

Original article

A novel study design to investigate the early-life origins of asthma in children (SAGE study)

This is a description of the Study of Asthma, Genes and the Environment (SAGE), a novel birth cohort created from provincial healthcare administrative records. It is a general population-based cohort, composed of children at high and low risk for asthma, living in urban and rural environments in Manitoba, Canada. The SAGE study captures the complete longitudinal healthcare records of children born in 1995 and contains detailed information on early-life exposures, such as antibiotic utilization and immunization, in relationship to the development of asthma. Nested within the birth cohort is a case-control study, which was created to collect information on home environmental exposures from detailed surveys and home dust sampling, to confirm asthma status in children and use this data to validate healthcare database measures of asthma, to determine differences in immune system responsiveness to innate and adaptive immune stimuli in asthma, to genotype children for genes likely associated with the development of asthma and to study the epigenetic regulation of pre-established protective vs allergic immune responses. The SAGE study is a multidisciplinary collaboration of researchers from pediatric allergy, population health, immunology, and genetic and environmental epidemiology. As such, it serves as a fertile, interdisciplinary training ground for graduate students, and postdoctoral and clinician fellows.

**A. L. Kozyrskyj^{1,2,3}, K. T. HayGlass⁴,
A. J. Sandford⁵, P. D. Paré⁵,
M. Chan-Yeung⁶, A. B. Becker³**

¹Department of Pediatrics, University of Alberta, Edmonton, AB; ²Department of Community Health Sciences, Manitoba Centre for Health Policy;

³Department of Pediatrics and Child Health;

⁴Department of Immunology, Faculty of Medicine, University of Manitoba, Winnipeg, MB; ⁵James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research, St Paul's Hospital;

⁶Occupational and Environmental Lung Diseases Unit, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

Key words: asthma; children; cohort; database records.

Anita Kozyrskyj
8226a Aberhart Centre
11 402 University Ave
Edmonton, Alberta
Canada

Accepted for publication 7 October 2008

Many insights into the origins of asthma and its phenotypes have been gained from birth cohort studies (1, 2). However, there are many challenges in executing studies which involve longitudinal data collection. As such, there is enormous diversity in the study design of existing birth cohorts of asthma and allergy (3). Some studies, such as the Canadian Asthma Primary Prevention Study (4), the Boston Home Allergens and Asthma Study (5) and the Madison Childhood Origins of Asthma Study (6) are naturalistic or intervention studies of high risk cohorts, which evaluate the offspring of parents with a history of atopy. Although family history is highly predictive of asthma, the majority of incident cases of asthma occur in children without a family history (7). Thus, there are cohorts which also recruit samples of children at low risk for asthma, such as the Manchester Asthma and Allergy Study (8) and the German Multicentre Asthma Study (9), as well as several general population cohort studies of all children born in a geographic region. The latter include, among others, the pregnancy cohort in northern Finland (10), the multidisciplinary cohort in Dunedin (11), the Tucson Children's Respiratory Study (12), the Childhood

Allergy Study in Detroit (13), the Isle of Wight cohort (14), the Western Australian Pregnancy Cohort (Raine) Study (15), and the Avon Longitudinal Study of Parents and Children (16). Questionnaires such as the ISAAC instrument have been the main source of data on asthma and allergy outcomes, as well as environmental exposures. Some studies have included more objective measures, such as skin prick tests, lung function assessments and house dust sampling. The follow-up rate in these studies has been as high as 94% and as low as 57%.

There exists another genre of the birth cohort study which is not dependent on primary data collection. This is the retrospective cohort study which can be implemented by linking healthcare records for whole populations from birth until date of disease outcome. These population-based healthcare databases can be found in Canada, the UK, Sweden, West Australia and the Netherlands (17–21). In Canada, they are the by-product of the universal care insurance system and capture records of all healthcare contacts for all residents in each province. Typically, databases of physician visits, hospitalizations and prescription medications have been linked to report on

disease prevalence, healthcare utilization and medication effectiveness or adverse events. Few have been used to create birth cohorts which follow children over time. Herein, we describe the creation of a birth cohort and nested case-control study, which were derived from provincial healthcare administrative records in Manitoba, Canada to study the origins of childhood asthma.

Overview of study design

The Study of Asthma, Genes and the Environment (SAGE) is a birth cohort study of 16 320 children born in Manitoba, Canada in 1995, which was created from the provincial healthcare registry. Children who died or moved out of province over the next 7 years were excluded from the cohort ($n = 2340$). In 2002, a one-page survey was mailed via Manitoba Health to the households of the cohort children to inquire about their health and home environment exposures. First Nations (also known as Canadian Aboriginal) communities were approached separately. Parents were asked for permission to link the survey data to their child's healthcare database records and to be contacted for future study. A total of 3598 surveys were returned, including children with asthma ($n = 398$, 11%), with hayfever or food allergy, but no asthma ($n = 192$, 5%) and with neither ($n =$

3008). A nested case-control study was populated from this data, with the intent of recruiting 400 children with asthma and 400 children without asthma (healthy children or children with allergies only). To achieve this, all children with asthma or with allergies but no asthma, were invited to participate. Children with neither asthma nor allergic disease were stratified by urban/rural residence and neighbourhood income. A random sample from each stratum (100 of children from rural areas, 50 from low income, urban areas and 50 from high income, urban areas with and without a family history of asthma) was contacted for participation. This method of sampling ensured representation from diverse environmental exposures by over-sampling children from rural, low income areas and First Nation communities. Following consent from parents to participate, children and parents living in Winnipeg came to the Pediatric Allergy laboratory at the Manitoba Institute for Child Health in Winnipeg. Families living in rural areas were seen by the SAGE research team in a clinic or hospital in the nearest town (see schema in Fig. 1).

Parents of 723 children agreed to participate in the study: 288 children with and 435 children without a parental report of asthma in the last 12 months. Following confirmation of asthma status by a pediatric allergist, the nested case-control study comprised of 246 cases and 477 controls. The following assessments were conducted

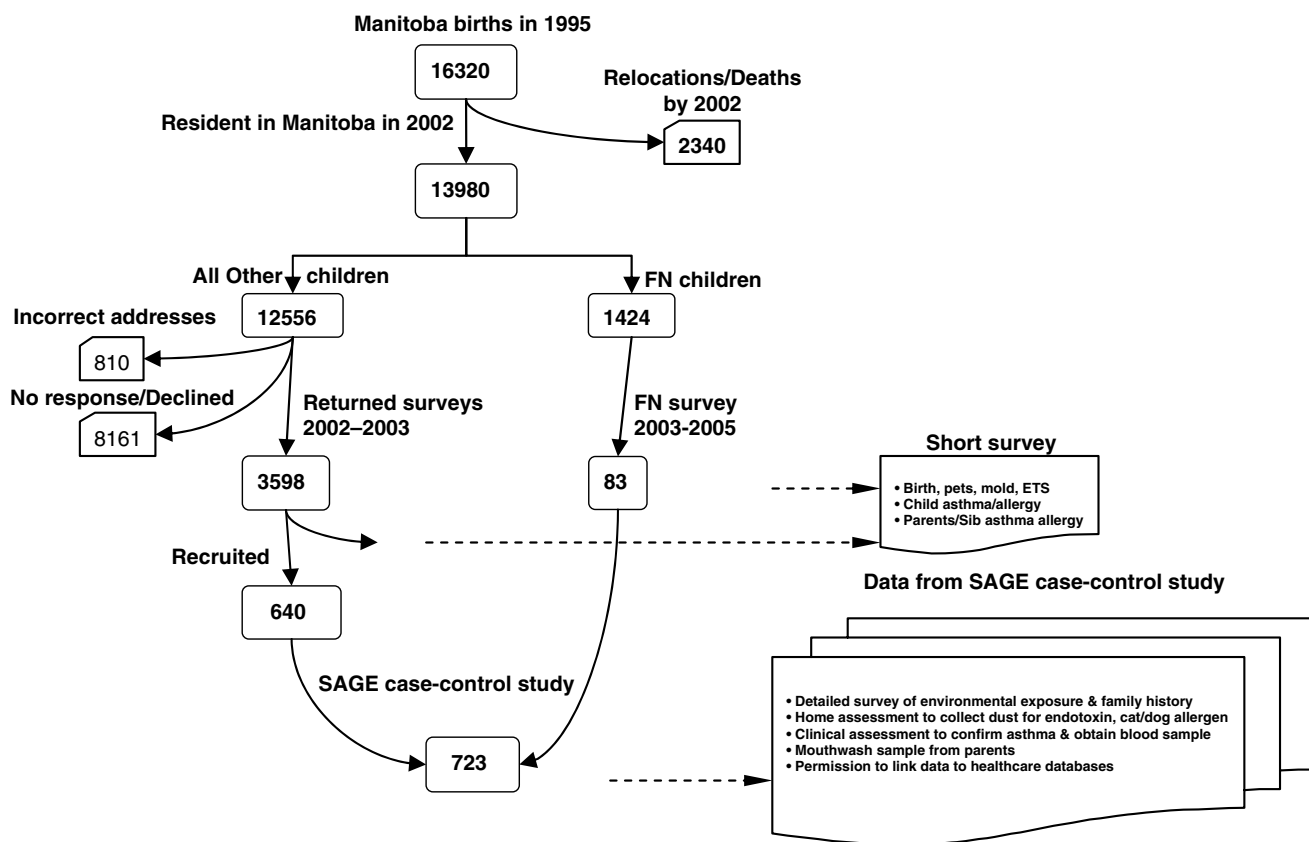


Figure 1. Design for SAGE cohort and case-control study.

in children by the SAGE team: (1) airway responsiveness, as measured by the methacholine challenge test using the Cockcroft method, and (2) atopy by skin-prick testing for common allergens (tree pollen mix, weed pollen mix, ragweed, grass pollen mix, *Alternaria*, *Cladosporium*, *Penicillium*, house dust mites *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*, cockroach, cat, dog, feathers and peanut). A detailed survey was administered to parents concerning their current and past home environment. A blood sample was obtained from children for DNA testing, determination of serum IgE and cortisol, and assessment of immune cell function. Parents were also surveyed with respect to their own respiratory health and they provided mouthwash samples for DNA. A research technician visited the homes of study children to do a visual inspection for mold, to assess whether the family had instituted allergen avoidance measures, and to collect dust samples from the child's mattress, the floor of the child's bedroom and the room most commonly used by the child. Dust samples were assayed for mite/cat/dog allergen and endotoxin levels, using the *Limulus* amoebocyte lysate and other commercial assays (22–25).

Processes established for the conduct of research in Manitoba First Nations communities were followed to recruit First Nations children living on reserve. First, the Health Information and Research Committee of the Assembly of Manitoba Chiefs approved the SAGE project. Next, the Chief and Council of individual First Nations communities were contacted to obtain approval to recruit families in their communities. An individual of Aboriginal descent with experience in doing research in First Nations communities was hired for this purpose. A community health worker in each participating community was then selected as the liaison for the study. The one-page surveys were administered by the community health worker, who also recruited families for participation in the case-control study. A visit by the SAGE team was scheduled to collect the same information obtained in other Manitoba children.

The study design resulted in the creation of three research platforms to investigate early-life determinants of asthma in schoolchildren: (1) a population-based longitudinal study using healthcare database records from birth until age 7–8 years in all 13 980 children, (2) a population-based mail-out survey on home environment exposures to parents of children at age 7–8 years ($n = 3598$) and (3) a nested case-control study of children with and without asthma recruited at age 7–10 years ($n = 723$).

Unique features of SAGE

There are several features unique to the SAGE design which provide enhanced capacity to study the origins of asthma and other atopic diseases. Each of these features is described in more detail below to

highlight the many facets and data capabilities of this approach.

Use of a provincial healthcare registry to create a general population birth cohort

Manitoba's healthcare registry records all new registrants with Manitoba Health from birth and immigrations in Manitoba, as well as all discontinuations when persons leave the province (26). Unlike health maintenance organizations (HMOs) which only capture data for persons enrolled in the HMO, this registry enumerates the complete population of Manitoba, regardless of whether persons accessed health care. From Manitoba's healthcare registry, we identified a 1995 birth cohort of children born in Manitoba who were resident in Manitoba until 2002, the time at which we initially assessed children for the presence of asthma. All Manitobans are assigned an individual and family registration number, which permitted linkage of parents to children. The availability of healthcare data on parents enabled epidemiologic, immunologic and genetic analyses, not usually possible with other healthcare databases.

Detailed and comprehensive healthcare records to allow measurement of early-life exposures to antibiotics and vaccinations, and of future atopic disease outcomes

Manitoba's healthcare databases have been shown to be comprehensive and valid data sources on physician diagnoses and prescription drug use (27, 28). They contain the complete healthcare utilization records for Manitobans from as early as the 1980s, including a record of every physician visit, hospitalization, immunization and prescription drug collected by the Manitoba Health Services Insurance Program. Records of physician reimbursement for medical care provided are submitted under a fee-for-service arrangement, and contain information on patient diagnosis at the 3-digit level of the ICD-9-CM classification system. Discharge abstracts for hospital services include information on up to 16 ICD-9-CM diagnostic codes, of which the first diagnosis is the primary diagnosis that is most responsible for the hospital stay. Prescription records submitted by retail pharmacies contain data on the date of prescription dispensing, drug name and identification number, dosage form and quantity dispensed. The Manitoba Immunizations Monitoring System is a computerized registry of the immunization records of Manitobans, which is collected from physician billing claims, manual entry of immunization record forms by public health nurses and hospital departments for immunizations that are not billed by physicians. Linkage among databases can be achieved via an encrypted health identification number.

Early-life exposures of the child, such as antibiotic prescriptions, are commonly assessed through questionnaire, but parental report is subject to recall bias (29).

Parent recall lacks detail on important elements, such as the timing and type of antibiotic or immunization. These events are accurately recorded in healthcare databases.

Administration of a mail-out survey to add information on early-life exposures to pets and tobacco smoke which can be linked to the healthcare database measures

A short, mail-out survey inquired into the presence of asthma, allergy and other health conditions in cohort children at age 7–8 and their parents and siblings, as well as birth and current home environmental exposures to tobacco smoke, pets and mold. While Manitoba's healthcare registry has a record of all persons born in the province, the copy of the registry housed in the Population Health Research Data Repository at the Manitoba Centre for Health Policy, is anonymized. Every Manitoban is denoted by their personal health identification number (PHIN), which has been encrypted by Manitoba Health. This ensures privacy of healthcare records during the research process. Names and addresses of Manitobans in the registry are maintained by Manitoba Health. As such, the list of scrambled PHINs for the SAGE cohort was sent to Manitoba Health, which located the names and addresses of the cohort children, and mailed the survey to their households.

In addition to the above questions, parents were asked for permission to be contacted and for permission to access their child's healthcare database records. Researchers had no knowledge of the identity of children in the complete cohort – only of the children for whom parents returned surveys. The returned surveys served as the sampling frame for the nested case-control study, and provided information on home environmental exposures and asthma outcomes, which could be analyzed on its own or could be linked to healthcare database records.

Nesting a case-control study within the birth cohort to test hypotheses concerning the relationships between exposures, growth, pulmonary function, atopy, immune cell function and genotype in children with and without asthma

A case-control study was nested within the birth cohort to collect information on home environmental exposures from detailed surveys and home dust sampling, to confirm asthma status in children and use this data to validate healthcare database measures of asthma for the whole cohort, to determine immune system responsiveness to innate and adaptive immune stimuli in asthma, to genotype children for genes likely associated with the development of asthma and to study the epigenetic regulation of pre-established protective *vs* allergic immune responses. Each of these procedures will be described individually.

Home environmental exposures from detailed surveys and home dust sampling. Most healthcare databases contain

information on contact with the healthcare system and do not routinely include data on exposure to pets and tobacco, or household endotoxin and allergen levels. We expanded the range of early childhood exposures that were evaluated through linkage of healthcare database records to primary data collected from parent surveys and home samples. This facilitated investigation into, for example, the combined or individual effect of immunization and lower respiratory tract infections (from database records), tobacco smoke (parent report), home allergen and endotoxin levels (house dust collections), and visual mold (home inspection), on the development of allergic disease.

All home environmental exposures were collected after the fact, either from parent survey of events during their child's early life or from home assessments conducted at age 7–10. We addressed recall bias with the former by, for example, comparing responses on extent of breast feeding in 40 SAGE mothers who also participated in the Canadian Asthma Primary Prevention Study (CAPPS) study in Winnipeg, to their same responses in CAPPS (correlation coefficient was 0.45, $P < 0.01$) (30). For other measures, we have relied on the experience of others. This includes the findings of Celedón et al., who found no statistically significant difference in exposure to endotoxin at age 2–3 months and at age 7 years (31). Further, as SAGE is an ongoing study, we will be able to investigate the temporal relationships between home environment measurements at age 7–10 and incident asthma during adolescence.

Confirmation of asthma status and validation of database measures of asthma. Healthcare databases are reliable and valid sources of information on diagnostic information, but this varies by disease, and requires data validation for specific diagnoses. Healthcare utilization bias (over- or under-reporting of diagnoses), is possible (32). This is especially a problem when childhood diagnoses of asthma may represent various phenotypes, such as early transient wheeze or persistent wheeze (1, 33).

Pediatric allergist-diagnosed asthma, methacholine challenge tests and skin tests were obtained for children recruited into the nested case-control study. These data were used as the gold standard for comparison to several database definitions of asthma. Candidate definitions varied according to number and type of healthcare visits, and prescription medications for asthma (34). The sensitivity, specificity, positive predicted value (PPV) and Youden index were calculated for each database measure (see Table 1), permitting the selection of database measures according to type of study (for example, asthma prevalence or determinants of asthma studies). The sensitivity of the asthma definitions ranged from 47% to 82%, the specificity from 83% to 92% and the PPV from 90% to 94%. Database measures of asthma with a high sensitivity and PPV are needed for studies of risk factors for asthma, while those with high sensitivities and Youden indices are ideal for asthma prevalence

Table 1. Validation parameters for database definitions of asthma at age 7–8 years in the SAGE cohort

Database definition	Sensitivity (95% CI)*	Specificity (95% CI)*	Positive predicted value (95% CI)	Youden index
At least one asthma hospitalization, or two physician visits, or four prescription medications	0.47 (0.35–0.60)	0.92 (0.78–0.98)	0.91 (0.76–0.98)	0.39
At least one asthma hospitalization, or two physician visits, or two prescription medications	0.67 (0.54–0.78)	0.92 (0.78–0.98)	0.94 (0.82–0.99)	0.58
At least one asthma hospitalization, or one physician visit, or two prescription medications	0.77 (0.65–0.87)	0.92 (0.78–0.98)	0.94 (0.85–0.99)	0.69
At least one asthma hospitalization, or one physician visit, or two bronchodilators, or one controller medication	0.80 (0.69–0.89)	0.89 (0.74–0.97)	0.93 (0.83–0.98)	0.69
At least one asthma hospitalization, or one physician visit, or two bronchodilators, or one bronchodilator and ketotifen or an oral steroid, or one controller medication	0.80 (0.69–0.89)	0.89 (0.74–0.97)	0.93 (0.83–0.98)	0.69
At least one asthma hospitalization, or one physician visit, or one bronchodilator, or one controller medication	0.82 (0.70–0.90)	0.83 (0.67–0.94)	0.90 (0.79–0.96)	0.65

*Gold standard is pediatric allergist diagnosis of asthma in the same year.

studies (35). Importantly, as all of our definitions included use of prescription medication, the PPV was higher than for database definitions derived only from medical diagnoses (36).

Determination of immune system responsiveness to Toll-like receptor stimuli and respiratory viruses. We undertook extensive immunologic characterization of children in the case-control study of 723 children. Innate immune cytokine and chemokine responses to a panel of Toll-like receptor (TLR) ligands were evaluated via short-term primary culture of peripheral blood mononuclear cells directly *ex vivo*. We hypothesized that individuals with current asthma and early-life bronchiolitis (or environmental tobacco smoke exposure) would exhibit quantitatively different regulatory responses in innate immune capacity. Concomitantly, adaptive immuno-regulatory responses to common respiratory viruses were evaluated following primary culture with respiratory syncytial virus, human metapneumovirus and reovirus. Food allergy was similarly targeted. Given the tremendous variability in immune responses of humans, and the broad diversity of environmental and genetic factors that interact in each individual to yield a given asthma phenotype, this cohort provided the necessary sample size of well-phenotyped subjects. It then became possible to dissect out the contributions of different immune, genetic and environmental influences that result in the widely accepted clinical endpoints known as ‘allergic asthma’, ‘nonallergic asthma’, ‘food allergic’ and ‘allergic rhinitis’.

In our initial studies, we found that levels of TLR agonists, typically used for *in vitro* assessment of immune capacity, yielded results distinct from those obtained using commonly occurring environmental levels of TLR ligands (37). These findings demonstrate the importance of using TLR ligands at concentrations more closely mimicking environmental levels when

evaluating immune capacity. We anticipate that subsequent studies will provide the opportunity to examine exquisitely well-defined (immunologically, genetically, environmentally and clinically) representatives of each asthma phenotype, and thereby, evaluate current hypotheses on the contributions for each of these influences towards maintaining an individual’s protective (in healthy individuals) or pathologic (in hypersensitive individuals) responses.

Genotyping for candidate genes/single nucleotide polymorphisms associated with asthma. In order to study the gene–environment interaction in the origins of atopic disease, we selected 160 candidate genes that had previously been implicated in allergy and asthma and/or had compelling biologic plausibility for association with atopic phenotypes (38). The majority of the genotyping was performed using Illumina technology in the Genome Quebec Genome Sciences Centre and the laboratory of Dr Thomas Hudson at the McGill University. The remaining genotyping was performed by TaqMan assays at the iCAPTURE Centre.

A set of ~1536 single nucleotide polymorphisms (SNPs) in the candidate genes was selected for genotyping using a combination of coding SNPs and a ‘tag-SNP’ approach (39). The tag-SNP approach utilizes the fact that alleles from closely linked polymorphisms are often associated with each other (i.e. are in linkage disequilibrium). The linkage disequilibrium between all known SNPs in each gene was assessed and this information was used to place the SNPs into bins. Within each bin, every SNP had a minimum level of linkage disequilibrium with the tag-SNP. These tag-SNPs were selected for genotyping along with SNPs that were not associated with any other SNP. This approach has the advantage of reducing the number of polymorphisms to be genotyped. Single nucleotide polymorphisms which code for changes in amino acids within the gene and SNPs previously

associated with atopic were also included among the 1536. As a result, a comprehensive survey of each gene was carried out because all common polymorphisms were tested either directly or because of their association with one of the tag polymorphisms.

Study of epigenetic regulation of pre-established protective vs maladaptive immunity. The field of epigenetics investigates modifications to DNA and to chromatin. Chromatin is comprised of DNA complexes and histone octamers, which are responsible for the coiling of chromatin and access of the transcriptional machinery to the DNA (40). As these modifications can be inherited, they provide a mechanism different from the classical mutation of DNA nucleotides to stably propagate enhanced or reduced gene activity from one generation of cells to the next. Many epigenetic mechanisms have been described, including DNA methylation and histone acetylation, methylation or phosphorylation. In marked contrast to changes in DNA coding sequence, epigenetic changes are reversible, suggesting that such changes may be key in the development and propagation of immune capacity.

The variable natural history of asthma phenotypes may be the result of epigenetic changes after early or later environmental exposures. Findings from several epidemiologic and experimental studies indicate that asthma risk may be modified by epigenetic regulation (41). Studies being carried out by Drs. Su and HayGlass at the University of Manitoba address the putative role that histone acetylation may play in maintenance of protective vs maladaptive immunity characteristic of healthy and allergic phenotypes in childhood (42).

Challenges in the execution of the SAGE study

The major limitation of the SAGE study design is that data on some exposures and child outcomes were not prospectively collected from birth. The clinical, biological and home environment data for SAGE children were obtained at age 7–10 years; so it is not possible to properly test the temporal relationships between earlier exposures and disease onset (31). Further, the SAGE study has little data on the nutritional intake of the children and information during pregnancy of the mothers is mostly limited to prescription medication or healthcare usage.

The 29% response rate to our mail-out survey was also troubling, although it is in keeping with population-wide mail surveys derived from health registries (43). Fortunately, the urban–rural distribution of asthma prevalence from survey responders was very similar to provincial statistics (44). This minimized sampling bias in the selection of cases and controls. Further, there were sufficient numbers of children in the

required stratifications to populate our case-control study.

Finally, we encountered some difficulties in the recruitment of children for the case-control study. Recruitment from rural areas, although it was costly and required the transport of the SAGE research team, was successful because of the high interest level in the study. However, recruitment of children living in First Nations communities was a challenge and participation rates were low. Out of 16 communities contacted by the research coordinator, six communities agreed to participate, comprising from 10 to 53 children born in 1995. The reasons provided by Chiefs and Councils for declining to participate were: lack of interest, parents not wanting children to go through the tests, no asthma or allergies in the community, no benefits seen from previous studies and lack of trust of researchers. One community declined to participate because of the DNA testing of children and parents. Our goal was to recruit 200 First Nations children; the final study number was 83. Even among communities who had agreed to participate, the response rate from parents varied from 38% to 90%.

Strengths of the SAGE study

A prospective cohort study with frequent collection of biological, clinical and survey data from early pregnancy until the child age of interest is ideal to study the origins of allergic disease. While the SAGE cohort lacks data on some of the early-life exposures, it is being conducted at a substantially lower cost and has the following strengths.

Cohort study

- Complete longitudinal healthcare records which measure early exposures such as antibiotic utilization and immunization, in relationship to the development of asthma.
- No loss to follow-up of children and little migration out of province.
- A population-based study with high- and low-risk children.
- Captures children living in urban and rural environments.
- First Nations children included in study.
- Survey information on home environmental exposures in a subset.

Nested case-control study

- Minimal study bias because cases and controls selected from the same population.
- Control children were over-sampled from rural, low income areas and First Nation communities to ensure representation from diverse environmental exposures.

- Numerous data sources for atopic phenotypes from parent report, pediatric allergist diagnosis, longitudinal healthcare records, bronchial responsiveness and skin-prick tests.
- Numerous data sources for home environment exposures which can be linked. This includes endotoxin, beta-glucan, cat and dog allergen levels in house dust samples; home inspection (for direct visualization of moisture damage, mold presence and use of allergen avoidance measures, such as mattress encasements); database information on early-life exposures such as antibiotic use; and parent report of breast feeding, tobacco use, pet ownership and daycare use during early life of the child.
- Assessment of innate and adaptive immune responses to a variety of immune stimuli from a large number of children to study the mechanisms of allergic disease *vs* clinical tolerance.
- Genotyping to study the genetic and gene-environment interactions in the origins of allergic disease.

The SAGE study is the product of a multidisciplinary collaboration which brings together the fields of pediatric allergy, population health, immunology, genetics and environmental epidemiology. The study has been built upon the existing expertise and research infrastructure of its investigators. Anita Kozyskyj, an epidemiologist with expertise in population health research with Manitoba's healthcare databases, created the cohort and conducted the linkage studies. Recruitment of children for the nested case-control study was facilitated by the existing research infrastructure of the Winnipeg arm of the Canadian Asthma Primary Prevention Study, led by pediatric allergist, Allan Becker. This infrastructure involved recruitment by research nurses who had experience in establishing trusting relationships with families. Moira Chan-Yeung, an environmental epidemiologist, coordinated the endotoxin assays of the dust samples. Kent HayGlass, an immunologist with a primary interest in human immune regulation, led the study arm examining the innate and adaptive immune responses in these children. Peter Pare, a clinician scientist with expertise in the genetics of asthma, and Andrew Sandford, a human geneticist, supervised the genotyping at the iCAPTURE Centre. The SAGE study has attracted funding from national agencies interested in supporting multidisciplinary research. It was initiated by the Canadian Institutes of Health Research (CIHR) New Emerging Team Program, and Operating and New Investigator grants to individual SAGE investigators. Funding has been extended by the CIHR Interdisciplinary Capacity Enhancement Program and AllerGen NCE Inc. to follow the SAGE case-control children longitudinally into adolescence. For this part of the study, the multidisciplinary team was expanded to include nutritionists and endocrinologists.

Moreover, the SAGE study serves as a fertile training ground for graduate students, postdoctoral fellows and clinician fellows in immunology, epidemiology and the social sciences. Many of the SAGE investigators are co-investigators of the CIHR-funded National Training Program in Allergy and Asthma, headed by Kent HayGlass. Trainees have acquired expertise in project management, including data collection and analysis, and have been given the opportunity to integrate expertise from their areas of specialty into a broad, multidisciplinary research team.

In conclusion, we have implemented a novel design to study the early-life origins of asthma. To date, our database cohort findings have identified several risk factors for the development of asthma at age 7–10, including: antibiotic use (45), transient tachypnea of the newborn (46), and maternal distress (47). Another database study of the same cohort found that premature infants are not more likely to develop food allergy (48). From our case-control study, we determined the risk of asthma to be highest in children who were both overweight and had low milk consumption (49) or were exclusively breastfed for <4 months (30). Our research is ongoing. SAGE study children were reassessed at age 10–11 years and most recently at age 12 to determine onset of new asthma. Studies are underway on the association between overweight and early thelarche on the development of asthma during adolescence, and its potential interconnection with insulin resistance and the metabolic syndrome. These investigations have been supplemented with biological markers such as leptin, adiponectin and estradiol. Qualitative interviews and quantitative surveys about body image and dieting patterns of children in preadolescence have also been added to provide context on the social environment of weight modification at this age. Finally, we have extended our studies on maternal distress by incorporating measures of cortisol levels at multiple time points over the lifespan of the child.

Acknowledgments

Sincere thanks to all study families without whose participation, this research could not have been conducted. Data were also accessed from the Population Health Research Data Repository at the Manitoba Centre for Health Policy. The results and conclusions are those of the authors and no official endorsement by Manitoba Health was intended or should be inferred. We would also like to acknowledge the computer analysis support of Shamima Huq, as well as the research support of the SAGE study team provided by Rishma Chooniedass, Donna Everette, Brenda Gerwing, Joel Liem, Marilyn Lilley, Tanya Lilley-Chan, Ingrid Loewen and Michelle Tillett. A. L. Kozyskyj is a CIHR New Investigator. K. T. HayGlass is a Canada Research Chair in Immunology. A. J. Sandford is a Canada Research Chair in Genetics. P. D. Paré is a Michael Smith Foundation and Jacob Churg Scholar. This study was funded by the Canadian Institutes of Health Research.

References

- Stein RT, Martinez FD. Asthma phenotypes in childhood: lessons from an epidemiological approach. *Paediatr Respir Rev* 2004;**5**:155–161.
- Lau S, Nickel R, Niggemann B, Grüber C, Sommerfeld C, Illi S et al. The development of childhood asthma: lessons from the German Multicentre Allergy Study (MAS). *Paediatr Respir Rev* 2002;**3**:265–272.
- Keil T, Kulig M, Simpson A, Custovic A, Wickman M, Kull I et al. European birth cohort studies on asthma and atopic diseases: II. Comparison of outcomes and exposures – a GA2LEN initiative. *Allergy* 2006;**61**: 1104–1111.
- Chan-Yeung M, Ferguson A, Watson W, Dimich-Ward H, Rousseau R, Lilley M et al. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* 2005;**116**: 49–55.
- Celedon JC, Wright RJ, Litonjua AA, Sredl D, Ryan L, Weiss ST et al. Day care attendance in early life, maternal history of asthma, and asthma at the age of 6 years. *Am J Respir Crit Care Med* 2003;**167**:1239–1243.
- Uekert SJ, Akan G, Evans MD, Li Z, Roberg K, Tisler C et al. Sex-related differences in immune development and the expression of atopy in early childhood. *J Allergy Clin Immunol* 2006;**118**:1375–1381.
- Burke W, Fesinmeyer M, Reed K, Hampson L, Carlsten C. Family history as a predictor of asthma risk. *Am J Prev Med* 2003;**24**: 160–169.
- Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A, NAC Manchester Asthma and Allergy Study Group. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002;**13**(Suppl):32–37.
- Lau S, Illi S, Sommerfeld C, Niggemann B, Bergmann R, von Mutius E et al. Early exposure to house-dust mite and cat allergens and development of childhood asthma: a cohort study. Multicentre Allergy Study Group. *Lancet* 2000;**356**:1392–1397.
- Xu B, Pekkanen J, Järvelin MR, Olsen P, Hartikainen AL. Maternal infections in pregnancy and the development of asthma among offspring. *Int J Epidemiol* 1999;**28**:723–727.
- Sears MR, Greene JM, Willan AR, Taylor DR, Flannery EM, Cowan JO et al. Long-term relation between breastfeeding and development of atopy and asthma in children and young adults: a longitudinal study. *Lancet* 2002;**360**:901–907.
- Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. *N Engl J Med* 1995;**332**: 133–138.
- Ownby DR, Johnson CC, Peterson EL. Exposure to dogs and cats in the first year of life and risk of allergic sensitization at 6 to 7 years of age. *JAMA* 2002;**288**:963–972.
- Arshad SH, Kurukulaaratchy RJ, Fenn M, Matthews S. Early life risk factors for current wheeze, asthma, and bronchial hyperresponsiveness at 10 years of age. *Chest* 2005;**127**:502–508.
- Oddy WH, Sherriff JL, de Klerk NH, Kendall GE, Sly PD, Beilin LJ et al. The relation of breastfeeding and body mass index to asthma and atopy in children: a prospective cohort study to age 6 years. *Am J Public Health* 2004;**94**: 1531–1537.
- Shaheen SO, Newson RB, Henderson AJ, Headley JE, Stratton FD, Jones RW et al. Prenatal paracetamol exposure and risk of asthma and elevated immunoglobulin E in childhood. *Clin Exp Allergy* 2005;**35**: 18–25.
- Roos LL, Brownell M, Lix L, Roos NP, Walld R, Macwilliam L. From health research to social research: privacy, methods, approaches. *Soc Sci Med* 2008;**66**:117–129.
- Helms PJ, Ekins Daukes S, Taylor MW, Simpson CR, McLay JS. Utility of routinely acquired primary care data for paediatric disease epidemiology and pharmacoepidemiology. *Br J Clin Pharmacol* 2005;**59**:684–690.
- Wigertz A, Westerling R. Measures of prevalence: which healthcare registers are applicable? *Scand J Public Health* 2001;**29**:55–62.
- Williams K, Leonard H, Tursan d'Espaignet E, Colvin L, Slack-Smith L, Stanley F. Hospitalisations from birth to 5 years in a population cohort of Western Australian children with intellectual disability. *Arch Dis Child* 2005;**90**:1243–1248.
- de Vries F, Pouwels S, Lammers JW, Leufkens HG, Bracke M, Cooper C et al. Use of inhaled and oral glucocorticoids, severity of inflammatory disease and risk of hip/femur fracture: a population-based case-control study. *J Intern Med* 2007;**261**:170–177.
- Tsuji K, Steindler KA, Harrison SJ. Limulus amoebocyte lysate assay for detection and quantitation of endotoxin in a small-volume parenteral product. *Appl Environ Microbiol* 1980;**40**:533–538.
- Luczynska C, Arruda LK, Platts Mills TA, Miller JD, Lopez M, Chapman MD. A two site monoclonal antibody ELISA for the quantification of the major *Dermatophagoides* spp. allergens, Der p 1 and Der f 1. *J Immunol Methods* 1989;**118**:227–235.
- de Blay F, Heyman P, Chapman M, Platts Mills TA. Airborne dust mite allergens: comparison of group II allergens with group I mite allergen and cat allergen. *J Allergy Clin Immunol* 1991;**88**:919–926.
- Schou C, Hansen GN, Lintner T, Lowenstein H. Assay for the major dog allergen Can f1. Investigation of house dust samples and commercial dog extracts. *J Allergy Clin Immunol* 1991;**88**:847–853.
- Roos L, Nicol J. A research registry: uses, development, and accuracy. *J Clin Epidemiol* 1999;**52**:39–47.
- Kozyrskyj AL, Mustard CA. Validation of an electronic, population-based prescription database. *Ann Pharmacother* 1998;**32**:1152–1157.
- Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease. Comparing administrative data and self-reports. *Med Care* 1997;**35**: 932–947.
- Metlay JP, Hardy C, Strom BL. Agreement between patient self-report and a veterans affairs national pharmacy database for identifying recent exposures to antibiotics. *Pharmacoepidemiol Drug Saf* 2003;**12**:9–15.
- Mai XM, Becker AB, Sellers EA, Liem JJ, Kozyrskyj AL. The relationship of breast-feeding, overweight, and asthma in preadolescents. *J Allergy Clin Immunol* 2007;**120**:551–556.
- Celedón JC, Milton DK, Ramsey CD, Litonjua AA, Ryan L, Platts-Mills TA et al. Exposure to dust mite allergen and endotoxin in early life and asthma and atopy in childhood. *J Allergy Clin Immunol* 2007;**120**:144–149.

32. Kozyrskyj AL, Mustard CA, Becker AB. Identifying children with persistent asthma from health care administrative records. *Can Respir J* 2004;**11**: 141–145.
33. Kozyrskyj AL, Mustard CA, Becker AB. Childhood wheezing syndromes and healthcare data. *Pediatr Pulmonol* 2003;**36**:131–136.
34. Kozyrskyj AL, Becker AB. Identifying childhood asthma from health care database records: is it valid? *Allergy Asthma Clin Immunol* 2005;**1**:108 (Abstract).
35. Pekkanen J, Pearce N. Defining asthma in epidemiological studies. *Eur Respir J* 1999;**14**:951–957.
36. Blais L, Lemi re C, Menzies D, Berbiche D. Validity of asthma diagnoses recorded in the medical services database of Quebec. *Pharmacoepidemiol Drug Saf* 2006;**15**:245–252.
37. Lissitsyn Y, Becker AB, Kozyrskyj AL, HayGlass KT. Level of Toll-like receptor agonist exposure differentially determines chemokine production in humans. *Can J Physiol Pharmacol* 2007;**85**: 739–746.
38. Hoffjan S, Nicolae D, Ober C. Association studies for asthma and atopic diseases: a comprehensive review of the literature. *Respir Res* 2003;**4**:14.
39. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004;**74**:106–120.
40. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007;**447**:433–440.
41. Miller RL, Ho SM. Environmental epigenetics and asthma: current concepts and call for studies. *Am J Respir Crit Care Med* 2008;**177**:567–573.
42. Su RC, Becker AB, Kozyrskyj AL, HayGlass KT. Epigenetic regulation of established human type 1 versus type 2 cytokine responses. *J Allergy Clin Immunol* 2008;**121**:57–63.
43. Angus VC, Entwistle VA, Emslie MJ, Walker KA, Andrew JE. The requirement for prior consent to participate on survey response rates: a population-based survey in Grampian. *BMC Health Serv Res* 2003;**3**:21.
44. Kozyrskyj AL, Becker AB. Rural-urban differences in asthma prevalence: possible explanations. *J Allergy Clin Immunol* 2004;**113**(Suppl):S306 (Abstract).
45. Kozyrskyj AL, Ernst P, Becker AB. Increased risk of childhood asthma from antibiotic use in early life. *Chest* 2007;**131**:1753–1759.
46. Liem JJ, Huq SI, Ekuma O, Becker AB, Kozyrskyj AL. Transient tachypnea of the newborn may be an early clinical manifestation of wheezing symptoms. *J Pediatr* 2007;**151**:29–33.
47. Kozyrskyj AL, Mai XM, McGrath P, Hayglass KT, Becker AB, Macneil B. Continued exposure to maternal distress in early life is associated with an increased risk of childhood asthma. *Am J Respir Crit Care Med* 2008;**177**:142–147.
48. Liem JJ, Kozyrskyj AL, Huq SI, Becker AB. The risk of developing food allergy in premature or low-birth-weight children. *J Allergy Clin Immunol* 2007;**119**:1203–1209.
49. Mai XM, Becker AB, Sellers EA, Liem JJ, Kozyrskyj AL. Infrequent milk consumption plus being overweight may have great risk for asthma in girls. *Allergy* 2007;**62**:1295–1301.