Kuo, Chang-Fu and Grainge, Matthew J. and Valdes, Ana M. and See, Lai-Chu and Luo, Shue-Fen and Yu, Kuang-Hui and Zhang, Weiya and Doherty, Michael (2015) Familial aggregation of systemic lupus erythematosus and coaggregation of autoimmune diseases in affected families. JAMA Internal Medicine, 175 (9). pp. 1518-1526. ISSN 2168-6114

Familial aggregation of systemic lupus erythematosus and co-aggregation of autoimmune diseases in affected families

Chang-Fu Kuo, MD PhD; Matthew J Grainge, PhD; Ana M Valdes, PhD; Lai-Chu See, PhD; Shue-Fen Luo, MD; Kuang-Hui Yu, MD; Weiya Zhang, PhD; Michael Doherty, MD

Author affiliations:

Division of Rheumatology, Orthopaedics and Dermatology, School of Medicine, University of Nottingham, Nottingham, UK (Drs Chang-Fu Kuo, Ana Valdes, Weiya Zhang and Michael Doherty); Division of Rheumatology, Allergy and Immunology, Chang Gung Memorial Hospital, Taoyuan, Taiwan (Drs Chang-Fu Kuo, Shue-Fen Luo and Kuang-Hui Yu); Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, UK (Dr Matthew J. Grainge); Department of Public Health, College of Medicine and Biostatistics Core Laboratory, Molecular Medicine Research Centre, Chang Gung University, Taoyuan, Taiwan (Dr Lai-Chu See)

Joint senior authors: Drs Weiya Zhang and Michael Doherty

Correspondence to Dr Chang-Fu Kuo

Email: zandis@gmail.com; mbxck2@nottingham.ac.uk. Telephone: +44 (0) 115 8231756; Fax: +44 (0) 115 8231757; Address: Academic Rheumatology, Clinical

Sciences Building, City Hospital, Nottingham, UK, NG51PB

Word count: 3,991

ABSTRACT

Importance Relatives of patients with systemic lupus erythematosus (SLE) appear to be at higher risk of SLE and other autoimmune diseases but estimates of individual familial risks are largely unavailable or unreliable. Furthermore, relative contributions of genetic, shared and non-shared environmental factors to SLE susceptibility remain unclear.

Objective To examine familial aggregation and heritability of SLE and the relative risks (RRs) of other autoimmune diseases in relatives of SLE patients.

Design and setting Population-based family study using the Taiwan National Health Insurance (NHI) Research Database.

Participants All individuals (n = 23,658,577) registered with the NHI in 2010, of whom 18,283 individuals had SLE. We identified 21,009,551 parent-child relationships, 17,168,340 full sibling pairs and 342,066 twin pairs. Diagnoses of SLE were ascertained up to December 31, 2010.

Main Outcomes and Measures The prevalence and relative risk (RR) of SLE and other autoimmune diseases in relatives and spouses of SLE patients and the relative contributions of heritability, shared and non-shared environmental factors to SLE susceptibility.

Results The RR (95% confidence interval [CI]) for SLE was 315.94 (210.66–473.82)

for twins of SLE patients; 23.68 (20.13–27.84) for siblings; 11.44 (9.74–13.43) for parents; 14.42 (12.45–16.70) for offspring and 4.44 (2.38–8.30) for spouses. The accountability for phenotypic variance of SLE was 43.9% for heritability, 25.8% for shared environmental factors and 30.3% for non-shared environmental factors. The RR (95% CI) in individuals with a first-degree relative with SLE was 2.66 (2.28–3.11) for rheumatoid arthritis; 5.40 (3.37–8.65) for systemic sclerosis; 5.87 (4.89–7.05) for primary Sjögren's syndrome; 2.77 (1.45–5.32) for idiopathic inflammatory myositis; 1.68 (1.22–2.32) for type 1 diabetes mellitus; 2.58 (1.16–5.72) for multiple sclerosis; 2.95 (2.04–4.26) for myasthenia gravis; 1.39 (0.66–2.91) for inflammatory bowel diseases and 0.86 (0.43–1.71) for vasculitis.

Conclusions and Relevance The individual risks of SLE and other autoimmune diseases were increased in families with SLE patients. Heritability of SLE was estimated to be approximately 44%. These data should be considered when counselling families with affected members.

INTRODUCTION

Systemic lupus erythematosus (SLE) is the prototype autoimmune disease with features of autoantibody production, immune complex deposition and multiple target organ damage. SLE can affect any part of the body[1] and the course of the disease is diverse and unpredictable.[2] The prevalence of SLE ranges from 0.02% to 0.15%.[3] Our group recently estimated that prevalence of SLE was 0.10% in the UK[4] and 0.07% in Taiwan.[5]

Early family studies have documented familial aggregation of SLE[6-13] and a classic twin study found a 10-fold increased concordance in monozygotic compared to dizygotic twins.[14] Furthermore, SLE is also reported to co-aggregate with other autoimmune diseases.[10,12] The tendency of SLE and other autoimmune diseases to cluster within families suggests a significant role for genetic or shared environmental factors in the pathogenesis of autoimmune diseases. Heritability, defined as the proportion of the phenotypic variance explained by genetic factors, is estimated to be 66% in SLE,[15,16] suggesting a strong genetic component to its pathogenesis. Consequently, efforts on defining pathogenesis of SLE focus on genetic factors and recent genome-wide association studies (GWAS) have successfully identified over 30 susceptibility loci for SLE.[17] However, these findings account for less than 10% of phenotypic variation observed.[18] Large unexplained heritability leads to the question

of the relative contribution of genetic factors to SLE susceptibility.

Two previous studies have reported the heritability of SLE,[15,16] however, both failed to differentiate genetic and shared environmental factors. Therefore they more accurately estimated 'familial transmission'—a measure of the combined contribution of genetic and shared environmental factors to disease susceptibility. Such estimates will overestimate the heritability of SLE. In contrast to heritability, quantitative estimates of an individual's risk of SLE and other autoimmune diseases if they have a positive family history of SLE are more useful for genetic counselling. However, reliable measures such as relative risks (RR) are largely unavailable or of limited reliability.

Therefore, we conducted this nationwide study comprising essentially the entire population of Taiwan in 2010. Using genealogy and linked health information derived from a comprehensive database, we determined familial clustering of SLE by estimating risks of SLE according to specific affected kinship and assessed the relative contribution of genetic, shared and non-shared environmental factors to SLE susceptibility. In addition we estimated the relative risks of other autoimmune diseases associated with a family history of SLE.

METHODS

Study population

This study was approved by the Institutional Review Board of the Chang Gung Memorial Hospital. A cohort of all individuals registered in the Taiwan National Health Insurance (NHI) in 2010 was established using data from the Registry for NHI beneficiaries, Registry for catastrophic illness patients and datasets of ambulatory care expenditures and details of ambulatory case orders, all of which are parts of the National Health Insurance Research Database. Individuals without valid insurance status were excluded from analysis. The NHI coverage rate is over 99.5% in 2010.[19]

Since 1995, the NHI Research Database has recorded gender, date of birth, place of residence, details of insurance (employment categories, sum of insurance amount, enrolment and discharge date), family relationships, vital status and details of clinical information including dates of inpatient and outpatient visits, medical diagnoses, medical expenditures, prescription details, vaccination status, examinations, operations and procedures. The National Health Research Institute acquires all the data from the Department of Health and Welfare and implemented the data into an electronic database. All the information is linked using a unique personal identification assigned to each resident in Taiwan. To ensure confidentiality, unique

personal identification is encrypted before releasing the data to researchers but the identification remains unique for each beneficiary in the database to facilitate internal linkage of records.

Methods to identify first-degree relatives have been reported before.[20] In brief, the registry of beneficiaries contains the identifiers of the relationships between the insured person (who paid the insurance fee) and his/her dependents. Only blood relatives and spouses are eligible for dependents of an insured person. A birth certificate issued by the medical facility who delivered the child or a DNA parentage testing for those who were not born in medical facilities is required for a child to register as a dependent of their parents. This allows us to establish family relationships (parents, offspring, full siblings, twins and spouse) using the identifiers and unique personal identification of parent, grandparent, children, grandchildren and spouse. In general, parent-offspring relationships and spouses can be identified directly. An algorithm allowing indirect identification of parent-offspring relationship is also used to maximise possible family links (for details please refer to supplement).[20] Full siblings of an individual were identified if they had the same parents. Twins were full siblings with the same date of birth (± 1 day) but twin zygosity cannot be derived from the database. To consider the correlation among subjects from the same family, individuals were grouped into families according to

their relationships (for details of pedigree assembly please refer to supplement).

Among 28,402,865 beneficiaries in the NHI (both alive and deceased during the period between 1995 and 2010) 8,186,069 individuals were registered alone without any identifiable relative. The remaining 20,216,796 individuals were classified into 4,229,301 families. Overall, 21,009,551 parent-child relationships, 17,168,340 full sibling pairs and 342,066 twin pairs were identified. Note that each individual may appear multiple times in different categories of family relationships depending on family structure.

Ascertainment of SLE and other autoimmune diseases

In Taiwan, patients with suspected autoimmune diseases are referred to specialists for diagnosis and treatment. Patients with diagnoses of SLE, and autoimmune diseases included in this study (rheumatoid arthritis, systemic sclerosis, primary Sjögren's syndrome, idiopathic inflammatory myositis, type I diabetes mellitus, multiple sclerosis, myasthenia gravis, inflammatory bowel diseases and vasculitis) are entitled to waive medical co-payment. Diagnostic information is sent to the insurance administration for a review by commissioned expert panels to confirm the diagnosis before approval of waivers. In general, the panel reviews the diagnosis in compliance with the updated classification criteria. For instance, the American College of Rheumatology (ACR) revised criteria for classification of SLE were used to assist the

review of certificate applications for SLE.[21,22] The Registry for Catastrophic Illness Patients contains information on these patients with unique personal identification, diagnosis, demographics, application date, diagnosing physician and hospital and other administration data. We used this registry to identify patients with SLE, rheumatoid arthritis, systemic sclerosis, primary Sjögren's syndrome, idiopathic inflammatory myositis, type I diabetes mellitus, multiple sclerosis, myasthenia gravis, inflammatory bowel diseases and vasculitis (for details please refer to supplements).

Covariates

We considered age, gender, occupation categories, income level quintiles and level of urbanisation of residence and family size that might confound or modify the familial associations. Details for socioeconomic factors were summarised in supplements.

Statistical analysis

The prevalence of SLE was calculated for the general population and for individuals with affected first-degree family members. We calculated relative risks (RRs) of SLE as the adjusted prevalence ratios between first-degree relatives of an individual with SLE and the general population. The RR estimated in this study is essentially relative recurrence risk according to the original Risch definition,[24] which was the prevalence ratio between individuals with a specific type of affected relative and the general

population. Several established methods are available for the estimation of prevalence ratios, including the Breslow-Cox proportional hazards models,[25] log-binomial[26] and robust Poisson methods.[27] Cox proportional hazards models are a well-recognised statistical technique to handle censored survival data and estimate instantaneous hazards ratios based on varying follow-up time. Breslow adapted the Cox models to estimate prevalence rate ratios in a cross-sectional study by applying an equal follow-up time for all subjects.[25] This method has been proved to produce consistent estimates for prevalence ratios close to true parameters.[26,28] The Cox model assumes independence between subjects. However, family members naturally cluster with each other. Both the marginal model and the shared frailty model are designed to handle bias caused by within-family clustering. While the shared frailty model estimates cluster-specific hazard functions before producing joint hazard function, the marginal model focuses on the population-averaged hazard function.[29] Previous studies have documented the comparability of RR and 95% confidence intervals (CIs) between the marginal model (given a robust sandwich method to adjust CIs)[30] and the frailty model.[29] In addition, a previous study suggests that the marginal model produces a more precise parameter if the proportion of families containing discordant pairs of disease is low as is the case in our study.[31] The RR was adjusted for age, sex, socioeconomic factors and family size. This approach has been applied before and

validated previously in other diseases.[32]

We calculated RRs and tetrachoric correlations for individuals with an affected first-degree relative of any kinship and also for individual kinship (parent, offspring, sibling and twin). As kinship and sex of the affected relative may also influence familial risk, we fitted models separately according to kinship and sex of affected relatives (mother, father, daughter, son, sister, brother, twin sister and brother). We excluded twins from the sibling analyses. In addition to first-degree relatives, we also estimated RR for spouses. The RR was estimated for the number of affected first-degree kinships (father, mother, son, daughter, brother, sister). In this model, we compared risk of SLE in individuals with one or two affected first-degree relatives with the risk in the general population. To measure the degree of similarity in different types of relatives, we estimated tetrachoric correlations for each category of first-degree relationships stratified by sex of SLE patients and their relatives, assuming that there is a continuous normally-distributed liability underlying the diagnosis of SLE.

Heritability was defined as the proportion of phenotypic variance that is attributable to genetic factors and the familial transmission is the proportion of genetic and shared environmental contribution. Familial transmission and heritability can be calculated using the polygenic liability model to calculate both measures.[33-35] This model assumes a normally distributed liability of disease resulting from small and additive

influences from a large number of unspecified genes and environmental factors. The liability of the affected individuals is greater than a critical threshold, the value of which can be determined with the information of the disease prevalence in the affected and the general population.

The familial transmission is the function of the difference of normal deviation of the threshold from the mean liability between individuals with affected relatives and the normal population (refer to supplement for full discussion of the model and methods).[16,20] The original model assumes zero common environmental variance and therefore familial transmission equals heritability. To account for contributions of shared environmental factors to phenotypic variance, we used the spouse as a control, assuming that spouses share the family environment but have no close genetic similarity with blood family members. We restricted family history to first degree relatives and assumed an average of two siblings in a family.

An alternative way to estimate heritability was based on comparing tetrachoric correlations, which were used as an index of phenotypic similarity, between siblings and spouses, assuming that they have similar shared environment but have 50% and 0% genetic similarity.[36] Heritability was calculated as:

$$\text{Heritability} = 2 \times (\text{tetrachoric correlation for full siblings} - \text{tetrachoric correlation for spouse}).$$

We further estimated the extent of familial co-aggregation of other autoimmune

diseases in SLE affected families by a marginal Cox proportional hazards regression model with an equal follow-up time for all subjects adjusting for age, sex, place of residence, income levels, occupation and family size. RRs of rheumatoid arthritis, systemic sclerosis, primary Sjögren's syndrome, idiopathic inflammatory myositis, type I diabetes mellitus, multiple sclerosis, myasthenia gravis, inflammatory bowel diseases and vasculitis were estimated as the adjusted prevalence ratio of specified autoimmune diseases between individuals with a first-degree relative with SLE and those without SLE family history.

All tests of statistical hypothesis were done on the 2-sided 5% level of significance. All analyses were performed using SAS v. 9.3 (SAS institute, Cary, NC).

RESULTS

SLE prevalence in individuals with affected first-degree family members versus the general population

The study population comprised of 23,658,577 individuals enrolled in NHI in Taiwan in 2010. Among them 18,283 patients had a diagnosis of SLE, giving a crude prevalence of 0.08%. Women had a significantly higher prevalence (0.14%) than men (0.02%) with a female:male ratio of 7 (Table 1). Overall 19,085,610 (80.67%) individuals had at least one known first-degree relative. The proportions of individuals in the study population with known parent, children, siblings and twins were 51.30%, 38.24%, 42.16% and 1.10%, respectively. In the general population of Taiwan in 2010, 45,718 (0.19%) individuals had at least one first-degree relative with SLE: 20,343 with affected parents, 12,435 with affected offspring, 13,115 with affected sibling and 101 with affected twins. Among the individuals with affected family members, 607 had SLE, giving a prevalence of 1.33%. For individuals with affected first-degree relatives with SLE the age-specific prevalence of SLE is significantly higher than the age-specific prevalence in the general population (Figure 1).

Relative risks for SLE in individuals with affected first-degree relatives

Prevalence (recurrence risk) of SLE in individuals with affected first-degree relatives

of specific types is shown in Table 2. Overall, having an affected first-degree relative with SLE was associated with an adjusted RR of 16.92 (15.23–18.80). Table 2 also presents adjusted RR for SLE and 95% CIs for different affected first-degree relatives stratified by sex. Overall, individuals with affected relatives of female, male and both genders had respective RRs (95% CIs) for SLE of 16.31 (14.60–18.23), 20.35 (16.01–25.87) and 65.24 (27.36–155.55). Although it seems that the gender of affected relatives did not influence RR, point estimates in table 2 suggests that there may be trends related to gender, in particular men with a male affected relative tend to have a higher RR.

The RRs (95% CIs) for SLE were associated with the degree of genetic distance between family relatives. RRs were 315.94 (210.66–473.82) for co-twins (with highest genetic similarity) of SLE patients; 23.68 (20.13–27.84) for siblings; 11.44 (9.74–13.43) for parents; 14.42 (12.45–16.70) for offspring and 4.44 (2.38–8.30) for spouses (without genetic similarity). In addition, the RRs increased with the number of types of affected first-degree relatives. Compared with the general population, individuals with one type of affected first-degree relative had a RR of 17.04 (95% CI, 15.31–18.96) and those with two or more had a RR of 35.09 (95% CI, 14.89–82.70) for SLE.

Familial resemblance and heritability of SLE

Overall, tetrachoric correlation for first-degree relatives was 0.33 (0.32–0.34). Tetrachoric correlations were substantially higher for first-degree relatives compared to those for spouses (table 2). Tetrachoric correlation (95% CI) was estimated to be 0.59 (0.54–0.64) for twins; 0.35 (0.33–0.36) for full siblings; 0.27 (0.25–0.29) for parents; and 0.25 (0.23–0.26) for offspring and 0.07 (0.02–0.11) for spouses. Using a threshold liability model we estimated the accountability for phenotypic variance of SLE was 43.9% for genetic factors (heritability), 25.8% for shared environmental factors and 30.3% for non-shared environmental factors. By comparing tetrachoric correlations between siblings and spouses, heritability was estimated to be 56.0%.

Co-aggregation of other autoimmune diseases

Table 3 presents adjusted RR (95% CI) other autoimmune diseases in individuals with affected first-degree relatives compared to the general population. The RR (95% CI) in individuals with a first-degree relative with SLE was 2.66 (2.28–3.11) for rheumatoid arthritis; 5.40 (3.37–8.65) for systemic sclerosis; 5.87 (4.89–7.05) for primary Sjögren's syndrome; 2.77 (1.45–5.32) for idiopathic inflammatory myositis; 1.68 (1.22–2.32) for type 1 diabetes mellitus; 2.58 (1.16–5.72) for multiple sclerosis; 2.95 (2.04–4.26) for myasthenia gravis; 1.39 (0.66–2.91) for inflammatory bowel diseases and 0.86 (0.43–1.71) for vasculitis.

Sensitivity analysis

We did a sensitivity analysis using rheumatologist-based diagnosis as an alternative diagnosis. Using this diagnosis, we identified 36,431 SLE patients and found a higher prevalence of SLE at 0.15%, probably due to the inclusion of patients with less severe disease of incomplete lupus. Overall, a family history of SLE was associated with a RR of 16.74 (95% CI, 15.77–17.77) for SLE. The RR (95% CI) for SLE was 22.35 (20.31–24.60) for siblings; 6.22 (4.77–8.12) for spouses. The heritability for SLE was 41.7% based on these parameters and the threshold liability model.

DISCUSSION

The pathogenesis of SLE is multifactorial including genetic, environmental factors and abnormalities of both the innate and the adaptive immunity.[37] Genetic predisposition plays a crucial role in susceptibility and environmental exposure can cause epigenetic change,[38] or trigger activation of innate and adaptive immune response to induce or accelerate the development of SLE in susceptible individuals.[39] Strong familial aggregation in SLE has been reported but to the best of our knowledge this is the first population-based study investigating the familial aggregation of SLE and co-aggregation of other autoimmune diseases in first-degree relatives of people with SLE. We found that first-degree relatives of people with SLE have a 17-fold increased risk of SLE compared to the general population and genetic relatedness is associated with the magnitude of risk of SLE. Gender differences in familial risks are not apparent despite men with a male affected relative tending to have a higher relative risk.

Heritability in this study was estimated to be 44%, which is significantly lower than previous estimates of 66%.[15,16] However, both previous studies did not find shared environmental contribution to the risk of SLE. The extensive family data in our study indicate that shared environmental factors also contribute to SLE. To estimate heritability of SLE, we compared liability threshold among individuals with an

affected sibling and individuals with an affected spouse to that of the general population. Since spouses only share the family environment but not genetic closeness, the differences in liability threshold between siblings and spouse are contributed to by heritability. We further estimated heritability using methods based on the tetrachoric correlation of disease status and this gave a slightly higher estimate of heritability (56%). The difference between estimates of heritability using different methods is probably due to the tetrachoric correlation coefficient not adjusting for other potential confounders. Therefore, our findings support the contention that familial factors are predominant contributors to SLE susceptibility and that genetic factors explain approximately half of the phenotypic variance of SLE.

Previous studies have documented familial aggregation of SLE. One study surveying 570 SLE patients in the US found that 27% of the patients had a family history of autoimmune diseases.[40] Another US study reported that 10% of SLE patients had affected first-degree relatives compared to only 1% of controls.[41] A tendency for familial aggregation of autoimmune diseases other than just SLE has also been suggested. In a multicentre study of 1,177 patients with SLE in 9 Latin American countries 97 had at least one relative with SLE and the sibling recurrence risk ratio was 29.[12] Furthermore, one of previous twin studies, comprising 107 twin pairs, found a concordance rate of 24% in monozygotic twins comparing to only 2% in dizygotic

twins.[14] The differences reported in the current study from earlier published research may be attributed to study design, including sampling, case ascertainment and analytical approach. Previous reports are often based on less robust sampling strategies and case ascertainment such as hospital records, self-reported diagnosis and disease registries, therefore limiting generalisability. In contrast, our study used the entire national population of Taiwan and the case definition of SLE and other autoimmune diseases are based on physician-diagnoses which were scrutinised by expert panels.

Although recent efforts using GWAS have identified over 30 susceptibility loci for SLE,[17] these account for less than 10% of phenotypic variation observed.[18] Previous studies generally attribute this apparent gap to (1) undiscovered genetic variances; (2) a heritable epigenetic component or structural variation;[42] and (3) gene-gene interactions among known or undiscovered loci.[43] Another possible explanation that has been suggested for complex diseases such as Crohn's disease suggests that a proportion of heritability may remain hard to detect because of contributions from rare variants.[44] Therefore our updated estimate of heritability, which is not as high as previously reported, could partly explain the gap between observed and theoretical variation.

A family history of SLE is also a risk factor for primary Sjögren's syndrome, systemic

sclerosis, rheumatoid arthritis, multiple sclerosis, myasthenia gravis and type 1 diabetes mellitus. Previous studies also report that families with SLE patients are enriched with cases of rheumatoid arthritis, autoimmune thyroiditis, systemic sclerosis and polymyositis.[12] These findings suggest that these autoimmune diseases share part of the pathogenesis of SLE but the extents of overlapping contributors to disease manifestation are different. This theory is partly supported by the findings that some immune-mediated disease risk single nucleotide polymorphisms (SNPs) are associated with multiple autoimmune diseases.[45]

There are several limitations to the present study. Firstly, the classification of cases was based purely on the diagnosis recorded in the registry of patients with catastrophic illnesses. We do not have detailed information on clinical findings, laboratory testing, and examinations to verify the diagnosis according to formal classification criteria for SLE. Nevertheless, issuance of a catastrophic illness certificate requires strong medical evidence for a diagnosis of SLE that is agreed by an expert panel, and applications for these certificates are submitted almost exclusively by rheumatologists. Therefore our case definitions are stringent. However, patients with less severe disease or incomplete lupus are not eligible for a certificate so could not have been identified as cases. Furthermore, an alternative case definition, using rheumatologist diagnosis as a case definition for SLE resulted in very similar

results to the primary analysis. Secondly, zygosity of twins is not recorded in the database; therefore we cannot estimate heritability using a classic twin study design. However, we utilised two of four methods most commonly used to estimate heritability[46] and found very similar results. Thirdly, because we estimated heritability using the threshold liability model the results are subject to the assumption that diseases result from underlying liability that is normally distributed in the population. Nevertheless, although this is a potential caveat, data on other diseases such as schizophrenia support the validity of this model.[47] Fourthly, we cannot account for the effects of assortative mating whereby spouses are more similar for a phenotype than they would be if mating occurred at random in the population. If this assortment is not negligible, heritability could have been underestimated.[48] However, our model has a theoretical limit of heritability, which cannot be higher than that of familial transmission. Therefore only if shared environmental factors were non-existent would the heritability of SLE approach previous estimates. Fifthly, this study was restricted to Taiwan and it is possible that different findings may occur in different populations and in different environments. Therefore further studies in other countries are required to determine the generalisability of the findings. Furthermore, cluster effect – correlation between family members may affect the estimate of RR and its 95% CI. There are two approaches to adjust this effect, the marginal and frailty models.[29] The

frailty model estimates a cluster-specific hazard function through latent variables common to the same cluster whereas the marginal model focuses on the marginal distribution of the function while separately modelling the association among responses from the same cluster. The current penalised partial likelihood approach to fit the frailty model fails due to a huge matrix caused by a very large number of families involved in our study (4.22 million). However, we randomly selected 10,000 families to compare the results from these two models. We found that the familial risks for SLE was 13.92 (95% CI, 6.23 – 31.14) by the marginal model and 12.24 (5.37– 26.29) by the frailty model, suggesting that the two models provide similar results. This is consistent with the literature concerning the comparability of these two models.[29,30]

In conclusion, this first nationwide family study confirms that in Taiwan a family history of SLE is one of the strongest risk factor for SLE. Differential risk associated with different kinships suggests a strong genetic component in the susceptibility of SLE. A family history of SLE also exerts an increased risk of other autoimmune diseases. These findings may help inform the design of future studies of familial and genetic risk of SLE and may also be useful in counselling families with SLE patients.

Acknowledgements section

Author contributions: Study concept and design: CK, WZ, MD; Acquisition of data: CK, LS, KY, SL; Analysis and interpretation of data: CK, MJG, LS, AMV, WZ, MD; Drafting of the manuscript: CK, WZ; Critical revision of the manuscript for important intellectual content: CK, MJG, LS, KY, SL, AMV, WZ, MD; Statistical analysis: CK, MJG, LS, WZ; Obtaining funding: CK, LS, KY, SL; Administrative, technical, or material support: MJG, LS, KY, SL, AMV, WZ, MD; Study supervision: WZ, MD, MJG.

Information on author access to data: Dr Chang-Fu Kuo had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Disclosure of potential conflicts of interest: The authors have no conflict of interest to declare.

Source of funding and support: This work was funded by the National Science Council of Taiwan (project 103-2314-B-182A-070-MY2 and 103-2314-B-182-043-MY2) and Chang Gung Memorial Hospital (project CMRPG3D1671) and was supported by the University of Nottingham in methodology and infrastructure.

Role of the sponsors: The sponsors of the study, the Chang Gung Memorial Hospital, the National Science Council and the University of Nottingham had no role in design

and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, approval of the manuscript or decision to submit the manuscript for publication.

Acknowledgement: This study is based in part on National Health Insurance Research Database data provided by the Administration of National Health Insurance, Ministry of Health and Welfare, and managed by the National Health Research Institutes. The interpretation and conclusions contained herein do not represent positions of the Administration of National Health Insurance or the National Health Research Institutes.

References

1. Tsokos GC. Systemic lupus erythematosus. *N Engl J Med.* 2011;365(22):2110-2121.
2. Alarcon GS, Roseman J, Bartolucci AA, et al. Systemic lupus erythematosus in three ethnic groups: II. Features predictive of disease activity early in its course. LUMINA Study Group. Lupus in minority populations, nature versus nurture. *Arthritis Rheum.* 1998;41(7):1173-1180.
3. Pons-Estel GJ, Alarcon GS, Scofield L, Reinlib L, Cooper GS. Understanding the epidemiology and progression of systemic lupus erythematosus. *Semin Arthritis Rheum.* 2010;39(4):257-268.
4. Rees F, Doherty M, Grainge M, Davenport G, Lanyon P, Zhang W. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999-2012. *Ann Rheum Dis.* 2014.
5. Chiu YM, Lai CH. Nationwide population-based epidemiologic study of systemic lupus erythematosus in Taiwan. *Lupus.* 2010;19(10):1250-1255.
6. Buckman KJ, Moore SK, Ebbin AJ, Cox MB, Dubois EL. Familial systemic lupus erythematosus. *Arch Intern Med.* 1978;138(11):1674-1676.
7. Cleland LG, Bell DA, Willans M, Saurino BC. Familial lupus. Family studies of HLA and serologic findings. *Arthritis Rheum.* 1978;21(2):183-191.
8. Reveille JD, Bias WB, Winkelstein JA, Provost TT, Dorsch CA, Arnett FC. Familial systemic lupus erythematosus: immunogenetic studies in eight families. *Medicine (Baltimore).* 1983;62(1):21-35.
9. Yeh HM, Chen JR, Tsai JJ, Tsai WJ, Yen JH, Liu HW. Prevalence of familial systemic lupus erythematosus in Taiwan. *Gaoxiong Yi Xue Ke Xue Za Zhi.* 1993;9(12):664-667.
10. Sestak AL, Shaver TS, Moser KL, Neas BR, Harley JB. Familial aggregation of lupus and autoimmunity in an unusual multiplex pedigree. *J Rheumatol.* 1999;26(7):1495-1499.
11. Koskenmies S, Widen E, Kere J, Julkunen H. Familial systemic lupus erythematosus in Finland. *J Rheumatol.* 2001;28(4):758-760.
12. Alarcon-Segovia D, Alarcon-Riquelme ME, Cardiel MH, et al. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. *Arthritis Rheum.* 2005;52(4):1138-1147.
13. Ramos PS, Kelly JA, Gray-McGuire C, et al. Familial aggregation and linkage analysis of autoantibody traits in pedigrees multiplex for systemic lupus erythematosus. *Genes Immun.* 2006;7(5):417-432.

14. Deapen D, Escalante A, Weinrib L, et al. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum.* 1992;35(3):311-318.
15. Lawrence JS, Martins CL, Drake GL. A family survey of lupus erythematosus. 1. Heritability. *J Rheumatol.* 1987;14(5):913-921.
16. Yang J, Visscher PM, Wray NR. Sporadic cases are the norm for complex disease. *Eur J Hum Genet.* 2010;18(9):1039-1043.
17. Deng Y, Tsao BP. Genetic susceptibility to systemic lupus erythematosus in the genomic era. *Nat Rev Rheumatol.* 2010;6(12):683-692.
18. Gateva V, Sandling JK, Hom G, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet.* 2009;41(11):1228-1233.
19. Bureau of National Health Insurance DoH, Executive Yuan. The National Health Insurance Statistics, 2010. http://www.nhi.gov.tw/English/webdata/webdata.aspx?menu=11&menu_id=296&webdata_id=1942&WD_ID=296. Accessed 15 May, 2012.
20. Kuo CF, Grainge MJ, See LC, et al. Familial aggregation of gout and relative genetic and environmental contributions: a nationwide population study in Taiwan. *Ann Rheum Dis.* 2015;74(2):369-374.
21. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;25(11):1271-1277.
22. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40(9):1725.
23. Liu CY HY, Chuang YL, Chen YJ, Weng WS, Liu JS, Liang KY. . Incorporating development stratification of Taiwan townships into sampling design of large scale health interview survey. *J health Manag.* 2006;14:1-22.
24. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet.* 1990;46(2):222-228.
25. Lee J, Chia KS. Estimation of prevalence rate ratios for cross sectional data: an example in occupational epidemiology. *Br J Ind Med.* 1993;50(9):861-862.
26. Skov T, Deddens J, Petersen MR, Endahl L. Prevalence proportion ratios: estimation and hypothesis testing. *International journal of epidemiology.* 1998;27(1):91-95.
27. Zou G. A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology.* 2004;159(7):702-706.
28. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence

ratio. *BMC medical research methodology.* 2003;3:21.

29. Biswas A. *Statistical advances in the biomedical sciences : clinical trials, epidemiology, survival analysis, and bioinformatics.* Hoboken, N.J.: Wiley-Interscience; 2008.
30. Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. *Stat Med.* 1994;13(21):2233-2247.
31. Whittemore AS, Halpern J. Logistic regression of family data from retrospective study designs. *Genetic epidemiology.* 2003;25(3):177-189.
32. Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet.* 2009;373(9659):234-239.
33. Falconer DS. The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Ann Hum Genet.* 1967;31(1):1-20.
34. Reich T, James JW, Morris CA. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Ann Hum Genet.* 1972;36(2):163-184.
35. Reich T, Rice J, Cloninger CR, Wette R, James J. The use of multiple thresholds and segregation analysis in analyzing the phenotypic heterogeneity of multifactorial traits. *Ann Hum Genet.* 1979;42(3):371-390.
36. Falconer DS. *Introduction to quantitative genetics.* 2d ed. London ; New York: Longman; 1981.
37. Gualtierotti R, Biggioggero M, Penatti AE, Meroni PL. Updating on the pathogenesis of systemic lupus erythematosus. *Autoimmun Rev.* 2010;10(1):3-7.
38. Ballestar E, Esteller M, Richardson BC. The epigenetic face of systemic lupus erythematosus. *J Immunol.* 2006;176(12):7143-7147.
39. Takvorian SU, Merola JF, Costenbader KH. Cigarette smoking, alcohol consumption and risk of systemic lupus erythematosus. *Lupus.* 2014;23(6):537-544.
40. Pistiner M, Wallace DJ, Nessim S, Metzger AL, Klinenberg JR. Lupus erythematosus in the 1980s: a survey of 570 patients. *Semin Arthritis Rheum.* 1991;21(1):55-64.
41. Hochberg MC. The application of genetic epidemiology to systemic lupus erythematosus. *J Rheumatol.* 1987;14(5):867-869.
42. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753.
43. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A.*

2012;109(4):1193-1198.
44. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet.* 2011;88(3):294-305.
45. Cotsapas C, Voight BF, Rossin E, et al. Pervasive sharing of genetic effects in autoimmune disease. *PLoS genetics.* 2011;7(8):e1002254.
46. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet.* 2013;14(2):139-149.
47. Wray NR, Gottesman, II. Using summary data from the danish national registers to estimate heritabilities for schizophrenia, bipolar disorder, and major depressive disorder. *Front Genet.* 2012;3:118.
48. Rice TK. Familial resemblance and heritability. *Adv Genet.* 2008;60:35-49.

Figure legend

Figure 1. Age specific prevalence of SLE in individuals with a first-degree relative with SLE (dashed line) and in the general population (solid line) in Taiwan in 2010.

Tables

Table 1. Baseline characteristics of individuals with affected relatives with SLE and the general population

Variables	Women			Men		
	≥1 affected relatives	General population	*p*	≥1 affected relatives	General population	*p*
No.	22,494	11,926,513	NA	23,224	11,732,064	NA
Age (years) (mean ± standard deviation)	33.9 ± 18.1	37.9 ± 20.4	<0.001	33.8 ± 18.4	37.1 ± 20.6	<0.001
SLE, No. (%)	516 (2.29)	16,385 (0.14)	<0.001	91 (0.39)	1,898 (0.02)	<0.001
Place of residence, No. (%)						
Urban	14,360 (63.84)	7,197,968 (60.35)	<0.001	13,729 (59.12)	6,737,087 (57.42)	<0.001
Suburban	6,120 (27.21)	3,209,020 (26.91)		6,663 (28.69)	3,372,637 (28.75)	
Rural	1,664 (7.40)	1,087,991 (9.12)		1,799 (7.75)	1,098,656 (9.36)	
Unknown	350 (1.56)	431,534 (3.62)		1,033 (4.45)	523,684 (4.46)	
Income levels, No. (%)						
Quintile 1	3,420 (15.20)	1,960,003 (16.43)	<0.001	3,653 (15.73)	2,117,136 (18.05)	<0.001
Quintile 2	3,524 (15.67)	1,839,576 (15.42)		3,088 (13.30)	1,495,341 (12.75)	
Quintile 3	5,536 (24.61)	3,161,293 (26.51)		6,107 (26.30)	3,135,633 (26.73)	
Quintile 4	5,089 (22.62)	2,252,173 (18.88)		5,152 (22.18)	2,294,886 (19.56)	
Quintile 5	4,569 (20.31)	2,274,656 (19.07)		4,190 (18.04)	2,163,222 (18.44)	
Unknown	356 (1.58)	438,812 (3.68)		1,034 (4.45)	525,846 (4.48)	
Occupation, No. (%)						
Dependents of the insured individuals	9,269 (41.21)	4,924,319 (41.29)	<0.001	8,977 (38.65)	4,285,015 (36.52)	<0.001
Civil servants, teachers, military personnel and veterans	776 (3.45)	401,734 (3.37)		946 (4.07)	582,717 (4.97)	
Non-manual workers and professionals	6,615 (29.41)	3,031,660 (25.42)		7,187 (30.95)	3,325,548 (28.35)	
Manual workers	4,197 (18.66)	2,612,534 (21.91)		3,925 (16.90)	2,272,550 (19.37)	
Other	1,637 (7.28)	956,266 (8.02)		2,189 (9.43)	12,66234 (10.79)	

Table 2. Relative risks and tetrachoric correlation for SLE in different kinships.

Type of affected relative	Sex of affected relative	Sex of individual	No. of cases	Prevalence (%)	Relative recurrence ratio (95% confidence interval)[a]	Tetrachoric correlation (95% confidence interval)
Any	Female	Female	455	2.27	15.98 (14.15–18.06)	0.34 (0.33–0.35)
		Male	72	0.35	20.51 (16.36–25.73)	0.32 (0.29–0.35)
		All	527	1.29	16.51 (14.78–18.42)	0.32 (0.31–0.34)
	Male	Female	66	2.62	18.25 (14.33–23.23)	0.30 (0.27–0.33)
		Male	22	0.85	49.23 (27.87–86.94)	0.36 (0.31–0.41)
		All	88	1.72	21.69 (17.17–27.40)	0.30 (0.28–0.33)
	All	Female	516	2.29	16.14 (14.44–18.04)	0.35 (0.33–0.36)
		Male	91	0.39	23.08 (18.57–28.68)	0.34 (0.31–0.36)
		All	607	1.33	16.92 (15.23–18.80)	0.33 (0.32–0.34)
Parent	Female (mother)	Female	125	1.46	13.28 (11.19–15.75)	0.25 (0.23–0.27)
		Male	25	0.26	18.10 (12.15–26.96)	0.25 (0.21–0.30)
		All	150	0.82	13.89 (11.86–16.27)	0.24 (0.22–0.26)
	Male (father)	Female	18	1.79	16.42 (10.16–26.54)	0.24 (0.18–0.29)
		Male	6	0.53	36.94 (16.75–81.44)	0.28 (0.20–0.37)
		All	24	1.12	19.02 (12.62–28.65)	0.23 (0.19–0.28)
	All	Female	143	1.50	13.60 (11.58–15.98)	0.26 (0.24–0.28)
		Male	31	0.29	20.09 (14.11–28.61)	0.27 (0.23–0.31)
		All	174	0.86	14.42 (12.45–16.70)	0.25 (0.23–0.26)
Offspring	Female (daughter)	Female	106	1.85	10.53 (8.73–12.71)	0.27 (0.25–0.30)
		Male	14	0.27	13.68 (8.12–23.06)	0.25 (0.19–0.30)
		All	120	1.10	10.89 (9.12–12.97)	0.26 (0.24–0.29)
	Male (son)	Female	20	2.36	13.62 (8.85–20.96)	0.26 (0.21–0.31)
		Male	5	0.66	32.93 (13.79–78.64)	0.30 (0.21–0.39)
		All	25	1.56	15.56 (10.61–22.83)	0.27 (0.22–0.31)
	All	Female	125	1.91	10.89 (9.17–12.94)	0.28 (0.26–0.30)
		Male	19	0.32	16.23 (10.40–25.34)	0.27 (0.22–0.32)
		All	144	1.16	11.44 (9.74–13.43)	0.27 (0.25–0.29)
Sibling	Female (sister)	Female	200	3.46	22.21 (18.40–26.80)	0.36 (0.34–0.38)
		Male	35	0.58	31.51 (22.75–43.64)	0.34 (0.30–0.38)
		All	235	2.00	23.22 (19.65–27.45)	0.34 (0.33–0.36)
	Male (brother)	Female	28	4.19	26.60 (18.57–38.12)	0.33 (0.28–0.37)
		Male	5	0.72	38.30 (11.84–123.91)	0.31 (0.22–0.39)
		All	33	2.42	27.92 (19.61–39.75)	0.31 (0.27–0.35)
	All	Female	228	3.54	22.71 (19.15–26.93)	0.37 (0.35–0.39)
		Male	39	0.58	31.55 (22.78–43.71)	0.34 (0.30–0.38)
		All	267	2.04	23.68 (20.13–27.84)	0.35 (0.33–0.36)
Twin	Female	Female	25	34.25	274.10 (177.87–422.39)	0.58 (0.52–0.64)

	(twin sister)	Male	0	0	N/A[b]	N/A[b]
		All	25	30.49	266.79 (173.36–410.58)	0.58 (0.53–0.64)
	Male	Female	0	0	N/A[b]	N/A[b]
	(twin brother)	Male	6	35.29	2682.75 (1130.46–6366.56)	0.68 (0.58–0.78)
		All	6	31.58	1329.02 (402.51–4393.21)	0.55 (0.45–0.66)
	All	Female	25	33.33	267.21 (173.07–412.58)	0.57 (0.52–0.63)
		Male	6	23.08	1381.43 (504.64–3781.62)	0.63 (0.53–0.73)
		All	31	30.69	315.94 (210.66–473.82)	0.59 (0.54–0.64)
Spouse	All	All	15	0.18	4.44 (2.38–8.30)	0.07 (0.07–0.22)

[a] Adjusted for age, gender, place of residence, quintiles of income levels, occupation and family size.

[b] Not applicable (N/A) because of no SLE cases with affected twin.

Table 3. Relative risks of other autoimmune diseases in individuals with affected first-degree relatives.

Autoimmune diseases	Sex	With affected relatives		General population		Relative risk
		No. of cases	Prevalence (%)	No. of cases	Prevalence (%)	(95% confidence interval)[a]
Rheumatoid arthritis	Female	126	0.56	29577	0.25	2.77 (2.33–3.30)
	Male	31	0.13	7905	0.07	2.29 (1.61–3.24)
	All	157	0.34	37482	0.16	2.66 (2.28–3.11)
Primary Sjögren's syndrome	Female	106	0.47	11462	0.10	5.90 (4.87–7.14)
	Male	11	0.05	1292	0.01	5.52 (3.07–9.94)
	All	117	0.26	12754	0.05	5.87 (4.89–7.05)
Systemic sclerosis	Female	14	0.062	1495	0.013	5.75 (3.42–9.66)
	Male	3	0.013	396	0.003	4.18 (1.36–12.82)
	All	17	0.037	1891	0.008	5.40 (3.37–8.65)
Idiopathic inflammatory myositis	Female	6	0.027	1260	0.011	2.74 (1.23–6.07)
	Male	3	0.013	548	0.005	2.84 (0.91–8.80)
	All	9	0.020	1808	0.008	2.77 (1.45–5.32)
Type I diabetes mellitus	Female	20	0.09	5416	0.05	1.66 (1.07–2.57)
	Male	19	0.08	4865	0.04	1.71 (1.07–2.73)
	All	39	0.09	10281	0.04	1.68 (1.22–2.32)
Multiple sclerosis	Female	6	0.027	965	0.008	3.40 (1.54–7.53)
	Male	0	0	287	0.002	N/A
	All	6	0.013	1252	0.005	2.58 (1.16–5.72)
Myasthenia gravis	Female	19	0.08	3472	0.03	3.10 (1.98–4.84)
	Male	11	0.05	2250	0.02	2.71 (1.50–4.87)
	All	30	0.07	5722	0.02	2.95 (2.04–4.26)
Inflammatory bowel diseases	Female	3	0.01	1027	0.01	1.68 (0.54–5.21)
	Male	4	0.02	1688	0.01	1.23 (0.46–3.26)
	All	7	0.02	2715	0.01	1.39 (0.66–2.91)
Vasculitis	Female	5	0.02	1843	0.02	1.40 (0.58–3.36)
	Male	3	0.01	2910	0.02	0.52 (0.17–1.61)
	All	8	0.02	4753	0.02	0.86 (0.43–1.71)

[a] Adjusted for age, gender, place of residence, quintiles of income levels, occupation and family size.

[b] Not applicable (N/A) because of no male multiple sclerosis cases with affected first-degree relatives.