The role of neonatal pulmonary morbidity in the longitudinal patterns of hospitalisation for respiratory infection during the first year of life

Kim S. Betts[1], Ricardo J. Soares Magalhães[2,3] and Rosa Alati[4]

[1]Institute for Social Science Research, The University of Queensland, Office 403, Cycad Building, Long Pocket Precent, 4068, Brisbane, Queensland, Australia; [2]UQ Spatial Epidemiology Laboratory, School of Veterinary Science, The University of Queensland, Gatton, Australia; [3]Children's Health and Environment Program, Child Health Research Centre, The University of Queensland, Brisbane, Australia and [4]Institute for Social Science Research, University of Queensland, Brisbane, Australia

Abstract

Respiratory infections among infants constitute a major burden to health care systems in developed nations, yet the course and risk factors leading to these conditions are poorly understood. We examine the longitudinal patterns of respiratory infection hospitalisation (RIH) and how these patterns are influenced by neonatal pulmonary morbidities. We included all live births ($n = 429\ 058$) occurring in the Australian state of Queensland between January 2009 and December 2015. Data were structured so that each participant had a record (present/absent) of RIH for each month from birth to 12 months. Initially, latent class growth analysis was used to identify the trajectories of RIH adjusted for spatial–temporal factors; using the identified trajectories of RIH as outcomes, we built a multinomial logistic regression model to identify neonatal predictors of RIH trajectories. Our results indicated that a four-class solution was the best fit to the data, comprising a 'no-risk' trajectory, a 'low-risk' trajectory, an 'early-risk' trajectory and a 'chronic-risk' trajectory. Compared with the no-risk trajectory, membership in the other trajectories was predicted by a range of neonatal pulmonary morbidities, with transient tachypnoea of newborn showing a specific relationship with the early-risk group and sleep apnoea showing a specific and strong risk with the chronic-risk group. Our findings suggest the possibility of identifying neonates at risk of recurrent RIH and implementing effective intervention strategies prior to neonatal discharge.

Introduction

Although no longer a major source of mortality in developed nations, respiratory infections continue to make the largest contribution to the burden of disease of any other group of infections [1], with little progress made in reducing this burden in recent decades [2]. Among infants, infectious diseases remain the main cause of hospitalisation in English-speaking countries [3], for which respiratory infections (pneumonia, influenza and other upper and lower respiratory infections) account for the largest proportion [4]. In the UK, acute respiratory infections (particularly respiratory syncytial virus, rhinovirus and influenza virus) account for 22% of hospitalisations and 59% of general practitioner consultations among this age group [5]. Furthermore, infants are at greater risk of severe outcomes and mortality as a result of respiratory infection [6], with respiratory infections in early life suggested as precursors for a number of health conditions in later life [7, 8]. In addition to health outcomes, acute respiratory infections among infants represent a major cost to the healthcare systems [5].

The majority of acute respiratory infections among infants are caused by viral agents. While most viral acute respiratory infections are confined to the upper respiratory tract (leading to lethargy and poor feeding), a third develop lower respiratory tract infections, which can result in tachypnoea, severe cough, respiratory distress and even death [5]. Importantly, most respiratory viruses can cause infections of varying severity and can accompany and even enhance bacterial coinfection [5].

Epidemiological studies have found that respiratory infections among otherwise healthy infants in developed nations are linked to minority status, lower socio-economic background, air pollution, parental smoking and lack of breast feeding [6, 9–11]. With regards to perinatal/neonatal risk factors, conditions such as premature birth, congenital heart disease, T-cell immunodeficiency [12], mode of delivery and spatial–temporal factors including birth season and geographic location [5, 13] are known to influence the risk of requiring hospitalisation for respiratory infections among infants. However, much more could be learnt about how other factors present at birth influence later respiratory infection hospitalisation (RIH) risk.

Recurrence of RIH is a crucial consideration when informing strategies aimed at reducing the burden of respiratory infections among infants, but one which has been ignored or poorly

measured in most previous studies [14–16]. Even less is known about the factors that predict recurring respiratory infections in infants, with studies suggesting recurrence is not simply a function of specific viral agents or immunodeficiency [14–16]. Transient tachypnoea of newborn (TTN), respiratory distress, neonatal aspiration and sleep apnoea (SAPN), while being relatively common, non-specific and often benign or self-limiting, may have a potential role in respiratory infections [17–23].

Prior to neonatal discharge, hospitalised mothers and their newborns represent a captive population for which preventive strategies can be implemented should certain groups of neonates be at a known risk of recurrent respiratory infections. In this study, we used hospital inpatient records to first identify empirical trajectories of RIHs during the first 12 months of life in neonates who were free from infection or major morbidity at discharge. We then quantify the association between different RIH trajectories and neonatal pulmonary morbidities including TTN, respiratory distress, neonatal aspiration and SAPN.

Methods

Sample

The Born in Queensland Study (BQS) is a birth cohort study comprised of linked health administrative records of all live births in the Australian state of Queensland (QLD) between 2009 and 2015. Individuals were matched across three data collections, including the perinatal data collection, hospital-admitted patient data collection and mental health outpatient data collection, resulting in a total sample of 429 058 babies and 306 799 mothers. All offspring and their mothers have complete data since birth until December 2015 (and mothers have data collected since 9 months prior to giving birth). Probabilistic linkage (Choicemaker software) is undertaken internally and routinely at the QLD Department of Health and stored as part of a master linkage file (MLF), for which a range of quality assurance procedures are undertaken both before, during and after linkage to ensure data quality [24]. Due to issues of confidentiality, dates of hospital admission were given to our research group only by month/year, thus the data were structured such that each participant had a record of hospital admission for respiratory infection for each month from birth to 1 year.

We made a number of exclusions to ensure participants health records across the first year of life were reflected in the data, and as such babies who were not QLD residents, who were not 12 months old by study conclusion or who died prior to reaching 1 year of age were excluded. Further exclusions were made to ensure the sample comprised of healthy newborns excluding those who: (i) were admitted to neonatal intensive care unit (ICU), (ii) had a major congenital malformation (Supplementary Table S1), (iii) were diagnosed with chronic respiratory disease originating in the perinatal period, intrauterine hypoxia, neonatal asphyxia or an immunological disorder diagnosed at birth or at any time in the first 12 months of life (Supplementary Section 2), (iv) resulted from a multiple birth and (v) resulted from assisted reproductive technologies. Lastly, participants were also excluded if they had a diagnosis of any infection prior to neonatal discharge or any congenital infection at any time in the study period (see Supplementary Text 3). After all exclusions, 73.3% ($n = 314 481$) of the offspring, born to 221 870 mothers were retained in the analysis.

Respiratory infections

Inpatient diagnoses codes conformed to the Australian version of the ICD-10 (ICD-10-AM versions 6, 7, 8 and 9), which corresponds to the WHO ICD-10 with some minor changes to suit the Australian context [25]. Classifying participants RIH status by primary diagnosis is problematic, as the presenting condition often does not describe the infection even if the infection was the only recorded disease (e.g. primary diagnosis – other specified respiratory disorder; secondary diagnosis – acute upper respiratory infection, unspecified) and because primary diagnoses are often selected by professional coding staff to suit the hospitals financial objectives. Instead, participants were classified as having an RIH for a given calendar month if any of the diagnoses from any inpatient episode in that month included a respiratory infection. Supplementary Table S4 shows the full list of infections included in the RIH variable. To account for the fact that an infection may have resulted from another underlying condition (i.e. a proximal risk factor), a second RIH variable was created for use in a sensitivity analysis that classified participants as having an RIH for a given calendar month only if the participant had experienced an inpatient episode in that month which included a respiratory infection and which did also not include a condition that may increase the risk of infection (e.g. cancer, diabetes, anaemia – see Supplementary Text 5 for the full list of exclusionary conditions).

Transient neonatal pulmonary morbidity

All perinatal and neonatal risk factors were recorded by hospital staff either during pregnancy at a hospital prenatal clinical visit, during birth or in the days after birth prior to discharge. Our main interest was in identifying non-specific and often transient pulmonary morbidity that have the potential to identify neonates who may benefit from more rigorous post-discharge follow-up or treatment. The conditions included TTN (ICD: P22.1), respiratory distress syndrome (RDS – ICD: P22.0), other respiratory distress unclassified (ORD – ICD:P22.8/22.9), neonatal aspiration of meconium (MAS – ICD: P24.0), SAPN (ICD: P28.3), apnoea of prematurity (PAPN – ICD: 28.4).

Potential confounders

Confounders were selected *a priori* for two reasons: (i) for their potential to be related to neonatal risk factors and respiratory infections and (ii) as indicators of socio-economic status. Confounders related to neonatal risk factors and respiratory infections included maternal pre-pregnancy overweight or obese (from self-reported height/weight and converted to body mass index categories), pregnancy smoking status (self-report: yes/no), pre-term birth (<37 weeks: yes/no) and low birth weight (<2500 g: yes/no). Confounders of socio-economic status included partner status (married or *de-facto*/no partner), hospital type (private/public), maternal age (years) and Socio-Economic Indexes for Areas (SEIFA) index of relative socio-disadvantage (IRSD). The IRSD takes into account a number of economic and social factors and calculates a score ranging from 1 to 10, which we recoded as low = 1, 2, 3, 4; middle = 5, 6, 7, 8; high = 9, 10 in accordance with established Australian Bureau of Statistics standards [26]. Lastly, we also adjusted for neonate sex and number of mother's previous pregnancies.

Spatial–temporal factors

We included three spatial–temporal covariates primarily to ensure that the longitudinal trajectories of RIH were not simply a function of place or time (i.e. exhibited generality). Baby birth months were categorised into the four southern-hemisphere seasons of summer (December–February), autumn (March–May), winter (June–August) and spring (September–November), and year of birth was categorised into three different periods (2009–10/2011–12/2013–14). Based on the broad area data, we divided the state of QLD into four geographic locations: (1) South East Queensland in which most of the state's population resides with sub-tropical and warm-temperate climatic conditions, (2) northern QLD which includes the coastal tropics in addition to the northern interior region with tropical and hot arid climatic conditions, (3) southern QLD which is mostly coastal but includes some inland areas with sub-tropical and warm arid climatic conditions, and (4) the interior with hot arid climatic conditions (Supplementary Section 6).

Statistical analysis

We used latent class growth analysis (LCGA) to identify longitudinal trajectories of RIH, using the Robust Maximum Likelihood estimator (MLR) available in Mplus version 7.4 [27]. We compared a total of 10 models, one set including only constant and linear terms to describe the growth with class enumeration between 2 and 6, and another set including quadratic terms with the same range of class enumeration. To assess which model best summarised the data, we used a combination of the Bayesian Information Criterion (BIC), the BIC-sample size adjusted (BIC-SSA), in addition to using the Lo–Mendell–Rubin adjusted likelihood ratio test (LMRT) to evaluate the fit between nested models [28]. After establishing the 'best-fitting' model, this model was first re-estimated in subsamples defined by the spatial–temporal factors, and second re-estimated in the full sample adjusted for these factors, to ensure the trajectories exhibited generality across place, time and birth season. Subsequently, multinomial logistic regression was used to assess the associations between the covariates and trajectories, with potential confounders added to the model using a method in which the covariates themselves do not influence class formation (using the Mplus 'Auxiliary = (R3step)' option), and with standard errors calculated to account for the clustering within families (i.e. offspring with the same mothers).

Results

Table 1 shows the prevalence of all the study variables among the 314 481 offspring included in the analyses, with TTN the most prevalent neonatal indicator of pulmonary morbidity (diagnosed among 1.71% of the sample) and PAPN the least prevalent (diagnosed among 0.08%). More than 40% of the maternal sample was overweight or obese prior to pregnancy and 16.2% of mothers smoked, with 4.8% and 3.3% of births recorded as premature or low birth weight, respectively. Only 12.9% of the maternal sample did not have a partner and only 1% were below the age of 18, while 20% were older than 35 years. With regards to the RIH outcomes, the prevalence was lowest in the first month of life and highest in the second, before stabilising to around 0.48% per person/month for the remainder of the study period. It is worth noting that the low prevalence of many of the perinatal variables

Table 1. Prevalence of study variables in the final sample ($n = 314\ 481$)

Covariates	Prevalence % (n)
Predictor variables	
Transient tachypnoea (TTN)	1.71 (5384)
Respiratory distress syndrome (RDS)	1.20 (3765)
Other respiratory distress (ORS)	1.13 (3562)
Aspiration of meconium (MAS)	0.20 (629)
Sleep apnoea (SAPN)	0.17 (543)
Apnoea of prematurity (PAPN)	0.08 (240)
Confounders	
Overweight	23.5 (73 831)
Obese	20.80 (65 420)
Smoking	16.20 (50 718)
Preterm	4.75 (14 928)
Low birth weight	3.32 (10 453)
Female offspring	48.98 (154 021)
Socio-economic status	
No partner	12.92 (40 633)
Private hospital	30.88 (96 951)
Mother age <18	1.00 (3069)
Mother age >35	20.71 (65 132)
Lower SES postcode	31.63 (99 476)
Middle SES postcode	54.58 (171 620)
Resp infection hospitalisation	
Month 1	0.20 (637)
Month 2	0.75 (2358)
Month 3	0.68 (2144)
Month 4	0.54 (1683)
Month 5	0.47 (1463)
Month 6	0.47 (1481)
Month 7	0.47 (1475)
Month 8	0.48 (1499)
Month 9	0.48 (1503)
Month 10	0.46 (1453)
Month 11	0.46 (1441)
Month 12	0.50 (1566)

Previous pregnancies was the only quantitative variable (range 1–10) and was highly skewed so was not included in the table. The median value was 1 with 0.25 and 0.75 percentiles of 1 and 2, respectively.

comes due to the fact that neonates requiring ICU admission or who had a major congenital malformation or infection were excluded (e.g. the prevalence of PAPN, premature birth and low birth weight in the full sample was 0.53%, 8.46% and 6.51%, respectively).

Table 2 shows the fit indices of the competing LCGA models, in which the models without the quadratic term achieve clearly

**Table 2.** Fit indices for the LCGA models (*n* = 314 481)

Model	BIC	BIC-SSA	Entropy	LMRT	LMRT-adj
Linear					
Two-class	229 646	229 646	0.92	<0.001	<0.001
Three-class	229 532	229 507	0.49	<0.001	<0.001
Four-class	229 539	229 504	0.59	0.007	0.008
Five-class	229 567	229 523	0.86	0.003	0.004
Six-class	229 605	229 551	0.86	<0.001	<0.001
Quadratic					
Two-class	229 196	229 173	0.6	<0.001	<0.001
Three-class	228 446	228 411	0.73	<0.001	<0.001
Four-class	228 173	228 125	0.61	<0.001	<0.001
Five-class	228 138	228 078	0.61	<0.001	<0.001
Six-class	228 149	228 076	0.64	0.285	0.291

Models including only constant terms were not considered realistic and therefore not included.

poorer BIC. Among the models including the quadratic term, the five-class model achieved the smallest BIC. However, this model was discarded in favour of the four-class solution for exhibiting evidence of class overextraction (one class from the four-class solution was split into two in the five-class solution with both classes following a substantively similar trajectory and exhibiting very poor classification probabilities – results available upon request). Prior to accepting the four-class solution, the model was re-run in the subsamples defined by season of birth, year of birth and geographic location. Graphs of the resulting trajectories are shown in Supplementary Section 6, indicating that the class pattern and membership proportions were generalisable across these spatial–temporal factors. Thus, we re-ran the four-class solution regressing both the growth factors and class probabilities on the spatial–temporal factors, so that they could inform class formation and membership. This final model is shown in Figure 1, which strongly resembles the unadjusted model (Supplementary Fig. S1a). The four classes included a 'no-risk' trajectory with no RIH risk throughout the study period (79.4% of the sample), a 'low-risk' trajectory exhibiting a weak and increasing risk across time (16.7%), an 'early-risk' trajectory peaking between 2 and 3 months with no risk after 6 months and 'chronic-risk' group peaking at around 7 months at which time members had around a one in four probability of RIH (0.3%).

Table 3 shows the associations among the neonatal pulmonary morbidities and the different RIH trajectories for the complete set of multinomial comparisons, with estimates adjusted for all variables included in the model. Compared with the no-risk group, PAPN was the strongest predictor of membership in one of the other three classes, with neonates diagnosed with PAPN at 25 times the risk of being members of the chronic-risk group (odds ratio (OR) 25.20; 95% confidence interval (CI) 7.12–89.23). RDS was the only other variable that placed neonates at significantly elevated risk of being in all other RIH groups
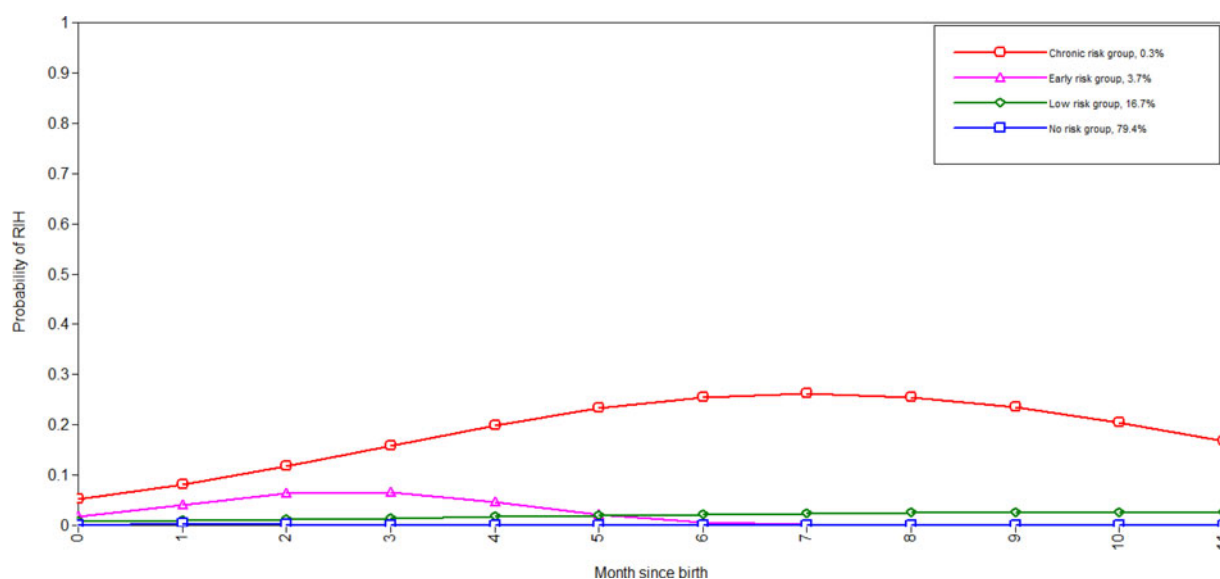
Fig. 1. Trajectories of RIH among infants adjusted for spatial–temporal factors (note the first month of birth is 0). BIC = 227 671; BIC-SSA = 227 414; entropy = 0.63.

Table 3. Multinomial logistic regression showing the associations among neonatal pulmonary morbidities and RIH trajectories (expressed as odds ratios (OR) with 95% confidence intervals (95% CI) and P-value (P)) (n = 314 481)

Predictors	Chronic/none			Early/none			Low/none			Chronic/low			Early/low			Chronic/early		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
TTN	1.35	(0.63–2.91)	0.442	2.66	(1.86–3.82)	<0.001	1.71	(1.40–2.09)	<0.001	0.79	(0.35–1.81)	0.578	1.56	(1.02–2.39)	0.041	0.51	(0.22–1.18)	0.115
RDS	4.74	(2.75–8.15)	<0.001	2.89	(1.87–4.48)	<0.001	2.17	(1.69–2.78)	<0.001	2.18	(1.14–4.16)	0.018	1.33	(0.81–2.20)	0.262	1.64	(0.82–3.26)	0.161
ORS	1.54	(0.62–3.84)	0.357	1.44	(0.77–2.69)	0.246	1.67	(1.32–2.11)	<0.001	0.92	(0.34–2.48)	0.873	0.87	(0.43–1.74)	0.690	1.06	(0.35–3.20)	0.913
MAS	1.53	(0.25–9.40)	0.646	0.94	(0.15–5.86)	0.949	1.58	(0.91–2.73)	0.105	0.97	(0.14–6.94)	0.976	0.60	(0.08–4.33)	0.610	1.62	(0.14–19.54)	0.702
PAPN	25.20	(7.12–89.23)	<0.001	6.95	(1.13–42.86)	0.037	9.42	(2.32–38.18)	0.002	2.68	(0.62–11.50)	0.186	0.74	(0.14–3.87)	0.719	3.63	(0.59–22.42)	0.165
SAPN	9.16	(3.83–21.92)	<0.001	1.80	(0.47–6.92)	0.390	1.33	(0.60–2.95)	0.477	6.87	(1.75–26.98)	0.006	1.35	(0.26–6.99)	0.718	5.08	(1.06–24.31)	0.042

This model is multivariable (i.e. estimates are adjusted for all included variables) but does not include the potentially confounding variables which are included in the analyses shown in Table 4. TTN, transient tachypnoea of newborn; RDS, respiratory distress syndrome; ORD, other respiratory distress unclassified; MAS, neonatal aspiration of meconium; SAPN, sleep apnoea; PAPN, apnoea of prematurity.

compared with the no-risk group, with the strongest risk once again found for the chronic-risk group. Neonates with SAPN were at increased risk of membership in the chronic-risk group when compared with all other RIH groups, and neonates with TTN were at increased risk of membership in the early-risk group when compared with the no- and low-risk groups.

After adjusting for the potential confounding variables (Table 4), the large PAPN estimates attenuated and were non-significant in the early-risk *vs.* no-risk comparison (presumably due to the inclusion of prematurity in the model), while the RDS estimates remained significantly elevated among all the trajectories when using the no-risk group as the reference group. Neonates with SAPN remained at a strongly increased risk of being in the chronic-risk group when compared with the no-risk (OR 7.01; 95% CI 3.02–16.51), low-risk (OR 5.60; 95% CI 1.67–18.81) and early-risk (OR 6.63; 95% CI 0.88–48.94) groups, respectively, although the latter of these estimates did not quite reach agreed standards for statistical significance (P = 0.063). TTN moderately increased the risk of being in the early-risk group compared with the no-risk and low-risk groups. Among the confounders, males were more likely than females to be in the chronic, low or early trajectories, and males were more likely than females to be in the chronic-risk group when compared with any other group. Further, prenatal smoking, not having a partner, preterm birth and being overweight or obese increased the risk of being in the chronic, low or early trajectories but did not differentiate among these trajectories. Lastly, low birth weight was the only other variable aside from SAPN and male gender that increased the likelihood of being in the chronic-risk trajectory compared with all other trajectories.

The results of the sensitivity analysis using the restricted RIH variable are shown in Supplementary Section 8, indicating that the findings did not differ substantively from our main analyses regarding both the trajectory formation and the associations with SAPN, TTN and RDS. In fact, the ORs for SAPN strengthened in the sensitivity analysis such that neonates receiving a diagnosis were at 10 times the risk of being in the chronic-risk *vs.* the low-risk groups (OR 10.19; 95% CI 2.16–48.01) and eight times the risk of being in the chronic-risk *vs.* the no-risk groups (OR 8.22; 95% CI 3.40–19.90). These results strongly suggest that our associations were not the result of other health problems that accompanied RIH.

Discussion

In this study, we identified for the first time empirical longitudinal trajectories of RIH across the first 12 months of life. Our results indicate that there are four distinct groups of children, which differ appreciably in their experience of respiratory infections. Two trajectories exhibited a recurrent course, with one confined to the first 6 months of life (early-risk group), while members of the second trajectory were at an increased risk of RIH across the study period (chronic-risk group). The chronic-risk group in particular represents an exceptional and unfortunate start to life, including only 0.3% of the sample of which members had as much as a one in four risk of RIH during the seventh month of life.

We were able to apply a number of exclusions that suggest that membership in these groups was not simply the result of congenital infection, malformation, immune disorder or major morbidity at birth. Further, the pattern of trajectories showed remarkable generality across place, time and birth season and were robust

Table 4. Multinomial logistic regression showing the associations among neonatal pulmonary morbidities and RIH trajectories (expressed as odds ratios (OR) with 95% confidence intervals (95% CI) and *P*-value (*P*)) (*n* = 314 481)

Predictors	Chronic/none			Early/none			Low/none			Chronic/low			Early/low			Chronic/early		
	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
TTN	1.03	(0.45–2.36)	0.935	2.24	(1.57–3.21)	0.000	1.43	(1.16–1.75)	0.001	0.72	(0.30–1.74)	0.472	1.57	(1.03–2.39)	0.035	0.46	(0.19–1.13)	0.090
RDS	2.61	(1.54–4.43)	0.000	1.96	(1.26–3.07)	0.003	1.53	(1.19–1.98)	0.001	1.70	(0.92–3.15)	0.089	1.28	(0.77–2.12)	0.337	1.33	(0.68–2.62)	0.409
ORS	1.17	(0.50–2.72)	0.713	1.16	(0.59–2.25)	0.669	1.43	(1.11–1.83)	0.005	0.82	(0.33–2.03)	0.666	0.81	(0.38–1.70)	0.577	1.01	(0.35–2.93)	0.982
MAS	1.40	(0.21–9.22)	0.723	1.01	(0.15–6.84)	0.993	1.62	(0.92–2.86)	0.093	0.87	(0.12–6.35)	0.887	0.62	(0.08–4.97)	0.653	1.39	(0.10–18.89)	0.803
PAPN	5.13	(1.60–16.51)	0.006	1.78	(0.20–16.14)	0.607	3.25	(1.07–9.92)	0.038	1.58	(0.41–6.14)	0.509	0.55	(0.05–5.90)	0.620	2.88	(0.27–31.23)	0.384
SAPN	7.01	(3.02–16.28)	0.000	1.06	(0.16–6.83)	0.954	1.25	(0.61–2.57)	0.544	5.60	(1.67–18.81)	0.005	0.84	(0.10–6.84)	0.874	6.63	(0.88–49.94)	0.066
Female	0.34	(0.26–0.45)	0.000	0.59	(0.50–0.68)	0.000	0.61	(0.58–0.65)	0.000	0.55	(0.41–0.74)	0.000	0.96	(0.80–1.14)	0.629	0.57	(0.42–0.79)	0.001
Smoke	2.04	(1.55–2.68)	0.000	1.56	(1.30–1.87)	0.000	1.54	(1.43–1.67)	0.000	1.32	(0.98–1.78)	0.069	1.01	(0.82–1.24)	0.922	1.30	(0.94–1.81)	0.110
Low SES	1.89	(1.14–3.14)	0.014	0.85	(0.69–1.06)	0.163	1.41	(1.27–1.56)	0.000	1.34	(0.78–2.29)	0.285	0.61	(0.47–0.79)	0.000	2.21	(1.27–3.83)	0.005
Middle SES	1.20	(0.71–2.01)	0.499	0.70	(0.57–0.87)	0.001	1.18	(1.07–1.30)	0.001	1.01	(0.59–1.75)	0.961	0.59	(0.46–0.76)	0.000	1.71	(0.98–2.99)	0.061
No partner	1.91	(1.45–2.52)	0.000	1.55	(1.28–1.88)	0.000	1.40	(1.28–1.52)	0.000	1.37	(1.01–1.85)	0.043	1.11	(0.89–1.38)	0.351	1.23	(0.88–1.72)	0.218
Private	0.82	(0.58–1.17)	0.272	1.03	(0.87–1.23)	0.702	0.95	(0.89–1.02)	0.184	0.86	(0.59–1.25)	0.426	1.08	(0.89–1.32)	0.424	0.79	(0.54–1.17)	0.246
Prev preg	1.21	(1.16–1.28)	0.000	1.26	(1.23–1.31)	0.000	1.16	(1.14–1.18)	0.000	1.05	(0.99–1.11)	0.082	1.09	(1.05–1.13)	0.000	0.96	(0.91–1.02)	0.158
Preterm	2.80	(1.88–4.18)	0.000	2.33	(1.77–3.05)	0.000	1.81	(1.57–2.09)	0.000	1.55	(0.99–2.41)	0.054	1.28	(0.94–1.74)	0.113	1.21	(0.75–1.94)	0.441
Low bwt	2.42	(1.59–3.69)	0.000	1.01	(0.68–1.49)	0.975	1.45	(1.23–1.71)	0.000	1.67	(1.04–2.68)	0.035	0.69	(0.45–1.07)	0.096	2.40	(1.37–4.21)	0.002
Overweight	1.47	(1.07–2.00)	0.017	1.24	(1.03–1.50)	0.026	1.22	(1.14–1.31)	0.000	1.20	(0.86–1.68)	0.280	1.02	(0.82–1.26)	0.871	1.18	(0.82–1.70)	0.369
Obese	2.29	(1.74–3.01)	0.000	1.78	(1.49–2.12)	0.000	1.52	(1.41–1.63)	0.000	1.51	(1.12–2.03)	0.006	1.17	(0.96–1.43)	0.117	1.29	(0.93–1.78)	0.127
Age < 18	2.39	(1.14–5.04)	0.022	1.08	(0.38–3.07)	0.888	2.34	(1.82–3.02)	0.000	1.02	(0.45–2.30)	0.961	0.46	(0.15–1.38)	0.166	2.22	(0.63–7.87)	0.217
Age > 35	0.62	(0.45–0.86)	0.005	0.73	(0.60–0.89)	0.002	0.68	(0.63–0.74)	0.000	0.91	(0.64–1.29)	0.592	1.07	(0.85–1.34)	0.574	0.85	(0.58–1.24)	0.403

TTN, transient tachypnoea of newborn; RDS, respiratory distress syndrome; ORD, other respiratory distress unclassified; MAS, neonatal aspiration of meconium; SAPN, sleep apnoea; PAPN, apnoea of prematurity.

to the exclusionary conditions placed on our sample. The identification of these two groups offers new insight into the course of recurrent respiratory infections in infancy and suggests much suffering could be alleviated if the precipitating causes of membership in these trajectories were understood.

The second part of our study was a first attempt in this regard, as we undertook an exploratory analysis to determine if indicators of neonatal pulmonary morbidity could predict class membership, with particular interest in the chronic- and early-risk RIH trajectories. After adjustment for potential confounders, a diagnosis of SAPN prior to neonatal discharge increased the risk of membership in the chronic-risk group compared with all other groups with ORs ranging between 5 and 7 (although the comparison with the early-risk group was borderline significant at $P = 0.063$). Further, PAPN remained a strong predictor of being in the chronic-risk *vs.* the no-risk groups (OR 5.13; 95% CI 1.60, 16.51) even after adjustment for prematurity and birth weight. TTN modestly increased the odds of being in the early-risk group compared with the low- and no-risk groups, and RDS exhibited a dose–response by which diagnosed neonates had increased the risk of being in the chronic-, early- and low-risk groups in that order (although the does–response was a trend as the CIs of the estimates overlapped). Lastly, aside from a weak association by which ORD predicted an increased risk of being in the low-risk trajectory compared with the no-risk trajectory, ORD and MAS were not associated with RIH course.

SAPN activates several pathophysiological mechanisms including sympathetic hyperactivity, pro-inflammatory states and endothelial and immune dysfunction via intermittent hypoxia and sleep fragmentation [29]. Studies have found that SAPN in adults [21, 22] and children [23] are associated with an increased risk of community-acquired pneumonia. Researchers have suggested increased aspiration of pathogen-containing pharyngeal secretions due to apnoea patients breathing against closed airways and the weakened immune system related to sleep loss may increase the infection risk [21, 22]. We were unable to find any previous research looking at SAPN specific to neonates and subsequent infection risk. Thus, to these earlier findings among adult and child samples, we present an important relationship between SAPN among both full-term and preterm neonates and chronic RIH in infancy, which requires further investigation.

Historically, TTN has been believed to have a good prognosis with symptoms usually resolving within 48 h [30]; however, more recent studies have detected associations with subsequent asthma and wheeze in childhood [17–19]. To this, our study suggests TTN may also be implicated in an early course of RIH, which resolves by 6 months of age. On the other hand, a previous study using a population of late pre-term births (32–36 weeks of gestational age) found RDS among newborns predicted lower respiratory tract infections in infancy [20]. Our findings are consistent with and extend this previous research by suggesting that RDS is associated with general respiratory infections with variable courses across the first year of life. Lastly, male gender predicted a worse RIH course, which is consistent with the previous studies finding that respiratory infections among male infants are more severe [31], and low birth weight also predicted a worse RIH course even after adjusted for prematurity.

In this study, we implemented statistical techniques to quantify the distribution of RIH trajectories and factors for the recurrence of RIH in infancy, with the resulting model scrutinised for generalisability across many settings and robustness across different sample compositions. However, a number of study limitations must be considered when interpreting our findings. Most importantly, it is well known that health administrative data lack clinical acuity meaning that our diagnostic groupings would have contained patients of varying severity. As one example, length of cessation in breathing among neonates with SAPN may vary considerably, a factor which is likely to influence the prognosis of the disorder, and one which was not available in our data. Likewise, our outcome variable combined a diverse range of RIH into a single outcome measure, and it is likely that the observed trajectories and associations were partly a function of different pathogens. Although, this fact alone does not invalidate either our trajectories or associations, pathogen-specific RIH trajectories comprise an important additional dimension to the knowledge produced in our study and future research with the capacity to investigate the trajectories of specific infections will likely be valuable. Second, due to confidentiality concerns, our research team was only provided admission dates for month/year, preventing a more detailed analysis including trajectories of re-admission by day (achieved with random-effects models). Thus, in this regard, our trajectories are lacking in detail (e.g. multiple admissions per month) and future studies with more specific date data are needed to improve upon our model.

Lastly, SAPN, PAPN, TTN and RDS can themselves be indicators of an underlying infection or of other serious pathology that may increase the risk for RIH. For example, a major cause of SAPN among children is adenotonsillar hypertrophy, though this contributing factor appears to be less common among infants [32]. Although we excluded neonates with a diagnosable infection prior to neonatal discharge, a major congenital malformation and those requiring ICU admission from all analyses, this does not preclude the risk factors as having arisen in some cases from unrecognised infection or pathology. As such, despite finding independent associations with RIH trajectories in our study, these risk factors may not necessarily constitute a causal link but rather may find use as a marker of increased risk.

In conclusion, we identified four RIH trajectories, associated with neonatal RDS, TTN and SAPN during the first 12 months of life. Our findings can be used to guide prognostic decisions of neonates diagnosed with these transient conditions. Purposively designed studies are needed to replicate our results in an effort to support prevention or intervention efforts to reduce RIH burden among vulnerable neonates.

What's known on this subject. Very little is known about either the longitudinal patterns of respiratory infections across infancy or the relationship with transient neonatal pulmonary morbidity. Previous research has focused on children and implemented crude analyses.

What this study adds. This is the first study to construct empirical trajectories of respiratory infection hospitalisation (RIH) across infancy, finding four different trajectories, two of which exhibited a recurrent course. Further, we identified two novel associations by which transient tachypnoea of newborn specifically predicted a recurrent course of RIH that concluded by age 6 months and sleep apnoea showed a strong specific relationship with chronic RIH.

References

1. **Mizgerd JP** (2006) Lung infection – a public health priority. *PLoS Medicine* **3**, e76.
2. **Armstrong GL, Conn LA and Pinner RW** (1999) Trends in infectious disease mortality in the United States during the 20th century. *JAMA* **281**, 61–66.
3. **Hobbs MR et al.** (2017) Ethnic disparities in infectious disease hospitalisations in the first year of life in New Zealand. *Journal of Paediatrics and Child Health* **53**, 223–231.
4. **Yorita KL et al.** (2008) Infectious disease hospitalizations among infants in the United States. *Pediatrics* **121**, 244–252.
5. **Tregoning JS and Schwarze J** (2010) Respiratory viral infections in infants: causes, clinical symptoms, virology, and immunology. *Clinical Microbiology Reviews* **23**, 74–98.
6. **Singleton RJ et al.** (2009) Risk factors for lower respiratory tract infection death among infants in the United States, 1999–2004. *Pediatrics* **124**, e768–e776.
7. **Hepworth SJ et al.** (2010) Early life patterns of common infection: a latent class analysis. *European Journal of Epidemiology* **25**, 875–883.
8. **Blomström Å et al.** (2015) Associations between maternal infection during pregnancy, childhood infections, and the risk of subsequent psychotic disorder – a Swedish cohort study of nearly 2 million individuals. *Schizophrenia Bulletin* **42**, 125–133.
9. **Brauer M et al.** (2002) Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. *American Journal of Respiratory and Critical Care Medicine* **166**, 1092–1098.
10. **Jones LL et al.** (2011) Parental and household smoking and the increased risk of bronchitis, bronchiolitis and other lower respiratory infections in infancy: systematic review and meta-analysis. *Respiratory Research* **12**, 5.
11. **Tromp I et al.** (2017) Breastfeeding and the risk of respiratory tract infections after infancy: the Generation R Study. *PLoS ONE* **12**, e0172763.
12. **Welliver RC** (2003) Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. *The Journal of Pediatrics* **143**, 112–117.
13. **Merenstein DJ, Gatti ME and Mays DM** (2011) The association of mode of delivery and common childhood illnesses. *Clinical Pediatrics* **50**, 1024–1030.
14. **Nokso-Koivisto J et al.** (2002) Viral etiology of frequently recurring respiratory tract infections in children. *Clinical Infectious Diseases* **35**, 540–546.
15. **Toivonen L et al.** (2016) Burden of recurrent respiratory tract infections in children: a prospective cohort study. *The Pediatric Infectious Disease Journal* **35**, e362–e369.
16. **Alho O-P et al.** (1990) Risk factors for recurrent acute otitis media and respiratory infection in infancy. *International Journal of Pediatric Otorhinolaryngology* **19**, 151–161.
17. **Birnkrant DJ et al.** (2006) Association of transient tachypnea of the newborn and childhood asthma. *Pediatric Pulmonology* **41**, 978–984.
18. **Liem JJ et al.** (2007) Transient tachypnea of the newborn may be an early clinical manifestation of wheezing symptoms. *The Journal of Pediatrics* **151**, 29–33.
19. **Çakan M et al.** (2011) Correlation between transient tachypnea of the newborn and wheezing attack. *Pediatrics International* **53**, 1045–1050.
20. **Szabo SM et al.** (2012) Respiratory distress syndrome at birth is a risk factor for hospitalization for lower respiratory tract infections in infancy. *The Pediatric Infectious Disease Journal* **31**, 1245–1251.
21. **Chiner E et al.** (2016) Association between obstructive sleep apnea and community-acquired pneumonia. *PLoS ONE* **11**, e0152749.
22. **Su VY-F et al.** (2014) Sleep apnea and risk of pneumonia: a nationwide population-based study. *Canadian Medical Association Journal* **186**, 415–421.
23. **Goldbart AD et al.** (2012) Sleep-disordered breathing is a risk factor for community-acquired alveolar pneumonia in early childhood. *CHEST Journal* **141**, 1210–1215.
24. **QLD Health** (2016) *Queensland Data Linkage Framework*. State of Queensland.
25. **Australian Consortium for Classification Development**. ICD-10-AM/ACHI/ACS 2017. Available at https://www.accd.net.au/Icd10.aspx.
26. **ABS**. IRSD 2014. Available at http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.
27. **Muthén LK and Muthén BO** (1998–2010) *Mplus User's Guide*. Los Angeles: Muthén & Muthén.
28. **Nylund K, Asparouhov T and Muthen BO** (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modelling* **14**, 535–569.
29. **Martínez-García MÁ and Chiner E** (2017) Sleep apnoea and risk of postoperative infection: beyond cardiovascular impact. *European Respiratory Society* **49**.
30. **Hagen E, Chu A and Lew C** (2017) Transient tachypnea of the newborn. *NeoReviews* **18**, e141–e148.
31. **Falagas ME, Mourtzoukou EG and Vardakas KZ** (2007) Sex differences in the incidence and severity of respiratory tract infections. *Respiratory Medicine* **101**, 1845–1863.
32. **Greenfeld M et al.** (2003) Obstructive sleep apnea syndrome due to adenotonsillar hypertrophy in infants. *International Journal of Pediatric Otorhinolaryngology* **67**, 1055–1060.