Dynamics of respiratory symptoms during infancy and associations with wheezing at school age

ONLINE SUPPLEMENT

BILD study group

Members of the BILD (Basel-Bern Infant Lung Development) study group that contributed data or intellectual input to this project (in alphabetical order):

Pinelopi Anagnostopoulou, Paediatric Respiratory Medicine, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

Urs Frey, University of Basel Children's Hospital, UKBB, Basel, Switzerland

Oliver Fuchs, Paediatric Respiratory Medicine, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

Olga Gorlanova, University of Basel Children's Hospital, UKBB, Basel, Switzerland

Insa Korten, Paediatric Respiratory Medicine, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

Claudia Kuehni, Institute of Social and Preventive Medicine, University of Bern, Switzerland

Philipp Latzin, Paediatrics Respiratory Medicine, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

Elena Proietti, University Children’s Hospital Zurich, Switzerland

Martin Röösli, University of Basel, Basel, Switzerland

Anne Schmidt, Barts Health NHS Trust London, England

Jakob Usemann, University of Basel Children's Hospital, UKBB, Basel, Switzerland

METHODS

Additional methods for study participants

This ongoing prospective birth cohort study comprised a group of unselected, healthy term-born neonates recruited antenatally.[1] Infants were enrolled in two centres in Switzerland (in Bern between 1999 and 2015, and in Basel between 2011 and 2015). Exclusion criteria for the study were preterm delivery (<37 week gestational age), major birth defects, respiratory distress after birth, other significant perinatal disease, or a later diagnosis of airway malformation, or specific chronic respiratory disease.

Additional methods on skin prick testing

Skin prick testing was performed in a subgroup of 270 children for the following allergens: dog and cat dander, two tree mixtures, grass mixture, *Alternaria alternate*, *Dermatophagoides pteronyssinus*, and *Dermatophagoides farinae* (Allergomed AG, Therwil, Switzerland). Positive control was histamine; negative control was the solution in which allergens were dissolved. The reaction was assessed 15 minutes after skin prick testing. Response was determined positive if a wheal diameter of any of the tested allergen was greater than the positive control.

Additional methods on lung function and fractional exhaled nitric oxide measurement

Spirometry was performed with the child in a seated position with the nose clipped using the MasterLab setup (Jaeger, Wuerzburg, Germany) according to standard guidelines. The best of at least three repeatable forced expiratory maneuvers (within 100 ml) was recorded.[2] Forced expiratory volume in 1 s (FEV_1), forced vital capacity (FVC), FEV_1/FVC ratio and forced expiratory flow at 25-75% of FVC ($FEF_{25-75\%}$) were expressed as sex- and height-standardized z-scores.[3] We measured the fraction of exhaled nitric oxide (FeNO) by the single-breath method with a rapid-response chemoluminescence analyser (CLD 88 sp; EcoMedics, Duernten, Switzerland). Flow was recorded using an ultrasonic flow meter (Spiroson; EcoMedics) according to current guidelines.[4]

Additional methods for the Markov matrix

Respiratory tract infections during infancy were prospectively assessed using the following scores: score zero for the healthy state and scores 1-4 (in ascending order of severity) for the symptomatic states (Table E1). For each possible symptom state *x* (i.e. initial state) out of the five states 0,1,2,3,4, we counted how often, during the time window of 52 weeks observation, a transition to any other state *y* (i.e. target state) (in symbols $x \rightarrow y$) took place within one week's time. For example, an infant is healthy (symptom score=0) for three consecutive weeks (transitions: 0→0→0); develops a severe infection during the fourth week with symptom score 3 (transitions: 0→0→0→3); this infection (symptom score=3) lasts another week (transitions: 0→0→0→3→3); and finally, the infant fully recovers during the subsequent week (symptom score=0) (transitions: 0→0→0→3→3→0). Hence, this 6-week observation period can be described by the following transitions (initial state → target state): 0→0, 0→0, 0→3, 3→3, 3→0. Therefore, during the time window of observation, the transition 0→0 was observed two times, and the transitions 0→3, 3→3, 3→0 once, respectively. All the other remaining possible transitions (there are a total of 5×5=25 possible transitions among the five states 0, 1, 2, 3, and 4), were not observed during this observation period. All these counts, or absolute frequencies, can be displayed in a 5×5-matrix in which the rows and columns are labelled using the symptom states (0, 1, 2, 3, and 4). Such a table is displayed in Figure 1B for one infant of the cohort.

The absolute frequencies from this table were then replaced by the corresponding relative frequencies (empirical probabilities). The entries in this table can be interpreted as conditional probabilities. For example, if a given infant is currently displaying symptoms corresponding to state *x*, then the probability that the symptoms will progress to symptom state *y* is given by the corresponding entry in this matrix (Figure 1C). Such a table of conditional probabilities is referred to in mathematics as a Markov matrix.[5] Each row in this matrix encodes a discrete probability distribution, namely the probability to progress from a given state to any other of the possible five states. We calculated the average estimated entropy[6] of all the conditional probability distributions defined by the rows of the participant's Markov matrix. This average estimated entropy gives a quantitative measure for the irregularity of the patterns of symptom deterioration and recovery of a given infant. The entropy was estimated using the maximum likelihood estimator, which results from the substitution of the empirical probabilities (i.e. the entries in a given row of a Markov matrix) into the Shannon entropy formula:

$$Entropy\ (row\ j) = -\sum_{i=0}^{4} p_{ji} \log(p_{ji})$$

Where p_{ji} denotes the entries in row *j* of the Markov matrix, and *i* is the symptom state score (0→4). When single empirical probabilities p_{ji} happen to be equal to zero, the corresponding term in the above formula is set equal to zero, which corresponds to taking the limit $\lim_{p \to 0} p \log p$. However, when all entries in an entire row of a given

Markov matrix are equal to zero, i.e., when a certain symptom state was never reported or observed during the time window of observation, then the above formula no longer applies, due to the complete lack of information about the corresponding conditional probability distribution. In such cases, the rows consisting of only zero's do not contribute to the calculation of the average entropy.

Additional methods for statistical analysis

The Shannon entropy is defined as the average amount of information needed to specify the state of a random variable governed by a given probability distribution. In the context of statistical physics it can be interpreted as a measure of disorder within a system; higher entropy values correspond to a higher irregularity.[7] For each infant, we calculated the average estimated entropy[6] of all the conditional probability distributions defined by the rows of the participant's Markov matrix. This average estimated entropy gives a quantitative measure of the irregularity in the patterns of symptom deterioration and recovery of a given patient. We used the parameter average entropy to identify phenotypes using the software R version 3.0.2 (http://www.R-project.org)[8] and applying hierarchical clustering. To this end, the separation between entropy values was calculated using the Euclidean distance, and the agglomeration procedure was done according to Ward's minimum variance method. The optimal number of phenotypes was determined based on the majority of indices using the R-package "NbClust".[9] Previous phenotyping studies used symptom occurrence (e.g. rhinitis and wheeze) rather than symptom dynamics during infancy.[10-12] Therefore, in order to compare our finding to previous studies, we defined similarly sized "reference phenotypes" based on the frequency distribution of the total number of symptom weeks during infancy.

Using logistic regression analysis, we studied the association of the symptom dynamic phenotypes, the reference phenotypes, and known asthma risk factors (e.g. sex, maternal asthma), with the child's outcomes at six years. For any wheezing, current wheezing, atopic disease, URTI, and positive prick test of the child, we adjusted the analyses for sex, maternal education, maternal asthma, ETS exposure, childcare, and presence of siblings. For the outcome FeNO, we additionally adjusted for the child's hay fever and inhaled corticosteroid use. For lung function, analyses were adjusted for maternal education, maternal asthma, ETS exposure, childcare, and presence of siblings.[13] Differences in the distribution of characteristics across phenotypes were assessed using Chi-squared tests for categorical variables, and Kruskal–Wallis for continuous variables, and a Bonferroni-corrected significance level was used to account for multiple pair-wise testing. We used the weighted kappa-statistic[14] to compare agreement between the symptom dynamic and reference phenotypes. For sensitivity analyses, we repeated the analysis in infants who had ≥one episode of symptom score ≥three, and also within an additional, independent sample from our cohort of 242 infants. In order to explore whether the entropy distribution found was an artefact of our analysis, we re-categorized the symptom states (0, 1, 2, (3+4)→3), simulated data, and perturbed the existing data. Analyses were done using Stata 11.2 software (StataCorp, College Station, Tex), figures using Matlab R2015b (The MathWorks, Inc., Natick, MA, USA).

Additional methods for sensitivity analysis

We performed several sensitivity analyses:

1) We repeated the analysis in infants having ≥one episode of symptom score ≥three.

2) Robustness of the identified phenotypes was investigated within an additional, independent sample from our cohort of 242 infants. This sample included infants not seen for follow-up, and hence, these infants were not included in the main analysis in the first place. This sample served as an independent data set for an external validation of the method.

3) To explore whether the distribution of entropy was an artefact of our analysis approach, we re-categorized the symptom states (0, 1, 2, (3+4)→3), simulated different data sets, and perturbed the existing data sets.

4) To the best of our knowledge, this is the first prospective cohort reporting respiratory symptoms during infancy at such a high resolution and accuracy. However, in theory, or in clinical practice, some symptomatic transitions may not be reported or acquired. The aim of this sensitivity analysis was to investigate the effect of correcting for potentially unobserved events. In particular, we hypothetically investigated the scenario of analysing – with our method – a population at higher risk for respiratory symptoms compared to our study population. To this end, we estimated the average entropy using an estimator that takes into account the issue of unobserved events (Schuerman and Grassberger).[15] Within this approach, unobserved or unreported symptom transitions are artificially introduced into the data set according to the expected value of a 'posterior' distribution of event occurrences resulting from a Bayesian analysis[16] of the data set. However, for the sake of consistency, when all entries in an entire row of a given Markov matrix were equal to zero, i.e., when a certain symptom state was never reported or observed during the time window of observation, then, such a row did not contribute to the calculation of the average entropy using the Schuerman and Grassberger[15] estimator. These calculations were done using the R-package "entropy".[17]

RESULTS

Additional results for sensitivity analysis

1) We obtained similar results in infants with ≥one episode of symptom score ≥three (data not shown).

2) In the validation dataset of 242 infants, symptom dynamic phenotypes included 80, 132, and 30 infants, respectively. Phenotypes 1, 2, and 3 had 0.23, 0.79, and 2.03 weeks with severe symptoms, respectively. Risk factors differed across phenotypes: phenotype 1 had an underrepresentation of infants who attended childcare. Phenotype 2 had fewer infants of asthmatic mothers. Phenotype 3, the smallest group, had an overrepresentation of infants born via Caesarean section, and more infants had atopic mothers. Reference phenotypes included 83, 108, and 51 infants, respectively. Phenotypes 1, 2, and 3 had 0.16, 0.68, and 1.91 weeks with severe symptoms, respectively. Phenotype 1 had an underrepresentation of infants with siblings and less infants attended childcare. In phenotype 2, infants were more likely born via Cesarean section. Phenotype 3 had an overrepresentation of males, and more infants had siblings (Table E5).

3) The outcomes of the following sensitivity analysis ruled out the possibility of an artefact in our methodology. Re-categorization of the states (0, 1, 2 (3+4)→3), did not systematically change the average entropy distribution and the identified phenotypes (data not shown). The distribution of the average estimated entropy from the original cohort (Figure E2A) did not show up in a simulated dataset in which all possible symptom transitions were assumed to be equally likely (Figure E2B). We perturbed the existing data by shuffling and thereby randomly altering the chronological order of each patient's reported symptoms. The resulting distribution of the average entropy remained qualitatively very similar to the one observed in the unperturbed data (Figure E2C). This is not surprising, as our method captures the patterns of symptom deterioration and recovery, and not necessarily the temporal symptom pattern. A multimodal distribution was observed when the data was simulated in such a way that the probability of any given symptom transition was determined by the pooled, existing data of the cohort. However, the shape of the distribution changed considerably in this simulated dataset (Figure E2D).

4) As elucidated above, we repeated the entire analysis taking into account the issue of unobserved events.[15] After this correction for unobserved events, the analysis of the resulting distribution of entropy values yielded four phenotypes. These phenotypes included 124, 79, 87, and 32 infants, respectively. Phenotypes 1, 2, 3, and 4 had 0.9, 1.2, 0.2, and 0 weeks with severe symptoms. Risk factors differed across phenotypes: phenotype 1 had less infants exposed to tobacco smoke exposure. Phenotype 2 had more males, more infants attended childcare, and more infants had siblings. Fewer infants in phenotype 3 had siblings, and infants were more likely born via Caesarean section. Phenotype 4 had less male infants, more infants were born to atopic mothers, and less infants attended childcare. We considered phenotype 2 a high-risk phenotype, since the prevalence of current wheezing (17%) at six years was the highest. However, compared to the high-risk dynamic phenotype originally obtained, the prevalence of wheezing was proportionally lower. Regarding the association with wheezing at six years, the following changes were observed: in the adjusted logistic regression model, compared to phenotype 1, there was no significant association with current wheezing (adjusted odds ratio; 95% CI) (1.05; 0.44–2.47) at six years of age, compared to significant associations in the high-risk phenotype originally obtained. The sensitivity analyses as described in the previous section 3 were equally applied in the dataset after taking into account the issue of unobserved events (Figure E3).

To summarize, we can conclude that the resulting phenotypes, and their characteristics, do depend on the distribution of symptom severity in the population studied, and on the quality and completeness of the data acquired. In light of this theoretical result, we believe that in a proof-of-principle study, like the present study, further validation using an additional cohort would not necessarily extend our insights. Such a validation in an additional cohort would most likely display a different distribution of symptom severity and may contain more missing values.

TABLES

Table E1 Characteristics of the cohort

	Follow-up study 322	Lost to follow-up 47	P-value
Anthropometrics at birth			
Gestational age, weeks	39.6 ± 1.2	39.8 ± 1.2	0.139
Birth weight, kg	3.3 ± 0.4	3.3 ± 0.4	0.622
Length, cm	49.5 ± 1.9	49.4 ± 2.2	0.821
Respiratory symptoms in the first year of life			
weeks with severe symptoms	0.72 ± 1.1	1.0 ± 2.04	0.169
weeks with symptoms	5.33 ± 4.61	5.74 ± 4.23	0.568
Risk factors			
Male sex	167 (51.8)	29 (61.7)	0.207
Siblings	159 (49.4)	28 (59.5)	0.192
Caesarean section	54 (16.7)	4 (14.8)	0.146
Maternal asthma	34 (10.5)	9 (19.2)	0.086
Maternal atopy	116 (36.1)	17 (36.2)	0.985
Childcare	62 (19.2)	7 (13.5)	0.474
Maternal smoking during pregnancy	27 (8.4)	9 (19.2)	0.020
Parental smoking during infancy	70 (21.7)	16 (34.1)	0.062
Breastfeeding >26 weeks	252 (78.3)	23 (65.2)	<0.001
Low maternal education	202 (62.7)	30 (65.2)	0.744
Season of birth			
Spring	87 (27.1)	10 (21.3)	0.404
Summer	83 (25.7)	27 (27.6)	0.783
Autumn	80 (24.8)	14 (29.8)	0.468
Winter	72 (22.4)	10 (21.3)	0.867

Data are given as mean ± standard deviation or absolute numbers (percent). Children younger than 5 years were not enrolled for follow-up at 6 years. P-values are for comparison between subjects lost to follow-up, and those who were followed-up. We used the Chi-squared tests for categorical variables, and Kruskal–Wallis for continuous variables.

Table E2 Association of symptom dynamic phenotypes, reference phenotypes, and risk factors with wheezing between 2 and 3 years of age

	Univariable association			Multivariable association[#]		
	OR	95% CI	P-value	OR	95% CI	P-value
Outcome wheezing between 2-3 years[+] (31/322)						
Reference phenotypes						
Phenotype 1 (baseline) (147)	1	reference		1	reference	
Phenotype 2 (128)	1.27	0.54 –3.01	0.573	1.31	0.53–3.19	0.565
Phenotype 3 (47)	2.53	0.95–6.74	0.062	2.44	0.78–7.63	0.122
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (145)	1	reference		1	reference	
Phenotype 2 (135)	0.61	0.23–1.58	0.308	0.65	0.23–1.82	0.422
Phenotype 3 (42)	4.43	1.81–10.82	0.001	4.86	1.74–13.56	0.003
Risk factors						
Male sex	1.78	0.82–3.85	0.142	1.57	0.69–3.53	0.142
Siblings	1.47	0.69–3.12	0.311	1.71	0.76–3.87	0.190
Maternal asthma	2.85	1.12–7.23	0.027	2.61	0.99–6.85	0.051
Childcare	1.25	0.51–3.05	0.622	1.36	0.51–3.65	0.538
Parental smoking during infancy	1.54	0.67–3.52	0.303	1.25	0.52–2.99	0.605
Low maternal education	1.51	0.67–3.39	0.321	1.43	0.59–3.47	0.421

CI: confidence interval; OR: odds ratio. Logistic regression for the outcome wheezing between two and three years of age. Symptom dynamic phenotypes were defined by entropy of transition states, reference phenotypes by weeks with any respiratory symptom. When considering phenotypes as exposure, phenotype one from the reference phenotype, or symptom dynamic phenotype served as baseline, respectively. Numbers in brackets indicate group sizes. [#]Adjusted for the following binary variables: male gender, low maternal education, maternal asthma, maternal smoking during pregnancy, childcare attendance during infancy, and presence of siblings. [+]Defined as any wheezing episode between two and three years of age.

Table E3 Association of symptom dynamic phenotypes, reference phenotypes and risk factors with positive prick test at six years

	Univariable association			Multivariable association*		
	OR	95% CI	P-value	OR	95% CI	P-value
Outcome positive prick test at 6 years (37/270)						
Reference phenotypes						
Phenotype 1 (baseline) (147)	1	reference		1	reference	
Phenotype 2 (128)	0.88	0.41–1.86	0.743	0.84	0.38–1.84	0.673
Phenotype 3 (47)	0.89	0.31–2.61	0.833	0.93	0.28–3.09	0.914
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (145)	1	reference		1	reference	
Phenotype 2 (135)	0.91	0.42–1.96	0.810	0.98	0.44–2.21	0.971
Phenotype 3 (42)	1.56	0.58–4.17	0.367	1.94	0.66–5.67	0.226
Risk factors						
Male sex	2.63	1.21–5.68	0.014	2.57	1.16–5.67	0.019
Siblings	0.99	0.49–1.98	0.986	0.92	0.44–1.92	0.841
Maternal asthma	0.43	0.99–1.91	0.272	0.41	0.09–1.86	0.252
Childcare	0.76	0.30–1.94	0.575	0.78	0.29–2.11	0.634
Parental smoking during infancy	0.96	0.41–2.22	0.925	0.77	0.32–1.87	0.577
Low maternal education	2.16	0.94–4.93	0.068	1.95	0.82–4.64	0.131

CI: confidence interval; OR: odds ratio. Logistic regression for the outcome positive skin prick test (SPT). Symptom dynamic phenotypes were defined by entropy of transition states, reference phenotypes by weeks with any respiratory symptom. When considering phenotypes as exposure, phenotype one from the reference phenotype, or dynamic phenotype, served as baseline, respectively. Numbers in brackets indicate group sizes. *Adjusted for the following variables: male gender, low maternal education, maternal asthma, maternal smoking during pregnancy, childcare attendance during infancy and presence of siblings.

Table E4 Association of symptom dynamic phenotypes and reference phenotypes with lung function and FeNO at six years

	Univariable association			Multivariable association*		
	OR	95% CI	P-value	OR	95% CI	P-value
Outcome $zFEV_1$						
Reference phenotypes						
Phenotype 1 (baseline) (105)	1	reference		1	reference	
Phenotype 2 (81)	0.01	-0.25–0.26	0.969	-0.01	-0.28–0.26	0.963
Phenotype 3 (36)	-0.02	-0.36–0.32	0.911	-0.01	-0.38–0.37	0.988
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (103)	1	reference		1	reference	
Phenotype 2 (89)	-0.02	-0.27–0.23	0.877	-0.01	-0.27–0.26	0.978
Phenotype 3 (30)	0.12	-0.24–0.48	0.518	0.14	-0.24–0.53	0.451
Outcome zFVC						
Reference phenotypes						
Phenotype 1 (baseline) (105)	1	reference		1	reference	
Phenotype 2 (81)	0.01	-0.25–0.26	0.969	-0.01	-0.28–0.26	0.963
Phenotype 3 (36)	-0.02	-0.36–0.32	0.911	-0.01	-0.38–0.37	0.988
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (103)	1	reference		1	reference	
Phenotype 2 (89)	-0.01	-0.27–0.23	0.877	-0.01	-0.27–0.26	0.978
Phenotype 3 (30)	0.12	-0.24–0.48	0.518	0.14	-0.24–0.53	0.451
Outcome $zFEV_1/FVC$						
Reference Phenotypes						
Phenotype 1 (baseline) (105)	1	reference		1	reference	
Phenotype 2 (81)	-0.11	-0.38–0.17	0.458	-0.04	-0.32–0.23	0.756
Phenotype 3 (36)	-0.11	-0.24–0.47	0.517	0.23	-0.15–0.61	0.241
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (103)	1	reference		1	reference	
Phenotype 2 (89)	0.01	-0.27–0.27	0.996	0.06	-0.21–0.34	0.649
Phenotype 3 (30)	-0.11	-0.51–0.26	0.544	-0.05	-0.45–0.34	0.793
Outcome $zFEF_{25-75\%}$						
Reference phenotypes						
Phenotype 1 (baseline) (105)	1	reference		1	reference	
Phenotype 2 (81)	-0.01	-0.25–0.24	0.962	0.04	-0.21–0.31	0.743
Phenotype 3 (36)	0.01	-0.28–0.41	0.728	0.17	-0.19–0.54	0.361
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (103)	1	reference		1	reference	
Phenotype 2 (89)	0.01	-0.22–0.27	0.868	0.08	-0.18–0.33	0.549
Phenotype 3 (30)	0.05	-0.29–0.41	0.749	0.15	-0.22–0.52	0.428
+Outcome FeNO ppb						
Reference Phenotypes						

Phenotype 1 (baseline) (101)	1	reference		1	reference	
Phenotype 2 (95)	0.96	0.81–1.14	0.667	1.01	0.85–1.21	0.845
Phenotype 3 (35)	0.98	0.77–1.24	0.376	1.05	0.81–1.36	0.678
Symptom dynamic phenotypes						
Phenotype 1 (baseline) (101)	1	reference		1	reference	
Phenotype 2 (95)	0.99	0.84–1.18	0.982	0.99	0.83–1.18	0.948
Phenotype 3 (35)	1.02	0.79–1.31	0.849	1.08	0.83–1.41	0.537

CI: confidence interval; FeNO: fractional exhaled nitric oxide; FVC: forced vital capacity; FEV_1: forced expiratory volume in one second; $FEF_{25-75\%:}$ forced expiratory flow from 25-75% of exhalation; OR: Odds ratio. Lung function parameters are expressed as z-scores according to Qanjer et al..[3] FeNO was log-transformed before analyses. Symptom dynamic phenotypes were defined by entropy of transition states, reference phenotypes by weeks with any respiratory symptom. Phenotype one from the reference phenotype, or symptom dynamic phenotype, served as baseline, respectively. Numbers in brackets indicate group sizes. [*]Adjusted for the following binary variables: maternal education, maternal asthma, maternal smoking during pregnancy, childcare attendance during infancy and presence of siblings. [+]Additionally adjusted for the following binary variables: sex, hay fever of the child, inhaled corticosteroid use.

Table E5 Characteristics of symptom dynamic phenotypes and reference phenotypes from the 242 study participants not seen for follow-up at six years

	Phenotype 1		Phenotype 2		Phenotype 3		P-value	
	Dyn.	Ref.	Dyn.	Ref.	Dyn.	Ref.	Dyn.	Ref.
	80	83	132	114	30	45		
Respiratory symptoms								
Weeks with severe symptoms	0.23 (0.45)	0.16 (0.41)	0.79 (1.31)	0.68 (0.87)	2.03 (1.93)	1.91 (2.13)	**<0.001**	**<0.001**
Weeks with symptoms	2.01 (1.57)	1.75 (1.08)	7.46 (3.41)	6.5 (1.83)	13.0 (4.21)	13.49 (2.91)	**<0.001**	**<0.001**
Respiratory symptoms transition states								
Entropy of transition states	0.09 (0.06)	0.15 (0.18)	0.57 (0.13)	0.54 (0.21)	1.07 (0.17)	0.85 (0.26)	**<0.001**	**<0.001**
Risk factors								
Male sex	55.0	53.1	52.3	50.1	50.0	57.8	0.876	0.734
Siblings	52.5	53.1	65.2	64.1	63.3	66.7	0.179	0.196
Cesarean section	20.0	18.1	26.5	31.5	33.3	22.2	0.313	0.086
Maternal asthma	12.5	13.3	8.3	7.9	10.0	8.9	0.616	0.447
Maternal atopy	42.5	38.5	32.6	38.6	36.7	26.7	0.346	0.325
Childcare	23.7	21.7	46.2	45.6	43.3	51.1	**0.004**	**<0.001**
Maternal smoking during pregnancy	8.7	8.4	8.3	8.7	3.3	4.4	0.613	0.639
Parental smoking during infancy	22.5	20.5	21.2	21.9	23.3	24.5	0.956	0.875
Breastfeeding <26 weeks	62.5	59.1	65.9	66.7	63.3	68.9	0.873	0.429
Low maternal education	41.6	46.2	45.1	42.1	40.0	41.2	0.823	0.803
Seasons at birth								
Spring	15.6	14.5	22.8	30.5	45.6	27.4	0.048	0.030
Summer	21.2	24.1	28.1	25.0	23.3	27.4	0.528	0.947
Autumn	40.0	39.7	25.7	25.9	20.0	21.6	0.041	0.045
Winter	23.7	21.7	18.2	18.5	23.3	23.5	0.579	0.672

Values are percent or means (SD). Symptom dynamic phenotypes (Dyn.) are shown on white background and reference phenotypes (Ref.) are shown on shaded grey background. Symptom dynamic phenotypes were defined by entropy of transition states, reference phenotypes by weeks with any respiratory symptom. Differences in the distribution of characteristics across phenotypes were assessed using Chi-squared tests for categorical variables, and Kruskal–Wallis for continuous variables. Significant p-values at the Bonferroni-corrected α level of 0.017 are shown in bold face.

FIGURES

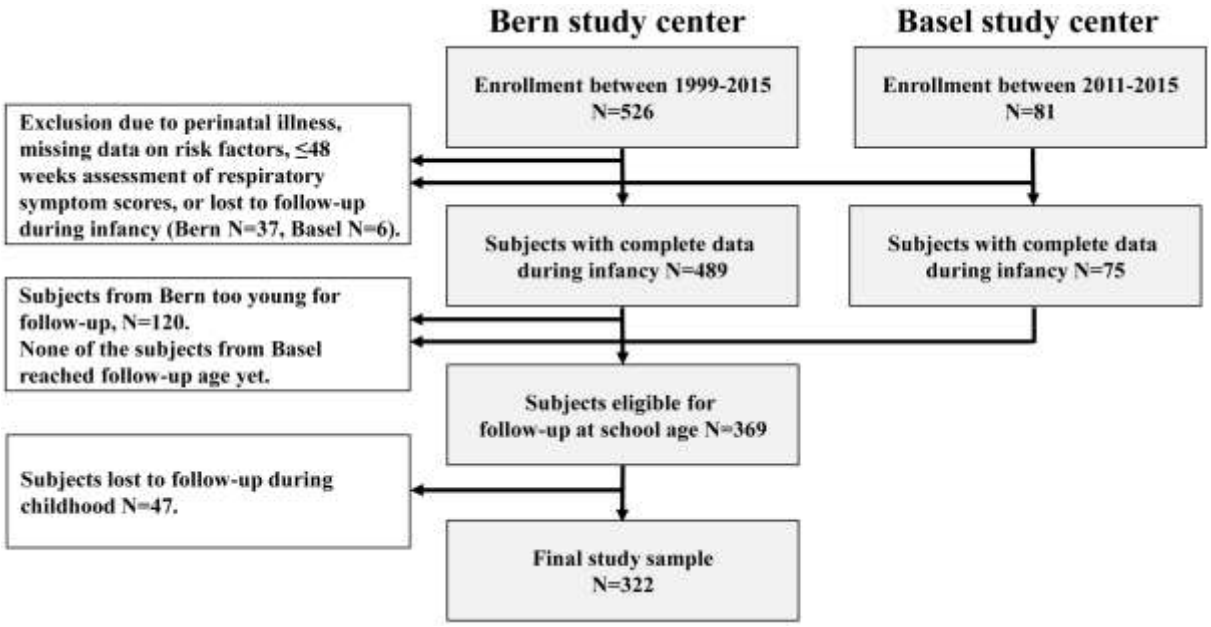

Figure E1 Flow Chart of the study population.

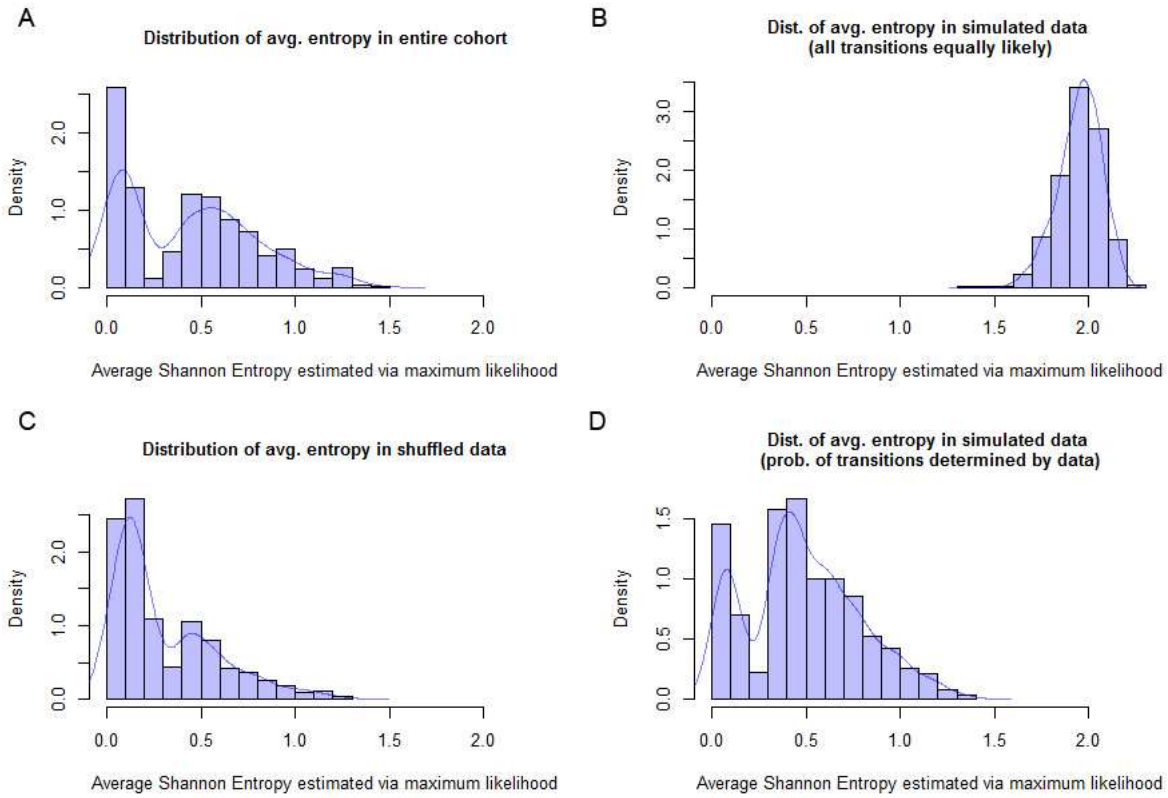

Figure E2 Histograms of estimated entropy in the cohort and in simulated and perturbed data sets. A) Histogram of the average estimated entropy of the rows of the Markov matrix for each participant in the cohort. B) Histogram of the average estimated entropy of the rows of the Markov matrix in a simulated data set in which all symptom transitions are assumed to be equally probable. C) Histogram of the average estimated entropy of the rows of the Markov matrix for each participant after each participant's chronological order of reported symptoms was randomly permuted. D) Histogram of the average estimated entropy of the rows of the Markov matrix in a simulated data set in which the probabilities of symptom transitions were determined by the pooled cohort data. The entropy in all four panels was estimated using the maximum likelihood estimator of the entropy.

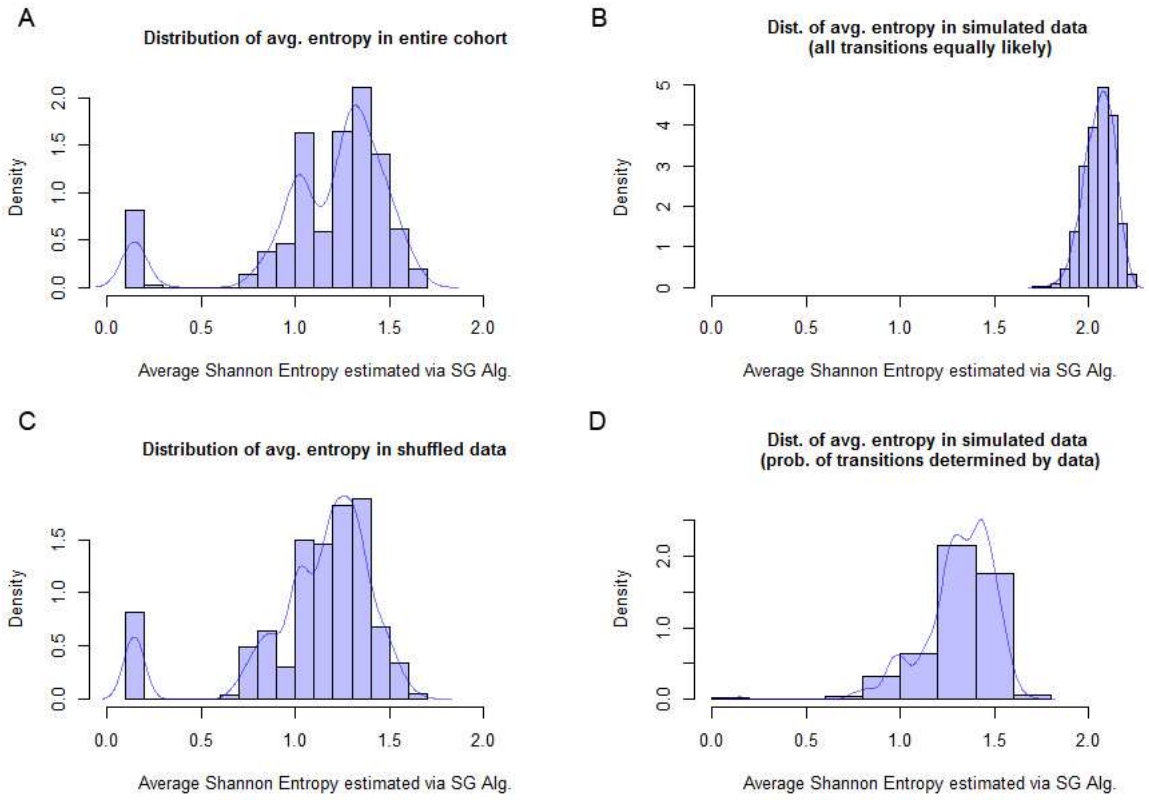

Figure E3 Histograms of estimated entropy in the cohort and in simulated and perturbed data sets taking into account the issue of unobserved events. A) Histogram of entropy of the rows of the Markov matrix for each participant in the cohort. Compared to Figure E2 A), the distribution is shifted towards higher entropy values. B) Histogram of the entropy of the rows of the Markov matrix in a simulated data set in which all symptom transitions are assumed to be equally probable. Compared to Figure E2 B), the entropy distribution is similar. C) Histogram of entropy of the rows of the Markov matrix for each participant after each participant's chronological order of reported symptoms was randomly permuted. Compared to Figure E2 C), a totally different distribution can be seen. D) Histogram of entropy of the rows of the Markov matrix in a simulated data set in which the probabilities of symptom transitions were determined by the pooled cohort data. Compared to Figure E2 D), a totally different distribution can be seen. The entropy in all four panels was estimated after applying the Schuerman and Grassberger (SG) algorithm.[15]

REFERENCES

1. Fuchs O, Latzin P, Kuehni CE, Frey U. Cohort profile: the Bern infant lung development cohort. *Int J Epidemiol* 2012; 41: 366-376.
2. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, Bisgaard H, Davis GM, Ducharme FM, Eigen H, Gappa M, Gaultier C, Gustafsson PM, Hall GL, Hantos Z, Healy MJ, Jones MH, Klug B, Lodrup Carlsen KC, McKenzie SA, Marchal F, Mayer OH, Merkus PJ, Morris MG, Oostveen E, Pillow JJ, Seddon PC, Silverman M, Sly PD, Stocks J, Tepper RS, Vilozni D, Wilson NM, American Thoracic Society/European Respiratory Society Working Group on I, Young Children Pulmonary Function T. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *Am J Respir Crit Care Med* 2007; 175: 1304-1345.
3. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, Initiative ERSGLF. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324-1343.
4. American Thoracic S, European Respiratory S. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med* 2005; 171: 912-930.
5. Asmussen S. Applied Probability and Queues, 2003.
6. Paninski L. Estimation of Entropy and Mutual Information. *Neural Computation* 2003; 15: 1191-1253.
7. Bishop CM. Pattern Recognition and Machine Learning. Springer, 2006.
8. Team RC. R: A Language and Environment for Statistical Computing. 2017 [cited 2017 01.10.2017]; Available from: http://www.R-project.org/.
9. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software* 2014; 61: 1-36.
10. Depner M, Fuchs O, Genuneit J, Karvonen AM, Hyvarinen A, Kaulek V, Roduit C, Weber J, Schaub B, Lauener R, Kabesch M, Pfefferle PI, Frey U, Pekkanen J, Dalphin JC, Riedler J, Braun-Fahrlander C, von Mutius E, Ege MJ, Group PS. Clinical and epidemiologic phenotypes of childhood asthma. *Am J Respir Crit Care Med* 2014; 189: 129-138.
11. Granell R, Henderson AJ, Sterne JA. Associations of wheezing phenotypes with late asthma outcomes in the Avon Longitudinal Study of Parents and Children: A population-based birth cohort. *J Allergy Clin Immunol* 2016; 138: 1060-1070 e1011.
12. Henderson J, Granell R, Heron J, Sherriff A, Simpson A, Woodcock A, Strachan DP, Shaheen SO, Sterne JA. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* 2008; 63: 974-980.
13. Dogaru CM, Strippoli MP, Spycher BD, Frey U, Beardsmore CS, Silverman M, Kuehni CE. Breastfeeding and lung function at school age: does maternal asthma modify the effect? *Am J Respir Crit Care Med* 2012; 185: 874-880.
14. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; 37: 360-363.
15. Schurmann T, Grassberger P. Entropy estimation of symbol sequences. *Chaos* 1996; 6: 414-427.
16. Jaynes ET. Probability Theory: The Logic of Science. Cambridge University Press, 2003.
17. Hausser J, Strimmer K. Estimation of Entropy, Mutual Information and Related Quantities. 2014 [cited 2017; Available from: https://CRAN.R-project.org/package=entropy