

Running Head: PRIMERO Birth Cohort

1 **Word counts**

2 Text: 3,980

3 Abstract: 200

4
5 **Title:** The Puerto Rican Infant Metagenomic and Epidemiologic Study of Respiratory Outcomes

6 (PRIMERO): Design and Baseline Characteristics for a Birth Cohort Study of Early-life Viral

7 Respiratory Illnesses and Airway Dysfunction in Puerto Rican Children

8 **Running head:** PRIMERO Cohort

9 **Author list:** Jonathan I. Witonsky^{*}, MD, MAS; Jennifer R. Elhawary^{*}, MS; Celeste Eng, BS;

10 Sam S. Oh, PhD, MPH; Sandra Salazar; Maria G. Contreras, BS; Vivian Medina, RN, BSN, SC;

11 Elizabeth A. Secor, MA; Priscilla Zhang, BS; Jamie L. Everman, PhD; Ana Fairbanks-Mahnke,

12 BS; Elmar Pruesse, PhD; Satria P. Sajuthi, PhD; Chih-Hao Chang, PhD, MPH; Tsunami Rosado

13 Guerrero, BS; Keyshla Canales Fuentes, BS; Natalie Lopez, MD; Chris Angely Montanez-

14 Lopez, BS; Richeliz Alfonso Otero, MD; Raymarie Colon Rivera, BS; Leysha Rodriguez, BA;

15 Gabriela Vazquez, BS; Donglei Hu, PhD; Scott Huntsman, MS; Nathan D. Jackson, PhD;

16 Yingchun Li, MD, PhD; Andrew Morin, MA; Natalie A. Nieves, BS; Cydney Rios, BS; Gonzalo

17 Serrano, BS; Blake J. M. Williams, MS; Elad Ziv, MD; Camille M. Moore, PhD; Dean

18 Sheppard, MD; Esteban Gonzalez Burchard[#], MD, MPH; Max A. Seibold[#], PhD; and Jose R.

19 Rodriguez Santana[#], MD

20 ^{*}*Shared first authorship*

21 [#]*Shared senior authorship*

22 **Affiliations:** From the Department of Pediatrics (Dr Witonsky) and Department of Medicine (Ms

23 Elhawary, Ms Eng, Dr Oh, Ms Salazar, Ms Contreras, Ms Zhang, Dr Chang, Dr Hu, Mr

24 Huntsman, Dr Ziv, Dr Sheppard, and Dr Burchard), University of California, San Francisco, San

25 Francisco, CA; Centro de Neumología Pediátrica (Ms Medina, Ms Fuentes, Ms Guerrero, Dr

Running Head: PRIMERO Birth Cohort

26 Lopez, Ms Montanez-Lopez, Dr Otero, Ms Rivera, Ms Rodriguez, Ms Vazquez, Ms Nieves, Ms
27 Serrano, and Dr Rodríguez Santana), San Juan, PR; Center for Genes, Environment & Health
28 (Ms Secor, Dr Everman, Ms Fairbanks-Mahnke, Dr Pruesse, Dr Sajuthi, Dr Jackson, Dr Li, Mr
29 Morin, Ms Rios, Mr Williams, Dr Moore, and Dr Seibold), National Jewish Health, Denver, CO;
30 and the Department of Pediatrics and Division of Pulmonary Sciences and Critical Care
31 Medicine, Department of Medicine (Dr. Seibold), University of Colorado, Aurora, CO.

32

33 **Correspondence to:** Jonathan Witonsky, MD, MAS, Department of Pediatrics, University of
34 California, San Francisco, 550 16th Street, 4th Floor, San Francisco, CA, 94143; e-mail:
35 jonathan.witonsky@ucsf.edu

36

37 **Funding information:**

38 This work was supported in part by the Sandler Family Foundation; the American Asthma
39 Foundation; the RWJF Amos Medical Faculty Development Program; Harry Wm. and Diana V.
40 Hind Distinguished Professor in Pharmaceutical Sciences II; the American Lung Association
41 [CAALA2023]; Wohlberg and Lambert Endowed Chair of Pharmacogenomics (MAS);
42 the National Institutes of Health, National Heart, Lung, and Blood Institute [5U01HL138626 and
43 1K23HL169911]. The content of this publication is solely the responsibility of the authors and
44 does not necessarily reflect the views or policies of the Department of Health and Human
45 Services, nor does mention of trade names, commercial products, or organizations imply
46 endorsement by the U.S. government.

Running Head: PRIMERO Birth Cohort

47 **Abstract**

48 Epidemiologic studies demonstrate an association between early-life respiratory illnesses (RIs)
49 and the development of childhood asthma. However, it remains uncertain whether these children
50 are predisposed to both conditions or if early-life RIs induce alterations in airway function,
51 immune responses, or other human biology that contribute to the development of asthma. Puerto
52 Rican children experience a disproportionate burden of early-life RIs and asthma, making them
53 an important population for investigating this complex interplay. PRIMERO, the *Puerto Rican*
54 *Infant Metagenomics and Epidemiologic Study of Respiratory Outcomes*, recruited pregnant
55 women and their newborns to investigate how the airways develop in early life among infants
56 exposed to different viral RIs, and will thus provide a critical understanding of childhood asthma
57 development. As the first asthma birth cohort in Puerto Rico, PRIMERO will prospectively
58 follow 2,100 term healthy infants. Collected samples include post-term maternal peripheral
59 blood, infant cord blood, the child's peripheral blood at the year two visit, and the child's nasal
60 airway epithelium, collected using minimally invasive nasal swabs, at birth, during RIs over the
61 first two years of life, and at annual healthy visits until age five. Herein, we describe the study's
62 design, population, recruitment strategy, study visits and procedures, and primary outcomes.

Running Head: PRIMERO Birth Cohort

63 INTRODUCTION

64 Asthma is the most common chronic condition in childhood, disproportionately affecting
65 children from racially and ethnically minoritized backgrounds (1). Asthma is over 40% more
66 prevalent among Black and Latino children compared to their non-Hispanic white counterparts,
67 and Puerto Rican children face a threefold higher likelihood of having asthma and an over
68 fivefold higher risk of asthma-related mortality (2–4). The development of asthma is influenced
69 by complex gene-by-environment interactions that remain poorly understood (5–13). It is
70 hypothesized that a key factor perpetuating asthma inequities among Puerto Rican children is
71 their increased early-life exposure to environmental risk factors, particularly to asthma-genic
72 respiratory viral pathogens, which are both ubiquitous and endemic in Puerto Rico (13–16).
73 Irrespective of race or ethnicity, viral infections causing lower respiratory tract symptoms in the
74 first two years of life are commonly linked to the subsequent development of recurrent wheezing
75 and childhood asthma (17,18). In particular, infants who had bronchiolitis or wheezing induced
76 by respiratory syncytial virus (RSV) face a 2.6-fold increased likelihood of developing asthma at
77 six years of age (14). Furthermore, a recent prospective study found that children who remained
78 free from RSV infection by one year of age had a 26% lower risk of developing asthma before
79 age five than those infected with RSV by the same age (19). Similarly, children with early-life
80 wheezing illnesses caused by human rhinoviruses (HRV) face a 9.8-fold increased risk of
81 developing asthma at age six years (14).

82
83 Despite these findings, more studies are needed to examine the asthma risk attributable to early-
84 life illnesses caused by RSV versus HRV, as well as the risk associated with other common
85 respiratory viruses infecting young children. Additionally, the causative mechanisms underlying

Running Head: PRIMERO Birth Cohort

86 these associations still need to be determined. One plausible explanation is that infections with
87 these respiratory viruses induce epithelial barrier dysfunction that leads to the development of
88 asthma (20–25). Conversely, children may be born with specific airway or immune defects
89 predisposing them to both poor responses to viral infections and an increased likelihood of
90 developing asthma (26). If the former holds, mitigating lower respiratory tract illnesses (LRIs)
91 could emerge as an effective strategy for preventing the onset of asthma. Supporting this notion,
92 intervention trials focused on reducing early-life RSV infections have demonstrated efficacy in
93 preventing asthma (27). This prospect has fueled the recent development of an RSV vaccine
94 administered during pregnancy and an RSV antibody treatment for infants born during RSV
95 season to block early-life acute RSV infections and inhibit asthma development (28,29).

96

97 *The Puerto Rican Infant Metagenomic and Epidemiologic Study of Respiratory Outcomes*, or
98 PRIMERO, is a prospective observational birth cohort study of Puerto Rican children established
99 to investigate the prevalence and role of all common respiratory viruses on the development of
100 LRIs and asthma. Specifically, PRIMERO seeks to ascertain whether the association between
101 early-life viral illness and asthma is rooted in deficiencies in the airway at birth or the subsequent
102 alterations of the airway and airway mucosal immune system in response to viral infections. To
103 achieve this aim, PRIMERO is designed to examine the molecular development and
104 pathobiology of the airway at birth and throughout early childhood, all while performing
105 surveillance for early-life respiratory illnesses, with viral and molecular response
106 characterization of the detected illnesses. In contrast to prior birth cohort studies, our study
107 utilizes cutting-edge multiplexed molecular assays, which allow us to identify illnesses caused
108 by both RSV and HRV, in addition to metapneumovirus, coronaviruses, parainfluenza, influenza,

Running Head: PRIMERO Birth Cohort

109 and other species that have been less studied in this context. Furthermore, leveraging Puerto
110 Rican's rich genetic admixture, PRIMERO children will be genetically characterized to
111 determine if the effects of these viruses on LRI and asthma risk are dependent on genetic
112 variation. Perhaps most innovative, the PRIMERO study will characterize molecular airway
113 dysfunction repeatedly at birth, yearly throughout childhood, and during respiratory illnesses,
114 through transcriptomic analyses of nasal airway mucosa, collected by swabbing the inferior nasal
115 turbinate. Importantly, our prior work in young children has shown this minimally invasive
116 technique collects both epithelia and airway immune cells, and that the cell types and expression
117 profiles of these upper airway samples mirror those in the lower (bronchial) airways in health
118 and asthma disease (30). Leveraging this rich dataset, our goals are fourfold: [1] identify the
119 genetic and viral determinants of early-life LRIs, [2] elucidate the influence of LRIs caused by
120 different viral species on the development of asthma, [3] delineate the trajectories of airway
121 molecular and cellular development from birth through childhood in both children who do and
122 do not develop asthma, and [4] ascertain whether the airway developmental trajectories leading
123 to asthma are predetermined at birth or influenced by early-life viral LRIs. Beyond these goals,
124 PRIMERO's unique and comprehensive characterization of respiratory illnesses and airway
125 pathobiology during childhood offers unprecedented opportunities for further investigation into
126 the impact of prenatal and postnatal social, environmental, and genetic determinants of childhood
127 health and disease.

128

129 PRIMERO, a National Heart, Lung, and Blood Institute (NHLBI) funded longitudinal birth
130 cohort consisting of 2,100 term healthy infants, is the result of a collaborative effort among
131 research groups from the University of California, San Francisco (UCSF), National Jewish

Running Head: PRIMERO Birth Cohort

132 Health (NJH) in Denver, CO, and Centro de Neumología Pediátrica (CNP) in Caguas, Puerto
133 Rico. The study meticulously tracks this cohort from birth through early childhood, incorporating
134 RI surveillance during the initial two years of life and annual airway sampling during healthy
135 child study visits. Herein, we provide a comprehensive overview of the objectives, study design,
136 procedures, specimen collections, and recruitment results.

137

138 METHODS

139 Primary Hypotheses

140 RIs caused by some viral species will independently increase the risk of acute lower respiratory
141 tract symptoms and asthma, while the effect of other species on these outcomes will be
142 dependent on host factors, including genetics. Early-life viral LRIs will directly contribute to the
143 risk of childhood asthma by modulating the early-life development of the airway epithelium and
144 immune system.

145

146 Primary Objectives

147 To address our hypotheses, PRIMERO pursues the following seven study objectives:

- 148 1. Identify the genetic and viral determinants of early-life upper RIs (URIs) and LRIs.
- 149 2. Determine the association between early-life RIs attributed to various viral species and both
150 the modified asthma predictive index (mAPI) and asthma disease in childhood.
- 151 3. Determine whether the risk of asthma conferred by different viral RIs is dependent on host
152 factors, including genetics.
- 153 4. Identify distinct early-life airway gene expression patterns and cellular developmental
154 trajectories that contribute to asthma in childhood.

Running Head: PRIMERO Birth Cohort

- 155 5. Determine whether early-life viral LRIs alter airway developmental trajectories.
- 156 6. Determine whether the molecular and cellular profiles of the airway at birth are associated
- 157 with the occurrence of early-life viral LRIs and/or the development of asthma in childhood.
- 158 7. Determine whether the gene expression responses to viral LRIs are associated with the
- 159 subsequent development of asthma.

160

161 Study Overview

162 To address these objectives, we designed PRIMERO as a prospective observational birth cohort

163 study, enrolling newborns and their mothers from March 2020 to June 2023. Surveillance for RIs

164 during the child's first two years involves weekly short message service (SMS) texts and email

165 messages, complemented by in-person visits and biospecimen collections specifically designated

166 for RIs. In addition, newborn and annual follow-up visits for the child's first five years of life are

167 used to assess respiratory health, environmental and social exposures, measure lung function,

168 and to collect biospecimens. The minimum duration of participant follow-up will be two years

169 with the aim to follow participants for a maximum of 10 years pending ongoing support of the

170 study.

171

172 Established Community Engagement

173 PRIMERO builds upon our extensive community engagement with the Puerto Rican community,

174 which commenced in 1998 with the initiation of the Genetics of Asthma in Latino Americans

175 (GALA) Study (31). This endeavor was designed to explore the heightened rates of asthma

176 prevalence and morbidity among Puerto Rican children. Expanding on this groundwork, our

177 research progressed with the Genes-environments & Admixture in Latino Americans (GALA II

Running Head: PRIMERO Birth Cohort

178 study) spanning from 2008 to 2014 (32). This case-control study focused on asthma in Latino
179 children, including Puerto Rican participants in our recruitment, epidemiological, and genetic
180 analyses. Our efforts in GALA and GALA II resulted in over an 85% recruitment rate and a
181 follow-up contact rate exceeding 90% among Puerto Rican study participants, with GALA II
182 providing much of the preliminary data for PRIMERO.

183

184 Study Population and Recruitment

185 PRIMERO recruited pregnant women who were receiving obstetric/gynecologic (OB/GYN)
186 services and giving birth at Hospital Interamericano de Medicina Avanzada-San Pablo (HIMA).
187 HIMA is located in Caguas, Puerto Rico, which neighbors San Juan, Puerto Rico's capital and
188 most populous municipality (**Figure 1**). Most babies born at HIMA (>95%) come from families
189 residing in the San Juan-Bayamón-Caguas metropolitan area, home to 63% of Puerto Rico's
190 population. Demographically, the San Juan-Bayamón-Caguas metropolitan area is similar to the
191 entire island in terms of age, marital status, birth rate, and educational attainment (33).

192

193 Healthcare providers from HIMA's OB/GYN services were introduced to PRIMERO via
194 presentations and information sessions. Medical staff were invited to join recruitment efforts, and
195 collaborating offices received PRIMERO brochures for distribution during routine prenatal
196 visits. Women receiving information packets had their details recorded in the custom electronic
197 database, PRIMERO-DB (**Supplemental Text 1**). Recruitment occurred in two stages, allowing
198 participants time to consider involvement. Initially, phone calls assessed interest, followed by
199 consent and eligibility determination. PRIMERO recruiters at the HIMA labor and delivery ward
200 used PRIMERO-DB to cross-reference consented participants. The second consent stage

Running Head: PRIMERO Birth Cohort

201 occurred after the child was born pre-discharge when mothers were allowed to reaffirm their
202 participation. Protocols and documents were approved by the UCSF Institutional Review Board
203 (IRB) and a National Institutes of Health-appointed Observational Study Monitoring Board
204 (OSMB), ensuring data safety, study progress, and confidentiality protocols.

205

206 **Enrollment, Study Visits, and Procedures**

207 Study visits and procedures are detailed in **Table 1**. Briefly, study participation consists of an
208 enrollment visit, annual visits for the first five years of life, and recurring RI visits during the
209 first two years of life.

210

211 Enrollment

212 Before birth, expectant mothers in the HIMA OB/GYN clinic expressing interest in the study
213 were assessed for eligibility and consented. Inclusion and exclusion criteria are provided in
214 **Table 2**. At birth, the newborns of consented mothers were assessed for eligibility, and cord
215 blood was collected. In the recovery ward of HIMA hospital, mothers of eligible infants were
216 approached with a re-consent. Re-consented mothers were then administered a baseline
217 questionnaire by trained staff who also collected peripheral blood from the mother and nasal
218 swabs of the inferior turbinates of the newborn. The cord blood collected from infants whose
219 mothers did not re-consent and those infants deemed not eligible was discarded, and infants were
220 not enrolled into the study. Blood and nasal samples were processed at CNP, the core facility for
221 our recruitment and follow-up procedures, located in the building adjacent to HIMA. After
222 processing, samples were shipped to each of the laboratory cores: blood to the UCSF Pediatric
223 Asthma Biobank and nasal swabs to the NJH Nasal Airway Biobank (**Supplemental Text 2**).

Running Head: PRIMERO Birth Cohort

224

225 Annual Visits

226 Initially, annual visits and questionnaires were aligned with the child's birthday. However, due to
227 the COVID-19 pandemic, the window was extended to 8 months post-birthday to accommodate
228 stay-at-home mandates and vaccination rollouts. For safety and efficiency, questionnaires are
229 conducted separately by trained PRIMERO staff via phone, utilizing the secure UCSF Research
230 Electronic Data Capture (REDCap) software (34,35). Questionnaires administered at newborn
231 and annual follow-up visits during the child's initial five years gather comprehensive data on
232 environmental factors, demographic variables, social aspects, and clinical risk factors associated
233 with recurrent wheezing and asthma. Biospecimens are collected annually during healthy
234 periods, and the mAPI is assessed at 2 years. Annual visits include recording anthropometrics, an
235 eczema assessment, and collection of two nasal swabs. In the second year, peripheral blood is
236 processed for allergen sensitization. In the third year, impulse oscillometry is conducted with the
237 plan for spirometry in subsequent years to evaluate lung function. Annual nasal swabs are sent to
238 the NJH Nasal Airway Biobank, and year two blood samples go to the UCSF Pediatric Asthma
239 Biobank after processing at CNP.

240

241 Respiratory Illness Surveillance

242 After discharge from the hospital following birth, participants enter the RI surveillance phase of
243 the study (**Figure 2**), which lasts for the first two years of life. Mothers of participating infants
244 receive a weekly automated SMS text or email that provides a hyperlink to indicate whether the
245 child has signs of RI (**Supplemental Figure 1**). Text and email messages that are unanswered
246 within 24 hours or have an affirmative response are compiled into a weekly call list, followed by

Running Head: PRIMERO Birth Cohort

247 a phone call from project staff. In addition to active surveillance, mothers can call PRIMERO
248 staff or visit the clinic for a phone-based or face-to-face assessment. Illness assessments, visits,
249 follow up, and severity determination are further described in **Supplemental Text 3**.

250

251 Biospecimen Collection

252 Cord and maternal blood are collected at enrollment to define genome-wide genetic variation in
253 mothers and children. Nasal swabs are collected from newborns at enrollment, at each annual
254 visit, and during in-person RI visits. Illness swabs are tested for SARS-CoV-2 and enterovirus
255 D68 using PCR assays and for 22 other common respiratory pathogens using the NxTAG
256 Respiratory Pathogen Panel. Nasal swabs are also used to evaluate longitudinal airway gene
257 expression in health and illness using whole transcriptome RNA-sequencing. In addition, blood
258 is collected at the year 2 visit to obtain complete blood counts with differentials and assess
259 allergen sensitization via specific IgE testing to perennial aeroallergens and foods.

260

261 Respiratory Illness Outcomes

262 RI outcomes of the study include the following:

- 263 1. Identification of early-life RIs.
- 264 2. Analysis of RI-associated airway gene expression.
- 265 3. Determination of respiratory virus infection at a species level.
- 266 4. The modified asthma predictive index (mAPI).

267 Secondary RI outcomes and the secondary research questions we aim to address with these
268 outcomes are described in **Supplemental Text 4**.

269

Running Head: PRIMERO Birth Cohort

270 Identification of Children with Early-life LRIs

271 Participants will be classified at the end of their 2-year surveillance period based on the severity
272 of LRI(s) experienced: did not experience an LRI; experienced at least one mild/moderate LRI,
273 but not a severe LRI; or experienced at least one severe LRI. In most cases, LRIs in the first two
274 years of life are determined by a physician assessment. However, due to missed visits for some
275 participants, a subset of LRIs is only defined by the symptoms reported during the screening and
276 tracking of an illness event.

277

278 Analysis of Respiratory Illness-Associated Airway Gene Expression

279 We will use RNA sequencing to measure whole transcriptome nasal airway epithelium gene
280 expression patterns at birth, during defined URI and LRI events, and from annual follow-up
281 visits from ages one through five.

282

283 Determination of Respiratory Virus Infection at a Species Level

284 Viral species-specific data generated from nasal airway swab RNA and metagenomic analysis of
285 RNA-sequence data will allow us to determine if a participant has been infected with a
286 respiratory virus and determine if viral species are associated with illness severity outcomes.

287

288 mAPI

289 Children undergo clinical assessment for asthma risk at two years of age via the mAPI, which is
290 an updated iteration of the asthma predictive index (API) and provides an assessment for future
291 asthma risk (36). The mAPI has greater sensitivity than the API and has been included in the
292 National Asthma Education and Prevention Program (NAEPP) Expert Panel Report 3 (EPR-3)

Running Head: PRIMERO Birth Cohort

293 guidelines as a criterion for the institution of long-term therapy to decrease asthma morbidity and
294 exacerbations (37).

295

296 Risk Mitigation Plan

297 Study enrollment commenced in March 2020, coinciding with the onset of the COVID-19
298 pandemic. After considering the risks of pandemic viral exposures to staff and participants,
299 PRIMERO leadership, in conjunction with HIMA, NHLBI, and the OSMB, decided to proceed
300 with the study, operating under a risk mitigation plan (**Supplemental Text 5**).

301

302 Participant Engagement

303 To enhance retention and engagement, mothers are initially reimbursed for parking costs at
304 HIMA during recruitment, and they receive a thank-you gift at the conclusion of the baseline and
305 each annual follow-up visit. Beyond study visits, participants are regularly contacted through
306 various channels, including SMS texts, email, and the study website (**Supplemental Figure 2**).
307 The study website and newsletters serve to keep participants informed about the study's progress
308 and provide prompts for updating contact information. Weekly contacts during the first two years
309 help identify RIs, and personalized SMS texts with a link to the study website are sent on the
310 child participants' birthdays. Additionally, participants receive annual holiday greetings via
311 email.

312

313 RESULTS

314 Recruitment

Running Head: PRIMERO Birth Cohort

315 Enrollment began in late February 2020 with the first eligible infant born in early March,
316 coinciding with the first reported cases of COVID-19 in Puerto Rico, at which point the governor
317 of Puerto Rico ordered an island-wide curfew and closure of non-essential businesses in mid-
318 March 2020 (38). As additional pandemic control measures ensued, PRIMERO remained one of
319 few studies that was successfully able to actively recruit through the height of the pandemic and
320 continued enrolling until June 2023. Over the 40-month recruitment period, HIMA delivered
321 3,543 babies. Mothers who delivered over the weekends or received prenatal care outside of
322 HIMA were not considered for recruitment. We approached 2,723 of those mothers with
323 welcome packets and screened them for potential enrollment. Of those, 318 mothers declined
324 participation, 183 mothers were ineligible, 12 did not reaffirm their consent, one birth was
325 missed, and 109 infants were not eligible, culminating in a total of 2,100 mother-child dyads
326 being enrolled in the study (**Figure 3**) representing 59% of all of HIMA births.

327

328 Descriptive Characteristics of Participants

329 Characteristics of PRIMERO mothers and children (**Table 3**) were derived from a baseline
330 questionnaire and infant eligibility form. The median age of mothers was 26.6 (IQR, 23.1–30.7)
331 years. Cesarean delivery was the predominant (72.6%) delivery method, with most (53.9%)
332 unplanned. The race of each child was derived from the parents' reported races, with 39.0%
333 White if both parents identified as White, 22.0% Black/African if both parents identified as
334 Black/African, and 39.0% Other/Mixed if one parent identified as White and the other as
335 Black/African, encompassing all possible combinations (39).

336

337 Annual Visits and Biospecimen Collection

Running Head: PRIMERO Birth Cohort

338 Among children eligible for annual visits, 57% of expected visits occurred and 62% of expected
339 questionnaires were completed as of February 3, 2024. Nasal swabs were successfully collected
340 from 99% of these children, and blood samples were obtained from 97%. Among children three
341 years and older who attended their third annual visit, 54% have completed impulse oscillometry.

342

343 Participant Retention

344 As of February 3, 2024, 1,940 study participants remain actively engaged, while 136 are on
345 standby status whereby the participant requested reduced contact frequency, relocated from
346 Puerto Rico, or informed us of an inability to commit to future follow-up. Additionally, 23
347 participants withdrew from the study and one child died due to reasons unrelated to the study.

348

349 Adverse Events

350 Related adverse events are defined as any unfavorable medical occurrence attributed to a
351 PRIMERO procedure, such as blood draw, nasal swab, or pulmonary function testing. A singular
352 adverse event was recorded, involving a syncopal episode during a child's blood draw.

353

354 DISCUSSION

355 *The Puerto Rican Infant Metagenomic and Epidemiologic Study of Respiratory Outcomes*
356 (PRIMERO) was designed to unravel the early life social, environmental, and genetic
357 contributors to severe respiratory wheezing and subsequent asthma. A primary goal is to
358 characterize the relationship between viral RIs in early life and the subsequent development of
359 childhood asthma, focusing on the pivotal role of airway epithelial dysfunction. PRIMERO sets
360 out to determine whether this dysfunction manifests at birth or emerges as a consequence of viral

Running Head: PRIMERO Birth Cohort

361 RIs during infancy and early childhood. Our prospective birth cohort study is poised to yield
362 groundbreaking insights into the genetic and viral risk factors associated with severe RIs.
363 Furthermore, it aims to pinpoint distinct airway cellular developmental trajectories observable in
364 children at high risk for asthma, both at birth, following RIs, and during periods free of
365 respiratory symptoms preceding the onset of asthma.

366
367 "*Primero*" means "first" in Spanish, signifying PRIMERO as the inaugural asthma birth cohort
368 study in Puerto Rico. The imperative for comprehensive research across diverse genomes cannot
369 be overstated as it is fundamental to advancing science and medicine (40,41). Despite significant
370 attention and resources directed toward genetic research, globally diverse populations remain
371 notably underrepresented (42). Here, PRIMERO is helping to bridge this gap towards equal
372 representation of study populations in biomedical science. The emphasis of PRIMERO on the
373 population of Puerto Rico is emblematic of the Strategic Vision of NHLBI, which underscores
374 the exploration of factors contributing to health disparities among diverse populations (43).
375 Additionally, recruitment and follow-up activities are exclusively conducted within Puerto Rico,
376 in close partnership with local investigators, clinicians, and other Puerto Rican professionals,
377 enriching the biomedical workforce with diversity and expertise.

378
379 A fundamental strength of PRIMERO lies in its innovative approach, capitalizing on the
380 minimally invasive nature of nasal airway swabs to procure critical respiratory samples at
381 multiple times: at birth, during early-life RIs, and annually for the first five years of life
382 (30,44,45). Our study utilizes cutting-edge multiplexed molecular assays, enabling the
383 identification of a spectrum of RI-causing viruses, including RSV, HRV, metapneumovirus,

Running Head: PRIMERO Birth Cohort

384 coronaviruses, parainfluenza, influenza, and other less studied species. Collected specimens from
385 PRIMERO participants undergo genetic characterization, allowing us to determine the role of
386 genetic variation in the effects of these viruses on LRI and asthma. Most notably, through
387 transcriptomic analyses of nasal airway mucosa, the PRIMERO study introduces a
388 groundbreaking approach by longitudinally characterizing molecular airway dysfunction at birth,
389 annually throughout childhood, and during RIs in the first two years of life. Extensive research
390 substantiates the effectiveness and reliability of nasal swabbing, underscoring critical findings: it
391 reveals that the upper and lower airways share similar cell types and expression patterns, and
392 upper airway insults can elicit lower airway symptoms and immune responses (30,46–49). The
393 inflammatory endotypes of asthma observed in both the nasal and bronchial airways exhibit a
394 high degree of correlation, emphasizing the importance of the nasal airway epithelium as a
395 primary site for infection with asthma-associated respiratory viruses (30). Furthermore, we
396 previously demonstrated the versatility of RNA-sequence data generated from nasal airway
397 epithelial swabs, serving a dual purpose: elucidating host transcriptional profiles and screening
398 samples for pathogenic respiratory viruses through metagenomic analysis (50). This single-swab
399 RNA-sequence approach offers a robust and tightly controlled assessment of the host response to
400 viral infection and stands as a central analysis method within our study.

401

402 As with any complex study, PRIMERO has its unique challenges. One challenge was the low
403 birth rate in Puerto Rico, which, at 0.90 births per woman, is approximately half that in the
404 mainland U.S. (51). Despite the relatively small source population of potential participants,
405 PRIMERO recruitment has been overwhelmingly successful. Another challenge was the lack of
406 available software to manage the recruitment, surveillance, and follow-up of 2,100 mother-child

Running Head: PRIMERO Birth Cohort

407 dyads and to interact with their electronic medical record (EMR) as they progress through the
408 study. A key innovation was the creation of PRIMERO-DB to address this need. PRIMERO-DB
409 has the same Health Insurance Portability and Accountability Act (HIPAA) securities as EMR
410 data and is available from any web-enabled device, providing user-specific permissions to
411 authorized users. Additionally, the COVID-19 pandemic coincided with the start of PRIMERO
412 recruitment. With pandemic control measures in place and OB/GYN offices limiting their indoor
413 capacity, recruiters lost the advantage of building the rapport provided through face-to-face
414 interactions. As recruitment continued throughout the evolving pandemic, the generalizability of
415 our findings to babies born before and after the pandemic will need to be evaluated.

416

417 Creating a birth cohort uniquely positions us to answer questions about the early-life origins of
418 health and disease. We aim to leverage this rare opportunity to study the broader health of
419 children over the next decades. Given adequate support, long-term prospective follow-up of
420 PRIMERO participants will enable studies of numerous exposures and outcomes and enable the
421 identification of biomarkers that may predict outcomes well before their traditional diagnosis.
422 This information collectively promises insights into potential pathways that may be targetable in
423 efforts to prevent illnesses and the subsequent development of childhood diseases such as
424 asthma.

425

426 ACKNOWLEDGMENTS

427 The authors acknowledge the families for their participation and thank the numerous data
428 collectors, recruiters, technicians, and hospital administrators for their support and participation
429 in PRIMERO. The authors would also like to thank Iliany Cubero, Peter Deford, George Halley,

Running Head: PRIMERO Birth Cohort

430 Michael Montgomery, James Nolin, Bhavika Patel, Elizabeth Plender, Lisa Butler, and Narjis
431 Kazmi. We also extend a special thank you to the San Francisco Chinatown Merchants
432 Association for their generous donations of personal protective equipment during the COVID-19
433 pandemic.

434

435 **Abbreviations**

436	API	asthma predictive index
437	CNP	Centro de Neumología Pediátrica
438	COVID-19	coronavirus disease 2019
439	EMR	electronic medical record
440	EPR-3	Expert Panel Report 3
441	GxE	gene-environment
442	HIMA	Hospital Interamericano de Medicina Avanzada
443	HIPAA	Health Insurance Portability and Accountability Act
444	HRV	human rhinovirus
445	IIAQ	Initial Illness Assessment Questionnaire
446	IRB	Institutional Review Board
447	ITFQ	Illness Tracking and Follow-up Questionnaire
448	LRI	lower respiratory tract illnesses
449	mAPI	modified asthma predictive index
450	MDIQ	Medical Doctor Illness Questionnaire
451	NAEPP	National Asthma Education and Prevention Program
452	NIH	National Institutes of Health

Running Head: PRIMERO Birth Cohort

453	NJH	National Jewish Health
454	non-RI	non-respiratory tract illness
455	OB/GYN	obstetric/gynecologic
456	OSMB	Observational Study Monitoring Board
457	PRAM	Pediatric Respiratory Assessment Measure
458	PRIMERO	Puerto Rican Infant Metagenomic and Epidemiologic Study of Respiratory
459	Outcomes	
460	qPCR	quantitative polymerase chain reaction
461	REDCap	Research Electronic Data Capture
462	RI	respiratory illness
463	RNA	ribonucleic acid
464	RSS	Respiratory Severity Score
465	RSV	respiratory syncytial virus
466	SMS	short message service
467	UCSF	University of California, San Francisco
468	URI	upper respiratory tract illness

469

470 REFERENCES

- 471 1. Johnson CC, Havstad SL, Ownby DR, et al. Pediatric asthma incidence rates in the United
472 States from 1980 to 2017. *J Allergy Clin Immunol*. 2021;148(5):1270–1280.
- 473 2. Table 4-1- Current Asthma Prevalence Percents by Age, United States: NHIS, 2018 | CDC.
474 2019;(https://www.cdc.gov/asthma/nhis/2018/table4-1.htm). (Accessed January 8, 2024)
- 475 3. Lugogo N, Judson E, Haight E, et al. Severe asthma exacerbation rates are increased among
476 female, Black, Hispanic, and younger adult patients: results from the US
477 CHRONICLE study. *J Asthma*. 2022;59(12):2495–2508.

Running Head: PRIMERO Birth Cohort

- 478 4. Miller RL, Schuh H, Chandran A, et al. Incidence rates of childhood asthma with recurrent
479 exacerbations in the US Environmental influences on Child Health Outcomes (ECHO)
480 program. *J Allergy Clin Immunol*. 2023;152(1):84–93.
- 481 5. Keet CA, Matsui EC, McCormack MC, et al. Urban residence, neighborhood poverty,
482 race/ethnicity, and asthma morbidity among children on Medicaid. *J Allergy Clin Immunol*.
483 2017;140(3):822–827.
- 484 6. Washington DM, Curtis LM, Waite K, et al. Sociodemographic Factors Mediate Race and
485 Ethnicity-associated Childhood Asthma Health Disparities: a Longitudinal Analysis. *J*
486 *Racial Ethn Health Disparities*. 2018;5(5):928–938.
- 487 7. Wohlford EM, Borrell LN, Elhawary JR, et al. Differential asthma odds following
488 respiratory infection in children from three minority populations. *PLoS One*.
489 2020;15(5):e0231782.
- 490 8. den Dekker HT, Jaddoe VWV, Reiss IK, et al. Fetal and Infant Growth Patterns and Risk of
491 Lower Lung Function and Asthma. The Generation R Study. *Am J Respir Crit Care Med*.
492 2018;197(2):183–192.
- 493 9. Örtqvist AK, Ullemar V, Lundholm C, et al. Fetal Growth and Childhood Lung Function in
494 the Swedish Twin Study on Prediction and Prevention of Asthma. *Ann Am Thorac Soc*.
495 2017;14(7):1147–1153.
- 496 10. Martinez FD, Guerra S. Early Origins of Asthma. Role of Microbial Dysbiosis and
497 Metabolic Dysfunction. *Am J Respir Crit Care Med*. 2018;197(5):573–579.
- 498 11. Bouzigon E, Corda E, Aschard H, et al. Effect of 17q21 variants and smoking exposure in
499 early-onset asthma. *N Engl J Med*. 2008;359(19):1985–1994.
- 500 12. Taussig LM, Wright AL, Holberg CJ, et al. Tucson Children’s Respiratory Study: 1980 to
501 present. *J Allergy Clin Immunol*. 2003;111(4):661–675; quiz 676.
- 502 13. McGuinness CB, Boron ML, Saunders B, et al. Respiratory syncytial virus surveillance in
503 the United States, 2007-2012: results from a national surveillance system. *Pediatr Infect*
504 *Dis J*. 2014;33(6):589–594.
- 505 14. Jackson DJ, Gangnon RE, Evans MD, et al. Wheezing rhinovirus illnesses in early life
506 predict asthma development in high-risk children. *Am J Respir Crit Care Med*.
507 2008;178(7):667–672.
- 508 15. Kusel MMH, de Klerk NH, Keadze T, et al. Early-life respiratory viral infections, atopic
509 sensitization, and risk of subsequent development of persistent asthma. *J Allergy Clin*
510 *Immunol*. 2007;119(5):1105–1110.
- 511 16. Molinari Such M, García I, García L, et al. Respiratory syncytial virus-related bronchiolitis
512 in Puerto Rico. *P R Health Sci J*. 2005;24(2):137–140.

Running Head: PRIMERO Birth Cohort

- 513 17. Castro-Rodriguez JA, Forno E, Rodriguez-Martinez CE, et al. Risk and Protective Factors
514 for Childhood Asthma: What Is the Evidence? *J Allergy Clin Immunol Pract.*
515 2016;4(6):1111–1122.
- 516 18. Castro-Rodríguez JA, Holberg CJ, Wright AL, et al. A clinical index to define risk of
517 asthma in young children with recurrent wheezing. *Am J Respir Crit Care Med.* 2000;162(4
518 Pt 1):1403–1406.
- 519 19. Rosas-Salazar C, Chirkova T, Gebretsadik T, et al. Respiratory syncytial virus infection
520 during infancy and asthma during childhood in the USA (INSPIRE): a population-based,
521 prospective birth cohort study. *Lancet.* 2023;401(10389):1669–1680.
- 522 20. Çalışkan M, Bochkov YA, Kreiner-Møller E, et al. Rhinovirus wheezing illness and genetic
523 risk of childhood-onset asthma. *N Engl J Med.* 2013;368(15):1398–1407.
- 524 21. Rosas-Salazar C, Tang Z-Z, Shilts MH, et al. Upper respiratory tract bacterial-immune
525 interactions during respiratory syncytial virus infection in infancy. *J Allergy Clin Immunol.*
526 2022;149(3):966–976.
- 527 22. Chirkova T, Rosas-Salazar C, Gebretsadik T, et al. Effect of Infant RSV Infection on
528 Memory T Cell Responses at Age 2-3 Years. *Front Immunol.* 2022;13:826666.
- 529 23. Connelly AR, Jeong BM, Coden ME, et al. Metabolic Reprogramming of Nasal Airway
530 Epithelial Cells Following Infant Respiratory Syncytial Virus Infection. *Viruses.*
531 2021;13(10).
- 532 24. Human S, Hotard AL, Rostad CA, et al. A Respiratory Syncytial Virus Attachment Gene
533 Variant Associated with More Severe Disease in Infants Decreases Fusion Protein
534 Expression, Which May Facilitate Immune Evasion. *J Virol.* 2020;95(2).
- 535 25. Rosas-Salazar C, Shilts MH, Tovchigrechko A, et al. Nasopharyngeal Microbiome in
536 Respiratory Syncytial Virus Resembles Profile Associated with Increased Childhood
537 Asthma Risk. *Am J Respir Crit Care Med.* 2016;193(10):1180–1183.
- 538 26. Brunwasser SM, Snyder BM, Driscoll AJ, et al. Assessing the strength of evidence for a
539 causal effect of respiratory syncytial virus lower respiratory tract infections on subsequent
540 wheezing illness: a systematic review and meta-analysis. *Lancet Respir Med.*
541 2020;8(8):795–806.
- 542 27. Fang L-C, Wang J-Y, Yu H-H, et al. Respiratory-syncytial virus immunoprophylaxis on
543 asthma symptoms development in prematurity with bronchopulmonary dysplasia. *J Allergy
544 Clin Immunol Glob.* 2023;2(4):100161.
- 545 28. Kampmann B, Madhi SA, Munjal I, et al. Bivalent Prefusion F Vaccine in Pregnancy to
546 Prevent RSV Illness in Infants. *N Engl J Med.* 2023;388(16):1451–1464.
- 547 29. Hammitt LL, Dagan R, Yuan Y, et al. Nirsevimab for Prevention of RSV in Healthy Late-
548 Preterm and Term Infants. *N Engl J Med.* 2022;386(9):837–846.

Running Head: PRIMERO Birth Cohort

- 549 30. Poole A, Urbanek C, Eng C, et al. Dissecting childhood asthma with nasal transcriptomics
550 distinguishes subphenotypes of disease. *J Allergy Clin Immunol.* 2014;133(3):670-678.e12.
- 551 31. Burchard EG, Avila PC, Nazario S, et al. Lower bronchodilator responsiveness in Puerto
552 Rican than in Mexican subjects with asthma. *Am J Respir Crit Care Med.*
553 2004;169(3):386–392.
- 554 32. Nishimura KK, Galanter JM, Roth LA, et al. Early-life air pollution and asthma risk in
555 minority children. The GALA II and SAGE II studies. *Am J Respir Crit Care Med.*
556 2013;188(3):309–318.
- 557 33. U.S. Census Bureau. San Juan-Carolina-Caguas, PR Metro Area.
558 ([https://censusreporter.org/profiles/31000US41980-san-juan-carolina-caguas-pr-metro-](https://censusreporter.org/profiles/31000US41980-san-juan-carolina-caguas-pr-metro-area/)
559 [area/](https://censusreporter.org/profiles/31000US41980-san-juan-carolina-caguas-pr-metro-area/))
- 560 34. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a
561 metadata-driven methodology and workflow process for providing translational research
562 informatics support. *J Biomed Inform.* 2009;42(2):377–381.
- 563 35. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international
564 community of software platform partners. *J Biomed Inform.* 2019;95:103208.
- 565 36. Chang TS, Lemanske RFJ, Guilbert TW, et al. Evaluation of the modified asthma predictive
566 index in high-risk preschool children. *J Allergy Clin Immunol Pract.* 2013;1(2):152–156.
- 567 37. Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-
568 Summary Report 2007. *J Allergy Clin Immunol.* 2007;120(5 Suppl):S94-138.
- 569 38. Puerto Rico COVID Resources | Government Actions: COVID-19 | Coronavirus COVID
570 19 Resource Center | Topics | Insights | Faegre Drinker Biddle & Reath LLP.
571 ([https://www.faegredrinker.com/en/insights/topics/coronavirus-covid-19-resource-](https://www.faegredrinker.com/en/insights/topics/coronavirus-covid-19-resource-center/government-actions-covid-19/puerto-rico-covid-resources)
572 [center/government-actions-covid-19/puerto-rico-covid-resources](https://www.faegredrinker.com/en/insights/topics/coronavirus-covid-19-resource-center/government-actions-covid-19/puerto-rico-covid-resources)). (Accessed September
573 19, 2023)
- 574 39. Puerto Rico's 2020 Race/Ethnicity Decennial Analysis | CentroPR.
575 ([https://centropr.hunter.cuny.edu/reports/puerto-ricos-2020-decennial-analysis-datasheet-](https://centropr.hunter.cuny.edu/reports/puerto-ricos-2020-decennial-analysis-datasheet-series/)
576 [series/](https://centropr.hunter.cuny.edu/reports/puerto-ricos-2020-decennial-analysis-datasheet-series/)). (Accessed January 8, 2024)
- 577 40. Oh SS, Galanter J, Thakur N, et al. Diversity in Clinical and Biomedical Research: A
578 Promise Yet to Be Fulfilled. *PLoS Med.* 2015;12(12):e1001918.
- 579 41. Choudhury A, Aron S, Botigué LR, et al. High-depth African genomes inform human
580 migration and health. *Nature.* 2020;586(7831):741–748.
- 581 42. Mills MC, Rahal C. A scientometric review of genome-wide association studies. *Commun*
582 *Biol.* 2019;2:9.

Running Head: PRIMERO Birth Cohort

- 583 43. Research Priorities | NHLBI, NIH. (<https://www.nhlbi.nih.gov/about/strategic->
584 vision/research-priorities). (Accessed September 19, 2023)
- 585 44. Sajuthi SP, DeFord P, Li Y, et al. Type 2 and interferon inflammation regulate SARS-CoV-
586 2 entry factor expression in the airway epithelium. *Nat Commun.* 2020;11(1):5139.
- 587 45. Chu C-Y, Qiu X, Wang L, et al. The Healthy Infant Nasal Transcriptome: A Benchmark
588 Study. *Sci Rep.* 2016;6:33994.
- 589 46. Harkema JR, Carey SA, Wagner JG. The nose revisited: a brief review of the comparative
590 structure, function, and toxicologic pathology of the nasal epithelium. *Toxicol Pathol.*
591 2006;34(3):252–269.
- 592 47. Corren J, Adinoff AD, Irvin CG. Changes in bronchial responsiveness following nasal
593 provocation with allergen. *J Allergy Clin Immunol.* 1992;89(2):611–618.
- 594 48. Littell NT, Carlisle CC, Millman RP, et al. Changes in airway resistance following nasal
595 provocation. *Am Rev Respir Dis.* 1990;141(3):580–583.
- 596 49. Schumacher MJ, Cota KA, Taussig LM. Pulmonary response to nasal-challenge testing of
597 atopic subjects with stable asthma. *J Allergy Clin Immunol.* 1986;78(1 Pt 1):30–35.
- 598 50. Wesolowska-Andersen A, Everman JL, Davidson R, et al. Dual RNA-seq reveals viral
599 infections in asthmatic children without respiratory illness which are associated with
600 changes in the airway transcriptome. *Genome Biol.* 2017;18(1):12.
- 601 51. World Bank Open Data. *World Bank Open Data.* (<https://data.worldbank.org>). (Accessed
602 January 8, 2024)
- 603

Running Head: PRIMERO Birth Cohort

604 **Table 1.** Schedule of study procedures and visits

Visit	Prenatal	Birth	Age 1	Age 2	Age 3	Age 4	Age 5
Recruitment							
• Eligibility and informed consent	X	X					
• Baseline questionnaire		X					
• Cord blood		X					
• Maternal blood		X					
• Nasal swabs		X					
Annual follow-up visits							
• Annual questionnaire			X	X	X	X	X
• Anthropometrics			X	X	X	X	X
• Nasal swabs			X	X	X	X	X
• Blood draw				X			
• Impulse oscillometry					X		
• Spirometry						X	
Respiratory illness surveillance (recurring)							
• Weekly SMS text/ or email messaging			X				
• In-person clinical illness assessments			X				
○ Respiratory illness questionnaire			X				
○ URI Nasal swabs			1 st URI per year				
○ LRI Nasal swabs			All LRI events				

605 Abbreviations: LRI, lower respiratory tract illness; SMS, short message service; URI, upper
 606 respiratory tract illness.

Running Head: PRIMERO Birth Cohort

607 **Table 2.** PRIMERO Enrollment Inclusion and Exclusion

	Inclusion	Exclusion
Mother	Residential address in Puerto Rico	
	At least 18 years of age	
	Have SMS text-enabled phone or working email address	
	Able to read and understand English or Spanish	
	Delivery at HIMA	
Newborn	Lives with biological mother	Sibling of a child already enrolled in PRIMERO
	At least 37 ^a weeks gestational age	Multiple birth ^c
	Birthweight of at least 2,500 g ^b	Requires mechanical ventilation
	Afebrile (<38°C)	Immune compromised
	No clinical signs of respiratory infection	Congenital ^d or neuromuscular ^e disorders

608 Abbreviations: SMS, short message service; HIMA, Hospital Interamericano de Medicina
609 Avanzada-San Pablo; g, grams; C, Celsius

610

611 ^a Newborns at least 36 weeks gestational age and ^b weighing within 10% of the 2,500 g cutoff
612 (i.e., weighing at least 2,250 g) at birth are evaluated for inclusion on a case-by-case basis by an
613 obstetrician/pediatrician who determined by physical examination that the child was otherwise
614 healthy.

615

616 ^c Only one child of a multiple birth is allowed to enroll.

617

618 ^d Examples include pulmonary abnormalities and Down syndrome.

619

620 ^e Impaired ability to clear airway secretions.

Running Head: PRIMERO Birth Cohort

621 **Table 3.** Characteristics of Enrolled Mothers and Children in PRIMERO, Puerto Rico, as of
622 February 3, 2024

Characteristic	No. of Participants	%	Median (IQR)
Maternal age, years			26.6 (23.1 – 30.7)
Cesarean deliveries	1,524	72.6	
Unplanned Cesarean	821	39.1	
Gestational age, weeks			38.7 (38.0 – 39.4)
Birth weight, grams			3,202 (2,915 – 3,466)
Child's sex, male	1,047	49.8	
Child's race ^a			
Black/African	461	22.0	
Other/Mixed	818	39.0	
White	820	39.0	
Not specified	1	0.0	
Perinatal antibiotics exposure ^b	433	20.7	
Maternal prenatal smoking ^c	22	1.1	
Maternal education status, low ^d	739	35.2	
Paternal education status, low ^d	946	45.1	

623
624 Abbreviations: IQR, interquartile range

625
626 ^a Child's race as a derivation of parents' reported races: "White" if both parents identified as
627 White, "Black/African" if both parents identified as Black/African, and "Other/Mixed" if one
628 parent identified as White and the other as Black/African.

629
630 ^b Perinatal antibiotic exposure if neonate was administered antibiotics in 3rd trimester or neonate
631 was administered antibiotics soon after birth.

632
633 ^c Maternal prenatal smoking if mother reported smoking during pregnancy or reported time from
634 quitting is less than child's gestation age.

635
636 ^d Low educational status if reported highest level of education as "Less than high school
637 education" or "High school education or GED."

Running Head: PRIMERO Birth Cohort

638 **Figure 1.** Municipalities and population density of Puerto Rico

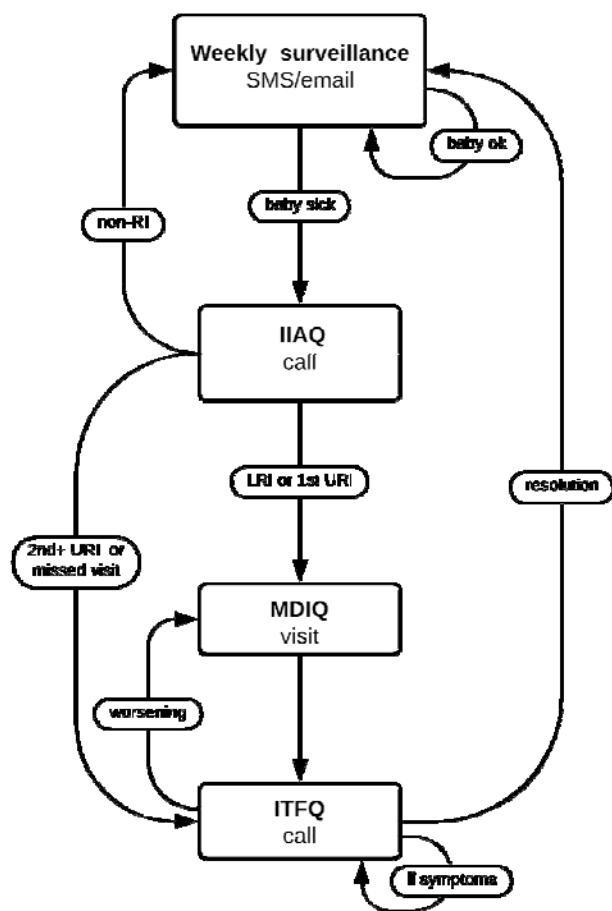
639



640 Redder areas represent municipalities with higher population densities. Greener areas represent
641 lower population densities. The San Juan-Bayamón-Carolina metropolitan area, bordered in blue,
642 represents HIMA's catchment area.
643

Running Head: PRIMERO Birth Cohort

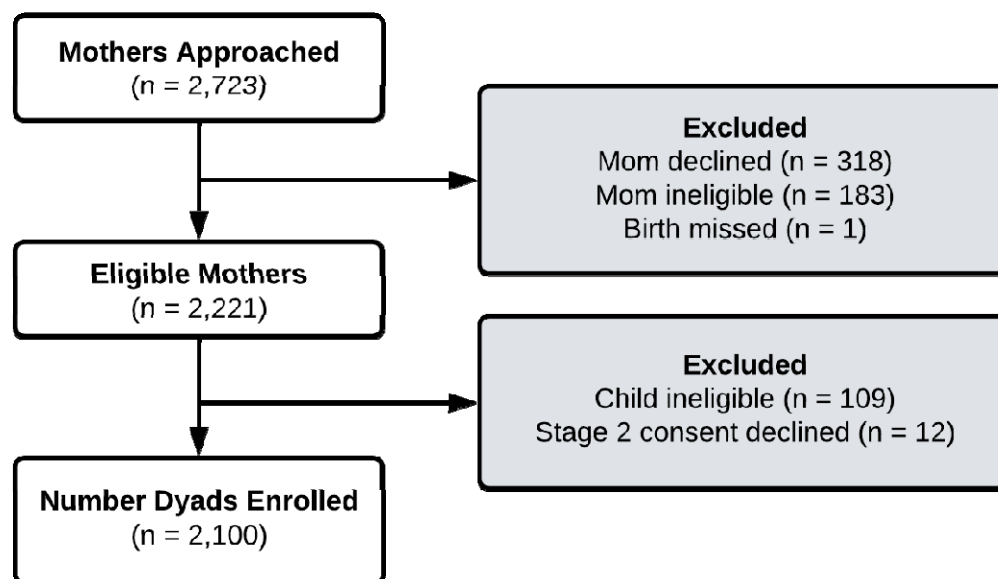
644 **Figure 2.** Illness Surveillance Flow Diagram



645
646 Abbreviations: IIAQ, Initial Illness Assessment Questionnaire; ITFQ, Illness Tracking and
647 Follow-up Questionnaire; LRI, lower respiratory tract illnesses; MDIQ, Medical Doctor Illness
648 Questionnaire; non-RI, non-respiratory tract illness; SMS, short message service; URI, upper
649 respiratory tract illness. Participants move through statuses assigned at each node to track the
650 illness from onset until resolution.

Running Head: PRIMERO Birth Cohort

651 **Figure 3.** Flowchart of PRIMERO enrollment spanning February 2020 to June 2023.



652