

1 The use of newborn foot length to identify low birth 2 weight and preterm babies in Papua New Guinea: A 3 diagnostic accuracy study

4 Alice Mengi¹ Lisa M Valley^{2,3}, Moses Laman⁴, Eunice Jally¹, Janeth Kulimbao¹, Sharon Ware¹, Regina
5 Enman¹, Jimmy Aipit⁵, Nicola Low^{6*}, Michaela Riddell^{1,3}.

- 6 1. Department of Infection and Immunity, Papua New Guinea Institute of Medical Research,
7 Madang, Papua New Guinea
- 8 2. Sexual and Reproductive Health Unit, Papua New Guinea Institute of Medical Research, Madang,
9 Papua New Guinea
- 10 3. Global Health Program, The Kirby Institute, University of New South Wales, Sydney, Australia
- 11 4. Department of Infection and Immunity, Papua New Guinea Institute of Medical Research,
12 Goroka, Papua New Guinea
- 13 5. Madang Provincial Health Authority, Madang, Papua New Guinea
- 14 6. Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

15 * Corresponding author

16 E-mail: nicola.low@ispm.unibe.ch (NL)

17

18 Abstract

19 Low birth weight (LBW, <2.50 kg) and preterm birth (PTB, <37 completed weeks of
20 gestation) are important contributors to neonatal death. Newborn foot length has been assessed to
21 identify LBW and PTB babies in Asia and Africa. The objectives of this observational study were to
22 determine the diagnostic accuracy of foot length to identify LBW and PTB and to compare foot
23 length measurements of a researcher with those of trained volunteers in Papua New Guinea.
24 Newborn babies were enrolled prospectively with written informed consent from their mothers,
25 who were participating in a clinical trial in Madang Province. The reference standards were birth
26 weight, measured by electronic scales and gestational age at birth, based on ultrasound scan and
27 last menstrual period at the first antenatal visit. Newborn foot length was measured within 72 hours
28 of birth with a firm plastic ruler. Optimal foot length cut-off values for LBW and PTB were derived
29 from receiver operating characteristic curve analysis. Bland-Altman analysis was used to assess inter-
30 observer agreement between researcher and volunteer. From 12 October 2019 to 6 January 2021,
31 we enrolled 342 newborns (80% of those eligible); 21.1% (72/342) were LBW and 7.3% (25/342)
32 were PTB. The area under the curve for LBW was 87.0% (95% confidence intervals 82.8-90.2) and for
33 PTB 85.6% (81.5-89.2). The optimal foot length cut-off was <7.7 cm for both LBW (sensitivity 84.7%,
34 74.7-91.2, specificity 69.6%, 63.9-74.8) and PTB (sensitivity 88.0% (70.0-95.8), specificity 61.8%
35 (56.4-67.0). In 123 babies with paired measurements, the mean difference between the researcher
36 and volunteer measurements was 0.07 cm (95% limits of agreement -0.55 to +0.70) and 7.3%
37 (9/123) of the pairs were outside the 95% limits of agreement. When birth at a health facility is not
38 possible, foot length measurement can identify LBW and PTB in newborns but needs appropriate
39 training for community volunteers and evaluation of its impact on healthcare outcomes.

40 Introduction

41 Globally, complications of preterm birth (PTB), including low birth weight (LBW), are the
42 leading cause of neonatal death in the first week of life [1]. Identifying the most vulnerable neonates
43 so that they can receive timely essential care is challenging in resource-limited settings, particularly
44 in rural areas or when births occur in the community [2]. Babies born LBW (<2.50 kg) [3] are either
45 preterm (<37 completed weeks gestation) or small for gestational age (birth weight below the 10th
46 percentile of the distribution for gestational age and sex) [3,4]. Many babies are, however, not
47 weighed at birth [5] or their mothers have not had an adequate assessment during pregnancy to
48 allow gestational age at birth to be calculated [2]. These assessments require reliable equipment and
49 skilled, trained health workers to conduct ultrasound scans during pregnancy, measure birth weight
50 accurately, or apply clinical assessments, such as Dubowitz or New Ballard scores [6].

51 Foot length is one of a range of anthropometric measures that have been used as proxies for LBW
52 and PTB in resource-limited settings [6-18]. In a systematic review and meta-analysis of studies
53 published between 2007 and 2020, Folger et al. found 19 studies, all conducted in south and south-
54 east Asia and sub-Saharan Africa [19]. The authors found that a newborn foot length of <7.7cm had
55 reasonable accuracy for identification of LBW babies in 3 studies in Asia (pooled sensitivity 87.6%,
56 95% confidence interval, CI 55.7-97.5%, specificity 70.9%, 23.5-95.1%). They noted, however, that
57 few studies used an accurate reference method for assessing gestational age and did not conduct a
58 meta-analysis of foot length and PTB because of heterogeneity in study methods [19]. The optimal
59 cut-off for newborn foot length may differ by country and region and performance characteristics
60 could also vary, according to whether measurements are taken by clinically trained staff or
61 community volunteers, field workers and care givers [11,20].

62 Papua New Guinea (PNG) is an island state in the South Pacific, which has one of the highest
63 neonatal mortality rates in the world, with 22 newborn deaths per 1000 live births in 2020 [21,22].
64 Although supervised delivery at a health facility is encouraged, uptake of antenatal care in PNG is

65 poor and high-risk babies cannot be easily identified and referred [21,23]. From 2016-2020, only
66 around half of pregnant women in PNG had at least one antenatal visit and just over a third gave
67 birth in a health facility. In Madang Province, about a third of pregnant women had at least one
68 antenatal visit and about a quarter gave birth in a health facility [21]. The objectives of this study
69 were to 1) determine the diagnostic accuracy of newborn foot length measurement to identify LBW
70 or PTB babies in health facilities and community settings and 2) to compare foot length
71 measurements taken by trained volunteers with those of a researcher to assess the feasibility of
72 newborn foot length measurement in a community setting.

73 **Materials and methods**

74 The Neofoot (neonatal foot length) study was a prospective single arm diagnostic test
75 accuracy study (S1 Protocol). We report the study using the Standards of Reporting of Diagnostic
76 accuracy 2015 (supplement S1 Checklist). The study was conducted in three primary care clinics in
77 Madang Province, PNG, which serve a total population of approximately 35,000, according to the
78 2011 PNG National Census [24]. The study population included both people living in urban shanty
79 settlements and rural subsistence farmers in areas with poor road conditions and difficult access to
80 health facilities. The Neofoot study was designed as a sub-study of the Women And Newborn Trial of
81 Antenatal Interventions and Management (WANTAIM) trial, a randomised controlled trial which
82 assessed the effect of point-of-care testing and treatment of curable sexually transmitted infections
83 and bacterial vaginosis during pregnancy on birth outcomes [25]. Women enrolled in the WANTAIM
84 study were aged 16 years or older at a gestational age below 26 completed weeks, based on the
85 date of the first day of the last menstrual period or an ultrasound scan performed by trained
86 midwives at the first antenatal visit [26]. Their newborn babies were eligible for inclusion if the
87 clinical researcher (AM) or senior study staff (EJ, JK, SW, RE) assessed them within 72 hours of birth,
88 their general condition was stable and the baby did not have abnormalities. There were no
89 additional exclusion criteria. At the first postnatal visit, the researcher or staff member used pictorial

90 diagrams to explain the Neofoot study in Tok-Pisin to consecutive women whom they attended,
91 assessed their understanding of the study and answered questions. The mother or guardian gave a
92 signature or thumbprint as informed consent for their baby to take part (S1 Protocol, page 20-21).
93 For women who were unable to read or write, an independent witness was present during the
94 consent procedures and also signed the consent form.

95 **Reference standards**

96 Birth weight was the average of two measurements taken at a 5 to 15 minute interval by
97 trained WANTAIM clinical staff within 72 hours of birth using an electronic infant scale (Charder
98 Cupid 1, Charder Medical, Taiwan) placed firmly on a hard surface and calibrated to zero [27]. Birth
99 weight was before the index test on a separate clinical record form. LBW was defined as below 2.50
100 kg [3]. Gestational age at birth was measured in completed weeks of pregnancy, calculated from the
101 best obstetric estimate of the due date [26], using the date of the first day of the last menstrual
102 period or an ultrasound scan (Logiq V2 portable ultrasound, GE Healthcare) performed by trained
103 midwives at the first antenatal visit. Gestational age at birth was calculated at the time of statistical
104 analysis. PTB was defined as birth before 37 completed weeks of pregnancy [4].

105 **Index test**

106 The index test was the length of the right foot, measured within 72 hours of birth. We
107 calculated the average of two measurements, taken 5 to 15 minutes apart by the clinical researcher.
108 Foot length was measured using a firm clear plastic 15cm ruler, bought from a local store and stuck
109 to an A4-sized cardboard box, which was cut open (Fig 1). The tip of the ruler was pushed into a slit
110 at the bottom of the box so that the zero level was at the base of the box. The researcher placed the
111 right heel at the zero level and held the plantar surface of the foot against the ruler. Foot length at
112 the tip of the hallux was measured to the nearest 1 mm, with a plastic setsquare perpendicular to
113 the ruler.

114 **Fig 1. Measurement of newborn foot length.** Measurements were made with a 15cm clear plastic
115 ruler stuck to the inside of a cardboard box. The zero mark of the ruler is at the base of the box. Foot
116 length is measured from the base of the heel to the tip of the hallux.

117 **Training of volunteers**

118 To assess the feasibility of community members or village health workers identifying at risk
119 babies, we enrolled nine volunteers, who were involved in the WANTAIM trial and were numerate
120 and literate. The clinical researcher gave a half-day training session at the postnatal ward of Madang
121 Provincial Hospital. The volunteers learned how to measure the baby's foot length, how to complete
122 the case record form and practised both tasks.

123 **Measures to reduce the risk of bias**

124 To reduce the risk of measurement bias in foot length, the researcher and the volunteer
125 recorded two measurements at a 5-to-15-minute interval. To ensure that the measurements were
126 independent, the researcher and volunteer recorded their measurements on separate pages of the
127 case record form and could not see the other person's record. Birth weight was measured before the
128 index test on a separate clinical record form. The clinician researcher had access to clinical
129 information about the newborn, but the best obstetric estimate of gestational age at birth [26] was
130 calculated at the time of statistical analysis.

131 **Ethics approval**

132 The Neofoot study protocol was approved by the PNG Institute of Medical Research
133 Institutional Review Board (IRB 1811) and the PNG Medical Research Advisory Committee (MRAC
134 18.18). The WANTAIM study (ISRCTN37134032) received ethical approval from the Institutional
135 Review Board of the PNG Institute of Medical Research (IRB number 1608); the Medical Research
136 Advisory Committee of the PNG National Department of Health (MRAC number 16.24); the Human

137 Research Ethics Committee of the University of New South Wales (HREC number 16708); and the
138 Research Ethics Committee of the London School of Hygiene and Tropical Medicine (REC number
139 12009) [25].

140 **Data management**

141 Data were recorded on printed case record forms designed using TELEform™ Elite version
142 10.5 (<https://teleform.software.informer.com/10.5/>). The case record forms were checked for
143 discrepancies or out of range values and corrected before scanning, after which they were machine-
144 read into a database (Microsoft). Variables collected from the mother at the first antenatal visit, as
145 part of the WANTAIM study, included demographic details (province of birth, age, and marital
146 status), gestational age at enrolment, estimated due date, haemoglobin, HIV test result
147 (Determine™ HIV 1/2 Test), syphilis (SD Bioline, Abbott Diagnostics) results and use of betel nut,
148 cigarettes, and alcohol during pregnancy. Data collected at the postnatal visit included date and
149 place of birth, reason for giving birth at that location, newborn birth weight and sex of the baby.
150 Data about foot length were the two measurements made by the researcher and the volunteers.

151 **Statistical analysis**

152 All statistical analyses used STATA (Stata 14.2 or 16, College Station, Texas, United States).
153 For each newborn we calculated the gestational age at birth using the estimated due date
154 established at the enrolment visit, and the average of the two measurements of birth weight and of
155 foot length. The distributions of continuous variables were examined using histograms and
156 described using the mean and standard deviation (SD) or median and interquartile range (IQR). For
157 the classification of LBW and PTB, birth weight and gestational age at birth were dichotomised in
158 accordance with the reference standard definitions.

159 We constructed receiver operating characteristics curves to display sensitivity against 1 minus
160 specificity for the classification of LBW and PTB at each 1 mm increment in foot length from the

161 lowest to the highest value. We calculated the overall area under the curve and test performance
162 characteristics sensitivity, specificity, positive and negative predictive values (as percentages with
163 95% CI). We selected the optimal foot length cut-off, separately for LBW and PTB, as the value with
164 the best balance of sensitivity and specificity, using the foot length cut-off with the highest area
165 under the curve. For each baby for whom paired foot length measurements were available from the
166 researcher and the volunteer, we assessed the level of agreement using Bland-Altman analysis [28].

167 **Sample size calculation**

168 The planned sample size was based on achieving an acceptable level of precision (95% CI +/-
169 4 to 5%) for values of sensitivity from 70 to 90%. Using projected WANTAIM trial enrolment at the
170 study clinics, we aimed to invite around 400 women to join the Neofoot study. Assuming 95% of
171 babies had a postnatal visit within the first 72 hours and that 90% of mothers agreed to foot length
172 measurement, the target was to have foot length measurements for 342 newborn babies.

173 **Results**

174 A total of 426 babies were born to 416 mothers enrolled in the WANTAIM trial from 12
175 October 2019 to 6 January 2021. Of these, 84 babies were excluded from the study for reasons
176 outlined in Fig 2. A total of 342 babies born (80.2% of those potentially eligible) to 335 mothers were
177 enrolled. The study population included 10 of 14 twins born to 7 mothers; 3 mothers had both twins
178 enrolled and four mothers had one twin enrolled (the other twin died before birth or before
179 neonatal assessment). The researcher recorded foot length measurements for all 342 babies. Owing
180 to disruptions during the COVID-19 pandemic and staff availability, a volunteer was available to
181 measure foot length at the time of the postnatal assessment for 123 (36.0%) babies.

182

183 **Fig 2. Flow chart indicating number of mothers and newborn babies at each stage of the Neofoot**
 184 **study.**

185 Among the 335 mothers, their mean age was 25.5 (SD 5.9, range 16-44) years, 43.9% (147/335) were
 186 primiparous and 89.3% (299/335) were married (Table 1). Sixty-seven women attended one of the
 187 primary care clinics covering urban shanty settlement areas and 268 attended clinics in rural areas.
 188 Most women reported that they currently chewed betel nut (281/335, 83.9%) and about a quarter
 189 were current smokers (81/335, 24.2%), but few (6/335, 1.8%) reported that they consumed alcohol
 190 during pregnancy. Forty-two women 12.5% (42/335) had a positive test result for syphilis and 2
 191 women (0.6%) had a positive test result for HIV. The mean haemoglobin concentration was 8.6 g/dL
 192 (SD 1.6), nearly two thirds of women (63.9%; 214/335) had mild anaemia (haemoglobin 7.0–9.9
 193 g/dL) and 13.7 % (46/335) had moderate to severe anaemia (haemoglobin <7.0 g/dL).

194 **Table 1. Characteristics of mothers whose newborn babies were enrolled in the study (N=335)**

Maternal characteristic		Total N = 335
Age, years	Mean (SD)	25.5 (5.9)
Birthplace, region n (%)	Momase	295 (88.1)
	New Guinea Island	19 (5.7)
	Highlands	16 (4.8)
	Southern Region	5 (1.5)
Marital status, n (%)	Married	299 (89.3)
	Single	29 (8.7)
	Divorced/separated/widowed	3 (0.9)
	Missing	4 (1.2)
Chew betelnut, n (%)	Currently chew	281 (83.9)
	Previously chew	5 (1.5)
	Never chew	49 (14.6)
Drink alcohol, n (%)	Currently drink	6 (1.8)
	Previously drink	31 (9.3)
	Never drink	298 (89.0)
Smoke cigarettes, n (%)	Currently smoke	81 (24.2)
	Previously Smoke	34 (10.2)
	Never smoke	220 (65.7)

Parity, n (%)	0	147 (43.9)
	1-3	141 (42.1)
	≥ 4	47 (14.0)
Current pregnancy, n (%)	Singleton	325 (97.0)
	Twin	10 (3.0)
Gestation in weeks, median (IQR)	USS	22 (19.0-24.0)
	Fundal height	22 (19.0-24.0)
	LMP	20 (17.0-23.0)
HIV test result, n (%)	Positive	2 (0.6)
	Negative	323 (96.4)
	Not tested	10 (3.0)
Syphilis test result, n (%)	Positive	42 (12.5)
	Negative	293 (87.5)
Haemoglobin, g/dL	Mean (SD)	8.6 (1.6)
Haemoglobin, category	≥ 10 g/dL	59 (17.6)
	7.0-9.9 g/dL	214 (63.9)
	< 7.0 g/dL	46 (13.7)
	Missing	16 (4.8)

195 Abbreviations: cm, centimetres; g/dL, grams per decilitre; HIV, human immunodeficiency virus; IQR,
 196 interquartile range; LMP, last menstrual period; N, total number of participants; n, number of
 197 participants in a category; SD, standard deviation; USS, ultrasound scan.

198 Table 2 describes the characteristics of the 342 babies. Most babies (330/342, 96.5%) were born by
 199 normal vaginal birth at a health facility (72.8%; 249/342), with 27.2% (93/342) births taking place at
 200 home or before arrival at a health facility.

201 Birth weight was normally distributed (mean 2.8 kg, SD 0.4), with a high frequency of birth weights
 202 recorded as 2.50 kg (known as digit preference or heaping, supplement S1 Fig S2 Fig).

203

204 **Table 2. Characteristics of newborn babies in the study (N=342)**

Newborn characteristic		Total N = 342	
Place of birth, n (%)	Health facility	249	(72.8)
	Born before arrival at health facility	5	(1.5)
	Home/village	88	(25.7)
Type of birth, n (%)	Normal vaginal birth	330	(96.5)

	Caesarean section	3	(0.9)
	Assisted birth	9	(2.6)
Sex, n (%)	Male	177	(51.8)
	Female	165	(48.3)
Birth weight, kg	Mean (SD)	2.8	(0.40)
	Median (IQR)	2.8	(2.5-3.1)
	Range, lowest-highest		1.3-4.2
Birth weight category n (%)	LBW (< 2.50 kg)	72	(21.0)
	Normal birth weight (\geq 2.50 kg)	270	(79.0)
Gestational age at birth, weeks	Mean (SD)	38.8	(1.8)
	Median (IQR)	39	(38-40)
	Range, lowest-highest		30-42
Gestational age category n (%)	Preterm < 37 weeks	25	(7.3)
	Term \geq 37 weeks	317	(91.7)
Foot length, cm	Mean (SD)	7.8	(0.5)
	Median (IQR)	7.8	(7.5-8.0)
	Range, lowest-highest		5.8-8.9

205 Abbreviations: IQR, interquartile range; LBW, low birth weight; N, total number of
 206 participants; n, number of participants in a category; SD, standard deviation.

207 Fig 3 shows scatterplots of birth weight, gestational age at birth and the foot length measured by the
 208 researcher. About one-fifth of all babies (21.1%, 72/342), were LBW (<2.50 kg). The mean
 209 gestational at birth was 38.8 weeks (SD 1.8) and 7.3% (25/342) were preterm. Of the babies born
 210 preterm (<37 completed weeks of gestation), 5/25 babies were of normal weight (\geq 2.50 kg) and
 211 amongst 317 babies born at term, 52 babies had a birth weight <2.50 kg (Fig 3). Mean foot length
 212 was 7.8 cm (SD 0.5) with no adverse events recorded as a result of measurement.

213

214 **Fig.3. Scatter plots of newborn measurements recorded within 72 hours of birth for 342**

215 **babies.** Panel A. Birth weight and foot length; B. gestational age at birth and foot length; C.

216 gestational age at birth and birth weight. r, Pearson correlation coefficient. Reference lines in

217 red indicate the cut-offs for low birth weight (2.50 kg) and preterm birth (37 completed

218 weeks of gestation).

219

220 **Optimal foot length cut-offs for classifying LBW and PTB**

221 The receiver operating curves summarise the overall accuracy of foot length measurement
 222 for the classification of newborn babies as LBW or PTB (Fig 4) at cut-off increments of 1 mm. The
 223 overall area under the curve for foot length was 87.0% (95% CI, 82.3-91.6) for LBW and 85.8% (77.1-
 224 94.6) for PTB.

225 **Fig. 4. Receiver operating characteristics curves of the accuracy of foot length for the classification**
 226 **of outcomes in 342 newborn babies.** Panel A, low birth weight, B preterm birth.

227 Optimal foot length cut-offs for the classification of LBW and PTB were determined as the value with
 228 the combination of the highest sensitivity and specificity (Table 3, supplement S1 Table, S2 Table).
 229 The optimal foot length cut-off for identifying LBW was <7.7 cm. At this foot length, 61/72 babies
 230 were correctly classified as LBW (sensitivity 84.7%, 74.7-91.2). The birth weights of 11 babies with
 231 foot length ≥7.7 cm, who were incorrectly classified as being of normal birth weight (false negatives),
 232 ranged from 2.15-2.43 kg. Among all babies with foot length <7.7 cm, 42.7 (34.8-50.9) were LBW
 233 (positive predictive value). The negative predictive value was 94.5% (90.4-96.9).

234 **Table 3. Diagnostic accuracy statistics at the optimal foot length cut-off for classification of babies**
 235 **as low birth weight or preterm**

Outcome	<7.7 cm	≥ 7.7 cm	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Positive predictive value, % (95% CI)	Negative predictive value, % (95% CI)
Birth weight						
< 2.50 kg	61	82	84.7	69.6	42.7	94.5
≥ 2.50 kg	11	188	(74.7-91.2)	(63.9-74.8)	(34.8-50.9)	(90.4-96.9)
Gestational age at birth						
< 37 weeks	22	121	88.0	61.8	15.4	98.5
≥ 37 weeks	3	196	(70.0 – 95.8)	(56.4 – 67.0)	(10.4 – 22.2)	(95.7 – 99.5)

236 The optimal foot length cut-off for PTB was <7.7 cm (Table 3), with 22/25 babies correctly classified
237 (sensitivity 88.0%, 70.0-95.8). Of the 3 preterm babies not identified as PTB at this cut-off (false
238 negatives), 2 were born at 36 weeks and 1 at 35 weeks gestation; none was very (28-32 weeks) or
239 extremely (<28 weeks) preterm. At this cut-off, 121 babies were false positives, being born at term
240 but with foot length <7.7 cm. The positive predictive value of a foot length of <7.7 cm was 15.4%
241 (10.4-22.2) and negative predictive value 98.5% (95.7-99.5).

242 Foot length measurements from both the researcher and a volunteer were available for 123 babies.
243 Values at full or half centimetre marks were over-represented for both sets of average
244 measurements (S2 Fig). Fig 5 shows the Bland-Altman plot, displaying the differences between the
245 123 paired measurements against the average of the measurements of the researcher and
246 volunteer. The mean difference between researcher and volunteer measurements (dashed line) was
247 0.07 cm. The 95% limits of agreement ranged from -0.55 cm to +0.70 cm. Differences between
248 measurements for 9/123 (7.3%) of the data pairs were outside the 95% limits of agreement.

249

250 **Fig 5. Bland-Altman plot for researcher and volunteers' newborn foot length measurements**

251 **(N=123)**. Size of circle is proportional to number of measurements.

252

253 Discussion

254 Summary of main findings

255 This prospective study enrolled 342 newborn babies in Madang Province, PNG. Overall, 21.1%
256 (72/342) of newborns were LBW and 7.3% (25/342) were born preterm. The optimal foot length cut-
257 off to classify LBW babies was <7.7 cm (sensitivity 84.7%, 95% CI 74.7-91.2, specificity 69.6%, 63.9-
258 74.8) and <7.7 cm to classify PTB (sensitivity 88.0%, 70.0-95.8, specificity 61.8%, 56.4-67.0%). The

259 95% limits of agreement between researcher and volunteer measurements were from -0.55 to +
260 0.70 cm and 7.3% (9/123) of the pairs were outside these limits.

261 **Strengths and limitations**

262 Strengths of this study included the use of accurate methods for determination of the reference
263 standards for measuring birth weight and gestational age at birth. All babies were weighed twice
264 using electronic scales [27] and gestational age at birth was calculated using published
265 recommendations for the best obstetric estimate of the delivery date, which are based on the date
266 of the last menstrual period and an ultrasound scan at the first antenatal visit [26]. Foot length was
267 measured using items readily available in the community and the study involved trained volunteers
268 from communities, with babies assessed at local health centres or at home. Also, the target sample
269 size was reached, and outcome variables LBW and PTB were available for all babies in the study,
270 even though a COVID-19 lock down was in place for part of the study.

271 There were also limitations of the study methods. First, the measurement of birth weight was taken
272 within the first 72 hours of birth. Ideally, birth weight should be measured within a few hours of
273 birth, before any postnatal weight loss has occurred [27]. However, Marchant et al. found that foot
274 length increased by only 0.2 cm (SD 0.3) between the first and fifth day after birth in a study in
275 Tanzania [10], so the delay of up to 72 hours in our study should not have affected the results.

276 Second, the researcher and volunteers could not be blinded from the clinical status of the newborn,
277 which might have influenced their foot length measurements. Knowledge of birthweight could have
278 resulted in overestimation of the accuracy of classification of LBW. Since the overall areas under the
279 curve for LBW and PTB were similar, we think risk of bias was low. Third, whilst mothers were
280 enrolled at primary care clinics and most gave birth at a health centre or at home/village, the study
281 was done in only one coastal province of PNG. The 2020 PNG National Department of Health Sector
282 Review Performance found that Madang Province was one of the provinces with the highest

283 percentage of low birth weight babies (13% in 2020) [21]. Optimal foot length cut-off values might
284 be different in other regions of PNG or other Pacific islands.

285 **Comparison with other studies**

286 Our study from PNG in the Pacific region adds to the evidence from previously published studies that
287 have explored foot length as a proxy to identify LBW and PTB and were conducted in countries in
288 Asia and Africa. The levels of LBW (around 20%) and PTB (around 7%) in the Neofoot study were
289 similar to those of several other published studies that have assessed the diagnostic accuracy of foot
290 length both in sub-Saharan Africa [10,15] and south Asia [29]. Our study in PNG, however, overcame
291 limitations of assessment of gestational age and of applicability of findings from tertiary care setting
292 that were identified in previous studies [19]. The Neofoot study shared several characteristics with a
293 study by Lee et al. among 710 babies in the community in Bangladesh, but with different findings
294 about the accuracy of foot length as a method to identify PTB [29]. Both studies were nested in
295 randomised controlled trials, the investigators assessed babies within 72 hours after birth, estimated
296 gestational age based on ultrasound scan and used a clear plastic ruler to measure foot length. Lee
297 et al. found that foot length measured at home visits by trained community health workers did not
298 classify PTB well, with an overall area under the curve of 51.9% and sensitivity at a cut-off of <7.7 cm
299 of 28%. In contrast, we found an overall area under the curve of 85.6% and sensitivity of 88.0% for
300 the same cut-off. Lee et al. attributed the poor performance of foot length (and other
301 anthropometric measurements) to the high level of foetal growth restriction (32.4% of babies were
302 assessed as being small for gestational age). In our study, however, levels of PTB and mean birth
303 weight were similar to those of Lee et al. and 12.3% (42/342) of term babies were LBW. One factor
304 that might have contributed to the different findings was that all babies in Lee et al.'s study were
305 assessed at home by community health workers, whilst more babies in our study were assessed at a
306 health facility or community by a paediatrician researcher, where the conditions for measuring foot
307 length might have been easier. Given the study setting, the proportion of supervised deliveries in

308 our study population was higher than the average for Madang Province [21]. Our study could be
309 repeated among women who do not give birth at a health facility to assess the performance of foot
310 length measurement by a community health worker.

311 **Interpretation of the findings**

312 In the Neofoot study, a foot length cut-off of <7.7 cm had similarly high overall accuracy for
313 classification of LBW and PTB (Table 3). A common cut-off for identifying LBW and PTB was also
314 found by Paulsen et al. (≤ 7.7 cm) [30] and Marchant et al. (<8 cm) in Tanzania [10]. The cut-off of
315 <7.7 cm for LBW concurs with the pooled result from the meta-analysis of 15 studies by Folger et al.,
316 which included studies from a range of countries [19]. The lower prevalence of PTB, compared with
317 LBW, results in a lower positive predictive value. At the optimal foot-length cut-off, around 15% of
318 babies were preterm, according to the reference standard (compared with nearly half of babies
319 classified as LBW). The referral of substantial numbers of babies for urgent care would increase the
320 workload for health systems and should be assessed. The use of a plastic ruler was an appropriate
321 technology for the measurement of foot length and taking the average of two measurements should
322 have reduced error, although we still observed digit preference for measurements at the half or full
323 centimetre (S2 Fig). A single training session for lay people and health facility staff to measure foot
324 length resulted in average foot length measurements that were similar to those of the researcher,
325 but paired measurements for individual babies differed by up to 1 cm. These findings were
326 comparable to those of Marchant et al. who assessed 142 paired measurements, using a plastic
327 ruler, between researchers and community volunteers in Tanzania [11]. As in our study, they found
328 no overall bias but volunteers' measurements were shorter than those of researchers.

329 **Implications for research and practice**

330 The goal of universal health coverage to ensure that all sick and small newborns survive and thrive
331 will take time to achieve; in the meantime, there is a need for innovative ways to simplify

332 assessment of gestational age at birth [2]. The findings of this study have implications for future
333 research and practice in settings such as PNG, where it is difficult to identify high risk babies in
334 populations with rural communities, high levels of home births and limited access to health facilities
335 with skilled staff and equipment. Most women in PNG attend antenatal clinic late in pregnancy,
336 often in the last trimester, so assessment of gestational age at birth is particularly challenging. Even
337 though the Neofoot study was nested in a trial following Good Clinical Practice guidelines, with
338 ultrasound-assisted pregnancy dating and efforts to encourage supervised delivery, more than a
339 quarter of women did not give birth at a health facility. Most of these women lived in hamlets, which
340 are an hour or more from the main road by foot, and from which access to the labour ward at the
341 health centre was difficult, especially at night. In settings such as these, validation of newborn
342 footlength as a surrogate for LBW and PTB is important. Our study adds to the published evidence
343 [19], showing that both LBW and PTB can be classified accurately and could be used by trained
344 health workers and community health volunteers. Newborn foot length has not, however, been
345 widely implemented. Implementation research studies should develop appropriate technologies to
346 apply optimal cut-off measurements, training programmes that reduce misclassification from digit
347 preference (heaping) and examine the impacts on referrals to health systems and health outcomes.
348 In this study in PNG, newborn foot length was a simple and feasible method for the identification of
349 LBW and PTB babies and could help to improve neonatal health outcomes.

350

351 **Acknowledgments**

352 This study was conducted during the Master of Medical Sciences, University of Papua New
353 Guinea of Alice Mengi. We sincerely thank all the parents of the 342 babies that participated, the
354 PNG Institute of Medical Research and WANTAIM staff, especially Kelly Masil for scanning and
355 verification of data and Andrew Vallely and William Pomat for acquisition of funding and oversight of

356 the WANTAIM trial. We thank Odile Stalder at the Clinical Trials Unit, University of Bern, who helped
357 with the statistical analyses. We thank Sister Hilda Marijembi, Sister Susie Yuangi, Naida Samson,
358 Sister Bridegette Nawali, Lina Puli, Josephine Fei, Christophilda Noubiri and Alvita Jimmy for
359 assistance with data collection.

360

361 **References**

- 362 1. Lawn J, Blencowe H, Oza S, You D, Lee AC, Waiswa P, et al. Every Newborn: Progress,
363 Priorities, and Potential Beyond Survival. *Lancet*. 2014;384:189-205.
- 364 2. UNICEF WHO. Survive and Thrive: Transforming Care for Every Small and Sick Newborn:
365 Geneva: World Health Organization; 2019 [Access Date; cited 2022 6/11/2022]. Available from:
366 <https://www.who.int/publications/i/item/9789241515887>.
- 367 3. World Health Organization. Newborns with Low Birth Weight (%) 2006 [Access Date; cited
368 2021 15/5/2021]. Available from:
369 <https://www.who.int/whosis/whostat/2006NewbornsLowBirthWeight.pdf>.
- 370 4. World Health Organization. Preterm Birth 2018 [Access Date; cited 2021 15/5/2021].
371 Available from: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>.
- 372 5. Blanc AK, Wardlaw T. Monitoring Low Birth Weight: An Evaluation of International Estimates
373 and an Updated Estimation Procedure. *Bulletin of the World Health Organization*. 2005;83(3):178-
374 85. PubMed Central PMCID: 15798841.
- 375 6. Lee A, Panchal P, Folger LV, Whelan H, Whelan R, Rosner B, et al. Diagnostic Accuracy of
376 Neonatal Assessment for Gestational Age Determination: A Systematic Review. *Pediatrics*. 2017;140.
- 377 7. Ashish KC, Nelin V, Vitrakoti R, Aryal S, Målqvist M. Validation of the Foot Length Measure as
378 an Alternative Tool to Identify Low Birth Weight and Preterm Babies in a Low-Resource Setting Like
379 Nepal: A Cross-Sectional Study. *BMC Pediatr*. 2015;15:43. Epub 2015/05/01. doi:
380 <https://doi.org/10.1186/s12887-015-0361-4>. PubMed PMID: 25928804; PubMed Central PMCID:
381 PMC4403672.

- 382 8. Feresu SA, Harlow SD, Gillespie BW, Welch K, Johnson TR. Birthweight-Adjusted Dubowitz
383 Methods: Reducing Misclassification of Assessments of Gestational Age in a Zimbabwean
384 Population. *The Central African journal of medicine*. 2003;49(5-6):47-53. PubMed PMID: 15214282.
- 385 9. Hadush MY, Berhe A, Medhanyie A. Foot Length, Chest and Head Circumference
386 Measurements in Detection of Low Birth Weight Neonates in Mekelle, Ethiopia: A Hospital Based
387 Cross Sectional Study. *BMC Pediatrics*. 2017;17.
- 388 10. Marchant T, Jaribu J, Penfold S, Tanner M, Schellenberg JA. Measuring Newborn Foot Length
389 to Identify Small Babies in Need of Extra Care: A Cross Sectional Hospital Based Study with
390 Community Follow-up in Tanzania. *BMC Public Health*. 2010;10(1):624. doi: 10.1186/1471-2458-10-
391 624.
- 392 11. Marchant T, Penfold S, Mkumbo E, Shamba D, Jaribu J, Manzi F, et al. The Reliability of a
393 Newborn Foot Length Measurement Tool Used by Community Volunteers to Identify Low Birth
394 Weight or Premature Babies Born at Home in Southern Tanzania. *BMC Public Health*.
395 2014;14(1):859. doi: 10.1186/1471-2458-14-859.
- 396 12. Mukherjee S, Roy P, Mitra S, Samanta M, Chatterjee S. Measuring New Born Foot Length to
397 Identify Small Babies in Need of Extra Care: A Cross-Sectional Hospital Based Study. *Iranian journal of*
398 *pediatrics*. 2013;23(5):508-12. PubMed PMID: 24800008.
- 399 13. Nabiwemba E, Marchant T, Namazzi G, Kadobera D, Waiswa P. Identifying High-Risk Babies
400 Born in the Community Using Foot Length Measurement at Birth in Uganda. *Child Care Health Dev*.
401 2013;39(1):20-6. doi: 10.1111/j.1365-2214.2012.01383.x. PubMed PMID: 22680250.
- 402 14. Nabiwemba E, Orach C, Kolsteren P. Determining an Anthropometric Surrogate Measure for
403 Identifying Low Birth Weight Babies in Uganda: A Hospital-Based Cross Sectional Study. *BMC Pediatr*.
404 2013;13:54. Epub 2013/04/17. doi: 10.1186/1471-2431-13-54. PubMed PMID: 23587297; PubMed
405 Central PMCID: PMC3639798.
- 406 15. Otupiri E, Wobil P, Nguah SB, Hindin MJ. Anthropometric Measurements: Options for
407 Identifying Low Birth Weight Newborns in Kumasi, Ghana. *PloS one*. 2014;9(9):e106712-e. doi:
408 10.1371/journal.pone.0106712. PubMed PMID: 25226505.
- 409 16. Pratinidhi AK, Bagade AC, Kakade SV, Kale HP, Kshirsagar VY, Babar R, et al. Action-Oriented
410 Colour-Coded Foot Length Calliper for Primary Healthcare Workers as a Proxy for Birth Weight &

- 411 Gestational Period. *The Indian journal of medical research*. 2017;145(3):347-52. doi:
412 10.4103/ijmr.IJMR_36_14. PubMed PMID: 28749397.
- 413 17. Singhal R, Jain S, Chawla D, Guglani V. Accuracy of New Ballard Score in Small-for-Gestational
414 Age Neonates. *J Trop Pediatr*. 2017;63(6):489-94. Epub 2017/10/05. doi: 10.1093/tropej/fmx055.
415 PubMed PMID: 28977559.
- 416 18. Thi HN, Khanh DK, Thu Hle T, Thomas EG, Lee KJ, Russell FM. Foot Length, Chest
417 Circumference, and Mid Upper Arm Circumference Are Good Predictors of Low Birth Weight and
418 Prematurity in Ethnic Minority Newborns in Vietnam: A Hospital-Based Observational Study. *PLoS*
419 *One*. 2015;10(11):e0142420. doi: 10.1371/journal.pone.0142420. PubMed PMID: 26555356;
420 PubMed Central PMCID: PMC4640656.
- 421 19. Folger LV, Panchal P, Eglovitch M, Whelan R, Lee AC. Diagnostic Accuracy of Neonatal Foot
422 Length to Identify Preterm and Low Birthweight Infants: A Systematic Review and Meta-Analysis.
423 *BMJ Global Health*. 2020;5(11):e002976. doi: 10.1136/bmjgh-2020-002976.
- 424 20. Hirve SS, Ganatra BR. Foot Tape Measure for Identification of Low Birth Weight Newborns.
425 *Indian Pediatr*. 1993;30(1):25-9. Epub 1993/01/01. PubMed PMID: 8406702.
- 426 21. Papua New Guinea National Department of Health. Sector Performance Annual Review:
427 Assessment of Sector Performance 2016-2020. National Report, August 2021. Port Moresby 2021.
- 428 22. The World Bank Data. Mortality Rate, Neonatal (Per 1,000 Live Births) - Papua New Guinea
429 2019 [Access Date; cited 2021 15/5/2021]. Available from:
430 <https://data.worldbank.org/indicator/SH.DYN.NMRT?locations=PG>.
- 431 23. Emori R, Vallely AJ, Gouda H, Siba P, Phuanukoonnon S, Mola G, et al. Supervised and
432 Unsupervised Birth and Early Newborn Care Practices in Papua New Guinea: Findings of a Maternal
433 and Infant Health Survey Conducted in Three Provinces. *Papua New Guinea Medical Journal*.
434 2017;60(1/2):27-40. doi: 10.3316/informit.301803940285652.
- 435 24. Papua New Guinea National Statistical Office. The 2011 National Population and Housing
436 Census, Ward Population: Momase Region: National Statistical office; 2014 [Access Date; cited
437 2022]. Available from: [https://sdd.spc.int/collection/2011-population-and-housing-census-](https://sdd.spc.int/collection/2011-population-and-housing-census-collection-papua-new-guinea)
438 [collection-papua-new-guinea](https://sdd.spc.int/collection/2011-population-and-housing-census-collection-papua-new-guinea).

- 439 25. Valley AJ, Pomat WS, Homer C, Guy R, Luchters S, Mola GDL, et al. Point-of-Care Testing and
440 Treatment of Sexually Transmitted Infections to Improve Birth Outcomes in High-Burden, Low-
441 Income Settings: Study Protocol for a Cluster Randomized Crossover Trial (the Wantaim Trial, Papua
442 New Guinea). *Wellcome open research*. 2019;4:53-. doi: 10.12688/wellcomeopenres.15173.2.
443 PubMed PMID: 32030356.
- 444 26. Committee Opinion No 700: Methods for Estimating the Due Date. *Obstet Gynecol*.
445 2017;129(5):e150-e4. Epub 2017/04/21. doi: 10.1097/aog.0000000000002046. PubMed PMID:
446 28426621.
- 447 27. Cutland CL, Lackritz EM, Mallett-Moore T, Bardaji A, Chandrasekaran R, Lahariya C, et al. Low
448 Birth Weight: Case Definition & Guidelines for Data Collection, Analysis, and Presentation of
449 Maternal Immunization Safety Data. *Vaccine*. 2017;35(48 Pt A):6492-500. Epub 2017/11/19. doi:
450 10.1016/j.vaccine.2017.01.049. PubMed PMID: 29150054; PubMed Central PMCID: PMC5710991.
- 451 28. Bland JM, Altman DG. Statistical Methods for Assessing Agreement between Two Methods
452 of Clinical Measurement. *Lancet*. 1986;1(8476):307-10. Epub 1986/02/08. PubMed PMID: 2868172.
- 453 29. Lee A, Mullany L, Ladhani K, Uddin J, Mitra D, Ahmed P, et al. Validity of Newborn Clinical
454 Assessment to Determine Gestational Age in Bangladesh. *Pediatrics*. 2016;138.
- 455 30. Paulsen C, Nielsen BB, Msemo O, Møller SL, Ekmann JR, Theander T, et al. Anthropometric
456 Measurements Can Identify Small for Gestational Age Newborns: A Cohort Study in Rural Tanzania.
457 *BMC Pediatrics*. 2019;19(1):120.

458

459 **Supporting information**

460 **S1 Checklist. Standards of Reporting of Diagnostic accuracy studies 2015 checklist.**

461 (<https://www.equator-network.org/reporting-guidelines/stard/>)

462 **S1 Fig. Bar chart of birth weight, measured on electronic scales (N=342).**

463 **S2 Fig. Bar chart of foot length measurements, average of two measurements with a plastic ruler.**

464 Panel A, researcher (N=342), B, volunteer (N=123).

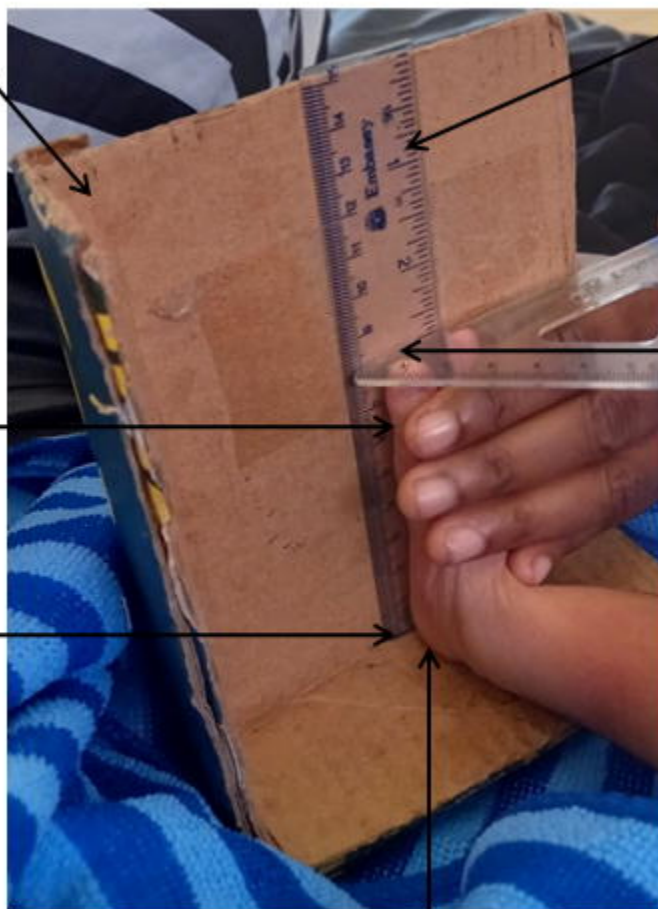
465 **S1 Table. Receiver operating curve statistics for accuracy of newborn foot length to classify low**
466 **birth weight, 1 mm increments (N=342).** Abbreviations: CI, confidence interval; LHR, likelihood ratio;
467 PV, predictive value.

468 **S2 Table. Receiver operating curve statistics for accuracy of newborn foot length to classify**
469 **preterm birth, 1 mm increments (N=342).** Abbreviations: CI, confidence interval; LHR, likelihood
470 ratio; PV, predictive value.

471 **S1 Protocol. Neofoot study protocol approved by the Papua New Guinea Institute of Medical**
472 **Research.**

Cardboard box

15 cm plastic ruler



Hallux (big toe)

Clear plastic set square at tip of hallux

Zero cm mark of ruler at base of box

Heel at base of box

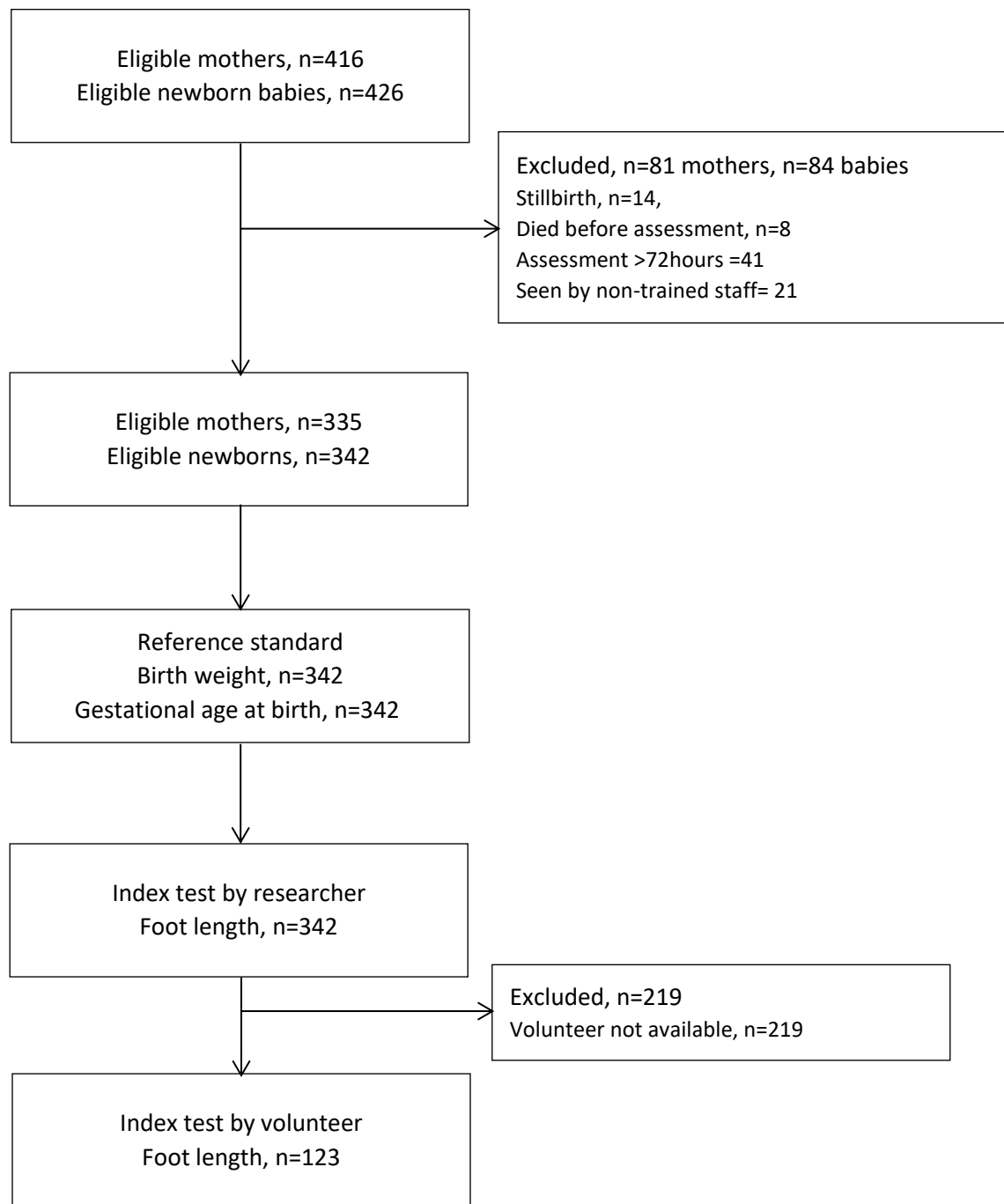


Fig. 2. Flow chart indicating number of mothers and newborn babies at each stage of the Neofoot study.

