Hongqi Niu ORCID iD: 0000-0002-9848-9574

Victor Oguoma ORCID iD: 0000-0001-9505-7197

Gabrielle McCallum ORCID iD: 0000-0002-0210-1468

Title: Latent Class Analysis to identify clinical profiles among Indigenous infants with bronchiolitis

Hongqi Niu, MS[1], Anne Bernadette Chang, PhD[1,2], Victor Oguoma, PhD[1], Zhiqiang Wang, PhD[1], Gabrielle Britt McCallum, PhD[1]

[1]Child Health Division, Menzies School of Health Research, Charles Darwin University, Darwin, Northern Territory, Australia; [2]Department of Respiratory and Sleep Medicine, Queensland Children's Hospital and Queensland University of Technology, Brisbane, Queensland, Australia.

Financial support

HN is supported by a Research Training Program (RTP) Fees Offset and Stipend Scholarship from the Australian Commonwealth Government; GBM is supported by a NHMRC early career fellowship (grant 1111705) and ABC is funded by a NHMRC practitioner fellowship (grant 1058213) and Queensland Children's Hospital Foundation (50286). No other funding was provided for this analysis.

Corresponding author's address*

Hongqi Niu

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ppul.25044.

Child Health Division,

Menzies School of Health Research

Rocklands Drive, Casuarina NT 0810 Australia

Ph: +61 8 86468565; Email: hongqi.niu@cdu.edu.au

Key words: bronchiectasis, bronchiolitis, hospitalisation, Indigenous, latent class analysis

Short title: LCA in Indigenous infants with bronchiolitis

ABSTRACT

Objective: Better phenotyping of the heterogenous bronchiolitis syndrome may lead to targeted future interventions. This study aims to identify severe bronchiolitis profiles among hospitalised Australian Indigenous infants, a population at risk of bronchiectasis, using Latent Class Analysis (LCA).

Methods: We included prospectively collected clinical, viral and nasopharyngeal bacteria data from 164 Indigenous infants hospitalised with bronchiolitis from our previous studies. We undertook multiple correspondence analysis (MCA) followed by LCA. The best-fitting model for LCA was based on adjusted Bayesian information criteria and entropy R^2.

Results: We identified five clinical profiles. Profile-A's (23.8% of cohort) phenotype was previous preterm (90.7%), low birth-weight (89.2%) and weight-for-length z-score <-1 (82.7% from combining those with z-score between -1 and -2 and those in the z-score of <-2 group) previous respiratory hospitalisation (39.6%) and

bronchiectasis on chest high-resolution computed tomography scan (35.4%). Profile-B (25.3%) was characterised by oxygen requirement (100%) and marked accessory muscle use (45.5%). Infants in profile-C (7.0%) had the most severe disease, with oxygen requirement and bronchiectasis in 100%, moderate accessory muscle use (85% vs 0-51.4%) and bacteria detected (93.1% vs 56.7-72.0%). Profile-D (11.6%) was dominated by rhinovirus (49.4%), mild accessory muscle use (73.8%) and weight-for-length z-score <-2 (36.0%). Profile-E (32.2%) included bronchiectasis (13.8%), RSV (44.0%), rhinovirus (26.3%) and any bacteria (72%).

Conclusion: Using LCA in Indigenous infants with severe bronchiolitis, we identified 5 clinical profiles with one distinct profile for bronchiectasis. LCA can characterise distinct phenotypes for severe bronchiolitis and infants at risk for future bronchiectasis, which may inform future targeted interventions.

INTRODUCTION

Bronchiolitis, a common acute lower respiratory infections (ALRIs) among infants worldwide[1] causes considerable morbidity and hospitalisations annually (e.g. >3 million hospitalisations)[2]. It is a heterogenous, multi-dimensional disorder from many perspectives, including clinical phenotypes (during acute disease[3] and future outcomes[4-7]), likely pathophysiology[8] and risk factors[9-11].

The importance of improving phenotyping of respiratory disorders (e.g. asthma[12] and bronchiectasis[13]) is increasingly appreciated as such data will likely improve targeted interventions. Indeed, to characterise clinical phenotypes (both during the acute disease and future outcomes) of infants with bronchiolitis using a multi-dimensional approach, Dumas and colleagues undertook two important studies[4]. Using latent class

analysis (LCA), Dumas and colleagues described 4 phenotypes in a large USA cohort, of which 2 were replicated (with the remaining 2 combined in a single phenotype) in a Finnish cohort[3]. Later, using LCA and they then described a profile (breathing problems/eczema in infancy and non-respiratory syncytial virus [RSV], mostly human rhinoviruses [hRV] infection) associated with increased risk of future asthma phenotype (doctor-diagnosed asthma Hazard Ratio=2.79; 95% confidence interval [CI] 1.78-4.39 by 3-years of age)[4]. These infants also had "higher eosinophil counts, higher cathelicidin levels, and increased proportions of Haemophilus-dominant or Moraxella-dominant microbiota profiles"[4].

These longer-term outcomes of infants hospitalised with bronchiolitis are of considerable interest with associations with asthma[14] and data linking viral agents (particularly RSV and hRV)[15] and/or nasopharyngeal bacteria[5] to future childhood asthma and allergy[16]. However, while asthma and allergy are relevant and important outcomes in mainstream settings, these are inconsistently reported among and across different populations and settings[2,17]. In settings where acute respiratory infections are prevalent and bronchiectasis unrelated to non-cystic fibrosis is relatively common, such as among Indigenous children living in high-income countries[10], bronchiolitis is more common and severe than their non-Indigenous counterparts[11]. In these settings, the long-term outcomes are also likely different. Indeed, Singleton and colleagues found that Alaskan Native children aged <2-years hospitalized for RSV infection had increased risk for chronic productive cough at 5 to 8 years of age and recurrent lower respiratory infections, but not asthma[18]. Further, we found that within 13-months post-hospitalisation for bronchiolitis, Aboriginal and/or Torres Strait Islander (from here referred as Indigenous) infants 19% (30/157) had bronchiectasis on chest high-resolution computed tomography (HRCT) scan[7].

A better understanding of clinical phenotypes, including outcomes in populations with a high incidence of bronchiectasis e.g. Indigenous settings are important, in the context of (a) Dumas and colleagues'[3,4] likely paradigm-changing findings for future intervention studies; (b) the absence of further studies using LCA to define these bronchiolitis clinical phenotypes; and (c) the differential long-term outcomes in populations at high-risk of bronchiectasis unrelated to non-cystic fibrosis. We used LCA to analyse data from Indigenous infants from Australia who participated in our previous prospective hospitalised bronchiolitis studies[19-21]. We aimed to identify distinct clinical profiles particularly those at-risk of future bronchiectasis.

METHODS

Study population

We used de-identified data sourced from three prospective studies undertaken between June 2008 and September 2013[19-21]. For this study, we included only the 232 Indigenous infants living in the Northern Territory (Australia). Here we briefly describe these studies; two were randomised controlled trials (RCT) where we found neither single or 3-weekly azithromycin doses, compared to placebo, improved any clinical outcomes for infants hospitalised with bronchiolitis[19,20]. The third was a cohort study to determine the validity and reliability of a bronchiolitis scoring system of infants hospitalised with bronchiolitis[21]. An overview of the studies is further described in the Figure. The local Human Research Ethics Committee approved each original study (HREC 07/60, HREC-2010-1324) and written informed consent obtained from the primary carer. In all studies[19-21], infants were aged ≤24-months hospitalised with a diagnosis of bronchiolitis using a standardised hospital protocol, with the same exclusion criteria (see E-text). As this current study was a reanalysis of

already collected data, additional ethics was not required.

Clinical Data

All clinical data and routine medical investigations during hospitalisation were recorded on standardised data collection forms[3,23,24]. Data for bronchiectasis was based on chest HRCT diagnosis (at median of 13 months (IQR 7-18) after their index hospitalisation for bronchiolitis) as described previously[7]. In this analysis, the most commonly detected co-morbidities were reported and combined as a single variable where appropriate (e.g. any otitis media or any skin infection). Length of stay (LOS) was defined in the original studies as duration from time from admission to "ready for discharge" (oxygen saturation >94% in air for >16-hours) and feeding adequately[19,20]. Accessory muscle use was taken from our modified Tal score[21] and weight-for-length z-score defined using WHO criteria (detailed in E-text).

Virus and bateria

Virus and bacteria data were obtained from nasopharyngeal swabs (NPS) collected at enrolment. NPS were placed into skim milk tryptone glucose glycerol broth (STGGB), stored at -80°C. NPS bacterial pathogens were cultured at our institution[22] and respiratory viruses using PCR undertaken at the Queensland Paediatric Infectious Diseases laboratory in Brisbane, as previously done[19-21]. For this study, we combined any bacteria detected as a single variable and only the two most frequently detected viruses (RSV and HRV) were analysed (so as to allow inclusion of more data). Further methods are detailed in the E-text file.

Statistical analysis

Sample size calculations were not undertaken for this study. Descriptive analyses of

patient characteristics are presented as median and interquartile range (IQR: 25th-75th percentile) for continuous variables or frequencies and percentages for categorical variables. Weight-for-length z-score categories were defined using Zanthro software in STATA v14.0. Data were analysed using R package version 3.4.1.

Multiple Correspondence Analysis (MCA)

In the original studies, a total of 104 variables were available for MCA analysis. Firstly, we removed variables if missing data was >15%, then further removed variables if there were less than 20 infants for each categorical result (e.g. present or absent)[23]. We then selected variables based on clinical importance for bronchiolitis severity which included; age (<12-months), preterm (<37-weeks), low birth weight (<2.5kg), weight-for-length z-score, exposure to household smoke or in utero, currently breastfed, lobar collapse/consolidation on chest x-ray, previous respiratory hospitalisation, sex, remote (defined as >100km from hospital with paediatric expertise)[7], any antibiotics during hospital, supplemental oxygen requirement, any co-morbidity, carer-reported cough or breathing difficulty in last 7-days, any NPS bacteria, bronchiectasis on HRCT in addition to including as many variables (RSV, HRV, accessory muscle use and LOS) as Dumas and colleagues included in their study[3]. These variables were then entered into the MCA[23] to identify the most relevant variables for LCA and to decompose the inertia by identifying a small number of mutually independent dimensions that represent the most important deviations from independence[24].

LCA analysis

To identify the best-fitting model for LCA (i.e. between 1-5 models), we chose the

minimum adjusted Bayesian Information Criteria (aBIC)[25,26] as this is known to be more robust with smaller datasets (<400 participants) than Bayesian Information Criterion (BIC)[26,27] and Akaike's Information Criterion (AIC)[26,28]. Entropy R^2 was used as the second criterion to determine the separation between the models using the highest value between models[29].

RESULTS

Participant characteristics

Clinical data of the final 164 infants are summarised in Table-1. During their hospital stay, 57.9% of infants received oxygen supplementation, RSV was detected in 42.1%, HRV in 28.0% and 68.9% had any bacteria detected. LOS varied but most infants (34.8%) were discharged within 48-hours. Bronchiectasis was detected on HRCT in 34 (20.7%) of infants at a later timepoint for clinical reasons.

MCA

Of 22 variables included in MCA (Table-2), the top three dimensions (of a possible 39) were chosen as they represented the largest number of independent variations between each dimension (Table-2). The top 4 variables (total n=12) from each dimension (e.g. LOS, low birth weight, preterm, weight-for-length z-score, supplemental oxygen requirement, accessory muscle use, RSV, HRV, any bacteria, any co-morbidity, bronchiectasis and previous respiratory hospitalisation) were then selected for LCA, including as many variables included in Dumas and colleagues' study[3].

LCA

Of the 232 infants, 68 were excluded due to missing data leaving a total of 164 infants for LCA (Figure). A Class-5 model was identified using LCA (Table-3). Class-5 model had the highest Entropy value and Class-3 model had the lowest aBIC value, however Class-5 model was chosen as the best-fitting model for LCA based on Entropy value and the clinic importance of the models. The five profiles included within the Class-5 model are described below and in Table-4.

Profile-A (23.8% of cohort) had the highest frequency of infants born preterm (90.7% vs 0.0-12.6%); low birth weight (89.2% vs 0.0-15.0% for other profiles), poor growth weight-for-length z-score of <-1 (82.7% from combining those with z-score between -1 and -2 and those in the z-score of <-2 group vs. 0.0-36.0% for other profiles), and previous respiratory hospitalisations (39.6% vs. 11.1-19.4%). Infants in this profile had the second highest frequency of bronchiectasis (35.4%).

Profile-B (25.3% of cohort) was characterised by highest detection of RSV (48.6%), marked accessory muscle use (45.5% vs. 13.4-26.2%), supplemental oxygen requirement (100% along with Profile C) and LOS [60-96-hrs (48.1 vs. 0.0-30.5%) and >96-hours (34.4% vs. 0.0-31.2%)], but none had bronchiectasis.

Profile-C (7% of cohort) was the smallest group characterised by bronchiectasis in 100% (compared to 0.0-35.4% for other profiles), the highest frequency of moderate accessory muscle use (84.7% vs 0.0-51.4%), presence of ‘any bacteria’ (93.1% vs 56.7-72.0%), any co-morbidity (75.4% vs. 33.9-68.3%) and supplemental oxygen requirement (100.0%).

Profile-D (11.6% of cohort) had the highest prevalence of hRV (49.4% vs 23.2-32.5%) and mild accessory use (73.8% vs 0.0-54.2%). Weight-for-length z-score <-2

was higher in this group (36.0% vs. 8.2-23.9%) with the least having bronchiectasis (7.0%).

Profile-E (32.2% of cohort) was characterised by RSV detected in 44.0%, HRV in 26.3%, any bacteria in 72.0%. None had low birth weight (0.0%); most had LOS <48-hours (39.8%), low requirement for oxygen supplementation (20.3%) but 13.8% had bronchiectasis.

Clinical profile for future bronchiectasis on LCA

We identified one (100.0%) distinct clinical profile for Indigenous infants at-risk for future bronchiectasis (Table-4) and one without any risk (0.0%) (Profile-B). Infants in Profile-C who all had bronchiectasis also had the highest frequency of markers of severity for bronchiolitis i.e. supplemental oxygen requirement (100.0%), moderate accessory muscle use (84.7%), any co-morbidity (75.4%) and any bacteria (93.1%). Profile-A was the second most severe group with bronchiectasis in 35.4%, preterm birth in 90.7%, low birth weight in 89.2% and previous respiratory hospitalisation in 39.6%. However, many infants in this profile lacked known markers of severity (e.g. prolonged LOS, supplemental oxygen requirement etc).

DISCUSSION

In this first study using LCA in a cohort at-risk of bronchiectasis rather than asthma, we identified 5 distinct clinical profiles among the 164 Indigenous infants hospitalised with bronchiolitis. The most striking characteristics were identified in Profile-A phenotype with high preterm birth (90.7%), low birth-weight (89.2%), whilst Profile-B phenotype were characterised by all (100.0%) requiring oxygen supplementation, yet none had future chronic symptoms necessitating chest CT scan (0.0%). In

contrast, all in Profile-C had bronchiectasis (100.0%) and NP bacteria (93.0%) at the point of hospitalisation. Profile-D was characterised by the absence of low birth-weight (100.0%) yet low weight for length z-score of <-2 (36.0%) and hRV infection (49.4%) and infants in Profile-E had low concurrent co-morbidity (33.9%).

Use of clinical phenotypes for various diseases have been appreciated increasingly over the last decade in a range of conditions. LCA has been generally used for this and to the best of our knowledge, there are only two such published studies[3,4] involving infants with bronchiolitis. Both these novel studies[3,4] were based in high-income settings where asthma is the outcome of interest. In the absence of any such studies in a setting where infants are at high risk of chronic suppurative lung disease[30], we undertook this study to determine of LCA can identify phenotypes that may inform future interventions relevant to children at high-risk of bronchiectasis.

Uniquely, we identified two main clinical profiles with future bronchiectasis i.e. all in Profile-C had bronchiectasis and the second most common group with bronchiectasis was Profile-A (35.0%). Over a third of infants were included in these profiles which included many known risk factors for severe disease. It was not surprising that these profiles included the greatest number of infants with bacterial carriage and current co-morbidities. In Australia, Indigenous children particularly from remote communities often live in overcrowded homes[30], are exposed to early and dense acquisition of respiratory bacteria (*Streptococcus pneumoniae, Haemophilus influenzae and Moraxella catarrhalis*)[31] which are associated with poorer clinical outcomes (chronic wet cough[18] and bronchiectasis[30]) in our setting. Whilst the above finding need confirmation in a large cohort from multi-centres, our novel findings highlight that LCA could be used for phenotyping children at-risk of future bronchiectasis and

subsequently inform targeted interventions for these infants that could possibly prevent future bronchiectasis.

Unsurprising, our phenotypes are very different from that of Dumas and colleagues'[3] for many reasons, of which the most important is the different target population. Further our study was substantially smaller, we excluded infants admitted to the intensive care unit, and did not include all the variables examined by Dumas and colleagues[3]. Also, known risk factors for bronchiectasis such as exposure to parental or household smoke were removed by MCA. Nevertheless, we identified 5 distinct clinical profiles by use of LCA. Profile-B was the third largest group with markers of severity (e.g. marked accessory muscle use, oxygen requirement and prolonged LOS). This group is similar to that of the USA cohort,[3] whereby Dumas and colleagues[3] reported a severe profile among infants with RSV, moderate to severe accessory muscle use and prolonged LOS (>3-days). The same result however was not replicated in the Finish study[3] where these markers of severity (e.g. retractions and LOS) were distributed across two profiles. However, our Profile-B included oxygen requirement, a factor that was not included in Dumas and colleagues'[3] LCA profile.

The limitations of the original studies whereby this cohort was obtained from, were previously published[19-21]. However, this study has further other important limitations. Firstly, we did not use BIC because our relatively small sample size restricted the use of BIC as cohorts of <400 that BIC underestimate classes[32]. We thus used aBIC, suitable for smaller cohort[26,32]. As aBIC still suggested that the sample is sufficiently large ($n \geq 200$)[32], we thus included entropy as a second information criterion[33] in class-models. Secondly, our data was cross-sectional and weakened by the lack of longer follow-up. Thirdly, our data were limited to Indigenous hospitalised for bronchiolitis

with the lack of community-based children and the absence of recurrent wheezing and asthma, important factors in studies from high-income settings[34].

In conclusion, our study is unique as, for the first time, we identified 5 clinical phenotypes in an at-risk population by using LCA. Two profiles were important for future bronchiectasis. Importantly, our data further highlights the heterogeneity of bronchiolitis phenotypes among infants. We now need confirmation of our novel findings in a large multi-centre cohort to determine such phenotypes in children at-risk of future bronchiectasis. Such work could subsequently inform targeted interventions for these infants that could possibly prevent future bronchiectasis, an increasingly recognised condition worldwide of which a significant proportion commences in childhood[13].

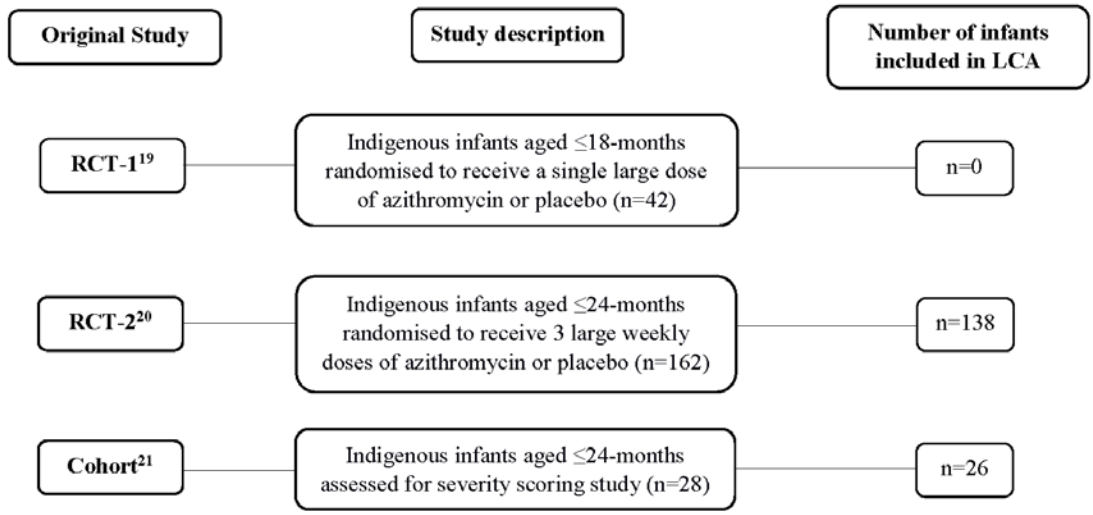
Acknowledgements

We are grateful for the children and families who participated in the original studies.

FIGURE LEGEND

Figure. Description of original studies where Indigenous infants were recruited from the Royal Darwin Hospital

Figure: Description of original studies where Indigenous infants were recruited from the Royal Darwin Hospital



REFERENCES

1. Nair H, Simoes EA, Rudan I, Gessner BD, Azziz-Baumgartner E, Zhang JSF, Feikin DR, Mackenzie GA, Moiisi JC, Roca A, et al. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. Lancet 2013; 381:1380-90.

2. Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, Johnson DW, Light MJ, Maraqa NF, Mendonca EA, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. Pediatrics 2014; 134:e1474-502.

3. Dumas O, Mansbach JM, Jartti T, Hasegawa K, Sullivan AF, Piedra PA, Camargo CA, Jr. A clustering approach to identify severe bronchiolitis profiles in children. Thorax 2016; 71:712-8.

4. Dumas O, Hasegawa K, Mansbach JM, Sullivan AF, Piedra PA, Camargo CA, Jr.

Severe bronchiolitis profiles and risk of recurrent wheeze by age 3 years. J Allergy Clin Immunol 2019; 143:1371-9 e7.

5. Bønnelykke K, Vissing NH, Sevelsted A, Johnston SL, Bisgaard H. Association between respiratory infections in early life and later asthma is independent of virus type. J Allergy Clin Immunol 2015; 136:81-6. e4.

6. Kusel MM, de Klerk NH, Kebadze T, Vohma V, Holt PG, Johnston SL, Sly PD. Early-life respiratory viral infections, atopic sensitization, and risk of subsequent development of persistent asthma. J Allergy Clin Immunol 2007; 119:1105-10.

7. McCallum GB, Chatfield MD, Morris PS, Chang AB. Risk factors for adverse outcomes of Indigenous infants hospitalized with bronchiolitis. Pediatr Pulmonol 2016; 51:613-23.

8. Cunningham S, Nair H, Campbell H. Deciphering clinical phenotypes in acute viral lower respiratory tract infection: Bronchiolitis is not an island. Thorax 2016; 71:679-80

9. Chawes BL, Poorisrisak P, Johnston SL, Bisgaard H. Neonatal bronchial hyperresponsiveness precedes acute severe viral bronchiolitis in infants. J Allergy Clin Immunol 2012; 130:354-61 e3.

10. Welliver RC. Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. J Pediatr 2003; 143:S112-7.

11. Bailey EJ, Maclennan C, Morris PS, Kruske SG, Brown N, Chang AB. Risks of severity and readmission of Indigenous and non-Indigenous children hospitalised for bronchiolitis. J Paediatr Child Health 2009; 45:593-7.

12. Pavord ID, Beasley R, Agusti A, Anderson GP, Bel E, Brusselle G, Cullinan P, Custovic A, Ducharme FM, Fahy JV, et al. After asthma: redefining airways diseases. Lancet 2018; 391:350-400.

13. Chang AB, Bush A, Grimwood K. Bronchiectasis in children: diagnosis and treatment. Lancet 2018; 392:866-79.

14. Balekian DS, Linnemann RW, Hasegawa K, Thadhani R, Camargo CA, Jr. Cohort Study of Severe Bronchiolitis during Infancy and Risk of Asthma by Age 5 Years. J Allergy Clin Immunol Pract 2017;5: 92-6.

15. Jackson DJ, Gangnon RE, Evans MD, Roberg KA, Anderson EL, Pappas TE, Printz MC, Lee WM, Shult PA, Reisdorf E, et al. Wheezing rhinovirus illnesses in early life predict asthma development in high-risk children. Am J Respir Crit Care Med 2008; 178:667-72.

16. Stein RT, Sherrill D, Morgan WJ, Holberg CJ, Halonen M, Taussig LM, Wright AL, Martinez FD. Respiratory syncytial virus in early life and risk of wheeze and allergy by age 13 years. Lancet 1999;354: 541-5.

17. Wright PF, Gruber WC, Peters M, Reed G, Zhu Y, Robinson F, Coleman-Dockery S, Graham BS. Illness severity, viral shedding, and antibody responses in infants hospitalized with bronchiolitis caused by respiratory syncytial virus. J Infect Dis 2002; 185:1011-8.

18. Singleton RJ, Redding GJ, Lewis TC, Martinez P, Bulkow L, Morray B, Peters H, Gove J, Jones C, Stamey D, et al. Sequelae of severe respiratory syncytial virus infection in infancy and early childhood among Alaska native children. Pediatrics 2003; 112:285-90.

19. McCallum GB, Morris PS, Chatfield MD, Maclennan C, White AV, Sloots TP, Mackay IM, Chang AB. A single dose of azithromycin does not improve clinical outcomes of children hospitalised with bronchiolitis: a randomised, placebo-controlled trial. PloS One 2013; 8:e74316.

20. McCallum GB, Morris PS, Grimwood K, Maclennan C, White AV, Chatfield MD, Sloots TP, Mackay IM, Smith-Vaughan H, McKay CC, et al. Three-weekly doses of azithromycin for indigenous infants hospitalized with bronchiolitis: a multicentre, randomized, placebo-controlled trial. Front Pediatr 2015; 3:32.

21. McCallum GB, Morris PS, Wilson CC, Versteegh LA, Ward LM, Chatfield MD, Chang AB. Severity scoring systems: are they internally valid, reliable and predictive of oxygen use in children with acute bronchiolitis? Pediatr Pulmonol 2013; 48:797-803.

22. Hare KM, Grimwood K, Leach AJ, Smith-Vaughan H, Torzillo PJ, Morris PS, Chang AB. Respiratory bacterial pathogens in the nasopharynx and lower airways of Australian indigenous children with bronchiectasis. J Pediatr 2010; 157:1001-5.

23. Di Franco G. Multiple correspondence analysis: one only or several techniques? Qual Quant 2016; 50:1299-315.

24. Sourial N, Wolfson C, Zhu B, Quail J, Fletcher J, Karunananthan S, Bandeen-Roche K, Béland F, Bergman H. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. J Clin Epidemiol 2010; 63:638-46.

25. Sclove SL. Application of model-selection criteria to some problems in

multivariate analysis. Psychometrika 1987; 52:333-43.

26. Dziak JJ, Coffman DL, Lanza ST, Li R. Sensitivity and specificity of information criteria. Brief Bioinform 2020; 21:553-565.

27. Schwarz G. Estimating the dimension of a model. Ann Stat 1978; 6: 461-4.

28. Akaike H. Information theory and an extension of the maximum likelihood principle. Selected papers of hirotugu akaike: Springer; 1998. p199-213.

29. Boeschoten L, Oberski DL, De Waal T, Vermunt JK. Updating Latent Class Imputations with External Auxiliary Variables. Struct Equ Modeling 2018; 25:750-61.

30. Singleton RJ, Valery PC, Morris P, Byrnes CA, Grimwood K, Redding G, Torzillo PJ, McCallum G, Chikoyak L, Mobberly C, et al. Indigenous children from three countries with non-cystic fibrosis chronic suppurative lung disease/bronchiectasis. Pediatr Pulmonol 2014; 49:189-200.

31. Smith-Vaughan H, Byun R, Nadkarni M, Jacques NA, Hunter N, Halpin S, Morris PS, Leach AJ. Measuring nasal bacterial load and its association with otitis media. BMC Ear Nose Throat Disord 2006; 6:10.

32. Usami S. Performance of information criteria for model selection in a latent growth curve mixture model. JJSCS 2014; 27:17-48.

33. Wraith D, Wolfe R. Classifying patients by their characteristics and clinical presentations; the use of latent class analysis. Respirology 2014; 19:1138-48.

34. Mansbach JM, Hasegawa K, Henke DM, Ajami NJ, Petrosino JF, Shaw CA,

Piedra PA, Sullivan AF, Espinola JA, Camargo CA, Jr. Respiratory syncytial virus and rhinovirus severe bronchiolitis are associated with distinct nasopharyngeal microbiota. J Allergy Clin Immunol 2016; 137:1909-13 e4.

Table-1. Demographic, medical and clinical characteristics of the 164 infants included in LCA

		n=164 (%)
Demographics		
Age (months), median (IQR)		5 (3-10)
Boys		103 (62.8)
Preterm (<37-weeks)		45 (27.4)
Low birth weight (<2.5kg)		42 (25.6)
Weight-for-length z-score[a]		
	>2	5 (3.0)
	+1 to +2	12 (7.3)
	-1 to +1	82 (50.0)

	-2 to -1	28 (17.1)
	<-2	28 (17.1)
Remote		137 (83.5)
Currently breastfed		142 (86.6)
Mother smoked during pregnancy		
	Yes	90 (54.8)
	No	71 (43.3)
Exposed to household smoke		
	Yes	104 (63.4)
	No	58 (35.4)
Previous respiratory hospitalisation		35 (21.3)
Enrolment observations		
Number required supplemental oxygen		95 (57.9)

Antibiotics prescribed prior to hospital		140 (85.4)
Any co-morbidity[b]		78 (47.6)
Lobar collapse/consolidation on chest x-ray		32 (20)
Carer reported cough (last 7-days)		161 (98.2)
Carer reported breathing difficulty (last 7-days)		163 (99.4)
Accessory muscle use[c 21]		
	+	55 (33.5)
	++	61 (37.2)
	+++	39 (23.8)
	None	9 (5.5)
Length of stay (hours)		
	<48	57 (34.8)
	48-60	32 (19.5)

	60-96	47 (28.7)
	>96	28 (17.1)
Viruses/Bacteria		
	Respiratory Syncytial Virus	69 (42.1)
	Human Rhinovirus	46 (28.0)
	Any bacteria[d]	113 (68.9)
Bronchiectasis on high-resolution computed tomography scan		
Bronchiectasis on HRCT		34 (20.7)
No bronchiectasis on HRCT		1 (0.6)

Abbreviations: HRCT: high-resolution computed tomography; IQR: interquartile range.

[a]Weight-for-length z-score categories are presented as standard deviation; [b]Any co-morbidity included presence of otitis media or any skin infection; [c]Accessory muscle use: None (no chest in-drawing); + (presence of mild intercostal in-drawing); ++ (moderate amount of intercostal in-drawing); +++ (moderate or marked intercostal in-

drawing with presence of head bobbing or tracheal tug); [d]Any bacteria detected on nasopharyngeal swab include: Streptococcus pneumoniae, Haemophilus influenza, Moraxella catarrhalis, Staphylococcus aureus.

Missing data: Weight-for-length z-score: n=9 (5.4%); Mother smoked in utero: n=3 (1.8%); exposure to household tobacco smoke: n=2 (1.2%).

Table-2. MCA result

	Dimension 1 (13.6%)	**Dimension 2 (11.5%)**	**Dimension 3 (10.1%)**
Accessory muscle use[a]	0.198	0.089	0.369
Sex	0.003	0.006	0.001
Preterm (<37-weeks)	0.075	0.467	0.256
Low birth weight (<2.5kg)	0.093	0.485	0.249
Weight-for-length z-score	0.243	0.418	0.338
Length of stay (hours)[b]	0.046	0.034	0.079

Age (<12-months)	0.057	0.022	0.001
Remote	0.003	0.114	0.06
Mother smoked during pregnancy	0.005	0.026	0.005
Exposed to household smoke	0.08	0.1	0.03
Currently breastfed	0.004	0.007	0.069
Lobar collapse/consolidation on chest x-ray	0.071	0.063	0.063
Previous respiratory hospitalisation	0.024	0.106	0.082
Supplemental oxygen required	0.178	0.022	0.204
Human Rhinovirus detected	0.707	0.031	0.092
Respiratory Syncytial Virus detected	0.707	0.152	0.085
Any co-morbidity[c]	0.003	0.143	0.044
Any antibiotics during hospital	0.024	0.028	0.001
Carer reported cough (last 7-days)	0.002	0.0	0.047

Carer reported breathing difficulty (last 7-days)	0.001	0.018	0.032
Any bacteria[d]	0.427	0.031	0.103
Bronchiectasis on HRCT	0.04	0.159	0.004

Abbreviations: HRCT: high-resolution computed tomography.

[a]Accessory muscle use: None (no chest in-drawing, i.e., absence of lower part of the chest moves in or retracts when inhalation occurs); + (presence of mild intercostal in-drawing [just visible], no head bobbing or tracheal tug); ++ (moderate amount of intercostal in-drawing, no head bobbing or tracheal tug); +++ (moderate or marked intercostal in-drawing with presence of head bobbing or tracheal tug); [b]LOS defined as duration from time from admission to "ready for discharge" (oxygen saturation >94% in air for >16-hours) and feeding adequately;[19,20] [c]Any co-morbidity includes, any otitis media or any skin infection; [d]Any bacteria includes, *Streptococcus pneumoniae*, *Haemophilus influenza*, *Moraxella catarrhalis*, *Staphylococcus aureus*.

Table-3. Information Criteria for Class 1-5 models for LCA

Model	aBIC	Entropy	BIC	cAIC	AWE	Log likelihood	Resid. df
1	3148.68	N/A	3215.17	3236.17	3385.26	-1554.03	143
2	3040.91	0.87	3177.05	3220.05	3525.34	-1478.88	121
3	3030.94	0.83	3236.73	3301.73	3763.22	-1452.62	99
4	3034.24	0.868	3309.67	3396.67	4014.36	-1432.99	77
5	3036.57	0.921	3381.66	3490.66	4264.55	-1412.89	55

Abbreviations: aBIC: adjusted Bayesian Information Criteria; Entropy: Entropy R2 information criteria; BIC: Bayesian Information Criteria; cAIC: consistent Akaike Information Criterion; AWE: Approximate Weight of Evidence; Resid. df: residual degrees of freedom.

Table-4. Clinical profiles within Class-5 LCA of 164 infants

	Profile					
	A (23.8% of cohort)	**B (25.3% of cohort)**	**C (7% of cohort)**	**D (11.6% of cohort)**	**E (32.2% of cohort)**	**Total n=164**
Respiratory Syncytial Virus detected	35.1	48.6	32.2	42.8	44.0	69
Human Rhinovirus detected	23.9	23.2	32.5	49.4	26.3	46
Any bacteria detected[a]	71.0	56.7	93.1	67.8	72.0	113
Weight-for-length z-score[b]						
>2	0.0	0.0	0.0	26.2	0.0	5
+1 to +2	0.0	9.4	8.6	31.7	2.0	12

-1 to +1	17.4	61.0	65.8	0.0	80.1	82
-2 to -1	58.8	6.8	0.0	0.0	4.2	28
<-2	23.9	13.2	17.0	36.0	8.2	28
Not Stated	0.0	9.6	8.6	6.1	5.5	9
Birth weight (kg)						
<2.5	89.2	15.0	8.2	0.0	0.0	42
≥2.5	10.8	85.0	91.8	100.0	100.0	122
Preterm (<37-weeks)	90.7	12.6	0.0	10.1	4.5	45
Any co-morbidity present[c]	59.8	36.2	75.4	68.3	33.9	78
Accessory muscle use[d]						
+	31.5	0.0	0.0	73.8	54.2	55
++	41.9	51.4	84.7	0.0	25.6	61

+++	16.1	45.5	15.3	26.2	13.4	39
None	10.5	3.1	0.0	0.0	6.9	9
Supplemental oxygen required	44.7	100.0	100.0	72.2	20.3	95
Previous respiratory hospitalisation	39.6	19.4	17.0	11.1	14.0	35
Length of stay (hours)						
<48	35.7	17.5	50.5	47.1	39.8	57
48-60	16.5	0.0	40.8	10.2	35.8	32
60-96	30.5	48.1	0.0	11.5	24.5	47
>96	17.3	34.4	8.7	31.2	0.0	28
Bronchiectasis on HRCT scan						
Did not have a HRCT	64.6	100.0	0.0	87.7	86.2	129

Bronchiectasis confirmed on HRCT	35.4	0.0	100.0	7.0	13.8	34
No bronchiectasis on HRCT	0.0	0.0	0.0	5.2	0.0	1

Abbreviations: HRCT: high-resolution computed tomography.

[a]Any bacteria include *Streptococcus pneumoniae*, *Haemophilus influenza*, *Moraxella catarrhalis*, *Staphylococcus aureus*; [b]Weight-for-length z-score categories are presented as standard deviation; [c]Any co-morbidity includes: any otitis media or any skin infection; [d]Accessory muscle use: None (no chest in-drawing); + (presence of mild intercostal in-drawing); ++ (moderate amount of intercostal in-drawing); +++ (moderate or marked intercostal in-drawing with presence of head bobbing or tracheal tug).