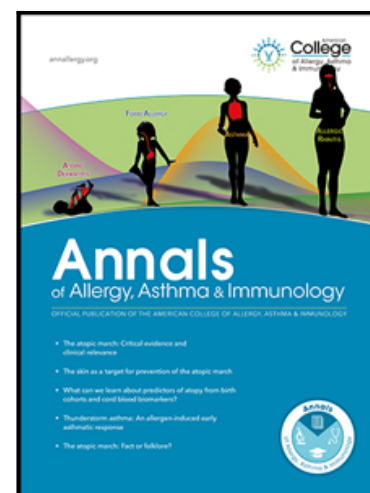


Accepted Manuscript

Asthma phenotypes based on health services utilization for allergic diseases in a province-wide birth cohort

Miceline Mésidor MPH , Andrea Benedetti PhD ,
Mariam El-Zein PhD , Dick Menzies MD , Marie-Élise Parent PhD ,
Marie-Claude Rousseau PhD

PII: S1081-1206(18)31199-2
DOI: <https://doi.org/10.1016/j.anai.2018.09.453>
Reference: ANAI 2696



To appear in: *Annals of Allergy, Asthma Immunology*

Received date: 21 April 2018
Revised date: 23 July 2018
Accepted date: 11 September 2018

Please cite this article as: Miceline Mésidor MPH , Andrea Benedetti PhD , Mariam El-Zein PhD , Dick Menzies MD , Marie-Élise Parent PhD , Marie-Claude Rousseau PhD , Asthma phenotypes based on health services utilization for allergic diseases in a province-wide birth cohort, *Annals of Allergy, Asthma Immunology* (2018), doi: <https://doi.org/10.1016/j.anai.2018.09.453>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Asthma phenotypes based on health services utilization for allergic diseases
in a province-wide birth cohort**

Miceline Mésidor, MPH^a, Andrea Benedetti, PhD^{b,c}, Mariam El-Zein, PhD^a, Dick Menzies, MD^{b,c}, Marie-Élise Parent, PhD^a, Marie-Claude Rousseau, PhD^a.

^a Epidemiology and Biostatistics Unit, INRS-Institut Armand-Frappier, Université du Québec, 531 boul. des Prairies, Laval (QC), Canada, H7V 1B7.

^b Respiratory Epidemiology and Clinical Research Unit, Research Institute of the McGill University Health Centre, 5252 boul. De Maisonneuve, Montreal (QC), Canada, H4A 3S5.

^c Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, 1020 Pine Ave. West, Montreal (QC), Canada, H3A 1A2.

Type of manuscript: Original article

Short title: Population-level asthma phenotypes

Corresponding author: Marie-Claude Rousseau
Epidemiology and Biostatistics Unit
INRS-Institut Armand-Frappier, Université du Québec
531 boul. des Prairies, Laval (QC), Canada, H7V 1B7
Email: marie-claude.rousseau@iaf.inrs.ca.

Telephone: 1-450-687-5010

Conflicts of interest: None

Funding source: This work was supported by an infrastructure grant from the Canada Foundation for Innovation & the Quebec Ministry of Education, Leisure and Sports (#12532), and research grants from the Canadian Institutes of Health Research (#MOP-97777, #MCH-97593) and Fonds de recherche du Québec-Santé (FRQS, #16227). When conducting the work described in this manuscript, A.B., M.E.P. and M.C.R. were recipients of Career Awards from FRQS. M.M. is recipient of scholarships from the Fondation Armand-Frappier (Beaulieu-Saucier scholarship) and the Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM).

Keywords: Administrative health databases; Allergic diseases; Asthma; Birth Cohort; Phenotypes

Abbreviations: QBCIH, Quebec Birth Cohort on Immunity and Health
RAMQ, Régie d'assurance maladie du Québec (Quebec Health Insurance Organization)
ICD, International Classification of Diseases

CAI, Commission d'accès à l'information (Access to Information Commission)

INRS, Institut national de la recherche scientifique (National Institute of Scientific Research)

ISQ, Institut de la statistique du Québec (Statistics Quebec)

FRQS, Fonds de recherche du Québec – Santé (Quebec Health Research Fund)

CCC, Cubic cluster criterion

SAS, Statistical Analysis System

Word count: 4,000 (Max 4,000)

Figures: 5

Tables: 2

Abstract

Background

Previous studies on asthma phenotypes were often conducted in selected clinical populations and overlooked changes over the life course.

Objective: In a population-based cohort, we identified asthma phenotypes based on utilization of health services for allergic diseases in three life periods and documented transitions between phenotypes across life periods.

Methods: In a cohort of 78,211 individuals born in 1974 in the province of Québec (Canada), we documented medical visits and hospitalizations for asthma and other allergic diseases until 1994. Phenotypes based on clusters of health services utilization in childhood (8-12 years), adolescence (13-17 years), and young adulthood (18-20 years) were identified using Ward's method among 9,989 (12.8%) subjects who had at least one health encounter for asthma during follow-up. Population level probabilities of transitioning between phenotypes were estimated in the full study population.

Results: In the subset with asthma, six phenotypes were identified during both childhood and young adulthood, and seven during adolescence. The most common phenotype was "no asthma or allergic diseases": 58% in childhood, 42% in adolescence, and 54% in adulthood. Second most common was the "mild asthma/no allergic diseases" phenotype, representing respectively 36%, 31%, and 21% in these three periods. In the study population, 87% of the subjects remained in the "no asthma" phenotype over the follow-up. The vast majority of subjects in the asthma phenotypes transitioned over time.

Conclusion: Our study uniquely contributes to a better understanding, at the population level, of the manifestations and transitions in asthma phenotypes over the life course.

INTRODUCTION

Approximately 235 million people suffer from asthma worldwide, and over 250,000 die from it yearly.¹ Disease onset is usually in childhood, but can also occur in adolescence or adulthood, whereby evolution over time varies greatly among affected individuals.²⁻⁴ Although a number of determinants of childhood asthma have been identified (e.g. genetic predisposition to atopic diseases, viral infections, eosinophilic inflammation, exposure to environmental allergens), most are still poorly defined.⁵ The difficulty in identifying specific risk factors could be related to the complexity of symptoms which are probably the result of different etiologies.⁶ In recent years, several sub-entities of asthma, called phenotypes, have been characterized to better understand etiology and natural history of the disease. Phenotype characteristics of asthma include exacerbations, eosinophilia, and also asthma related to obesity, exercise or infection.⁷

Studies aimed at describing asthma phenotypes in children were mainly based on cohort or cross-sectional designs, and most used questionnaires and objective measures. Only less than

half of the studies reviewed by Pinart et al. have considered the presence of other allergic diseases when identifying asthma phenotypes in children.⁸ Studies including both children and adolescents did not separately investigate these two subgroups.^{9,10} Transitions in asthma phenotypes over time were only addressed in two studies among children^{4,11} and one among adults.² So far, no study has focused on transitions in phenotypes from childhood to young adulthood.

Previous studies on asthma phenotypes were clinically oriented, informed with detailed clinical and physiological data, but restricted to selected subgroups of persons with asthma or considered at higher risk. As such, asthma phenotypes have not been well studied in samples representative of the general population. Here we propose an approach that is complementary to previous work on asthma phenotypes. It is relevant in an epidemiological perspective, since it provides population level, as opposed to clinic-based, evidence on the distribution and dynamics of asthma phenotypes over time.

In a large population-based cohort study, we aimed to: 1) identify asthma phenotypes during three periods of life (childhood, adolescence, young adulthood) based on the use of health services for asthma and other allergic diseases and 2) describe the probabilities of transitioning between phenotypes over the follow-up, from age 8 to 20.

METHODS

Study design

The Quebec Birth Cohort on Immunity and Health (QBCIH) is a retrospective cohort (n=81,496) established by linking a birth registry with administrative health databases. It

includes 90.5% of individuals born in the province of Quebec in 1974 at or after 32 weeks of gestation. Subjects were followed for their utilization of health services until 1994, at age 20. The cohort has been described in detail elsewhere.^{12,13} Pertinent data were extracted from the birth registry and the Quebec Health Insurance (“Régie d’assurance maladie du Québec”, RAMQ) billing databases of medical visits and hospitalizations. The latter data sources were available in an electronic format since 1983 and 1987, respectively. Since we aimed to compare phenotypes across periods and their transition probabilities over time, analyses were restricted to subjects with complete follow-up (n=78,211, 96% of the QBCIH). As seen in Figure 1, we excluded 328 subjects due to death and 2,957 due to loss to follow-up in provincial health coverage (Figure 1).

The subset of individuals (n=9,989) who had at least one asthma-related medical visit or hospitalization between 1983 and 1994 (12.8% of the QBCIH) was used to identify asthma phenotypes based on health services utilization (first objective), whereas the probabilities of transitioning between asthma phenotypes over time (second objective) were estimated in the study population (n=78,211).

Description of variables

The birth registry was used to obtain perinatal information (e.g., birth weight, gestational age at birth) and individual-level demographic characteristics (sex and parental age). Area of residence was defined for each subject using the postal code corresponding to the address of residence in 1991, available in the Healthcare Registration File. The second character of the postal code indicates the type of area based on the classification of postal delivery: rural or urban.¹⁴ Income was estimated based on the median family income in the 1991 Canadian census using the subjects’ 3-digit residential postal code in that year.

All variables on the utilization of health services related to asthma and allergic diseases were defined for three periods of life: childhood (8-12 years), adolescence (13-17 years) and young adulthood (18-20 years). For each period, four variables were created for asthma. First, the number of medical visits for asthma (International Classification of Diseases codes ICD-9: 493) describes the number of times a patient had a physician visit for asthma. Second, the number of days of hospitalization for asthma compiles the duration of hospitalization with a primary diagnosis of asthma. The dates of diagnosis and of birth were used to determine the subject's age at each health encounter. Third, we created one binary variable, indicating if the subject had health services for asthma during the period. Fourth, we created another binary variable indicating whether the subject was using health services for asthma for the first time during the period. For allergic diseases, medical visits and hospitalizations were considered together under the denomination of "health services". One variable indicating the total number of health services during the period was created for each of the following conditions: rhinitis (ICD-9: 477), eczema (ICD-9: 691, 692), urticaria (ICD-9: 708), and "other allergies" (ICD-9: 995, V07.1, V15.0) which encompassed desensitization to allergens, unspecified allergies, and personal history of allergies (including to food), but excluded allergic reactions to medicinal substances.

Ethics

The constitution of the cohort, access to data, and study were approved by the "Commission d'accès à l'information" of Quebec (CAI, Access to Information Commission) and the ethics committees of the National Institute of Scientific Research (INRS), the Institute of Statistics of Quebec (ISQ) and the Quebec Health Research Fund (FRQS).

Statistical analyses

Eight variables were considered for the identification of asthma phenotypes, among the subset with asthma (n=9,989): 1) an indicator of health services utilization for asthma during the period (Y/N), 2) an indicator of first health care encounter for asthma during the period (Y/N), 3) the number of asthma-related medical visits, 4) the days of asthma-related hospitalization, and 5) the number of health services for rhinitis, 6) for eczema, 7) for urticaria, and 8) for other allergies. All variables were treated as quantitative, and were standardized to ensure that they had the same weight in the analysis. Pearson correlations between variables were examined.

Ward's method, a hierarchical approach, was used to identify asthma phenotypes based on clusters of health services utilization for allergic diseases. At each generation of clusters, Ward's method measures the distance between the subjects to combine them. Plots of fit statistics such as the pseudo-F, cubic cluster criterion (CCC), and pseudo- t^2 were used to determine the number of groups. Each of these statistics measures different aspects of the classification and we used all of them to ensure a certain robustness to the number of clusters selected. The pseudo-F and the pseudo- t^2 measure the distance between clusters whereas the CCC helps to determine if the classification deviates from that which would be obtained by a uniform distribution of the scatter point of the individuals. Maximal values for CCC and pseudo-F were selected. For the pseudo- t^2 , the value after the maximum was considered as the optimal number of clusters.¹⁵ As a sensitivity analysis, analyses were repeated using K-means, a partitioning method based on between-group similarity.¹⁶ We considered the same criteria as for Ward's method to identify the number of groups.

In order to describe the probabilities of transitioning between phenotypes over the follow-up, we calculated, in the study population ($n=78,211$), the transition probabilities between phenotypes from childhood to adolescence, and to young adulthood. To simplify this analysis, we combined some of the phenotypes identified by Ward's method so that the same phenotypes were present across all periods. Thus for each period, subjects were regrouped into the four following classes: 1) no asthma; 2) mild asthma, no allergic diseases; 3) mild asthma and allergic diseases; 4) moderate or severe asthma. The first phenotype combined people who had no health services for asthma during the entire follow-up and people from the *asthma subgroup* who were classified by Ward's method in the *no asthma or allergic diseases* for that specific period. The second phenotype corresponded to *mild asthma, no allergic diseases*. The third phenotype was a combination of all phenotypes with *mild asthma* and some allergic disease. The last phenotype covered individuals with moderate or severe asthma, regardless of allergic diseases. All analyses were conducted using Statistical Analysis System (SAS) Enterprise Guide version 4.2 (SAS Institute Inc., Cary, North Carolina, USA). The procedures PROC CLUSTER and PROC FASTCLUS were used respectively for Ward's and K-means methods.

RESULTS

Characteristics of included and excluded subjects, in the study population and the subset with asthma, are presented in Table 1. In the study population, included subjects were generally similar to excluded ones according to all characteristics except for median family income and health services for allergic diseases. Among excluded subjects, as compared to included ones, median family income was higher and the proportion of subjects with health services for allergic diseases was lower. The latter is congruent with subjects being excluded due to incomplete follow-up, leading to an inability to capture all health services during the study period. Fifty-one

percent of subjects in the study population were male and 67% had a normal weight. Similar observations were made when comparing included and excluded subjects in the subset with asthma. Forty-six percent of the subjects with asthma were male and 68% had a normal birth weight.

For the subset with asthma, health services utilization for asthma and other allergic diseases in the three life periods is presented in Table 2. The proportion of subjects who had at least one medical visit or hospitalization for asthma was highest in adolescence compared to childhood and young adulthood. In childhood, 4,016 (40%) had at least one asthma-related medical visit or hospitalization. During adolescence, 5,449 individuals (55%) used health services for asthma, and 3,742 (38%) had their first asthma-related health service. During young adulthood, 4,445 (45%) subjects consulted for asthma, and 2,231 (22%) had their first health service for asthma. The proportion hospitalized for asthma was 0.3% in childhood, 3% in adolescence and 1% in young adulthood. Among those having been hospitalized for asthma, the average duration was 4.7 days in childhood, 5.4 days in adolescence and 3.8 days in young adulthood. The number of medical services and hospitalizations for rhinitis, eczema, urticaria, and other allergies was, on average, equivalent to one service or fewer in each period.

Phenotype identification

In the subset with asthma between 1983 and 2014, we identified six phenotypes for the childhood period (Figure 2), seven during adolescence (Figure 3), and six in young adulthood (Figure 4). As shown in Figure 2, asthma phenotypes identified in childhood were: 1) *No asthma, no allergic diseases*: This is the largest group with 5,748 subjects (58%), whereby most did not

use health services for asthma or allergic diseases from ages 8-12; 2) *Mild asthma, without allergic diseases*: Almost all of the 3,589 subjects (36% of subset with asthma) had asthma-related medical visits or hospitalizations receiving, on average, four health services for asthma; 3) *Mild asthma with urticaria*: Among these 318 subjects (3% of subset with asthma), there was an average of one health service for urticaria and two medical visits for asthma, without hospitalization; 4) *Mild asthma, rhinitis, and severe other allergies*: These 176 persons (2% of subset with asthma) received 39 health services on average for other allergies. They also presented for asthma- and rhinitis-related medical visits (four services on average for each); 5) *Severe asthma with rhinitis and other allergies*: Only 107 persons (1% of subset with asthma) have used several health services for asthma, comprising on average 58 medical visits and one day of hospitalization. They have also received an average of two health services for rhinitis and other allergies; 6) *Mild asthma with severe rhinitis and other allergies*: These 51 subjects had an average of six medical visits for asthma, 82 services for rhinitis, and three health services for other allergies.

Phenotypes in adolescence (Figure 3) and young adulthood (Figure 4) presented some similarities and differences with those found in childhood. Specifically, we found three phenotypes that were similar for all three periods: 1) *No asthma, no allergic diseases*; 2) *Mild asthma without allergic diseases*; 3) *Mild asthma with urticaria*. The proportion of subjects in the *no asthma, no allergic diseases* phenotype was smaller in adolescence (42% in adolescence versus 58% in childhood and 54% in young adulthood). Some differences were observed between the three periods. Twenty-one percent of adolescents were in the *mild asthma, rhinitis, other allergies and eczema* phenotype, a phenotype not identified in childhood. Only 0.3% of the subjects with asthma were classified in the *severe asthma and rhinitis* phenotype at adolescence,

as opposed to 1% in childhood. For the young adulthood period, we identified two phenotypes that were not present in the previous periods: *mild asthma with severe rhinitis and severe other allergies* (1.4% of the subjects with asthma) and *severe asthma without allergic diseases* (0.7% of the subjects with asthma).

When applying K-means method, similar phenotypes were found [eFigures 1 (childhood), 2 (adolescence), 3 (young adulthood)] and the proportion of subjects in each phenotype was very similar, except for the adolescence period. In this period, we found nine phenotypes with K-means, compared with seven when applying Ward's method. With the latter, we found that 21% of subjects were grouped in *mild asthma, rhinitis, other allergies and eczema*, whereas K-means method classified some of these subjects in the *mild asthma (late onset) without allergic diseases* and others in smaller separate phenotypes.

Phenotype transitions

The probabilities of transitioning between asthma phenotypes from childhood to young adulthood in the study population are presented in Figure 5. In childhood, 95% of subjects were classified as asthma-free, most of whom remained in this phenotype over life. Indeed, 87% of the study population (n=68,238) was classified in the *no asthma* phenotype during the entire follow-up. Overall, in phenotypes other than *no asthma*, the majority of subjects changed phenotype over time. Among subjects who were in the *mild asthma without allergic diseases*, *mild asthma with allergic diseases*, and *moderate/severe asthma* phenotypes in childhood, the probabilities of transitioning to another phenotype at adolescence were 1.0, 0.6, and 0.9 respectively. Among those who were in these phenotypes at adolescence, the probabilities of transitioning to another phenotype in young adulthood were respectively 0.7, 0.9, and >0.9 (>26 out of 30 subjects, exact probability cannot be stated to preserve confidentiality). Nonetheless, some subjects remained in

the same phenotype over time, notably: 1) 245 subjects (0.3% of the study population) who were in the *mild asthma with allergic diseases* and 10 subjects (0.01% of the study population) in the *moderate/severe asthma* phenotypes in childhood remained in these phenotypes at adolescence; 2) 780 subjects (1% of the study population) in the *mild asthma without allergic diseases*, 182 subjects (0.2% of the study population) in the *mild asthma with allergic diseases*, and less than 5 subjects (0.0% of the study population) in the *moderate/severe asthma* phenotypes during adolescence remained in these respective phenotypes in young adulthood.

Subjects with moderate or severe asthma during childhood had a 79% probability of having mild asthma with allergic diseases during adolescence; and similarly most of those who had moderate or severe asthma during adolescence (73%) had mild asthma, with allergic diseases in young adulthood.

As a sensitivity analysis, we calculated the transition probabilities to the phenotypes identified with K-means. Some transition probabilities between phenotypes differed from those based on Ward's method (eFigure 4). For instance, subjects with moderate/severe asthma in childhood had a higher probability of staying in the same phenotype over life: 36% with K-means versus 9% with Ward's method in adolescence, whereas for young adulthood these values were 29% versus less than 20% respectively.

DISCUSSION

Our first objective was to identify asthma phenotypes based on the utilization of health services for asthma and allergic diseases during three periods of life. Six phenotypes were identified in childhood and adulthood, and seven in adolescence. In the subset with asthma, the

asthma and allergic disease-free and *mild asthma without allergic diseases* phenotypes were the most common, representing more than 70% of those with asthma in every period. These results are in line with those from other investigators who identified two main asthma phenotypes using clinical characteristics: 1) people without asthma and no allergic diseases, 2) people with asthma and no allergic diseases.^{2,3,10,17} In each period of life, we found a phenotype of asthma with urticaria. Few studies have focused on urticaria in childhood.^{8,18,19} Ours seems to be the first to identify this phenotype in adolescence and young adulthood.

Our second objective was to describe the population-based probabilities of transitioning between phenotypes over the follow-up. To our knowledge, only three studies have addressed transitions in asthma phenotypes over time. Soto-Ramirez et al. used latent transition analysis to identify transitions in asthma phenotypes based on physician diagnosed asthma and wheeze in a birth cohort (n=1,456) followed at 4 time points from age 1 to 18 years.⁴ Garden et al. examined transitions in asthma and asthma-related traits among 370 children, following them from years 1.5 to 11.5 of age.¹¹ Boudier et al. used latent transition analysis to identify transitions in asthma phenotypes among 3,320 adults who reported having asthma at baseline or during a ten-year follow-up.²

In our population-based study, we observed that subjects without asthma in childhood (8-12 years) had more than a 90% probability of staying in this phenotype until age 20. Among children without asthma or wheeze at 1-2 years, Soto-Ramirez et al. reported that approximately 60% remained free of asthma and wheeze at 18 years.⁴ In children at risk of asthma, Garden et al. observed only a 35% probability for those in the *nonatopic, few symptoms* at 1.5 years to remain in this phenotype until 11.5 years.¹³ Adults in the *nonallergic, few symptoms, no treatment* phenotype at baseline had a 78% probability of expressing the same phenotype after 10 years,

according to Boudier et al.² These probabilities differ greatly, but this is somewhat expected given the differences in phenotype definitions, populations, and follow-up. Our results are complementary to those already published, and may provide a more accurate picture of population level transitions in asthma phenotypes than those observed in children¹¹ or adults² at risk of asthma.

In our study, the majority of subjects who were classified in a phenotype other than *no asthma* changed phenotype over time. Garden et al. found that children between ages 1.5 and 3 years mostly transited from one phenotype to another. They also concluded that most of the children in the *nonatopic, asthma and rhinitis symptoms* in early childhood transited to the *nonatopic, few symptoms* phenotype.² Soto-Ramirez et al. observed that the severity of asthma increased until age 10 then decreased thereafter.⁴ Similarly, we found that subjects with *moderate/severe asthma* in childhood most often had *mild asthma* during adolescence and young adulthood. The transition probabilities obtained in our study showed some stability over time. Indeed, we observed similar patterns across the two transition periods (from childhood to adolescence and from adolescence to young adulthood). As suggested by Boudier and al., allergies may play a role in the evolution of asthma.²

Many studies that identify asthma phenotypes have focused on clinical characteristics, physiological measurements, and environmental allergies.^{9,20-22} Such clinical and physiological data are likely the best information to consider when studying asthma phenotypes, but it is only feasible to collect them in small populations considering their cost. Indeed, it would not be feasible to use these data at the population level, except if comprehensive national electronic health records were available, which is not the case in Canada and in most countries. Despite the lack of relevant clinical and physiological characteristics (e.g., family history of asthma), our

approach was innovative in that it enabled us to generate a true population-based description of asthma phenotypes. It allowed us to characterize asthma phenotypes for an entire birth cohort, from childhood until young adulthood, by capitalizing on administrative health data gathered in a public health system covering the vast majority of the population (96% in 2011-2012).^{23,24}

The choice of clustering method and of the variables considered in the identification of phenotypes may have influenced the observed results. We have chosen Ward's method, a clustering approach, because we wanted to define empirical phenotypes based on the data without using a priori definitions. We also conducted the analyses with K-means, and presented the results as supplemental material. Some results differed between the two methods. For example, subjects with *moderate/severe asthma* in childhood had a higher probability of remaining in the same phenotype over time with K-means than Ward's method. These differences may partly be due to our analyses not having considered uncertainty when estimating the probabilities. Consequently, if we were able to calculate the confidence intervals of the transition probabilities of the phenotypes, it is possible that they may be very wide or overlap, which would suggest no statistical difference between the values.

One of the main strengths of this study is the use of data from a large population-based birth cohort with a long follow-up. This birth cohort included most people born in the province of Quebec in 1974 at or after 32 weeks of gestation (90%).¹² These results represent the use of health services for asthma and allergic diseases for this entire birth cohort, and can thus be complementary, from a public health perspective, to phenotypes derived from highly selected clinical populations. This clearly enhances the external validity of such a phenotypic approach to characterizing asthma. Since we used a highly representative birth cohort in a country with universal health care, these results are generalizable to other industrialized countries with a

similar coverage. The large sample size compared to the number of variables used in the determination of phenotypes represents another strength.²⁵

Some limitations should be considered when interpreting these findings. The phenotypes were based on administrative health data generated from the public health system, thus they can be subject to some measurement errors. Further, these data did not allow identifying persons who never consulted for their wheezing or asthma. However, for Quebec patients 11-44 years-old, when comparing an asthma definition of one or more medical services for asthma over a one year period with medical records, sensitivity was 0.97 (95% CI: 0.94-1.00) and specificity 0.85 (95% CI: 0.78-0.94).²⁶

Administrative health data did not allow us to differentiate wheezing due to infectious bronchitis or respiratory infections from asthma. Thereby, some of the transitions observed could be explained by transient wheezing. Also, some variables were not as detailed as we would have liked. For instance, we could not distinguish urticaria subtypes or separate food allergies from other allergies, aspects that should be investigated further. Data on medical services and hospitalizations were not available before 1983 and 1987 respectively, and information on medications was not available. Consequently, asthma-related health care encounters before age 8 could not be considered in our study, which may have resulted in underestimating the proportion of subjects with asthma during childhood and may have led to some data skewing. This lack of coverage impacts our ability to conclude about the entire history of health care utilization, but in no way invalidates our conclusions or interpretations from age 8 to 20 years.

Our study is among the few ones which characterized asthma phenotypes and described their dynamics over time at the population level. Based on administrative health data, these results are generalizable to the population and can therefore greatly contribute to a better

understanding of the manifestations and transitions in this disease over time, which is relevant for public health practitioners, health researchers, as well as clinicians. Moreover, the identification of factors associated with the worsening of asthma can help clinicians achieve better prevention of disease progression. Indeed, our findings can be a guide for building prediction models in order to improve clinical management of asthma.

Acknowledgements

We gratefully acknowledge Isabelle Leroux and Jimmy Baulne from Statistics Quebec, as well as François Blouin from the Régie de l'assurance maladie du Québec for their contribution to various aspects of the establishment of the Quebec Birth Cohort on Immunity and Health. We are also grateful to Florence Conus from INRS-Institut Armand-Frappier for data management, data documentation, facilitating data access, and providing technical support. These results are based on data from the © Gouvernement du Québec, Institut de la statistique du Québec (ISQ). ISQ is not responsible for compilations or interpretation of results.

REFERENCES

1. de Nijs SB, Venekamp LN, Bel EH. Adult-onset asthma: is it really different? *Eur Respir Rev.* 2013;22(127):44–52.
2. Boudier A, Curjuric I, Basagaña X, et al. Ten-year follow-up of cluster-based asthma phenotypes in adults. A pooled analysis of three cohorts. *Am J Respir Crit Care Med.* 2013;188(5):550–60.
3. Shaaban R, Zureik M, Soussan D, et al. Rhinitis and onset of asthma: a longitudinal population-based study. *Lancet.* 2008;372(9643):1049–57.
4. Soto-Ramírez N, Ziyab AH, Karmaus W, et al. Epidemiologic methods of assessing asthma and wheezing episodes in longitudinal studies: measures of change and stability. *J Epidemiol.* 2013;23(6):399–410.

5. Scherzer R., Grayson MH. Heterogeneity and the origins of asthma. *Ann. Allergy Asthma Immunol.* 2018, doi: 10.1016/j.anai.2018.06.009.
6. Wenzel S. Severe asthma: from characteristics to phenotypes to endotypes. *Clin Exp Allergy.* 2012;42(5):650–8.
7. Bostantzoglou C, Delimpoura V, Samitas K, et al. Clinical asthma phenotypes in the real world: opportunities and challenges. *Breathe.* 2015; 11: 186–93.
8. Pinart M, Albang R, Maier D, et al. Systematic Review on the Definition of Allergic Diseases in Children: The MeDALL Study. *Int Arch Allergy Immunol.* 2015;168(2):110–21.
9. Fitzpatrick AM, Teague WG, Meyers DA, et al. Heterogeneity of severe asthma in childhood: confirmation by cluster analysis of children in the National Institutes of Health/National Heart, Lung, and Blood Institute Severe Asthma Research Program. *J Allergy Clin Immunol.* 2011;127(2):382–389-13.
10. Kelley CF, Mannino DM, Homa DM, Savage-Brown A, Holguin F. Asthma phenotypes, risk factors, and measures of severity in a national sample of US children. *Pediatrics.* 2005;115(3):726–31.
11. Garden FL, Simpson JM, Mellis CM, Marks GB, CAPS Investigators. Change in the manifestations of asthma and asthma-related traits in childhood: a latent transition analysis. *Eur Respir J.* 2016;47(2):499–509.

12. Rousseau M-C, El-Zein M, Conus F, Benedetti A, Parent M-É. Cohort Profile: The Québec Birth Cohort on Immunity and Health (QBCIH). In Press. *Int J Epidemiol*.
13. El-Zein M, Conus F, Benedetti A, Parent M-E, Rousseau M-C. Evaluating the Validity of a Two-stage Sample in a Birth Cohort Established from Administrative Databases. *Epidemiology*. 2016;27(1):105–15.
14. Canada Post. Addressing Guidelines. 2014. <https://www.canadapost.ca/tools/pg/manual/PGaddress-e.pdf>. Accessed 21 June 2015.
15. Ward JH. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc*. 1963;58(301):236–44.
16. Tan PN, Steinbach M, Kumar V. Cluster analysis: Basic concepts and algorithms. In: Tan PN, Steinbach M, Kumar V. Introduction to data mining. Boston : Addison-Wesley Longman Publishing Co. 2006. pp. 495–7.
17. Simpson A, Tan VYF, Winn J, et al. Beyond atopy: multiple patterns of sensitization in relation to asthma in a birth cohort study. *Am J Respir Crit Care Med*. 2010;181(11):1200–6.
18. Reddy MB, Covar RA. Asthma phenotypes in childhood. *Curr Opin Allergy Clin Immunol*. 2016; 16(2):127–134.
19. Ban GY, Kim MA, Yoo HS, Ye YM, Park HS. Letter to the Editor: Two Major Phenotypes of Sulfite Hypersensitivity: Asthma and Urticaria. *Yonsei Med J*. 2014; 55(2), 542–44.
20. Garden FL, Simpson JM, Marks GB, CAPS Investigators. Atopy phenotypes in the Childhood Asthma Prevention Study (CAPS) cohort and the relationship with allergic disease: clinical mechanisms in allergic disease. *Clin Exp Allergy*. 2013;43(6):633–41

21. Siroux V, Basagaña X, Boudier A, et al. Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J*. 2011;38(2):310–7.
22. Haldar P, Pavord ID, Shaw DE, et al. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med*. 2008;178(3):218–24.
23. Régie de l'assurance maladie du Québec. Rapport annuel de gestion 2011-2012. 2012.
<http://www.ramq.gouv.qc.ca/SiteCollectionDocuments/citoyens/fr/rapports/rappann1112-fr.pdf>. Accessed 20 June 2015.
24. Statistics Canada. Table 051-0001 - Estimates of population, by age group and sex for July 1, Canada, provinces and territories. 2014. <http://www5.statcan.gc.ca/cansim/a26?id=510001>. Accessed 10 Jul, 2015.
25. Dolnicar S. A Review of Unquestioned Standards in Using Cluster Analysis for Data-Driven Market Segmentation. Faculty of Commerce - Papers 2002. <http://ro.uow.edu.au/commpapers/273>. Accessed 10 Jul, 2015.
26. Blais L, Lemièrre C, Menzies D, Berbiche D. Validity of asthma diagnoses recorded in the Medical Services database of Quebec. *Pharmacoepidemiol Drug Saf*. 2006;15(4):245–52.

Figure 1 The initial population consisted of 91,675 individuals born in Quebec in 1974 according to Statistics Quebec. Among the 81,496 eligible persons, 3,285 died or were lost to follow-up.

Figure 2 The “No asthma, no allergic diseases” subgroup had an average of <1 health service for these conditions during the period.

Figure 3 The “No asthma, no allergic diseases” subgroup had an average of <1 health service for these conditions during the period.

Figure 4 The “No asthma, no allergic diseases” subgroup had an average of <1 health service for these conditions during the period.

Figure 5 The content of the boxes refers to observed phenotypes and proportions of the analytical sample in each phenotype. The number on top of each arrow is the probability of transitioning.

Table 1 Characteristics of Included and Excluded Subjects in the Study Population and the Subset with Asthma, the Quebec Birth Cohort on Immunity and Health (n=81,496), 1974-1994

Characteristics	Study population	Subset with asthma
-----------------	------------------	--------------------

	Included n=78,211 ^a n (%)	Excluded n=3,285 ^a n (%)	Included n=9,989 ^a n (%)	Excluded n=339 ^a n (%)
Sex				
Male	39,976 (51.1)	1,711 (52.1)	4,595 (46.0)	161 (47.5)
Female	38,230 (48.9)	1,574 (47.9)	5,392 (54.0)	178 (52.5)
Birth Weight				
Very low (<1,500 g)	11,964 (15.3)	486 (14.8)	1,641 (16.4)	56 (16.5)
Low (1,500-2,499 g)	6,523 (8.3)	256 (7.8)	812 (8.1)	27 (8.0)
Normal (2,500-4,200 g)	52,619 (67.3)	2,195 (66.8)	6,707 (67.2)	217 (64.0)
High (≥4,200g)	7,088 (9.1)	348 (10.6)	827 (8.3)	39 (11.5)
Maternal age at childbirth (years)				
14-24	28,229 (36.7)	1,290 (40.1)	3,794 (38.8)	117 (35.9)
25-34	43,850 (56.9)	1,790 (55.7)	5,462 (55.8)	201 (61.7)
≥ 35	4,933 (6.4)	135 (4.2)	532 (5.4)	8 (2.4)
Paternal age at childbirth (years)				
14-24	13,233 (17.6)	665 (21.4)	1,792 (18.9)	66 (21.1)
25-34	50,748 (67.6)	2,052 (65.2)	6,423 (67.6)	215 (68.7)
≥ 35	11,147 (14.8)	384 (12.4)	1,287 (13.5)	32 (10.2)
Area of residence^b				

Rural	22,692 (29.7)	529 (24.0)	2,264 (23.1)	55 (20.2)
Urban	53,842 (70.4)	1,680 (76.1)	7,544 (76.9)	218 (79.9)
Median family income (\$/year)^c				
< 35,333	19,620 (25.8)	429 (20.6)	2,135 (21.9)	44 (17.4)
35,334-40,667	18,464 (24.3)	460 (22.0)	2,245 (23.1)	47 (18.6)
40,668-47,289	19,048 (25.1)	451 (21.6)	2,653 (27.3)	63 (24.9)
≥ 47,290	18,894 (24.9)	747 (35.8)	2,693 (27.7)	99 (39.1)
≥1 health service ^d for rhinitis	9,827 (12.6)	248 (7.6)	2,792 (28.0)	79 (23.3)
≥1 health service ^d for eczema	20,552 (26.3)	534 (16.3)	3,360 (33.6)	103 (30.4)
≥1 health service ^d for urticaria	4,943 (6.3)	113 (3.4)	987 (9.9)	24 (7.1)
≥1 health service ^d for other allergies	9,296 (11.9)	219 (6.7)	2,405 (24.1)	69 (20.4)

^a Due to missing values, the number of subjects for every variable does not necessarily equate to the total number of subjects.

^b Determined using the second character of subjects' residential postal codes (0: rural, ≠0: urban).

^c From the 1991 Canadian census using the first three characters of subjects' residential postal codes.

^d Either medical visits or hospitalizations.

Table 2 Use of Health Services for Asthma and Other Allergic Diseases by Life Period among the Subset with Asthma, the Quebec Birth Cohort on Immunity and Health (n=9,989), 1974-1994

Variables	Life Period		
	Childhood 8-12 years	Adolescence 13-17 years	Young adulthood 18-20 years
Asthma, n (%)^a	4,016 (40.2)	5,449 (54.6)	4,445 (44.5)
≥1 medical visit for asthma, n (%)	4,015 (40.2)	5,441 (54.5)	4,443 (44.5)
≥1 hospitalizations for asthma, n (%)	29 (0.3)	307 (3.1)	107 (1.1)
First health service for asthma, n (%)^b	4,016 (40.2)	3,742 (37.5)	2,231 (22.3)
Number of health services, mean (SD)			
Medical visits for asthma	2.19 (8.09)	2.71 (7.57)	1.68 (4.25)
Medical visits for asthma (among those with ≥1 visit)	5.44 (12.05)	4.97 (9.69)	3.79 (5.72)
Days of hospitalizations for asthma	0.01 (0.35)	0.17 (1.39)	0.04 (0.54)
Days of hospitalizations for asthma (among those with ≥1 day)	4.69 (4.45)	5.38 (5.87)	3.78 (3.61)
Health services ^c for rhinitis	0.94 (6.93)	0.91 (6.17)	0.31 (2.84)
Health services ^c for eczema	0.17 (0.56)	0.31 (1.18)	0.23 (1.51)
Health services ^c for urticaria	0.04 (0.31)	0.07 (0.37)	0.04 (0.26)
Health services ^c for other allergies ^d	1.06 (6.57)	1.08 (7.21)	0.44 (3.58)

SD: standard deviation

^a At least one medical visit or one hospitalization for asthma in the period.

^b Subjects can have health services for asthma and allergic diseases in any or in all of the periods. However, by definition, they will have their first health service for asthma only once. Thus, the numbers for the first health service for asthma sum up to 9,989 over the three periods, whereas the totals for the other lines are not expected to sum to the total number of subjects.

^c Either medical visits or hospitalizations.

^d Other allergies include desensitization to allergens and unspecified allergies, but exclude allergic reactions to a medicinal substance.

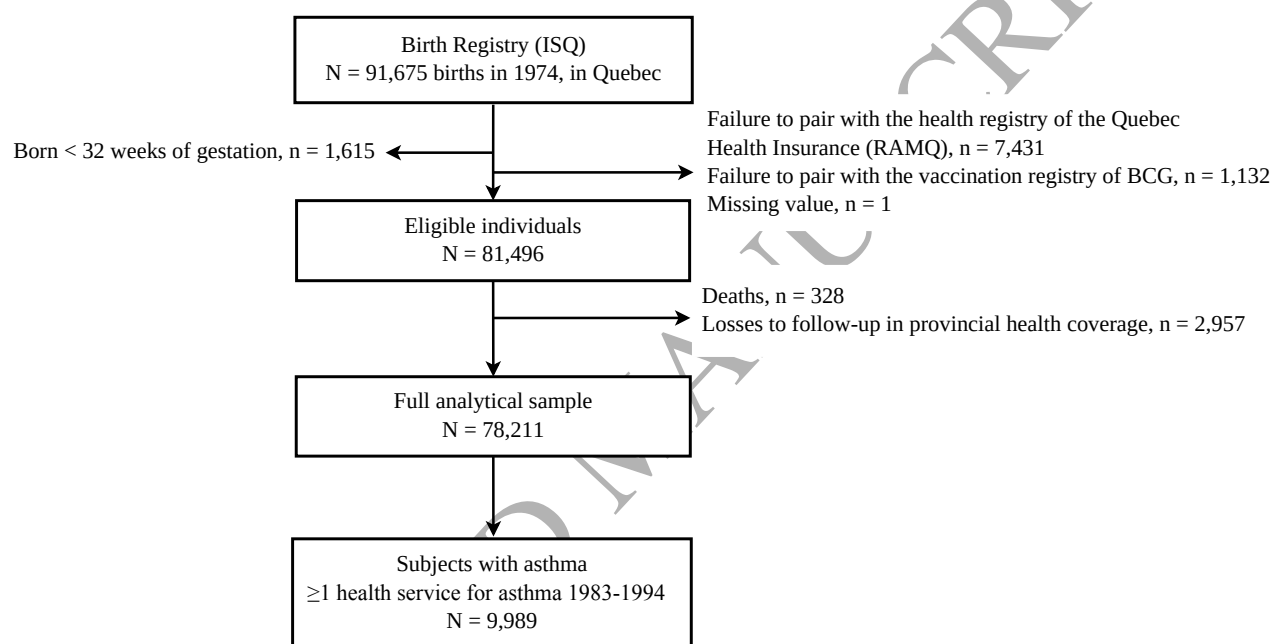


Figure 1 Establishment of the analytical samples from the Quebec Birth Cohort on Immunity and Health, Quebec, Canada, 1974-1994

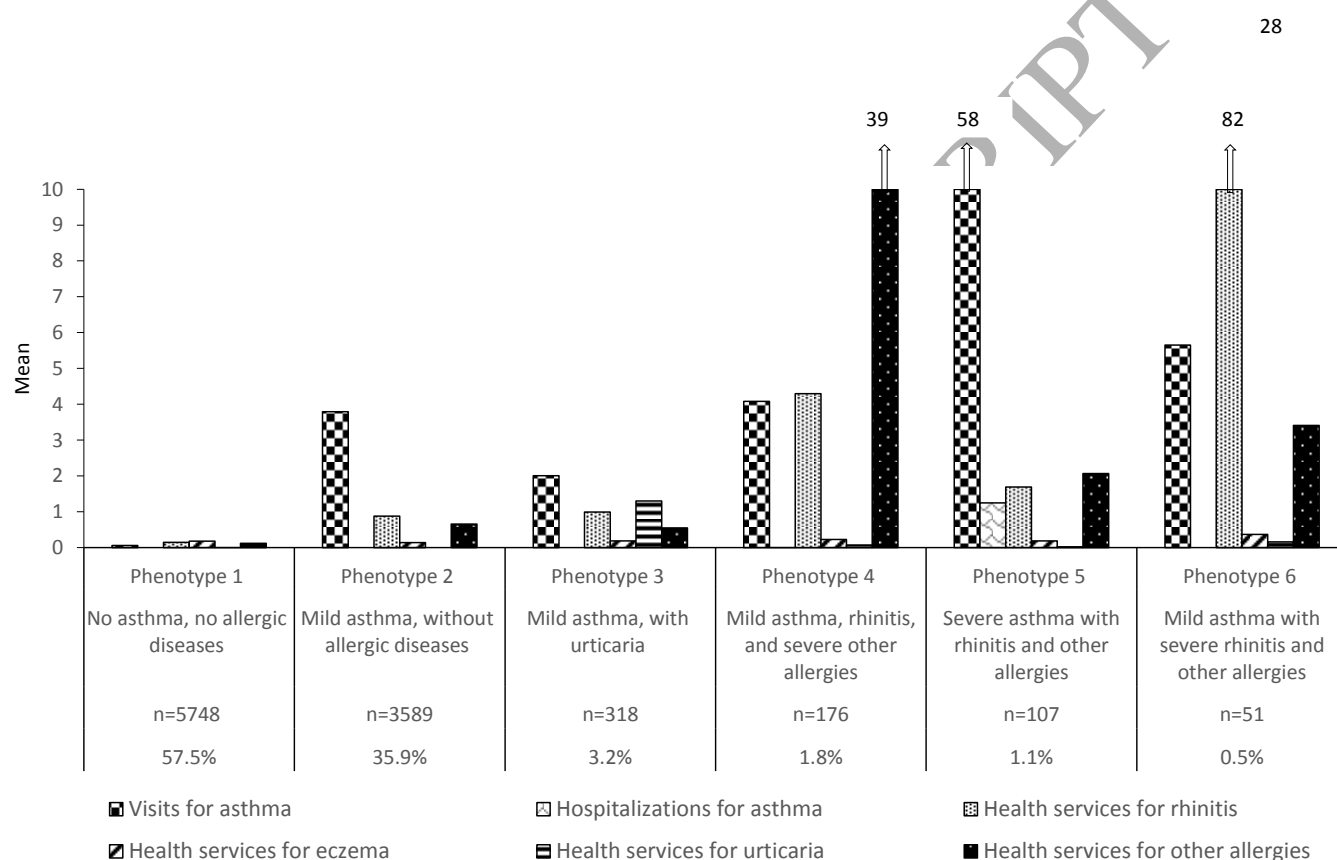


Figure 2 Asthma Phenotypes Identified by Ward's Method According to the Utilization of Health Services in Childhood (8-12 years) in the Subset with Asthma, the Quebec Birth Cohort on Immunity and Health (n=9,989), 1974-1994

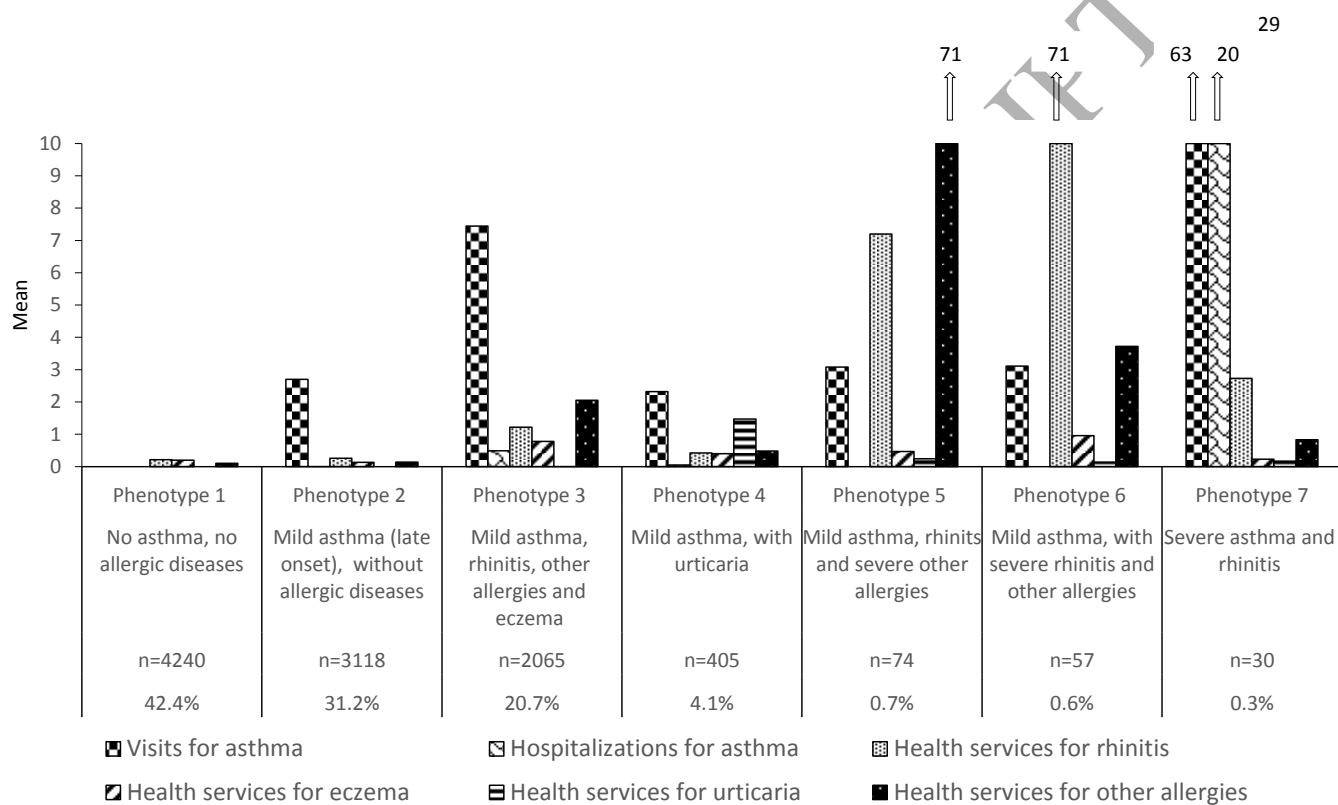


Figure 3 Asthma Phenotypes Identified by Ward's Method According to the Utilization of Health Services in Adolescence (13-17 years) in the Subset with Asthma, the Quebec Birth Cohort on Immunity and Health (n=9,989), 1974-1994

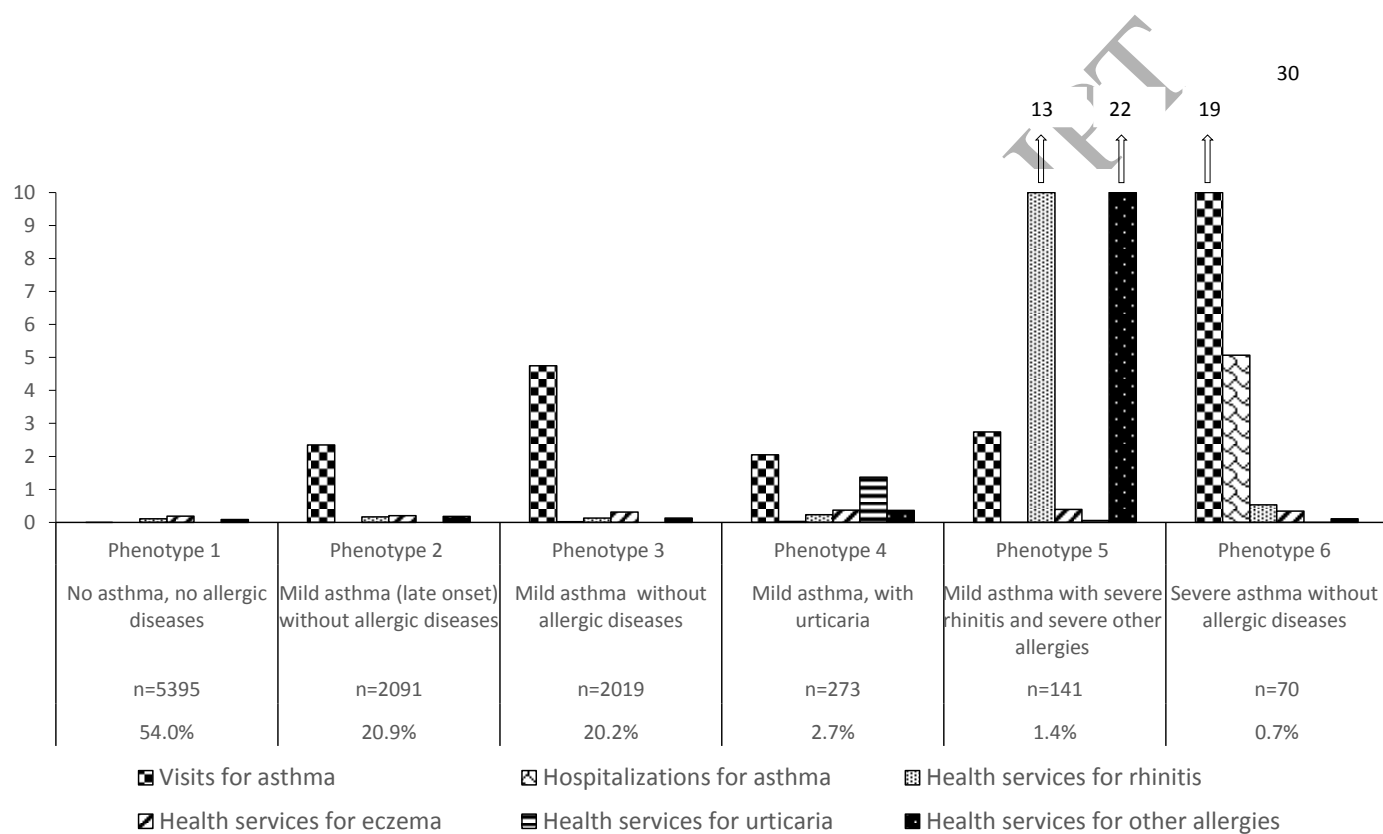


Figure 4 Asthma Phenotypes Identified by Ward's Method According to the Utilization of Health Services in Young Adulthood (18-20 years) in the Subset with Asthma, the Quebec Birth Cohort on Immunity and Health (n=9,989), 1974-1994

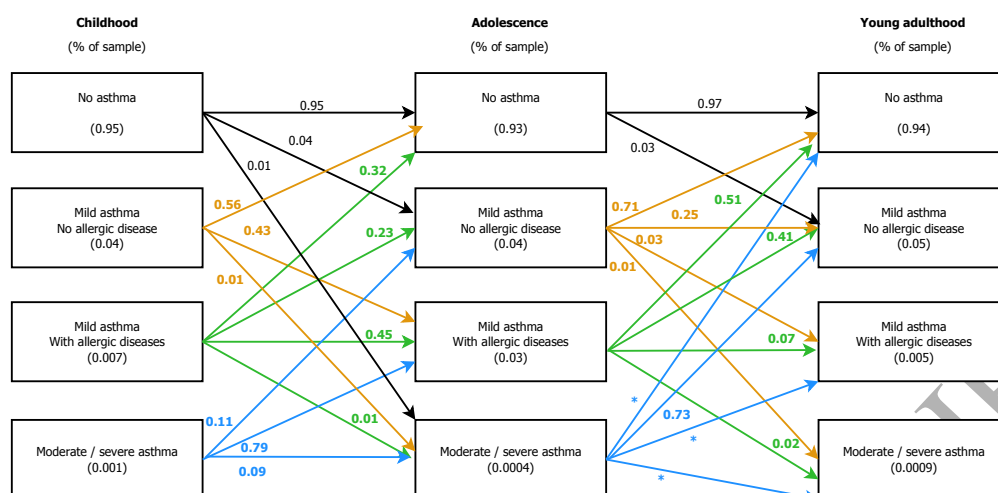


Figure 5 Transition probabilities between asthma phenotypes from childhood to young adulthood, identified from Ward's method, the Quebec Birth Cohort on Immunity and Health, (n=78,211), Quebec, Canada, 1974-1994