2019) 4:2:04

International Journal of Population Data Science

Journal Website: www.ijpds.org

The Centre for Data and Knowledge Integration for Health (CIDACS): Linking Health and Social Data in Brazil

Barreto, ML[1,2*], Ichihara, MY[1,2], Almeida, BA[1], Barreto, ME[1,3], Cabral, L[1], Fiaccone, RL[1,4], Carreiro, RP[1], Teles, CAS[1], Pitta, R[1], Penna, GO[1,5,6], Barral-Netto, M[1], Ali, MS[1,7,8], Barbosa, G[1], Denaxas, S[9], Rodrigues, LC[1,8], and Smeeth, L[1,8]

TO INVESTIGATE HEALTH INEQUALITIES

Ethno-racial inequalities on adverse birth and neonatal outcomes: a nationwide, retrospective cohort study of 21 million Brazilian newborns

Poliana Rebouças,[a,*] Enny S. Paixão,[a,b] Dandara Ramos,[a,d] Julia Pescarini,[a] Elzo Pereira Pinto-Junior,[a] Ila R. Falcão,[a] Maria Yury Ichihara,[a] Samila Sena,[a] Rafael Veiga,[a] Rita Ribeiro,[a,e] Laura C. Rodrigues,[c] Maurício L. Barreto,[a,d] and Emanuelle F. Goes[a,d]

[a]Center for Data and Knowledge Integration for Health (CIDACS), Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Bahia, Brazil
[b]Maternal, Adolescent, Reproductive & Child Health (MARCH) Centre, London School of Hygiene & Tropical Medicine, London, UK
[c]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK

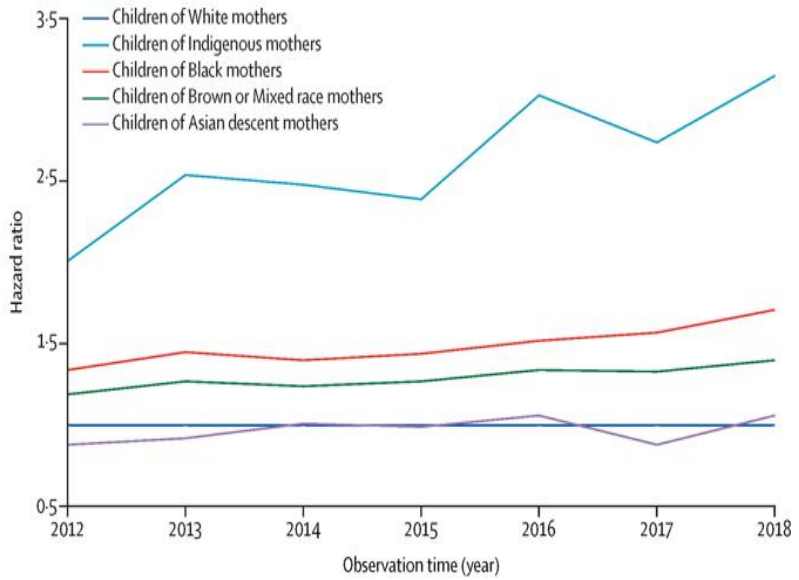
THE LANCET
Global Health

ARTICLES | VOLUME 10, ISSUE 10, E1453-E1462, OCTOBER 01, 2022

Ethnoracial inequalities and child mortality in Brazil: a nationwide longitudinal study of 19 million newborn babies

Poliana Rebouças, PhD • Emanuelle Goes, PhD • Julia Pescarini, PhD • Dandara Ramos, PhD • Maria Yury Ichihara, PhD • Samila Sena, MSc • Rafael Veiga, PhD • Laura C Rodrigues, PhD • Prof Maurício L Barreto, MD PhD • Enny S Paixão, PhD • Show less

Open Access • Published: October, 2022 • DOI: https://doi.org/10.1016/S2214-109X(22)00333-3 •



The Lancet Regional Health - Americas
2024;37: 100833

Published Online xxx
https://doi.org/10.1016/j.lana.2024.100833

ETHNICITY & HEALTH
2024, VOL. 29, NO. 1, 46–61
https://doi.org/10.1080/13557858.2023.2245183

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

The intersection of race/ethnicity and socioeconomic status: inequalities in breast and cervical cancer mortality in 20,665,005 adult women from the 100 Million Brazilian Cohort

Emanuelle F. Góes[a*], Joanna M. N. Guimarães[a*], Maria da Conceição C. Almeida[b], Ligia Gabrielli[c,d], Srinivasa Vittal Katikireddi[e], Ana Clara Campos[a], Sheila M. Alvim Matos[d], Ana Luísa Patrão[f], Ana Cristina de Oliveira Costa[g], Manuela Quaresma[h], Alastair H. Leyland[e], Mauricio L. Barreto[a,d], Isabel dos-Santos-Silva[h**] and Estela M. L. Aquino[d**]

LANGLH-D-23-00410R3
S2214-109X(23)00405-9
Embargo: [add date when known]
Doctopic: Primary Research

23tlgh0410
Articles
Poppy C
This version saved: 10:41, 05-Sep-23

Maternal and congenital syphilis attributable to ethnoracial inequalities: a national record-linkage longitudinal study of 15 million births in Brazil

Enny S Paixao*, Andréa J F Ferreira*, Julia M Pescarini, Kerry L M Wong, Emanuelle Goes, Rosemeire Fiaccone, Guilherme Lopes de Oliveira, Poliana Rebouças, Andrey Moreira Cardoso, Liam Smeeth, Maurício L Barreto, Laura C Rodrigues†, Maria Yury Ichihara†

Summary

Background This study estimated ethnoracial inequalities in maternal and congenital syphilis in Brazil, understanding race as a relational category product of a sociopolitical construct that functions as an essential tool of racism and its manifestations.

Methods We linked routinely collected data from Jan 1, 2012 to Dec 31, 2017 to conduct a population-based study in Brazil. We estimated the attributable fraction of race (skin colour) for the entire population and specific subgroups compared with White women using adjusted logistic regression. We also obtained the attributable fraction of the intersection between two social markers (race and education) and compared it with White women with more than 12 years of education as the baseline.

Findings Of 15 810 488 birth records, 144 564 had maternal syphilis and 79 580 had congenital syphilis. If all women had the same baseline risk as White women, 35% (95% CI 34·89–36·10) of all maternal syphilis and 41% (40·49–42·09) of all congenital syphilis would have been prevented. Compared with other ethnoracial categories,

Lancet Glob Health 2023

*Contributed equally
†Contributed equally
Department of Infectious Disease Epidemiology and International Health, London School of Hygiene & Tropical Medicine, London, UK
(E S Paixao PhD, J M Pescarini PhD, K L M Wong PhD, Prof L Smeeth XX, Prof L C Rodrigues XX); Centro de Integração de Dados e Conhecimentos para Saúde, hia, Brazil
i J F Ferreira PhD, . E Goes PhD, tD, Ilveira PhD, hD, Prof O(, Prof i, M Y Ichihara PhD Center on Racism ments & Health Equity, ool of Public el University, i, PA, USA); Instituto de e Estatística, e Federal da Bahi hia, Brazil Centro Federal d cnológica de Min

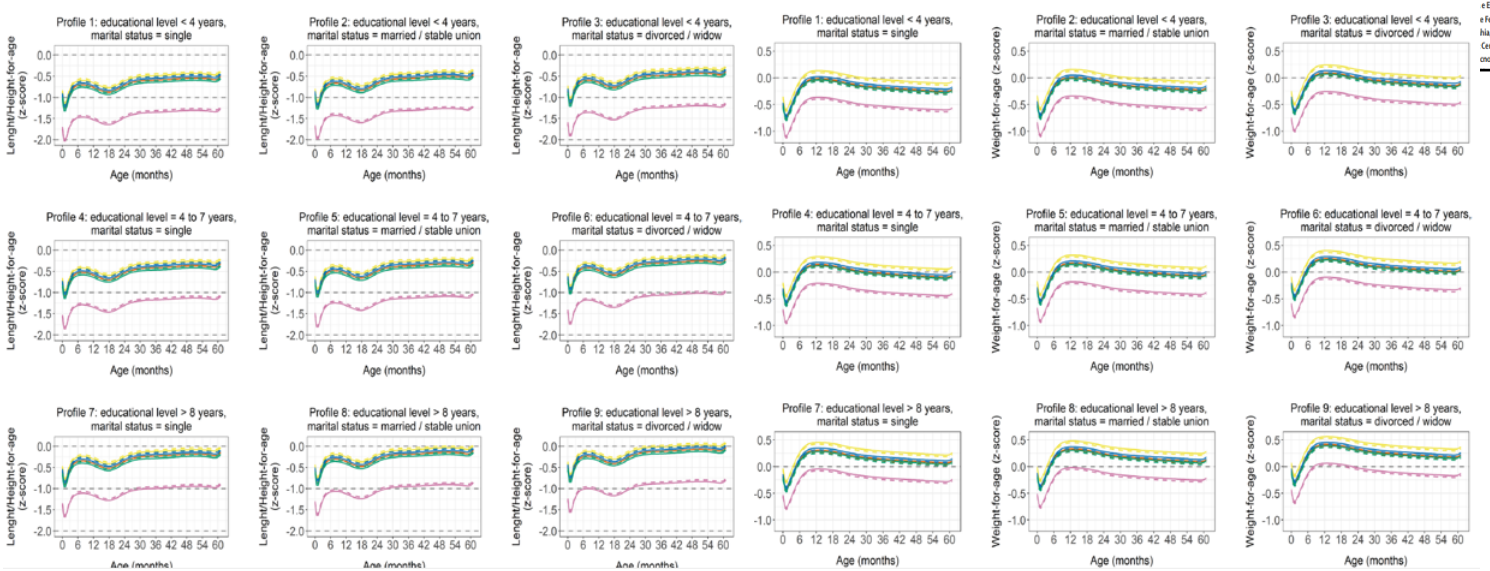
Benes Matos da Silva et al. BMC Pediatrics (2024) 24:103
https://doi.org/10.1186/s12887-024-04550-3

BMC Pediatrics

RESEARCH Open Access

Ethnoracial disparities in childhood growth trajectories in Brazil: a longitudinal nationwide study of four million children

Helena Benes Matos da Silva[1,2*†], Rita de Cássia Ribeiro-Silva[1,2], Juliana Freitas de Mello e Silva[2†], Irina Chis Ster[3], Poliana Rebouças[2], Emanuelle Goes[2], Maria Yury Ichihara[2], Andréa Ferreira[2,4], Julia M. Pescarini[5], Rosemeire Leovigildo Fiaccone[6], Enny S. Paixão[2] and Maurício L. Barreto[2]



TO EVALUATE THE IMPACT OF POVERTY REDUCTION POLICIES (CASH-TRANSFER) ON HEALTH

JAMA Network Open

Original Investigation | Health Policy

Participation in Conditional Cash Transfer Program During Pregnancy and Birth Weight–Related Outcomes

Ila R. Falcão, PhD; Rita de Cássia Ribeiro-Silva, PhD; Rosemeire L. Fiaccone, PhD; Flávia Jôse Oliveira Alves, PhD; Aline dos Santos Rocha, PhD; Naiá Ortelan, PhD; Natanael J. Silva, MSc; Poliana Rebouças, PhD; Elzo Pereira Pinto Júnior, PhD; Marcia Furquim de Almeida, PhD; Enny S. Paixao, PhD; Júlia M. Pescarini, PhD; Laura C. Rodrigues, PhD; Maria Yury Ichihara, PhD; Mauricio L. Barreto, PhD

PLOS MEDICINE

RESEARCH ARTICLE

Conditional cash transfer program and child mortality: A cross-sectional analysis nested within the 100 Million Brazilian Cohort

Dandara Ramos[1,2]*, Nívea B. da Silva[1,3], Maria Yury Ichihara[1,2], Rosemeire L. Fiaccone[1,3], Daniela Almeida[1,4], Samila Sena[1], Poliana Rebouças[1,2], Elzo Pereira Pinto Júnior[1], Enny S. Paixão[1,5], Sanni Ali[1,5], Laura C. Rodrigues[1,5], Maurício L. Barreto[1,2]

American Journal of Epidemiology

Vol. 00, No. 00
DOI: 10.1093/aje/kwaa127
Advance Access publication:

Original Contribution

Conditional Cash Transfer Program and Leprosy Incidence: Analysis of 12.9 Million Families From the 100 Million Brazilian Cohort

ccone,
stino Strina,
O. Penna, and

nature communications

Article
https://doi.org/10.1038/s41467-024-44975-z

Income determines the impact of cash transfers on HIV/AIDS: cohort study of 22.7 million Brazilians

Received: 4 September 2023
Accepted: 10 January 2024
Published online: 12 February 2024
Check for updates

Andréa F. Silva[1,2], Inês Dourado[1], Iracema Lua[1,2], Gabriela S. Jesus[1,3], Nathalia S. Guimarães[1], Gabriel A. S. Morais[1], Rodrigo V. R. Anderle[1], Julia M. Pescarini[2], Daiane B. Machado[2,4], Carlos A. S. T. Santos[2], Maria Y. Ichihara[2], Mauricio L. Barreto[1,2], Laio Magno[1,5], Luis E. Souza[1], James Macinko[6] & Davide Rasella[1,2,7]

PLOS MEDICINE

RESEARCH ARTICLE

Relationship between the Bolsa Família national cash transfer programme and suicide incidence in Brazil: A quasi-experimental study

Daiane Borges Machado[1,2]*, Elizabeth Williamson[3], Julia M. Pescarini[1,3], Flavia J. O. Alves[1], Luís F. S. Castro-de-Araujo[1,4], Maria Yury Ichihara[1], Laura C. Rodrigues[1,3], Ricardo Araya[5], Vikram Patel[2,6], Maurício L. Barreto[1,7]

JAMA Network Open

Original Investigation | Public Health

Association of Conditional Cash Transfers With Maternal Mortality Using the 100 Million Brazilian Cohort

Flávia Jôse O. Alves, PhD; Dandara Ramos, PhD; Enny S. Paixão, PhD; Ila R. Falcão, PhD; Rita de Cássia Ribeiro-Silva, PhD; Rosemeire Fiaccone, PhD; Davide Rasella, PhD; Camila Teixeira, PhD; Daiane Borges Machado, PhD; Aline Rocha, PhD; Marcia F. de Almeida, PhD; Emanuelle F. Goes, PhD; Laura C. Rodrigues, PhD; Maria Yury Ichihara, PhD; Estela M. L. Aquino, PhD; Maurício L. Barreto, PhD

TO INVESTIGATE LOW-FREQUENCY BUT RELEVANT HEALTH EVENTS

The Lancet Regional Health - Americas 3 (2021) 100045

Contents lists available at ScienceDirect

The Lancet Regional Health - Americas

journal homepage: www.elsevier.com/locate/lana

Research paper

Risk of mortality for small newborns in Brazil, 2011-2018: A national birth cohort study of 17.6 million records from routine register-based linked data

Enny S. Paixao, Ph.D[1,2,*], Hannah Blencowe, MD, Ph.D[2], Ila Rocha Falcao, Ph.D[1,3], Eric O. Ohuma, Ph.D[2], Aline dos Santos Rocha[1,3], Flávia Jôse Oliveira Alves[1,4],

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Mortality from Congenital Zika Syndrome — Nationwide Cohort Study in Brazil

Enny S. Paixão, Ph.D., Luciana L. Cardim, Ph.D., Maria C.N. Costa, M.D., Ph.D., Elizabeth B. Brickley, Ph.D., Rita C.O. de Carvalho-Sauer, M.Sc., Eduardo H. Carmo, M.D., Ph.D., Roberta F.S. Andrade, Ph.D.,

Incidence and risk factors of tuberculosis among 420 854 household contacts of patients with tuberculosis i the 100 Million Brazilian Cohort (2004–18): a cohort study

Priscila F P S Pinto, Camila S S Teixeira, Maria Yury Ichihara, Davide Rasella, Joilda S Nery, Samila O L Sena, Elizabeth B Brickley, ulia M Pescarini*

JAMA Dermatology | Original Investigation

Incidence of and Factors Associated With Among Household Contacts of Patients V

Camila Silveira Silva Teixeira, MsC; Júlia Moreira Pescarini, PhD; Flávia Jôse Oliveira A Joilda Silva Nery, PhD; Mauro Niskier Sanchez, PhD; Carlos Teles, PhD; Maria Yury Ti Anna Ramond, PhD; Liam Smeeth, PhD; Maria Lucia Fernandes Penna, PhD; Laura C Elizabeth B. Brickley, PhD; Gerson Oliveira Penna, PhD; Maurício Lima Barreto, PhD;

Figure 2. Cumulative Incidence of Subsequent Leprosy Cases Among Households of Patients With Leprosy

A Subsequent leprosy cases among household contacts

Risk of death following chikungunya virus disease in the 100 Million Brazilian Cohort, 2015–18: a matched cohort study and self-controlled case series

Thiago Cerqueira-Silva, Julia M Pescarini, Luciana L Cardim, Clémence Leyrat, Heather Whitaker, Carlos Alexandre Antunes de Brito, Elizabeth B Brickley, Manoel Barral-Netto, Maurício L Barreto, Maria G Teixeira, Viviane S Boaventura, Enny S Paixão

Lancet Infect Dis 2024; 24: 504-13

Published Online

Summary

Background Chikungunya virus outbreaks have been associated with excess deaths at the ecological level. Previous studies have assessed the risk factors for severe versus mild chikungunya virus disease. However, the risk of death following chikungunya virus disease compared with the risk of death in individuals without the disease remains

Received: 31 August 2021 | Revised: 4 November 2021 | Accepted: 11 November 2021 | First published online: 13 December 2021

DOI: 10.1002/ijgo.14053

CLINICAL ARTICLE

Obstetrics

GYNECOLOGY OBSTETRICS FIGO WILEY

Recurrence of preterm births: A population-based linkage with 3.5 million live births from the CIDACS Birth Cohort

Aline S. Rocha[1,2] | Rita de Cássia Ribeiro-Silva[1,2] | Enny S. Paixao[2,3] | Ila R. Falcão[1,2] | Flavia Jôse. O. Alves[2,4] | Naiá Ortelan[2] | Marcia F. de Almeida[5] | Rosemeire L. Fiaccone[2,6] | Laura C. Rodrigues[3] | Maria Yury Ichihara[2] | Mauricio L. Barreto[2,4]

Causes of death in children with congenital Zika syndrome in Brazil, 2015 to 2018: A nationwide record linkage study

Maria da Conceição N. Costa[1,2✉], Luciana Lobato Cardim[1✉], Cynthia A. Moore[3], Eliene dos Santos de Jesus[2,4], Rita Carvalho-Sauer[2,5], Mauricio L. Barreto[1,2], Laura C. Rodrigues[1,6], Liam Smeeth[6], Lavínia Schuler-Faccini[7], Elizabeth B. Brickley[6], Wanderson K. Oliveira[8], Eduardo Hage Carmo[1,9], Julia Moreira Pescarini[1,6], Roberto F. S. Andrade[1,10], Moreno M. S. Rodrigues[1], Rafael V. Veiga[1], Larissa C. Costa[1], Giovanny V. A. França[9], Maria Gloria Teixeira[1,2‡], Enny S. Paixão[1,6‡*]

1 Center of Data and Knowledge Integration for Health (CIDACS), Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Bahia, Brazil. 2 Collective Health Institute, Federal University of Bahia, Salvador

TO INVESTIGATE MIGRATION EFFECTS ON HEALTH

Mortality among over 6 million internal and international migrants in Brazil: a study using the 100 Million Brazilian Cohort

Julia M. Pescarini,[a,b,*] Emanuelle F. Goes,[b,c] Priscila Fernanda Porto Scaff Pinto,[b] Beatriz Pinheiro Schindler Dos Santos,[b] Daiane B. Machado,[b,d] Ibrahim Abubakar,[e] Laura C. Rodrigues,[a] Elizabeth B. Brickley,[a,g] Liam Smeeth,[a,f,g] and Mauricio L. Barreto[b,c,g]

[a]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[b]Centre of Data and Knowledge Integration for Health (CIDACS), Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Bahia, Brazil
[c]Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, Brazil
[d]Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA
[e]Faculty of Population Health Sciences, University College London (UCL), London, UK
[f]Health Data Research (HDR), London, UK

Summary

Background To understand if migrants living in poverty in low and middle-income countries (LMICs) have mortality advantages over the non-migrant population, we investigated mortality risk patterns among internal and international migrants in Brazil over their life course.

Methods We linked socio-economic and mortality data from 1st January 2011 to 31st December 2018 in the 100 Million Brazilian Cohort and calculated all-cause and cause-specific age-standardised mortality rates according to individuals' migration status for men and women. Using Cox regression models, we estimated the age- and sex-adjusted mortality hazard ratios (HR) for internal migrants (i.e., Brazilian-born individuals living in a different Brazilian state than their birth) compared to Brazilian-born non-migrants; and for international migrants (i.e., people born in another country) compared to Brazilian-born individuals.

Findings The study followed up 45,051,476 individuals, of whom 6,057,814 were internal migrants, and 277,230 were international migrants. Internal migrants had similar all-cause mortality compared to Brazilian non-migrants (aHR = 0.99, 95% CI = 0.98–0.99), marginally higher mortality for ischaemic heart diseases (aHR = 1.04, 95% CI = 1.03–1.05) and higher for stroke (aHR = 1.11, 95% CI = 1.09–1.13). Compared to Brazilian-born individuals, international migrants had 18% lower all-cause mortality (aHR = 0.82, 95% CI = 0.80–0.84), with up to 50% lower mortality from interpersonal violence among men (aHR = 0.50, 95% CI = 0.40–0.64), but higher mortality from avoidable causes related to maternal health (aHR = 2.17, 95% CI = 1.17–4.05).

Interpretation Although internal migrants had similar all-cause mortality, international migrants had lower all-cause mortality compared to non-migrants. Further investigations using intersectional approaches are warranted to understand the marked variations by migration status, age, and sex for specific causes of death, such as elevated maternal mortality and male lower interpersonal violence-related mortality among international migrants.

Funding The Wellcome Trust.

The Lancet Regional Health - Americas 2023;20: 100455

Published Online 27 February 2023
https://doi.org/10.1016/j.lana.2023.100455

Contents lists available at ScienceDirect

Travel Medicine and Infectious Disease

journal homepage: www.elsevier.com/locate/tmaid

Perinatal health outcomes of international migrant women in Brazil: A nationwide data linkage study of the CIDACS birth cohort (2011–2018)

Julia M. Pescarini [a,b,*], Ila R. Falcao [b], Poliana Reboucas [b], Enny S. Paixao [a,b], Nuria Sanchez-Clemente [a,c], Emanuelle F. Goes [b], Ibrahim Abubakar [d], Laura C. Rodrigues [a,b], Elizabeth B. Brickley [a], Liam Smeeth [a,e], Mauricio L. Barreto [a,f]

CIDACS NEW INITIATIVES

Data Platform to study environmental and climate effects on Health

cidacsClima
Plataforma de Dados Climáticos, Ambientais e de Saúde

Objective 3. Expand and update the *100 Million Brazilian* and *CIDACS Birth Cohorts* and develop the *CIDACS Climate and Environmental Platform*

- Extend and expand the geocoding of data in the *100 Million Brazilian* and *CIDACS Birth Cohorts*
- Update the the *100 Million Brazilian* and *CIDACS Birth Cohorts* to include new data up to 2023.
- Develop a fully functional *CIDACS Climate and Environmental Platform*
  - Hydrometeorological, reanalysis, and satellite imagery data

Plataforma de dados climáticos ambientais no CIDACS: Exemplos

Dados medidos in situ pelo INMET	Modelo de Reanálise ERA 5 LAND	Modelo de Reanálise VAN DONKELLAR	MAPBIOMAS
Período: 2000-presente (diário). Download: município e UF. Variáveis: • Temperatura: max - min - média - ponto orvalho • Precipitação Total • Vel e direção do vento • Pressão à Superfície • Umidade Relativa	Período: 2000-presente (diário). Download: município e UF. Variáveis: • Temperatura: max - min - média - ponto orvalho • Precipitação Total • Comp do vento: u e v • Pressão à Superfície • Umidade Relativa	Período: 2000-presente (mensal). Download: município e UF. Modelo: V5GL01.HybridPM25 Variável: • PM2.5 • NO2	Período: 2000-presente (mensal/anual). Download: município e UF. Variáveis: Uso e ocupação do solo Cicatriz de fogo Cobertura de água

Agnes Soares, coordenadora geral de Vigilância em Saúde Ambiental do Ministério da Saúde

AIMS:

- Develop a Climate and Environmental Data Platform at CIDACS.
- Integrate Brazil's georeferenced environmental and climate data into an accessible platform, interoperable with existing national health and socioeconomic data.
- Generate new knowledge on the links between climate and health that are currently unavailable from existing dissipated and unlinked resources and inform mitigation and adaptation responses.

cidacsClima
Plataforma de Dados Climáticos, Ambientais e de Saúde

Coorte de 100 milhões de brasileiros

Coorte de Nascimentos

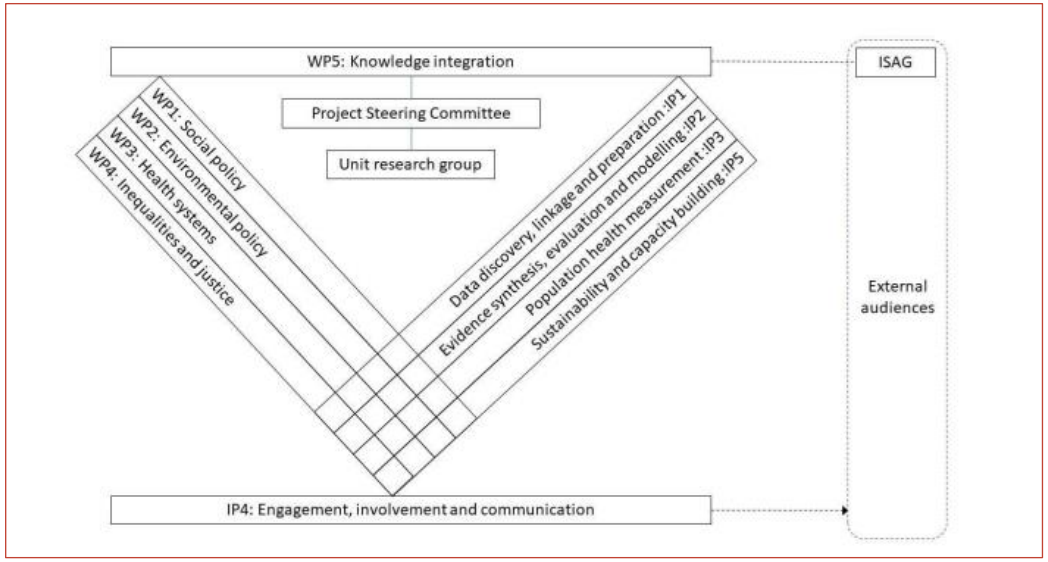
CIDACS NEW INITIATIVES

NIHR GLOBAL HEALTH RESEARCH UNIT ON SOCIAL AND ENVIRONMENTAL DETERMINANTS OF HEALTH INEQUALITIES (2022-2027)

Our vision

- Over the next five years, we plan to be global leaders, focused on Latin America, in harnessing existing databases by integrating them to evaluate the impact of social and environmental policies on health and health inequalities, and the extent to which these can be affected by the health system.

SEDHI
NIHR Unit in Social and Environmental Determinants of Health Inequalities



Ministério da Saúde FIOCRUZ Fundação Oswaldo Cruz Instituto Gonçalo Moniz
cidacs Centro de Integração de Dados e Conhecimentos para Saúde
MRC/CSO Social and Public Health Sciences Unit
MRC Medical Research Council
CHIEF SCIENTIST OFFICE
University of Glasgow
FUNDED BY NIHR | National Institute for Health Research
UIDE
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE
St George's University of London

PROBLEMS AND WEAKNESSES OF USING LINKED ADMINISTRATIVE DATA FOR BUILDING LARGE NATIONWIDE COHORTS

- **Data not collected for research purposes**

  Lack of relevant variables

- **Changes in how the data is collected over time**

  While the data are standardized across the country, some collection instruments have changed at some point, creating a need for harmonization.

- **Missing Values**

  Depending on the database lots or few missing data. For instance, in CADU, a considerable proportion of missing values in some variables, such as the occupation of family members and family characteristics, however, the description of all individuals in the household (e.g., gender, age, and ethnicity) and variables such as income, are highly complete.

PROBLEMS AND WEAKNESSES OF USING LINKED ADMINISTRATIVE DATA FOR BUILDING LARGE NATIONWIDE COHORTS

- **Misclassification and Selection biases**

  Selection and misclassification biases may be introduced in exposure and outcomes in three stages in the course of the assemblage of the cohort: registering and recording people's life events and use of services, linkage across administrative databases, and cleaning and coding variables from derived datasets.

- **Do not include all population or a random sample of it, but the poorest part of the population**

  In the Brazilian cohorts, it covers only the poorest part of the Brazilian population.

- **Register occurs when contact with service is provided**

  Mild outcomes or outcomes not routinely registered will not be included

- **Data Governance** (access, security, privacy, ethical and legal issues)

- **Computational Resources**

ADVANTAGES AND STRENGTHENS OF USING LINKED ADMINISTRATIVE DATA FOR BUILDING LARGE NATIONWIDE COHORTS

- **Data quality and standardized data collection throughout the country**

  1.The quality and coverage of health data has improved; 2. Brazil, despite being a large and federative country, the data collection forms for different purposes are standardized(death, birth, hospitalization etc)

- **Nationwide Data Coverage, including all ages and ethnical groups**

  Possible to study minoritized population groups as well as traditional groups

- **Possibility to link data from multiple sources (health, social, environmental, climate)**

  Routine data is very often spread in silos with a connection between them. It links health, social, and environmental data from multiple sources. This adds enormous value to existing health data. Answer questions not possible with an isolated database

- **Quality of the Data Linkage**

  The linkage (deterministic or probabilistic) is conducted with very powerful and accurate software developed in-house and adapted to the Brazilian data ecosystem (Cidacs-RL)

ADVANTAGES AND STRENGTHENS OF USING LINKED ADMINISTRATIVE DATA FOR BUILDING LARGE NATIONWIDE COHORTS

Large Sample size

Large sample sizes allow several analytical advantages, such as the analysis of rare events; It is possible to study rare exposure and rare outcomes simultaneously; Increase the statistical power of the study, allowing for more precise estimates of associations and enabling subgroup analyses such as age, gender, ethnicity, geographies, etc.

•Overcome classical epidemiological challenges

- Eliminates the risk of **recall bias** for important variables, which is a problem if data collection relies on self-reports of service use (e.g., hospitalization or childbirth). Participants and the data collector do not know the research question to be put for the data
- Very low attrition rates – immigration from Brazil is negligible among the poorest populations

•Using longitudinal individual data could overcome limitations of the time-series analysis

•Cost reduction

Costs involved in data collection with comparable-sized primary studies are reduced. Cost of the infrastructure.

cidacs
Centro de Integração de Dados e Conhecimentos para Saúde

Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Gonçalo Moniz

Obrigado!
Thank you!

COLABORADORES CIENTÍFICOS:

APOIADORES:

Parque Tecnológico da Bahia
Rua Mundo, 121, Trobogy
Salvador - Ba, CEP 41745-715

www.cidacs.bahia.fiocruz.br
cidacs@bahia.fiocruz.br

fiocruzbahia.cidacs
cidacs_fiocruz
Cidacs Fiocruz