Improving the Design & Analysis of Prognosis Studies in Rare Diseases

by

Siok Hoon Lily Lim

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Institute of Health Policy Management & Evaluation
University of Toronto

© COPYRIGHT BY SIOK HOON LILY LIM 2017
Abstract

Prognosis studies examine outcomes and identify factors associated with or predictive of observed outcomes. In rare diseases, prognosis studies play an especially important role in informing patients and guiding treatment. Yet the current prognosis literature is plagued by design issues limiting clarity of information.

Objectives: This thesis aims to improve on the design and analysis of prognosis studies in rare diseases. The objectives are: 1) to evaluate the quality of current Systemic Lupus Erythematosus (SLE) prognosis literature and identify areas of deficits; 2) to apply the longitudinal study design to a childhood-onset SLE (cSLE) cohort to study longitudinal evolution of organ damage and disease activity; and 3) to explore the application of modern longitudinal analytic methods in analyzing an observational cohort.

Methods: A systematic review of the literature was performed. A cSLE cohort was studied longitudinally. Longitudinal modelling of an irregular-visit-schedule observational cohort was
performed using marginal and random effect models, along with Bayesian latent class growth mixture modelling.

**Results:** I have identified study design elements at high risk of bias within the SLE literature. Using the longitudinal design, I demonstrated that cSLE patients accrued damage throughout their disease courses and determined prognostic factors with clear temporal predictive relationships with damage evolution. I applied latent class technique to group patients’ disease activity trajectories into more homogeneous classes and identified early factors that predicted patterns of disease activity evolution.

**Conclusions:** I have demonstrated that the design and analysis of prognosis studies can be improved in rare diseases. I identified areas of deficits in prognosis study design. I have successfully applied the longitudinal study design and modern longitudinal analytic techniques to provide more precise prognostication in cSLE. Given this body of work, we now have the tools to design robust prognosis studies to achieve maximal gain in clarity of information.
Acknowledgements

First, I thank God who opened the first door that led me on this path. I thank God for bringing me to SickKids in Canada 9 years ago and the opportunities to come into contact with so many amazingly intelligent and inspiring friends and mentors. I thank Him for finding me a church family who supported me every step of the way through many trying times and who prayed for me.

I am thankful to my mentor Dr Earl Silverman. You taught me meticulousness and integrity in research and supported me in so many ways throughout this journey.

I am grateful to my thesis supervisor, Dr Brian Feldman, for always challenging me to think and innovate. You calmed me down every Friday at our meetings by refocusing me, and teaching me to balance work and life.

I am thankful to my most excellent thesis committee members, Dr Eleanor Pullenayegum and Dr Dafna Gladman. Thank you, Eleanor, for helping me with all the complex statistics and coding in this thesis. You have a gift for explaining advanced statistics simply. Thank you Dr Gladman for being the voice of reason during my committee meetings, and for helping me get access to the adult data in this thesis.

I am thankful to the Division of Rheumatology for training and supporting me. Thank you to Dr Deborah Levy, who supported me in many ways during my years in the Lupus program. Thank you to the generations of Lupus fellows: Senq Lee, Sylvia Kamphuis and Julie Barsalou, for being great friends and compatriots in research, clinical work and play. Thank you Dr Rahim Moinnedin who taught me longitudinal modelling, and took time to impart not only academic knowledge but also life wisdom. I especially thank you for jumping in and giving assistance when I was stuck with coding!
Finally, I thank my family in Singapore, my mother and sisters, who supported my decision to pursue my dreams in Canada. This dissertation is dedicated to you.
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................................................ IV

LIST OF TABLES .................................................................................................................................................. VIII

LIST OF FIGURES .................................................................................................................................................... X

LIST OF APPENDICES ........................................................................................................................................ XI

CHAPTER 1 .......................................................................................................................................................... 1

BACKGROUND AND RATIONALE .................................................................................................................. 1

1.1 THE ROLE OF PROGNOSIS RESEARCH IN CLINICAL PRACTICE .......................................................... 1
1.2 SYSTEMIC LUPUS ERYTHEMATOSUS: CHALLENGES IN STUDYING A DISEASE OF A THOUSAND FACES .... 5
1.3 CHILDHOOD-ONSET SLE: A MODEL OF RARE DISEASE FOR IMPROVING THE DESIGN OF PROGNOSIS STUDIES .... 7

CHAPTER 2 .......................................................................................................................................................... 12

AIMS AND HYPOTHESES ................................................................................................................................. 12

OVERARCHING AIM ............................................................................................................................................. 12
2.1.1 HYPOTHESIS 1 ......................................................................................................................................... 12
2.1.2 METHODOLOGICAL AIM 1 ................................................................................................................... 12
2.2.1 HYPOTHESIS 2 ......................................................................................................................................... 13
2.2.2 METHODOLOGICAL AIM 2 ................................................................................................................... 13
2.2.3 CLINICAL AIM 2 ..................................................................................................................................... 13
2.3.1 HYPOTHESIS 3 ......................................................................................................................................... 14
2.3.2 METHODOLOGICAL AIM 3 ................................................................................................................... 14
2.3.3 CLINICAL AIM 3 ..................................................................................................................................... 14

CHAPTER 3 .......................................................................................................................................................... 16

A SYSTEMATIC REVIEW OF THE QUALITY OF PROGNOSIS STUDIES IN SYSTEMIC LUPUS ERYTHEMATOSUS ................................................................................................................................. 16

CHAPTER 4 .......................................................................................................................................................... 36

FROM CHILDHOOD TO ADULTHOOD: THE LONGITUDINAL TRAJECTORY OF DAMAGE IN CHILDHOOD-ONSET SYSTEMIC LUPUS ERYTHEMATOSUS PATIENTS .................................................. 36

CHAPTER 5 .......................................................................................................................................................... 65

FROM CHILDHOOD TO ADULTHOOD: LATENT CLASSES OF DISEASE TRAJECTORIES IN CHILDHOOD-ONSET SYSTEMIC LUPUS ERYTHEMATOSUS PATIENTS ........................................... 65
CHAPTER 6 ....................................................................................................................... 86

SYNTHESIS .................................................................................................................... 86

6.1 WHAT HAS BEEN ACHIEVED? ................................................................................ 86
6.2 WHAT HAVE I LEARNED? ...................................................................................... 87
6.3 LIMITATIONS AND FUTURE DIRECTIONS ...................................................... 90
6.4 CONCLUSIONS ..................................................................................................... 96

7 APPENDICES .............................................................................................................. 97

7.1: APPLYING MODERN LONGITUDINAL ANALYTIC METHODS TO AN IRREGULAR VISIT SCHEDULE OBSERVATIONAL COHORT ........................................................................... 97

7.2: SUPPLEMENTARY MATERIAL FOR APPENDIX 1 ......................................... 124

7.3 SLE MEASURES .................................................................................................... 141

7.3.1 THE AMERICAN COLLEGE OF RHEUMATOLOGY/ SLE INTERNATIONAL COLLABORATING CLINICS SLE DAMAGE INDEX (SDI) ......................................................................................................................... 141
7.3.2 THE SLE DISEASE ACTIVITY INDEX 2000 (SLEDAI2K) ................................................................. 142

7.4: SUPPLEMENTARY MATERIAL FOR CHAPTER 3 ............................................. 143

7.5: SUPPLEMENTARY MATERIAL FOR CHAPTER 4 ............................................. 145

7.6: SUPPLEMENTARY MATERIAL FOR CHAPTER 5 ............................................. 151

8 REFERENCES FOR ALL CHAPTERS ....................................................................... 156
List of Tables

Table 3.1: Types of SLE prognosis studies reviewed

Table 3.2: Risk of bias of SLE prognosis studies

Table 4.1: Demographics and baseline characteristics of 473 childhood-onset SLE patients

Table 4.2: Damage incurred by childhood-onset SLE patients (by domains)

Table 4.3: Time invariant/baseline predictors of the average population damage trajectory in
patients with childhood-onset SLE

Table 4.4: Time varying/disease course predictors of the average population damage trajectory
in patients with childhood-onset SLE

Table 4.5: Multivariable model predicting the longitudinal damage trajectory of childhood-onset
SLE patients

Table 5.1: Demographics and baseline clinical characteristics of patients in the most probable
latent classes of SLE activity (DisAct)

Table 5.2: Results of univariable screen for predictors of membership in latent SLE DisAct
classes (class 5 as reference)

Table 5.3: Specific effects of univariable membership predictors of SLE DisAct trajectory
classes (variable class of reference)

Table 5.4: Results of multivariable model for membership predictors in latent SLE (DisAct)
trajectory classes (Class 5 is the reference class)

Table 7.1.1: Predictors identified from the joint multivariate model of MDAS and SDAS

Table 7.1.2: Overview of the 4 modern longitudinal analytic methods

Table 7.2.1: Baseline characteristics and follow-up information of study JDM patients

Table 7.2.2: Mean posterior probabilities of belonging in the 3 latent classes of JDM
Table 7.4.1: Basic characteristics of reviewed and non-reviewed studies

Table 7.5.1: List of items considered in deriving weights for the weighted GEE analysis

Table 7.5.2: Results of univariable analysis of individual SLEDAI2K items for the damage trajectory

Table 7.6.1: Ethnicities of cSLE patients by latent classes

Table 7.6.2: Results of univariable screen for baseline predictors of SLE DisAct trajectory (different reference classes)

Table 7.6.3: Results of multivariable model predicting class membership in latent SLE DisAct trajectory classes (class 5 as reference class)
List of Figures

Figure 1.1: Comparing the Information from Before-After Studies and True Longitudinal Studies

Figure 3.1: Flow diagram for selection of articles

Figure 4.1: Study population of childhood-onset SLE patients

Figure 5.1: Plots of latent SLE disease activity (DisAct) in cSLE by subclasses

Figure 5.2: Damage trajectories of cSLE by disease activity classes

Figure 7.1.1: Population-averaged modified DAS trajectory in JDM patients (representation of GEE)

Figure 7.1.2: Plot of all JDM patients’ individual and the population-averaged modified DAS trajectories (representation of MRM)

Figure 7.1.3: Effect of baseline (time-invariant) modified DAS on the modified DAS trajectory in JDM patients (MRM)

Figure 7.1.4: Latent classes of JDM disease activity trajectories (LCTA)

Figure 7.1.5: Musculoskeletal and skin disease activity trajectories in JDM (joint multivariate modeling)

Figure 7.2.1: Plot of all JDM patients’ modified DAS trajectories

Figure 7.5.1: Plot of damage trajectories in cSLE patients

Figure 7.5.2: Effects of time-invariant predictors in predicting the population-averaged trajectory in childhood-onset SLE patients

Figure 7.6.1: Individuals’ and mean population trajectories of latent SLE disease activity (DisAct) by latent classes
List of Appendices

**Appendix 7.1:** Applying Modern Longitudinal Analytic Methods to an Irregular Visit Schedule Observational Cohort

**Appendix 7.2:** Supplementary material accompanying “Methods for Analyzing Observational Longitudinal Prognosis Studies for Rheumatic Diseases: A Worked Example Using a Clinic-Based Cohort of Juvenile Dermatomyositis Patients”

**Appendix 7.3:** SLE measures

**Appendix 7.4:** Supplementary material accompanying “A Systematic Review of Quality of Prognosis Studies in Systemic Lupus Erythematosus”

**Appendix 7.5:** Supplementary material accompanying “From Childhood to Adulthood: The Longitudinal Trajectory of Damage in Childhood-onset Systemic Lupus Erythematosus”

**Appendix 7.6:** Supplementary material accompanying “From Childhood to Adulthood: The Latent Classes of Disease Trajectories in Childhood-onset Systemic Lupus Erythematosus”
Chapter 1

Background and Rationale

1.1 The Role of Prognosis Research in Clinical Practice

Prognosis research describes and explains future outcomes of a disease in relation to current diagnostic and treatment practices\(^1\). High quality prognosis research should address three issues: causes of disease progression, prediction of risk in individuals, and individual response to treatment\(^2\). Prognosis research is research in “real life”, reflecting the spectrum of patients whom clinicians see in practice. In contrast, trial patients are often very unlike the regular patients in our clinical practice. Studies have shown that only 6-34\% of rheumatoid arthritis patients in routine care would have met the criteria to participate in the early biologics trials\(^3\)-\(^5\). Information gleaned from trial patients therefore cannot necessarily be generalizable to the whole patient population. By potentially involving the full spectrum of patients, prognosis research can provide more generalizable information about patients’ outcomes.

The scope of prognosis research is very wide. In fact, prognosis research can be thought of as encompassing many questions from the perspectives of the patient (e.g., what are the disease outcomes after diagnosis, including patient-reported outcomes), the clinician (e.g., what treatment to give, who to give specific treatments to, what outcomes to expect), the scientist (e.g., what kind of cellular or molecular variations lead to different disease phenotypes or responses to treatment) and the policymaker (e.g., how cost effective are current healthcare provisions). In short, prognosis research encompasses a spectrum of studies from the classical outcomes of prognostic factor identification studies to biomarker research studies to health services research studies. This wide scope of prognosis research might have contributed to a less
developed visibility for prognosis research as compared to therapeutic (especially randomized trials) or diagnostic studies.

Although prognosis studies are very common, they are not considered robust evidence. In systematic reviews of evidence, observational evidence (other than the prospective cohort design), has been relegated to lower levels in the hierarchy of evidence. As prospective longitudinal cohorts such as the Framingham cohort are rare and very expensive to maintain, most prognosis research is considered lower level evidence. This neglect of prognosis research is perhaps best reflected in funding allocations. The National Institutes of Health (NIH) is one of the largest global funders of health research. A search of the NIH intramural funding history shows that for every prognosis study funded, another 2.5 therapeutic studies and another 11.5 diagnosis studies were funded.

There is evidence for poor quality in prognosis research. Cancer literature is one of the most active fields of prognosis research. The roles of oncogenes and tumor suppressor genes are well-recognized in the causative pathways of cancers. However, their roles in prognosis, e.g., in affecting outcomes such as treatment response and recurrence, are less clear. Changes in p53, a tumor suppressor gene in bladder cancer, are well-known. Many studies have evaluated p53 changes as a prognostic factor for outcomes of bladder cancer, but the evidence has been conflicting. A group of investigators subsequently conducted a meta-analysis of p53 prognosis studies in bladder cancer, including more than 10,000 patients from 168 reports, to summarize results from this body of research. In the end, the investigators could not conclude that p53 contributed to progression, recurrence, or mortality in bladder cancer. The quality of evaluated studies was poor on average. The most basic patient demographic and clinical factors were not reported in 60-83% of studies. Just one-third used multivariable analysis, and few among those
reported the factors used in the models. Sample sizes were too small for the kind of differences that the studies purportedly sought to demonstrate. *Such poor quality in prognosis studies impedes progress in our understanding, rendering data unacceptable for use in helping patients in clinical practice.*

**Several reasons for poor quality in prognosis research have been postulated.** It has been suggested that many prognosis studies *have been created opportunistically*, based on sample or data availability. As prognosis studies attract less funding and are often set up opportunistically, they are *less likely to have undergone and benefitted from external peer view to improve study rigor, prior to conduct of the studies.* In their 2009 recommendations to improve the quality of prognosis studies, Hemingway et al recommended the publication of protocols for prognosis studies to outline clearly the questions, the biomarkers to be assessed, the outcomes of interest, and the quality of data and pre-specified analytic plans. Biomed Central has provided such an open resource, inviting submission of any kind of study protocol (including prognosis studies) for the possibility of peer review. There, study protocols without funding or ethics approval will be peer reviewed; proposals submitted can be for proposed or ongoing studies. In 2016 (up to the summer), only 9% (26 of 300 submitted studies) were prognosis studies— and none from rheumatology. Despite the burgeoning number of biomarker studies, only two were biomarker studies (0.6% of the total, 7% of all prognosis studies). Perhaps as a result of the nature of study set-ups and the lack of external peer review, *many prognosis studies may not even have study protocols.* In an analysis of 83 studies that evaluated the predictive effect of C-reactive protein (CRP) in cardiovascular disease, only two alluded to a study protocol and none had pre-specified analytic plans. Without a set protocol or analytic plan, investigators might change their questions, analyses, or reports as driven by the data. This can only be exacerbated by the
potential bias of publishers to accept studies that report positive results in literature\textsuperscript{10}.

In the past 15 years, there has been an increased funding and research interest in using biomarkers or “omics” technology to predict outcomes and personalize treatment across many fields, including rheumatology\textsuperscript{11}. Unfortunately, biomarker studies often neglect basic tenets of good clinical epidemiology study designs and neglect reporting basic clinical information (i.e., history, physical exam, investigations)\textsuperscript{8,12,13}. Pragmatically, before further work or investment, any novel prognostic biomarker should at least demonstrate a performance superior to basic clinical predictors, either in terms of earlier or more precise prediction of outcomes. How else can we justify paying more if the simple basic clinical information will provide equivalent prognosis information? All of the above factors contribute to the common non-reproducibility and poor value of information derived from prognosis research. As a result, the potential of prognosis studies to change practice has not been fully realized.

As a pediatric rheumatologist, I work regularly with rare childhood-onset diseases. Since randomized trials are uncommon for most of the diseases that I treat, prognosis literature can be very important to my practice, whether by identifying individuals at risk of worse outcomes or by informing possible treatment modifications. However, the literature has been challenging as a source of evidence to inform practice. Problems associated with low quality in prognosis research in cancer literature might be present in rheumatology as well. In addition, sample sizes are often small. I have therefore undertaken this thesis, using childhood-onset SLE (cSLE) as a disease model, to improve the design and interpretability of prognosis studies in rare diseases.
1.2 Systemic Lupus Erythematosus: Challenges in Studying a Disease of a Thousand Faces

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disease. In Canada, SLE afflicts an average of three (range 2.6-5.0) per 1000 individuals of all age groups\textsuperscript{14}. Therefore, approximately 110,000 Canadians (2016 population census) are currently living with SLE\textsuperscript{15}. There is a predilection for females, with an average of five times more females affected than males across all age groups\textsuperscript{14}.

**SLE is a disease of a thousand faces.** SLE has a tremendously wide spectrum of possible clinical manifestations, involving almost any organ or system in the body, e.g., skin, joints, blood, brain, heart and kidneys. Individuals have different combinations of manifestations in different degrees of severity, resulting in diverse clinical phenotypes. As well, SLE disease activity changes throughout a patient’s lifetime; it can be chronically active, remitting-relapsing, clinically quiescent but serologically active, clinically active but serologically quiescent, or in remission\textsuperscript{16-18}. Patients with similar disease phenotypes can also have highly dissimilar clinical activity trajectories. **SLE is therefore a highly heterogeneous disease.**

**SLE is associated with premature morbidities.** SLE and its treatment are associated with irreversible damage to the function of the underlying afflicted organ systems. Lupus nephritis can result in end-stage renal failure. Studies have demonstrated that the risks of myocardial infarction in young SLE women are increased 50 fold compared to unaffected women within the 35-44 years old stratum\textsuperscript{19}. As well, chronic inflammation in SLE has been mechanistically linked to accelerated atherosclerosis\textsuperscript{20-22}. Treatment of SLE is commonly associated with significant adverse effects, e.g., steroid-induced osteoporosis and osteonecrosis\textsuperscript{23-30}. **SLE patients therefore suffer a high burden of morbidity because of the underlying disease and its treatment.**
**Prognosis studies are the most commonly published clinical studies in SLE.** In a MEDLINE search (1996-2016) of systemic lupus erythematosus (SLE) and “diagnosis,” “therapy,” and “prognosis,” using filters with the best balance of sensitivity and specificity, the number of articles were 1782, 692 and 5217 respectively. However, despite the multitude of prognosis studies, most aspects of SLE management are not stratified by prognostic factors.

**The current SLE prognosis literature has several common design problems.** Most SLE prognosis studies are *cross-sectional* in design. Outcomes of patients with widely varying disease durations are reported as summary measures; as such, results are challenging to interpret and apply\textsuperscript{25,28,31,32}. Even when durations of disease are considered, outcomes are reported from a *single occasion of outcome ascertainment*, e.g., the 5 or 10-year outcomes\textsuperscript{33,34}. Such an approach runs the risk of immortality bias, i.e., patients had to have survived or continued follow-up to the time of outcome ascertainment to be included in the study\textsuperscript{35}. While the single-occasion outcome ascertainment may be sufficient when studying an irreversible state, e.g., mortality or end-stage renal failure, this is *insufficiently informative for outcomes that fluctuate and change over time*, e.g., disease activity or health-related quality of life. Longitudinal changes in such outcomes constitute patients’ disease trajectory experiences. This disease trajectory experience constitutes the reality for our patients who live with their diseases daily. *Patients’ disease trajectories cannot be inferred through outcome measurements on a single occasion*. However, this information about longitudinal disease course has been conspicuously absent in the prognosis literature.

**Current SLE prognosis studies have not been able to assess the prognostic impacts of time-varying factors**, such as treatment exposure. Prognostic factors that emerge later in disease course, e.g., exposure to cyclophosphamide, either become simplified into “ever
exposed” or limited to exposed within a short, defined duration before the single-occasion outcome ascertainment. Neither approach is ideal, as the former reduces the specificity and precision, and the latter may run the risk of missing the possible prognostic factor completely. *Lack of evidence in the current SLE literature might be due to deficiencies in study designs commonly used and not due to genuine deficits of evidence.*

**The intrinsic heterogeneity of SLE in individuals’ clinical phenotypes and disease trajectories has not been addressed in current prognosis studies.** Implicit in the design of many SLE prognosis studies is that the patients have similar disease experiences. Results are usually presented as a summary measure. Analytic methods do not account for individuals’ heterogeneity or the possibility of subpopulations with different disease trajectories or outcomes beyond the different clinical manifestations. *This results in reduced clarity of the evidence obtained from the current SLE prognosis literature.*

**1.3 Childhood-onset SLE: A Model of Rare Disease for Improving the Design of Prognosis Studies**

**Childhood-onset SLE (cSLE) is an example of a rare disease**\textsuperscript{14,36}. SLE has a prevalence of about 3 per 1000 individuals in Canada, while cSLE makes up 10-15% of all SLE; this puts the prevalence of cSLE at approximately 3 per 10,000 individuals\textsuperscript{14}.

**Survival of cSLE has improved tremendously in the last 50 years**\textsuperscript{37}. Most cSLE children in the developed world (85%) will now survive > 10 years after diagnosis\textsuperscript{38-41}. It therefore has become the norm for most cSLE children to survive into adulthood. Information about the outcomes of cSLE patients in the long-term, especially into adulthood is therefore important to inform patients and their families as well as adult rheumatologists taking over their care.

**Information about the adult outcomes of cSLE patients is sparse in the current**
literature. The largest reports to date of disease-specific outcomes included fewer than 100 cSLE patients from a group of investigators in California\textsuperscript{42,43}. There have been other studies of cSLE patients, including patients in adulthood, but sample sizes were much smaller than the California studies\textsuperscript{38,44,45}.

**cSLE studies in the current literature share common design problems that limit interpretations. Previous studies were likely biased.** Study patients were often recruited only from tertiary centers where they were treated\textsuperscript{44,45}. This might have created *selection bias*: only sicker patients who needed continuing follow-up at a tertiary center would have been studied. Some studies examined only those with more than one to five years of follow-up, introducing *immortality bias*\textsuperscript{38,44}. The California cohort was recruited from community practices, tertiary clinics as well as advertisements. This is still problematic as compared to working with an inception cohort from childhood. In the case of the California cohort, it is *impossible to quantitate the direction of bias*, as the total population of possible eligible patients is unknown. Individuals who have manageable mild disease might not enter the study, those who have died before study recruitment would not have been accounted for, and there might be those too sick to participate in the study.

**Previous studies were cross-sectional, ascertaining outcomes of patients with varying durations of disease, limiting interpretability of results**\textsuperscript{38,42-45}. All studies presented prior disease course information as summary data. Most studies used retrospective physician documented chart data\textsuperscript{38,44}. In the California studies, pediatric-age disease information and course prior to study enrolment were assessed by patient report (only renal disease information was validated against adult physicians’ charts)\textsuperscript{42,43}. *The longitudinal disease trajectories of patients up to the occasion of outcome ascertainment were not available*. It was therefore
impossible to use the information from these studies to inform cSLE patients about their future outcomes.

**cSLE is an excellent disease model for this thesis.** *cSLE is a rare disease.* Like in many rare diseases, there have been many problems and methodological issues in studying prognosis in cSLE (as noted above). Therefore, the insights gained from studying cSLE will be relevant to help improve prognosis study designs in other rare diseases. *Little is known about the longitudinal outcomes of cSLE patients; this is a gap of clinical knowledge.* Like its adult counterpart, cSLE is a chronic and highly heterogeneous disease in clinical manifestations and disease course. *To maximize the yield of prognosis information in such a rare and heterogeneous disease, a longitudinal view is necessary to provide a complete overview of disease evolution.* Information from a longitudinal view of their outcomes will be very helpful in informing patients, adult rheumatologists and policymakers, and therefore improve the long-term care of cSLE patients.

**In this thesis, I used the longitudinal design to improve the information yield of prognosis studies in rare diseases.** While many studies have labelled themselves as “longitudinal” when comparing a remote outcome to a baseline status, true longitudinal design (as defined here) involves repeated measurements of the outcomes of interest over at least three occasions in time\(^{46}\). Most “longitudinal” studies in literature are before-after or two time-point studies. Before-after studies cannot provide disease trajectory information as there is only one way of joining two points separated in time (Figure 1.1). With three or more occasions, the shape of disease trajectory becomes increasingly clear. Very importantly, true longitudinal studies require specialized longitudinal analytic methods. These methods have advanced tremendously in recent years and are now able to deal with the many problems inherent in observational
longitudinal studies (Appendix 7.1). This thesis uses the longitudinal design and cutting-edge longitudinal analytic methods to answer different kinds of prognosis questions in rare diseases (using cSLE as a disease model).

Considering the lessons learned from previous literature, I designed my prognosis studies to address current gaps of knowledge, specifically that of the longitudinal evolution of cSLE. I studied the evolution of cSLE disease course from childhood into adulthood using an inception cohort of cSLE patients, established since 1984 at The Hospital for Sick Children, Toronto, Canada. This center sees 90% of all cSLE patients within its catchment area. Patient outcomes have been collected prospectively at every clinic visit, allowing me access to childhood disease trajectory information. With the help of my collaborators from both academic and community adult rheumatology practices, I obtained their outcomes in adulthood, and reconstructed their entire disease trajectories for my thesis. I then applied advanced longitudinal analytic methods with due considerations for the underlying irregular visit schedule and problems of longitudinal studies—e.g., missing data and attrition—to clarify the prognosis of cSLE patients from childhood to adulthood (Chapters 4 & 5).
Figure 1.1: Comparing the Information from Before-After Studies and True Longitudinal Studies

1.1a) Before-after studies have two occasions of outcomes ascertainment. There is only one way of joining these outcomes. Before-after study designs do not provide disease course information.

1.1b) True longitudinal design has at least three occasions of outcomes ascertainment. Disease course information can be derived from longitudinal design. The greater the number of occasions of outcome ascertainment, the more refined is the information about outcome trajectories.

Blue and red lines denote examples of different patterns of disease courses.
Chapter 2

Aims and Hypotheses

Overarching aim

The overarching aim of this thesis is to identify the areas of deficits in rare disease prognosis studies and to suggest—through application—improved methods of study design, drawing from recent advancements in methods for studying prognosis, that can be applied in the rare disease setting to gain a more refined and complete understanding of rare diseases.

2.1.1 Hypothesis 1

1a) Using SLE as an example, I hypothesize that rare/uncommon disease prognosis studies have common methodological issues resulting in significant risks for bias.

1b) Using SLE as an example, I hypothesize that rare disease prognosis studies currently do not present longitudinal information on disease course evolution, as would be suggested by the lack of appropriate longitudinal design or analysis.

2.1.2 Methodological Aim 1

1c) To assess the quality of prognosis studies in a rare/uncommon disease and identify common methodological issues, using the example of SLE.

1d) To assess the use of longitudinal design in prognosis studies of a rare/uncommon disease using the example of SLE.

To address the above hypotheses, I conducted a systematic review of the quality of SLE prognosis studies (Chapter 4). As a primary aim, I assessed the risk of bias of SLE prognosis studies. I used an established tool, the QUality of In Prognosis Studies (QUIPS), aiming to assess
Chapter 2

Rationales & Hypotheses

the risk of bias of prognosis studies according to each of six study domains: study population, study attrition, study outcomes, prognostic factors, confounders and statistical analysis. To address the second hypothesis, I analyzed the kind of study designs – with an emphasis on longitudinal design and analytic methods – used in SLE prognosis papers.

2.2.1 Hypothesis 2

2a) Modern longitudinal analytic methods (weighted generalized estimating equation) can be successfully applied to delineate the average population disease trajectory of a rare disease cohort, with an irregular, individualized visit schedule.

2b) By applying a modern longitudinal analytic method, a precise determination of the period of prediction between a time-varying prognostic factor and disease trajectory evolution can be achieved.

2.2.2 Methodological Aim 2

2c) To demonstrate longitudinal study methods that can be applied to a rare chronic disease, using cSLE as an example.

2d) To demonstrate that the predictive relationships of time-varying prognostic factors can be studied and show their influences on the evolution of disease outcome trajectory in a rare chronic disease (cSLE).

2.2.3 Clinical Aim 2

2e) To determine the longitudinal trajectory of damage – as measured by the American College of Rheumatology/ SLE International Collaborating Clinics SLE damage index – in an inception cohort of cSLE patients.

2f) To identify the effects of baseline demographic and clinical factors, disease activity and
treatment factors as potential predictors of the longitudinal trajectory of damage in cSLE.

Using an inception longitudinal cohort of cSLE patients, I constructed the average population damage trajectory from diagnosis in childhood into adulthood. With the longitudinal data, I studied the prognostic factors that occurred over the entire disease course (i.e., time-varying factors) in addition to traditional baseline factors. By applying appropriate longitudinal analytic methods (Chapter 3), I determined not only the longitudinal trajectory of damage evolution but also the predictive influence of time-varying factors within precise periods of time (Chapter 4).

2.3.1 Hypothesis 3

3a) Application of a latent class growth mixture model to a heterogeneous rare disease will help identify more homogeneous subgroups within the total patient population.

3b) Membership in different latent disease subgroups can be predicted by baseline factors.

2.3.2 Methodological Aim 3

3c) To demonstrate that latent subclasses of disease trajectories can be identified in a heterogeneous rare disease, using cSLE as an example.

3d) To demonstrate that baseline factors can predict membership in different latent subclasses of disease trajectories, using cSLE as an example.

2.3.3 Clinical Aim 3

3e) To identify subgroups of latent disease activity trajectories— as jointly measured by the SLE disease activity index 2000 (SLEDAI2K) and prednisone exposure— in a population of cSLE patients followed from childhood to adulthood.
3f) To determine prognostic factors predicting membership in different latent cSLE disease activity subgroups.

From clinical experience, I know that even though each individual is unique, patients do seem to cluster into groups of more similar disease courses. These subclasses of disease activity trajectories are *latent* as they are *not directly observable or measurable*.

To identify the subclasses of cSLE disease activity trajectories, I applied Bayesian growth mixture modelling to the SickKids inception longitudinal cohort (Chapter 5). After identification of the latent classes, I went on to identify baseline factors that predicted membership in the different latent classes of disease trajectories.
Chapter 3

A Systematic Review of the Quality of Prognosis Studies in Systemic Lupus Erythematosus

Printed with permission, Wiley Limited.


Running head: Quality of Lupus Prognosis Studies

Lily SH Lim, MBBS, MRCPCH, FRCPC, PhD(c)\(^1,4\)

Senq J Lee, MBBS, FRACP\(^1\)

Brian M Feldman, MD, MSc, FRCPC\(^1,4\)

Dafna D Gladman, MD, FRCPC\(^2\)

Eleanor Pullenayegum, PhD\(^3\)

Elizabeth Uleryk, BA, MLS\(^5\)

Earl D Silverman, MD, FRCPC\(^1\)

\(^1\) Division of Rheumatology, Department of Pediatrics, The Hospital for Sick Children, Hospital for Sick Children Research Institute, University of Toronto, Toronto, Canada

\(^2\) Division of Rheumatology, The University of Toronto, Toronto, Canada

\(^3\) Division of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Canada

\(^4\) Departments of Pediatrics, Medicine, Institute of Health Policy Management and Evaluation, and University of Toronto, Toronto, Canada

\(^5\) Hospital Library & Archives, The Hospital for Sick Children, Toronto, Canada

Funding:
The Canadian Institute Health Research (CIHR) and the SickKids Clinician-Scientist Training Program (Canada) support Dr Lim’s salary.

**Corresponding author:**

Brian Feldman

Head, division of Rheumatology, The Hospital for Sick Children

Senior Scientist, Child Health Evaluative Sciences, The Hospital for Sick Children

Professor Pediatrics, Medicine, Health Policy Management & Evaluation, University of Toronto

555 University Ave

Toronto, ON M5G 1X8

Canada

Email: brian.feldman@sickkids.ca

Tel: 4168137654

Fax: 4168134989

**Word count** 3000

The authors do not have any financial disclosures pertaining to the above work.
Abstract

**Objective:** Prognosis studies examine outcomes and/or seek to identify predictors or factors associated with outcomes. Many prognostic factors have been identified in Systemic Lupus Erythematosus (SLE) but few have been consistently found across studies. We hypothesized that this is due to lack of rigor of study designs. This study aimed to systematically assess the methodological quality of prognosis studies in SLE.

**Methods:** A search of prognosis studies in SLE was performed using MEDLINE and EMBASE, from January 1990 to June 2011. A representative sample of 150 articles was selected using a random number generator and assessed by 2 reviewers. Each study was assessed by a risk-of-bias tool according to 6 domains: study participation, study attrition, measurement of prognostic factors, measurement of outcomes, measurement/adjustment for confounders and appropriateness of statistical analysis. Information about missing data was also collected.

**Results:** A cohort design was used in 71% of studies. High risk of bias was found in 65% for confounders, 57% of studies for study participation, 56% for attrition, 36% for statistical analyses, 20% for prognostic factors and 18% for outcome. Missing covariate or outcome information was present in half of the studies. Only 6 studies discussed reasons for missing data and 2 imputed missing data.

**Conclusions:** Lack of rigorous study design—especially in addressing confounding, study participation and attrition, and inadequately handled missing data—has limited the quality of prognosis studies in SLE. Future prognosis studies should be designed with consideration of these factors to improve methodological rigor.
Significance

- The quality of prognosis study design in SLE has been suboptimal.

- Study population, study attrition and (accounting for) confounding are 3 areas demonstrating the highest risk of bias among study domains in prognosis studies.

- Missing data has not been reported well, potentially adding on to study bias.

- Future prognosis studies in SLE should be designed with careful attention to the above domains.
Prognosis studies examine the course and outcomes of a disease \(^{51}\). They may seek to determine predictors of future events or identify etiological factors \(^{51}\). Established prognostic factors can be used in clinical trials to guide development of criteria for subject inclusion or to stratify study patients during analysis \(^{52,53}\). Prognosis studies therefore may significantly alter clinical practice.

The design of a prognosis study requires a high degree of rigor in order to generate valid results. In naturalistic cohorts, patients enter into an observational study due to the presence of several baseline characteristics, e.g. having an exposure of interest \(^{54}\). Many factors other than the exposure of interest may contribute to the occurrence of the observed outcomes. As these patients are not randomized, it is critical to consider possible sources of bias and confounding, and incorporate methods to handle these biases and confounding. Poorly designed prognosis studies can distort the apparent effect of prognostic factors and result in conflicting evidence.

Systemic Lupus Erythematosus (SLE) is a disease with a wide spectrum of manifestations and severity \(^{55}\). Prognostic factors that identify SLE patients at particular high or low risk to develop specific outcomes may help clinicians optimize their management strategies. Many prognostic factors have been reported in SLE studies but few have been consistently identified from one study to another \(^{56-66}\).

We hypothesized that the observed inconsistency among identified prognostic factors in SLE is due largely to lack of rigor of study methods. The aim of this study was to systematically assess the risk-of-bias in designs of prognosis studies in SLE.

METHODS

*Eligibility criteria and search strategy*
For this study, we focused on prognosis studies that dealt with SLE-specific clinical outcomes. We excluded the following types of studies: 1) animal studies, 2) studies investigating only cutaneous lupus (i.e., not SLE), 3) pregnancy outcome studies, 4) neonatal lupus erythematosus (NLE) studies, 5) studies in which SLE was not the main focus, 6) SLE susceptibility studies, 7) case reports or qualitative research, 8) studies in which the main aim of the study was for development or validation of measurement scales, 9) studies in which the main aim of the study was to investigate effectiveness of diagnostic or screening tests, 10) randomized controlled trials, 11) studies evaluating psychosocial issues only, 12) non-original research publications (i.e., editorials, reviews), and 13) abbreviated reports (i.e., letters to editors).

A research librarian performed the literature search for prognosis studies from 1st January 1990 to 30th June 2011. As there is no optimal strategy for identification of prognosis studies, we used 2 different strategies: the first more defined and the second less restrictive, so as not to miss possible studies. The defined search strategy used terms as suggested for the search of prognosis studies by the Altman group 51 and McMaster University Health Information Research Unit, (Hedges team)- Clinical Epidemiology & Biostatistics (CE&B) unit. We also conducted a less restrictive search by augmenting our search with specific keywords to ensure broad coverage of articles. We performed the searches in both MEDLINE and EMBASE. The search terms were adjusted according to the database used; the search strategies are included in the appendix.

After combining the searches from MEDLINE and EMBASE, duplicates, and studies meeting exclusion criteria (above) were eliminated. The remaining publications were screened by title and aim(s) to determine if it was prognostic; any reference with an ambiguous title or aim was forwarded to the abstract screening round. As it was not possible to review all eligible prognosis articles, we reviewed 150 articles for assessment of the general methodological quality.
in this field. A sample of this size allowed us to quantify the proportion of articles at high risk of bias with a precision of 8% (95% confidence interval). We used the Mersenne-Twister algorithm to select a random representative sample of 150 papers.

Data extraction and definitions

From each study, the following data was extracted directly into a database: year, journal, study design, topic, method of handling confounding, types of statistical models used and methods of reporting missing data.

The main study designs were: cohort, case-control, case-series, and cross-sectional. We defined a cohort study as one that was assembled on prognostic factors of interest and further classified as prospective (study begun before the outcomes had occurred) or retrospective (if the outcomes had already occurred at the time of study). An inception cohort referred to one in which subjects were assembled at a similar and early period of their disease. For inception cohorts, subjects needed to be included within 1 year of the occurrence of a prognostic factor or exposure of interest. A case-control study was assembled based on the presence or absence of the outcomes of interest. A case-series was assembled primarily on outcome. For our study, longitudinal refers to study designs where outcomes were measured on more than 2 occasions.

For examining whether the type of journal where papers were published was associated with methodological quality, we grouped journals into general medical or rheumatology specific journals. The 5 general medical and 5 rheumatology journals with top-ranking impact factors were identified using their most recent average 12-year impact factor rating available from the Journal Citation Report. The top 5 general medical journals (in order) were: New England Journal of Medicine, Journal of the American Medical Association, Lancet, British Medical Journal and Annals of Internal Medicine (see appendix for more information). The top 5
rheumatology journals were: Arthritis & Rheumatism, Arthritis Care & Research, Annals of the Rheumatic Diseases, Arthritis Research & Therapy and Rheumatology.

**Data assessment**

We assessed every study for clarity of the research question(s) by looking for clear definitions of the study population, prognostic factor(s) of interest and the outcome(s) of interest. This was assessed in 2 locations— 1) abstract and introduction and 2) methods. Each domain was graded as yes, partial or no.

Every study was evaluated for risk-of-bias in 6 study domains, using the QUality In Prognosis Study (QUIPS) tool. Domains assessed for bias were: study participation, study attrition, prognostic factor measurement, confounding measurement and handling, outcome measurement and statistical analysis, and presentation. Each domain was assessed with the help of 3-6 prompting questions and a final grading for the risk-of-bias, as low, moderate or high was assigned. Two reviewers (LSH, SJL) evaluated each study independently. Any discrepancy in grading was discussed and if no consensus was reached, a third reviewer (BF) was consulted for final arbitration.

For consistency of assessment, we tested the QUIPS instrument for agreement between the 2 reviewers before beginning the study. Our aim was to have an inter-rater correlation coefficient (ICC2,1) of 0.9, but a minimal agreement of 0.7 was acceptable. We calculated that we would need 17 articles to achieve this and therefore assessed 17 random rheumatic disease (not SLE) prognosis articles. After the first round of articles, we did not reach the minimum agreement for all the domains. However, after discussion and further assessment standardization, in this second round we reached our target ICC of 0.9 on 10 additional articles.
Missing data is a common methodological problem and although QUIPS considers missing data, the evaluation is not comprehensive. We therefore, adapted a recent guideline for reporting missing covariate data in prognosis studies that divided the reporting of missing data into 3 main domains: quantitation of completeness of data, approaches for handling missing data and exploration of missing data \textsuperscript{76}. In addition, we assessed whether any data exploration had been performed to discern the pattern of missing outcomes (i.e., missing at random or missing not at random) \textsuperscript{77}.

\textbf{Statistics}

The individual domains of QUIPS and other items assessed for missing data were reported as proportions. Proportions were compared using a chi-square test. Relative risk was computed for significant comparisons. All statistical tests were performed using SAS 9.2 (Cary, North Carolina, USA).

\textbf{RESULTS}

After reviewing the abstracts, 888 articles were eligible for review (figure 4.1). Of these, 150 were randomly selected for full review. This sample was fully representative of the larger group (see appendix).

\textbf{Study Designs and Journal Publications}

The cohort design was the most commonly used design (71%): retrospective (57%) more commonly than prospective (14%) (Table 4.1). Only 19% were inception cohorts. Twelve cohort studies measured longitudinal outcomes repeatedly. Of the 150 studies reviewed, 9% studied children only, 2% compared adults to children, and 89% studied adults only. Thirty-eight percent of studies were multicenter; 33% were publications from lupus collaborative research networks.
Twenty-seven percent of studies were published in one of the 10 top ranking general medical or rheumatology journals.

**Data assessment**

The study population was easily identifiable in the research question, in either the abstract or introduction section (93%) and the methods section (95%). The main prognostic factor(s) of interest was clearly identifiable in about half of the studies: 54% for the abstract or introduction and 57% for the methods section. The outcome was clearly named in the abstract or introduction for 61% and in the methods for 73%.

**QUIPS**

Of all the domains assessed by QUIPS, the 3 domains at highest risk of bias were study confounding, study population and study attrition (Table 4.2).

Study design was used to deal with confounding, through matching or stratification, in 21% of studies. During analyses, the majority (61%) did not name any confounder nor made any adjustment for the effects of confounders. Only 10% explicitly named and adjusted for confounder(s) in methods. Two percent named confounders but did not explore or adjust for their effects. Twenty-seven percent did not name confounders but explored or adjusted for their effects. Thirteen percent dealt with confounding in both the study design and analysis.

The study population domain was deemed to be at high-risk of bias in 57% of studies. The reasons for high-risk ratings were inadequate reporting of: the source population; sampling strategy for assembling population; and/or the proportion of eligible population who participated.

Study attrition was deemed to be at moderate to high risk of bias in 56% of studies. Most of the cohort studies reviewed did not report information about whether patients were lost to follow-up. Forty-two percent of cohort studies used the last known follow-up date as a study
end-point; in another 14%, this information could not be ascertained. Therefore, attrition rate could be falsely low as some of the “observed” outcomes would have been missing – due to subjects leaving a study – had a defined study end date been imposed.

The risk of bias in statistical analysis was assessed to be high in 36% of studies. Only 4 studies reported sample size calculation. Fifty-one percent of studies used statistical models to assess prognostic relationships. Of these, a conceptual framework and/or method of prognostic factor selection were defined in 71%. Fourteen studies (9%) used more than 1 kind of statistical model. In 1 study, there was a discrepancy between the model named in the methods section and in the presentation of statistical data. Longitudinal (repeated measures) modeling methods were used in 4 studies. Forty-eight studies used survival methods but 75% of these studies did not show the numbers at risk over time or report proportion of patients lost to follow-up or censored.

We assessed the effects of time-period of publication, ranking of journal and type of underlying study group in affecting study quality. The risk categories of bias for each study domain were grouped into low/moderate risk or high-risk. To assess whether the time-period of publication affected the quality of study design, we divided the publications into 1990-2000 and 2001-2011, and compared the risk of bias of study design by domain. There was a significant difference only in the confounding domain. Studies published from 1990-2000 were 32% more likely to be rated as being at high-risk of bias (RR 1.32, 95% CI 1.06—1.64) in the domain of confounding than studies published from 2001-2011. We compared studies published in the top 5 medical or rheumatology journals against those that were not, by study domains. There were significant differences in the study population and study confounding domains only. When published in top-ranking journals, study population was 34% less likely to be graded as high-risk of bias (RR 0.66, 95% CI 0.45-0.98, p= 0.02) and 36% less likely (RR= 0.64, 95% CI 0.45-0.91,
p=0.002) to be graded as at high-risk of bias for the handling of confounding. Lastly, studies performed by research networks were 55% less likely to be at high-risk of bias for the handling of confounding compared to those not conducted by research networks (RR=0.45, 95% CI 0.23-0.88). There were no other domains showing statistically significant differences.

**Missing covariate**

Fifty-two percent of studies had no missing covariate data. Of those with missing data, 14 had made apparent their means of dealing with missing data, commonly by reporting reduced values in denominators for individual covariates. Imputation was used in 1 of the studies. The possible reasons for missing covariate data was discussed in only 6 studies. Comparisons of characteristics of subjects with and without missing covariates were reported in only 2 studies.

**Missing outcomes**

Twenty-one (14%) studies used the availability of measurement of an outcome as study inclusion criterion. Of these, 71% reported the number of subjects excluded because of a lack of availability of data. The frequency of complete outcome data was reported clearly in 67%. About half (48%) of studies had no apparent missing outcome.

Handling of missing outcome data was unclear in 47% of studies with missing outcomes. Potential missing outcome data were treated as censored using actuarial methods in 27%. Missing data was omitted from the analysis in 23%. Imputation was used in 1 study. The reason(s) for missing outcome data was not discussed in any study with missing data. No study explored available data for suggestion of non-ignorable dropout.

**DISCUSSION**

We found 3 domains in SLE prognosis studies – confounding, study population, study attrition– to have the greatest proportions of studies at high-risk of bias. Missing data was present
in about half of the studies. When missing data were observed, few studies gave reasons for the missing data. Our findings provided evidence for a lack of the highest standards in the current SLE prognosis literature.

We were unable to find other studies that reviewed the quality of prognosis studies in specific diseases (SLE or others). Studies like ours can potentially help researchers improve their study designs and avoid common methodological pitfalls.

As prognosis studies do not randomize patients, they have the potential for unbalanced co-distribution of confounders with prognostic factors. In the presence of unadjusted confounding, any risk estimate could be misleading. This could explain at least partially the variability of prognostic factors identified from different studies. Possible confounders should be taken into account in study design and presentation of results.

Information about the source population from which a study population was drawn was often lacking, resulting in difficulty in assessing whether the information was applicable to the reader’s population. Sampling strategy was frequently unclear. This is problematic as certain strategies, could result in the assembly of a biased study population. Many studies were not explicit in defining possible exclusions, this can bias the study in uncertain directions and not allow for assessment of the robustness of the results.

Attrition was a common problem and it may be a greater problem than we have found as a number of studies assessed the outcome at "last follow-up", resulting in an apparent lack of attrition. While the proportion of the cohort available for evaluation at study end is important, the reason for being unavailable for follow-up is also important. Few studies reported the reasons for lost to follow-up or explored for differences between those who left and those who stayed. This information is important because if the reasons for being lost to follow-up were related to the
outcome, this type of missing data pattern—missing not at random (MNAR)—can invalidate study conclusions. This could contribute to the wide variations in prognostic risk estimates and non-reproducibility of prognostic factors between studies.

Missing data is a common problem. Many statistical modeling procedures will exclude subjects with missing data resulting a decreased population to derive risk estimates. We found the extent of missing covariates difficult to assess, as frequently the information was not reported. Prognostic models derived from a reduced population may not be representative of the population. Although many methods are now available, e.g. multiple imputation, to deal with missing data (covariate), these techniques were rarely used in SLE prognostic research. Investigators should consider using these techniques to mitigate the effects of missing covariate data.

We evaluated for possible reasons for observed differences in the rigor of study designs. Studies published before 2000 were more likely to be at high-risk of bias than those published after 2000. This might be related to increasing awareness of good research methods in recent years. As may be expected, studies published in the top-ranking journals were less likely to be at high-risk of bias in the domains of study population and confounding. Studies published by research networks were less likely to be high risk for bias in confounding likely reflecting involvement of epidemiologists and/or experts in the field.

Our study must be considered in the light of possible limitations. We have limited the language of publications to English. However, a systematic review of randomized controlled trials reported that the language of publication did not affect significantly the quality of studies. We included some publications that might have been concise reports and therefore constrained by word count. However, even in concise reports key methodological features need
to be reported as they are essential in determining the validity of the results. We did access all appendices, web-only data or references of methods where cited and considered those when assessing the risk of bias in all studies.

Lupus prognosis studies can provide highly valuable information for the care of SLE patients and information that cannot be derived from randomized trials. We identified 4 areas of study design—study population, attrition, handling of confounding and reporting of missing data—that investigators frequently did not pay sufficient attention to in the design and reporting of SLE prognosis study. Although STROBE guideline deals with study reporting, we suggest that STROBE guidelines may serve as a reference during design phase of observational studies to help improve the quality of future lupus prognosis studies. By improving the quality of prognosis studies, we will improve the management of SLE patients.
Figure 3.1: Flow diagram for selection of articles

EMBASE & MEDLINE search 12020 (English only)

Excluded at title and aim screen: 10989

1031 for abstract screen

Excluded: Reviews: 20 Not prognosis study: 123

888 eligible prognosis papers

150 random prognosis papers for review
Table 3.1: Types of SLE prognosis studies reviewed

<table>
<thead>
<tr>
<th>Study Features</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Types of Study Design</strong></td>
<td></td>
</tr>
<tr>
<td>Cohort (retrospective and prospective)</td>
<td>106 (71)</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>23 (15)</td>
</tr>
<tr>
<td>Case-control</td>
<td>9 (6)</td>
</tr>
<tr>
<td>Others</td>
<td>12 (8)</td>
</tr>
<tr>
<td><strong>General topics</strong></td>
<td></td>
</tr>
<tr>
<td>Nephritis</td>
<td>37 (25)</td>
</tr>
<tr>
<td>Cardiovascular outcomes</td>
<td>15 (10)</td>
</tr>
<tr>
<td>Thrombosis/ Antiphospholipid syndrome related</td>
<td>14 (9)</td>
</tr>
<tr>
<td>Central Nervous System</td>
<td>6 (4)</td>
</tr>
<tr>
<td>Bone-related outcomes</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Miscellaneous topics</td>
<td>73 (49)</td>
</tr>
<tr>
<td><strong>Distribution of years of publication</strong></td>
<td></td>
</tr>
<tr>
<td>1990-1995</td>
<td>22 (15)</td>
</tr>
<tr>
<td>1996-2000</td>
<td>31 (21)</td>
</tr>
<tr>
<td>2001-2005</td>
<td>44 (29)</td>
</tr>
<tr>
<td>2006-2011</td>
<td>53 (35)</td>
</tr>
<tr>
<td><strong>Types of journals</strong></td>
<td></td>
</tr>
<tr>
<td>Top-ranking medical journals</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Other medical journals</td>
<td>45 (30)</td>
</tr>
<tr>
<td>Top-ranking rheumatology journals</td>
<td>39 (26)</td>
</tr>
<tr>
<td>Other rheumatology journals</td>
<td>64(43)</td>
</tr>
</tbody>
</table>
Table 3.2: Risk of bias of SLE prognosis studies

<table>
<thead>
<tr>
<th>Study Domains and Risk of Bias</th>
<th>Number of studies (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Population</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>12(8)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>53(35)</td>
</tr>
<tr>
<td>High risk</td>
<td>85(57)</td>
</tr>
<tr>
<td>Study Attrition*</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>35(33)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>12(11)</td>
</tr>
<tr>
<td>High risk</td>
<td>59(56)</td>
</tr>
<tr>
<td>Prognosis Factor Definition &amp; Measurement</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>51(34)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>69(46)</td>
</tr>
<tr>
<td>High risk</td>
<td>30(20)</td>
</tr>
<tr>
<td>Confounding Measurement &amp; Handling</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>16(11)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>36(24)</td>
</tr>
<tr>
<td>High risk</td>
<td>98(65)</td>
</tr>
<tr>
<td>Outcome Definition &amp; Measurement</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>44(29)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>79(53)</td>
</tr>
<tr>
<td>High risk</td>
<td>27(18)</td>
</tr>
<tr>
<td>Statistical Presentation &amp; Analysis</td>
<td></td>
</tr>
<tr>
<td>Risk Level</td>
<td>Count (%)</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Low risk</td>
<td>44(29)</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>52(35)</td>
</tr>
<tr>
<td>High risk</td>
<td>54(36)</td>
</tr>
</tbody>
</table>

* Only assessed for cohort studies, N=106
Chapter 4

From Childhood to Adulthood: The Longitudinal Trajectory of Damage in Childhood-onset Systemic Lupus Erythematosus Patients

Printed with permission, Wiley Limited.

Arthritis Care & Research 2017 Jan 24. DOI: 10.1002/acr.23199.

*Lily Siok Hoon Lim*, MBBS, MRCPCH, FRCPC, PhD(c)

*Eleanor Pullenayegum*, PhD

*Lillian Lim*, MD

*Dafna Gladman*, MD, FRCPC

*Brian Feldman*, MD, MSc, FRCPC

*Earl Silverman*, MD, FRCPC

The SickKids clinician-scientist training program and the Canadian Institutes of Health Research (CIHR) jointly funded Dr Lim’s salary for her doctoral program.

**Corresponding author:** Earl Silverman
Address: 8253A- 555 University Ave, Toronto, Ontario M5A1X8, Canada
Email: earl.silverman@sickkids.ca
Telephone: 4168137550  Fax: 4168134989

**Word count:** 2882

The authors have not received any commercial financial support for this work. The authors do not have any disclosures to make regarding potential conflicts of interest for this work.
ABSTRACT

Objectives

1) To determine the longitudinal damage trajectory of patients with childhood-onset SLE (cSLE). 2) To identify baseline and disease course predictors of damage trajectory.

Methods

This is a retrospective inception cohort. Longitudinal pediatric-age data was obtained from a cSLE research database while adult-age data was obtained from either a research database or patients’ charts. Baseline factors were tested as predictors. Time-varying factors were lagged 6 to 24 months before a visit for testing their predictive effects. The longitudinal damage trajectory was modeled using a weighted generalized estimating equation (WGEE).

Results

This study cohort consisted of 473 subjects with follow-up to 26 years. 65% of patients were >18 years old at last follow-up. Cataract (14%), avascular necrosis (10%) and osteoporosis (5%) were the most common items of damage. Two patients had myocardial infarctions. Baseline features—self-reported ethnicity (Afro-Caribbean), earlier time-periods of diagnosis and presence of a life-threatening major organ manifestations (lupus nephritis III-V, cerebrovascular accidents, major organ vasculitis, pulmonary hemorrhage or myocarditis)—were associated with greater damage. Throughout the disease course, acute confusion state, lupus headache and fever predicted subsequent increases in the damage trajectory. Higher prednisone dose and exposure to cyclophosphamide also predicted subsequent increases in the damage trajectory. Antimalarial exposure was protective against an increase in damage trajectory.

Conclusions
Patients with cSLE accrue damage steadily into adulthood. Baseline factors predict greater damage and/or influence the evolution of the damage trajectory. Additionally, SLE clinical features and therapies during the course of disease predict additional changes in the damage trajectory.

Abstract word count: 250
Significance & Innovations

1) First study to use the longitudinal design to study disease evolution of cSLE patients from diagnosis in childhood into adulthood.

2) Damage accrual continued steadily throughout the disease course of cSLE patients into adulthood.

3) Throughout the disease course, higher prednisone exposure and cyclophosphamide exposure from 6-24 months before predicted increase in damage trajectory at current visit. This was the first time that specific windows of effects for disease course predictors of longitudinal trajectory was studied.

4) Hydroxychloroquine exposure from 6 months before protected against an increase in damage trajectory at current visit independently. This gives a more specific window of time to expect protective effects from hydroxychloroquine.
Patients with childhood-onset systemic lupus erythematosus (cSLE) are now surviving longer\textsuperscript{44,88-90}; however, longer survival has been associated with increased morbidity from accruing organ damage. In most cSLE studies, 40-65% of patients have sustained some degree of organ damage— as measured by the SLE International Collaborating Clinics/ American College of Rheumatology (SLICC/ACR) SLE damage index (SDI) within 10 years of diagnosis \textsuperscript{44,88,91-94}.

Little is known about adult outcomes of cSLE\textsuperscript{95}. Two studies from the United States suggest a poor prognosis with persistent disease activity requiring treatment at a mean of 13-16 years after diagnosis\textsuperscript{43,45}. However, both study populations were small, not inception cohorts and might be highly selected\textsuperscript{43,45}. More information on long-term outcomes of cSLE into adulthood is therefore needed.

Outcome studies in cSLE to date have only studied patients on a single occasion and do not provide information about the course of disease. Two patients with the same outcome might have arrived at that outcome via very different disease courses. Differences in disease courses might have vastly different disease burdens or impacts on patients’ lives. Studying outcomes at a single occasion is thus limiting. We need a more complete picture of the outcomes of cSLE patients, from childhood into adulthood.

The longitudinal study design—where patients’ outcomes are measured repeatedly (\(\geq 3\) occasions)—provide better information about how a disease evolves over time \textsuperscript{46}. Previous studies have shown faster and greater damage accrual within the first 2 to 3 years after diagnosis in cSLE patients (compared to adult SLE (aSLE) patients) \textsuperscript{90,96}. Whether this rapid accrual continues throughout the disease course or plateaus after the initial period can only be answered
by the longitudinal design. By understanding how outcomes change longitudinally over time, clinicians may modify their management strategies.

The aims of the study were: 1) to determine the longitudinal trajectory of outcome of an inception cohort of cSLE patients; 2) to identify the effects of disease activity and treatment as potential predictors of the longitudinal trajectory of damage.

METHODS

Study population and attrition

This is a single center longitudinal inception cohort study. All subjects were diagnosed and followed in the pediatric lupus clinic at SickKids from 1\textsuperscript{st} January 1985 to 30\textsuperscript{th} September 2011. This cSLE clinic follows about 60-70\% of all cSLE patients in Ontario, Canada. However, as general pediatricians and adult rheumatologists within our catchment area rarely treat or follow cSLE patients, and our inception cohort does not include patients followed and treated outside SickKids, we believe our cohort contains 90\% of patients within our catchment area. Patients were included if all the following conditions were satisfied: 1) American College of Rheumatology or the SLE International Collaborating Clinics (SLICC) classification criteria for SLE\textsuperscript{97,98}; 2) Age less than 18 years old at diagnosis; 3) Seen at SickKids ≥ 3 visits \textsuperscript{46}; 3) Not treated with continuous steroid or immunosuppressant for ≥ 3 months as another autoimmune disease, prior to initial visit to cSLE clinic.

Since the inception of our cohort, we have systematically transferred the care of our graduated patients to the care of the Lupus clinic at the Toronto Western Hospital (TWH). For patients who did not enter into the care of TWH, we: 1) Contacted their rheumatologists (if known); 2) Contacted their family doctor to ascertain their current or past rheumatologists; 3) Mailed patient self-report questionnaires to their last known addresses and 4) Advertised this
study on provincial chapters of Canadian lupus patients' websites. We attempted to obtain information on the complete disease trajectory of this population of patients from diagnosis to the study end-date.

Clinical data were collected prospectively at every visit at the cSLE clinic. A visit included a full history, examination, completion of a standardized data collection form (including disease activity and damage) and routine laboratory tests. These data have been prospectively entered into a clinical research database since the inception of this clinic.

The adult disease course information was obtained using different methods. The majority of our patients (66%) had been transferred to the TWH. Patients at the TWH Lupus clinic (85% of cSLE patients at TWH) had their clinical data entered into their research database. For patients not followed by TWH Lupus clinic, information was extracted from clinic charts (LSHL).

The research ethics boards at SickKids (1000028143), the University of Toronto, and all 5 participating institutions/ health regions (for adult data) approved this study.

**Outcome**

The primary outcome measure for this study was the SLICC/ACR SDI, a validated measure of irreversible damage in SLE. The attribution of damage is not considered relevant to scoring. This was scored at every visit for the pediatric database and when data was extracted from the charts. Patients in the pediatric cohort that preceded the publication of the SDI had it scored retrospectively in the past. The TWH database scored the SDI annually.

Patients were considered lost to follow-up if their last date of follow-up was more than 18 months from the study end date or time of death, whichever earlier (see appendix). The reasons for loss to follow-up were: moved out of the province or Canada, unable to contact, not followed by a rheumatologist, or refused to allow a review of their charts. For the 67 patients
who were lost to follow-up, we obtained one-occasion damage scores reported by 5 patients (using Lupus Damage Index Questionnaire) and for 1 patient, the most current rheumatologist scored the SDI (after study end date)\textsuperscript{102} (see Appendix).

**Predictors**

We collected baseline and longitudinal (time-varying) information for the purpose of prognostication. Time-varying information was collected at every visit throughout the course of each patient's follow-up.

The following baseline information was collected: sex, ethnicity (self-reported), age at diagnosis, time periods of diagnosis (1985-1990, 1991-2000, 2001-2011), manifestations of lupus at diagnosis, number of ACR classification criteria at diagnosis, life-threatening major organ involvement at diagnosis and autoantibody profile at diagnosis. The periods of diagnosis were divided into the decades that this cohort spanned, but also coincided with differences in management practice, from a predominantly high dose and chronic steroid regime with or without cyclophosphamide (1980s), to the more common use of second-line immunosuppressants e.g., azathioprine for major organ diseases (1990s), and to the widespread use of mycophenolate mofetil in the last decade of this cohort. There might be other unmeasured auxiliary factors such as increased expertise in the management or prevention of complications arising from treatment, that might contribute to the patients’ outcomes in the different periods of time that this factor might act as a surrogate for. We defined potentially life-threatening major organ involvement as any of the following: lupus nephritis classes III-V, cerebrovascular accident, major organ vasculitis, pulmonary hemorrhage, or myocarditis. The baseline organ manifestations were missing in 10/473 individuals and could not be checked as their earliest charts were missing.
We collected time-varying information in these categories: the SLE disease activity index 2000 (SLEDAI-2K) and therapeutic information. Therapeutic information included exact prednisone doses and use of immunosuppressant/antimalarial (yes/no). We defined significant immunosuppression as being exposed to any of the following: cyclophosphamide, azathioprine and mycophenolate mofetil. For azathioprine and mycophenolate mofetil, these were counted in the significant immunosuppression category only if used to treat serious manifestations, e.g., nephritis, CNS disease.

**Statistical Methods**

As this was a longitudinal study, there were missing data. We used multiple methods to handle and explore the data, so as to discern the nature of the missingness (see Appendix).

**Statistical analysis**

We started by plotting all patients' damage trajectories (see appendix) using the LOESS plot. We used fractional polynomials to fit the time form of the damage trajectory. The best time form of the trajectory was defined by 2 fractional polynomial terms labelled as time terms 1 and 2 in this paper (see Appendix).

We explored the nature of missingness in this cohort by modelling time-to-loss-to-follow-up using Cox Proportional Hazards. The damage trajectory was used as a time-varying predictor of loss-to-follow-up. If damage predicted lost to follow-up, this would be an indication of possibly data Missing At Random (MAR) or Missing Not At Random (MNAR). The data would be MAR if loss to follow-up was predicted by observed parameters. In MNAR, the missing outcome could be the cause for loss to follow up, or the missing outcome might be systematically different for those lost to follow-up compared to those who remained in the study, which may then result in a biased conclusion. MNAR is always a concern but cannot be
definitely proven unless the missing outcome data is available. We tried to collect the most recent damage data from those lost to follow-up to assess the nature of loss-to-follow-up (Appendix) 107.

For prediction, time-varying factors: SLEDAI-2K, laboratory tests, prednisone and immunosuppressant use, were lagged by 6, 12, 18 and 24 months prior to current visit (i.e., values from 6 to 24 months before the visit of interest were tested for their effect on subsequent outcome trajectory). We chose to lag by a minimum of 6 months as most damage items can only be scored if they persisted for at least 6 months. The maximum lagged effects were up to 2 years as we felt that candidate predictors could have delayed effects on damage as late as 2 years. Prednisone was tested as a continuous variable as exact doses were available.

The inverse-intensity-weighted Generalized Estimating Equation (WGEE) model was used to determine the overall damage trajectory (Appendix) 108. Candidate predictors were chosen based on clinical understanding of the disease and from review of current literature. Predictors were retained for testing in the multivariable model if p< 0.1 on initial screening. In the final multivariable model (other than the time form parameters), predictors were kept if p<0.05.

RESULTS

Study population and attrition

Four hundred and seventy-three patients were included in the study (Figure 1). Of these, 9 died (2%), 67 were lost to follow-up (14%). The median duration of follow-up was 5.63 (25th-75th percentile: 3.00-10.82) years, maximum follow-up was 26.3 years. The median number of visits was 18 (25th-75th percentile: 8-32). A total of 14097 visits, 3289.73 patient years was
studied. There were 25 patients with ≥ 20 years, 63 patients with ≥ 15 years, 131 patients with ≥ 10 years, and 259 patients with ≥ 5 years of follow-up.

Sixty-five percent (308/473) of subjects were >18 years at the last follow-up: median age at last follow-up was 22.2 (25th-75th percentile: 19.5–27.7, max: 41.3) years. Forty-four percent (202/463) had potentially life-threatening organ manifestations at baseline (Table 1).

Outcomes

The predicted mean population SDIs were 0.64 (5 years), 1.16 (10 years), 1.67 (15 years), 2.16 (20 years) and 2.64 (25 years).

There was evidence that those lost-to-follow-up were different from those who remained in the cohort. Increasing damage protected against loss to follow-up (HR 0.71, 95% CI 0.57-0.90). Baseline factors such as sex, ethnicity and major organ involvement were not significantly associated with loss to follow-up.

The most common damage was cataract (Table 2). The second commonest was avascular necrosis; 68 events were documented in 48 patients (maximum 2 events/patient). Thirteen instances of cardiovascular damage were documented in 10 patients: 11 cerebrovascular accidents and 2 myocardial infarctions. The median age at time of cardiovascular damage was 16.8 (25th percentile-75th percentile: 12.5-22.8, range 4.2-39.6) years. The 2 myocardial infarctions occurred at 26.3 and 39.6 years. No patient had angina, coronary bypass or peripheral vascular claudication by the end of follow-up period.

Predictors of damage trajectory

Univariable predictors of damage trajectory

At baseline, the presence of a potentially life-threatening disease manifestation predicted more rapid early increase in damage (i.e., a steeper trajectory) (Table 3, Appendix). Afro-
Caribbean patients had higher early damage compared to Caucasian and Asian patients (results not shown). Increased age and the number of baseline ACR criteria were associated with persistently greater total damage but did not affect the rate of the damage trajectory. Diagnosis in the 2000s (versus 1980s) was associated with less damage but also did not affect the rate of the damage trajectory.

Over the course of disease, each 10 mg increment of prednisone exposure from 6 to 24 months before each visit predicted an increase in damage (Table 4). Significant prior immunosuppression, in particular cyclophosphamide, also predicted an increase in damage. Prior antimalarial exposure (6-24 months) protected against an increase in damage.

Longitudinal SLEDAI-2K score did not predict changes in damage trajectory. However, the following individual SLEDAI-2K items before each visit predicted worsening in damage trajectory: a) Psychosis or acute confusional state 6-24 months before; b) Visual changes 12 months before; c) Lupus headache 6-18 months before; d) Pleurisy 18 months before; and e) Proteinuria, haematuria or fever 12-24 months before. The individual SLEDAI-2K items before each visit protected against an increase in damage trajectory were: a) New rash; b) Mucosal ulcers (6-24 months before for both); and c) Pericarditis 6 months before (see Appendix).

**Multivariable model for damage trajectory**

All significant predictors from the initial screen were then tested concurrently (Table 4). The patients diagnosed in the most recent era (2001-2011) accrued lower damage early on in their course. Patients of Afro-Caribbean ethnicity, as compared to Caucasian and Asian patients, had higher early damage and persistently higher damage trajectory. Patients with a potentially life-threatening organ manifestation at baseline, compared to those without, accrued damage more rapidly initially and continued to have more damage accrual (see Appendix).
The lagged time-varying factors that predicted an increased damage trajectory over the course of disease were: acute confusional state, lupus headache and fever. Prior mucosal ulcers or pericarditis protected against an increase in damage trajectory. Each prior 10 mg increment of prednisone (12, 24 months) and cyclophosphamide exposure (6-24 months) were both associated with increased damage trajectory. An increase in damage trajectory was averted by antimalarial exposures 6 months before each visit.

**DISCUSSION**

Prior to this study, our literature review revealed only 4 papers examining the outcomes of cSLE patients in adulthood cross-sectionally. By using the longitudinal design, we examined the trajectory of damage rather than cross-sectional views of damage accrual. We found that damage continued to accrue over the entire disease course for cSLE patients; the damage trajectory did not plateau. Although it has been shown that Afro-Caribbean patients have more damage, we showed that Afro-Caribbean patients (compared to Caucasians and Asians) had persistently greater damage trajectory starting early. Baseline potentially life-threatening major organ manifestations were associated with more rapid early damage accrual. Throughout the disease course, acute confusional state, lupus headache or fever predicted a subsequent increase in the damage trajectory. Increased doses of prednisone and use of cyclophosphamide predicted a subsequent increase in the damage trajectory. Antimalarials protected against an increase in the damage trajectory even in face of high dose prednisone and cyclophosphamide.

In a paper that examined outcomes of adults with cSLE, 8% had myocardial infarction (mean 32 years old). In our inception cohort, there were only 2 cases of myocardial infarction (0.4%). Our study designs and study populations were very different and not directly comparable. The previous study used a cross-sectional design, studying outcomes from patients
with widely varying disease durations and recruited adult cSLE patients from many different settings. Their patients also had longer disease duration (16.5 years) and were much older (mean age 31 years) when studied.

Race has been associated with differences in disease manifestations, in particular, major organ manifestations and outcomes in aSLE patients. We found Afro-Caribbean race to be predictive of a higher initial and persistently greater damage accrual explaining the higher damage observed in previous cross-sectional studies. Previous cSLE studies (smaller numbers, shorter disease durations) had not consistently found race to influence outcomes. However, some support for Afro-Caribbean race and damage in cSLE can be found in the higher rates of intensive care admissions, end-stage renal disease and death compared to patients of other races. Afro-Caribbean cSLE patients also have higher rates of renal involvement and a trend towards higher central nervous system involvement; as such that it is not surprising that they have a higher damage trajectory.

Diagnosis in the 2000s was associated with less damage but did not change the rate of the damage trajectory. Our longitudinal model has already accounted for duration of follow-up. This time period term was created to capture unmeasured differences in management in the different eras. This higher initial damage in earlier eras might be related to the following unmeasured differences such as patients being sick for longer before diagnosis (less recognition), fewer therapeutic options and less expertise in the earliest eras.

SLEDAI-2K items predicted a subsequent change in the damage trajectory within defined intervals. The lupus headache was likely due to the severity of the underlying disease as in our cohort, the majority of our patients had underlying CNS vasculitis (data not shown). Acute confusional state and fever have been identified to be associated with an increased risk of death.
in aSLE patients\textsuperscript{118,119}. A previous study found statistically significant lower likelihood for patients with mucosal ulcers to acquire pulmonary damage compared to those without mucosal ulcers\textsuperscript{120}.

Increasing prednisone dose from 12 and 24 months before increases the damage trajectory. Although some authors have attributed several items within the SDI, e.g., avascular necrosis and cataracts, to corticosteroid exposure, this might be confounding by indication as patients with severe SLE receive high dose and/or prolonged corticosteroids\textsuperscript{121}. Active SLE independent of steroid had been identified as a risk factor for avascular necrosis\textsuperscript{122}. Cataracts have been seen in individuals on low-dose steroids or who were steroid naïve\textsuperscript{123-125}. Cyclophosphamide had been previously identified as being predictive of damage\textsuperscript{94}, our study expanded this by establishing the temporal effect of cyclophosphamide.

We found that antimalarial exposure from the 6 months prior to the visit protected against an increase in damage trajectory. The beneficial effect of antimalarials in SLE has been well established but the window of protection had not been addressed\textsuperscript{118,126,127}. We were able to show that the protective effects were continued only 6 months after exposure. Therefore, in order to achieve sustained benefits, clinicians need to encourage long-term adherence to anti-malarials.

Our results must be interpreted in the light of several potential limitations. We could not obtain adult outcomes in all our patients. Those patients who were lost to follow-up could have had significantly different outcomes. Our exploration showed that those who were lost-to-follow-up showed significantly less damage accrual (before being lost) compared to those who continued. This finding suggested data was at least missing at random (missingness predictable by observed data, i.e., damage up to the point of being lost to follow-up) and this was concordant with the assumptions of the WGEE\textsuperscript{104,128}. We tried to assess for missingness not at random by
obtaining the most recent damage score of those who were lost to regular follow-up. We managed to obtain damage data on 9% of patients lost-to-follow-up and did not find any difference in the final damage accrued after adjusting for disease duration, suggesting that the data was possibly not missing not at random (missingness due to unobserved outcomes)\(^\text{104}\). However, as the numbers are small, they may not be truly reflective of all lost to follow-up. Some adult outcomes information was retrospectively extracted from patient charts. Milder damage items e.g., mild cataracts, might not have been documented and the total damage underestimated. However, the majority of our adult outcomes information was derived from data prospectively collected in a database so we are confident that information on damage accrual was accurate in the majority. We could not capture fully all possible avascular necrosis and cardiovascular events as these are capped at a maximum of 2 in SDI. Socioeconomic status and medication adherence can also contribute to disease outcomes \(^\text{129}\). However, we did not collect data specifically to address these factors. Clearly, these would be important information to collect prospectively for future studies.

This was the first inception cohort of cSLE patients longitudinally followed into adulthood. We studied the damage accrual trajectory over the entire disease course of cSLE patients and the mean damage trajectory continued to increase over time. A major life-threatening organ manifestation at diagnosis predicted more rapid initial damage accrual. Afro-Caribbean race predicted a higher damage trajectory. Throughout the course of disease, prior clinical features and therapeutic exposures predicted subsequent changes in the damage trajectory within definite timeframes. Our study has identified prognostic factors that following validation in independent cohorts, clinicians may use in their patient management strategies to modify damage trajectory evolution.
Figure 4.1: Study population of childhood-onset SLE patients

Of the 61 patients with <3 visits: 10 had only 1 visit (2 of whom died), 29 were outside the eligible study timeframe, 22 were within the study timeframe.
Table 4.1: Demographics and baseline characteristics of 473 childhood-onset SLE patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Number (%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>87 (18)</td>
</tr>
<tr>
<td>Females</td>
<td>386 (82)</td>
</tr>
<tr>
<td>Race#</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>153 (32)</td>
</tr>
<tr>
<td>Asian</td>
<td>181 (38)</td>
</tr>
<tr>
<td>Afro-Caribbean</td>
<td>60 (13)</td>
</tr>
<tr>
<td>Others</td>
<td>78 (17)</td>
</tr>
<tr>
<td>Age at diagnosis in years</td>
<td></td>
</tr>
<tr>
<td>median (25th-75th percentile)</td>
<td>14.1 (11.5-15.8)</td>
</tr>
<tr>
<td>Age at last follow-up in years</td>
<td></td>
</tr>
<tr>
<td>median (25th-75th percentile)</td>
<td>19.7 (17.2-24.4)</td>
</tr>
<tr>
<td>range</td>
<td>6.0-41.9</td>
</tr>
<tr>
<td>Time periods of diagnosis</td>
<td></td>
</tr>
<tr>
<td>1985-1990</td>
<td>39 (8)</td>
</tr>
<tr>
<td>2001-2011</td>
<td>317 (67)</td>
</tr>
<tr>
<td>No of ACR criteria at diagnosis</td>
<td></td>
</tr>
<tr>
<td>median (25th-75th percentile)</td>
<td>6 (5-7)</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Major organ involvement at diagnosis^</td>
<td></td>
</tr>
<tr>
<td>Lupus nephritis</td>
<td></td>
</tr>
<tr>
<td>Class III, IV or V</td>
<td>170 (36)</td>
</tr>
<tr>
<td>Cerebral vascular accident</td>
<td>83 (18)</td>
</tr>
<tr>
<td>Myocarditis</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Major organ vasculitis</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Pulmonary hemorrhage</td>
<td>4 (1)</td>
</tr>
<tr>
<td>Family history§</td>
<td></td>
</tr>
<tr>
<td>Autoimmune diseases</td>
<td>157 (34)</td>
</tr>
<tr>
<td>Lupus</td>
<td>45 (10)</td>
</tr>
</tbody>
</table>

*Unless otherwise stated. #South Asians were included under Asians. ^ Detailed Information available for 463 patients. §Family history was available for 463 patients for autoimmune diseases.
Table 4.2: Damage incurred by childhood-onset SLE patients (by domains) *

<table>
<thead>
<tr>
<th>Damage by domains</th>
<th>Number of patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eye</strong></td>
<td></td>
</tr>
<tr>
<td>Cataract</td>
<td>65 (14)</td>
</tr>
<tr>
<td>Retinal changes or optical atrophy</td>
<td>6 (1)</td>
</tr>
<tr>
<td><strong>Central and Peripheral Nervous System</strong></td>
<td></td>
</tr>
<tr>
<td>Cognitive impairment</td>
<td>20 (4)</td>
</tr>
<tr>
<td>Seizures</td>
<td>15 (3)</td>
</tr>
<tr>
<td>Cerebrovascular accident</td>
<td>8 (2)</td>
</tr>
<tr>
<td>Cranial/ peripheral neuropathy</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Transverse myelitis</td>
<td>2 (0.4)</td>
</tr>
<tr>
<td><strong>Cardiovascular System</strong></td>
<td></td>
</tr>
<tr>
<td>Angina or coronary artery bypass grafting</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Myocardial infarct</td>
<td>2 (0.4)</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Valvular lesion</td>
<td>6 (1)</td>
</tr>
<tr>
<td>Condition</td>
<td>Count (Percentage)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Pericarditis (chronic)</strong></td>
<td>1 (0.2)</td>
</tr>
<tr>
<td><strong>Peripheral Vascular System</strong></td>
<td></td>
</tr>
<tr>
<td>Claudication</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Venous embolism</td>
<td>3 (0.6)</td>
</tr>
<tr>
<td>Minor tissue loss (pulp space)</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Significant tissue loss</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Pulmonary</strong></td>
<td></td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Pulmonary fibrosis</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Shrinking Lungs</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Pleural fibrosis</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Pulmonary infarct or resection</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td><strong>Renal</strong></td>
<td></td>
</tr>
<tr>
<td>Reduced GFR</td>
<td>12 (3)</td>
</tr>
<tr>
<td>Persistent proteinuria</td>
<td>10 (2)</td>
</tr>
<tr>
<td>End-stage renal failure</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Infarct or resection of intraabdominal organs</td>
<td>8 (2)</td>
</tr>
<tr>
<td>Mesenteric insufficiency</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Chronic peritonitis</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Stricture or upper gastrointestinal surgery</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Pancreatic insufficiency</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Musculoskeletal</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscle atrophy or weakness</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Deforming or erosive arthritis</td>
<td>11 (2)</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>26 (5)</td>
</tr>
<tr>
<td>Avascular necrosis</td>
<td>48 (10)</td>
</tr>
<tr>
<td>Osteomyelitis</td>
<td>2 (0.4)</td>
</tr>
<tr>
<td>Ruptured tendon</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skin and Integument</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopecia</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Extensive scarring</td>
<td>5 (1)</td>
</tr>
<tr>
<td>Skin ulcer (not thrombosis)</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Condition</td>
<td>Value</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Endocrine</td>
<td></td>
</tr>
<tr>
<td>Premature Ovarian Failure</td>
<td>10 (2)</td>
</tr>
<tr>
<td>Diabetes Mellitus</td>
<td>16 (3)</td>
</tr>
<tr>
<td>Malignancy</td>
<td>6 (1)</td>
</tr>
</tbody>
</table>

*As defined in the SLICC/ ACR damage index.
Table 4.3: Time-invariant/ baseline predictors of the average population damage trajectory in patients with childhood-onset SLE

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Parameter estimate (standard error) *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline effects</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (Males)</td>
<td>0.010(0.154)</td>
</tr>
<tr>
<td>Sex (Males)</td>
<td>-0.163(0.232)</td>
</tr>
<tr>
<td>Race</td>
<td>--</td>
</tr>
<tr>
<td>Caucasian (ref)^</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>0.008(0.138)</td>
</tr>
<tr>
<td>Afro-Caribbean</td>
<td>0.628(0.272)*</td>
</tr>
<tr>
<td>Others</td>
<td>-0.259(0.144)*</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Caucasian (ref)^</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>-0.089(0.231)</td>
</tr>
<tr>
<td>Afro-Caribbean</td>
<td>-0.413(0.326)</td>
</tr>
<tr>
<td>Others</td>
<td>-0.249(0.205)</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>0.042(0.019)*</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>-0.046(0.033)</td>
</tr>
<tr>
<td>Time period of diagnosis</td>
<td>--</td>
</tr>
<tr>
<td>1985-1990 (ref)^</td>
<td>0</td>
</tr>
<tr>
<td>1991-2000</td>
<td>-0.416(0.250)*</td>
</tr>
<tr>
<td>2001-2011</td>
<td>-0.632(0.213)*</td>
</tr>
</tbody>
</table>
### Time period of diagnosis

<table>
<thead>
<tr>
<th>Time period of diagnosis</th>
<th>Number of ACR criteria at diagnosis</th>
<th>Life-threatening major organ at diagnosis</th>
<th>Renal disease at diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985-1990 (ref)^</td>
<td>0.104 (0.051)*</td>
<td>0.326 (0.127)*</td>
<td>0.302 (0.131)*</td>
</tr>
<tr>
<td>1991-2000</td>
<td>-0.021 (0.064)</td>
<td>0.176 (0.102)*</td>
<td>-0.114 (0.194)</td>
</tr>
<tr>
<td>2001-2011</td>
<td>0.152 (0.261)</td>
<td>-0.122 (0.103)</td>
<td>0.091 (0.096)</td>
</tr>
</tbody>
</table>

* Indicates p<0.1; # The SDI trajectory slope was modelled with the 2nd order fractional polynomial, FP (0.5,1). First record was the result of crossing time-invariant predictor with time form1–FP (0.5)– and vice versa for the second record with time form 2 which represented FP (1). When crossed with the time-invariant predictors, they represented the effects of that predictor on the overall damage trajectory. ^Ref indicates reference category. ♯ Each individual item (except otherwise stated) of the SLEDAI2K was tested by examining scores from 6,12,18,24 months before. The following items of SLEDAI2K were not tested: seizures, cerebrovascular accidents and cranial neuropathy.
Table 4.4: Time-varying/disease course predictors of the average population damage trajectory in patients with childhood-onset SLE

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Parameter estimate (standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6 months</td>
</tr>
<tr>
<td>Lagged SLEDAI</td>
<td>-0.004(0.007)</td>
</tr>
<tr>
<td>Lagged Prednisone</td>
<td>0.011(0.002)*</td>
</tr>
<tr>
<td>Lagged Significant immunosuppression</td>
<td>0.437(0.097)*</td>
</tr>
<tr>
<td>Lagged Cyclophosphamide</td>
<td>0.655(0.140)*</td>
</tr>
<tr>
<td>Lagged Antimalarials</td>
<td>-0.505(0.137)*</td>
</tr>
</tbody>
</table>

* Indicates p<0.1. All predictors were lagged from 6 to 24 months before a current visit.
Table 4.5: Multivariable model predicting the longitudinal damage trajectory of childhood-onset SLE patients

<table>
<thead>
<tr>
<th>Predictors*</th>
<th>Effects of Predictors (SE)#</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time period of diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985-1990 (ref)$^o$</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>1991-2000</td>
<td>-0.022(0.113)</td>
<td>0.846</td>
</tr>
<tr>
<td>2001-2011</td>
<td>-0.201(0.080)</td>
<td>0.012</td>
</tr>
<tr>
<td>Race**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian (ref)$^o$</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>Asian</td>
<td>0.005(0.145)</td>
<td>0.972</td>
</tr>
<tr>
<td>Afro-Caribbean</td>
<td>0.601(0.268)</td>
<td>0.025</td>
</tr>
<tr>
<td>Others</td>
<td>-0.191(0.142)</td>
<td>0.178</td>
</tr>
<tr>
<td>Life-threatening manifestations at diagnosis</td>
<td>-0.643(0.278)</td>
<td>0.021</td>
</tr>
<tr>
<td>Life-threatening manifestations at diagnosis/Time form 1</td>
<td>0.272(0.119)</td>
<td>0.022</td>
</tr>
<tr>
<td>Life-threatening manifestations at diagnosis/Time form 2</td>
<td>-0.016(0.008)</td>
<td>0.052</td>
</tr>
</tbody>
</table>
### Chapter 4

#### cSLE Damage Trajectory

<table>
<thead>
<tr>
<th>Condition</th>
<th>Days Since Diagnosis</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute confusion state 24 months^</td>
<td>0.533(0.184)</td>
<td>0.004</td>
</tr>
<tr>
<td>Lupus headache 6 months^</td>
<td>0.809(0.356)</td>
<td>0.023</td>
</tr>
<tr>
<td>Mucosal ulcers 18 months^</td>
<td>-0.312(0.101)</td>
<td>0.002</td>
</tr>
<tr>
<td>Mucosal ulcers 24 months^</td>
<td>-0.216(0.087)</td>
<td>0.013</td>
</tr>
<tr>
<td>Pericarditis 6 months^</td>
<td>-0.567(0.235)</td>
<td>0.016</td>
</tr>
<tr>
<td>Fever 12 months^</td>
<td>0.307(0.122)</td>
<td>0.012</td>
</tr>
<tr>
<td>Fever 18 months^</td>
<td>0.681(0.270)</td>
<td>0.012</td>
</tr>
<tr>
<td>Prednisone 12 months^¶</td>
<td>0.039(0.016)</td>
<td>0.015</td>
</tr>
<tr>
<td>Prednisone 24 months^¶</td>
<td>0.051(0.015)</td>
<td>0.001</td>
</tr>
<tr>
<td>Antimalarial 6 months^</td>
<td>-0.402(0.145)</td>
<td>0.006</td>
</tr>
<tr>
<td>Cyclophosphamide 6 months^</td>
<td>0.331(0.111)</td>
<td>0.003</td>
</tr>
<tr>
<td>Cyclophosphamide 12 months^</td>
<td>0.305(0.115)</td>
<td>0.008</td>
</tr>
<tr>
<td>Cyclophosphamide 18 months^</td>
<td>0.431(0.102)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cyclophosphamide 24 months^</td>
<td>0.424(0.143)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

#SE, standard error. *Time in this longitudinal model– months since diagnosis– was specified by 2nd order fractional polynomials of (0.5,1). ** Afro-Caribbean patients also had more early damage compared to Asian patients (p= 0.020). This was modelled with Asians as the reference group, unlike the above table where Caucasians were the reference group. °Ref, indicates
reference category. ¶ Depicted change in damage trajectory for every 10 mg increase of prednisone. ‡Time indicator referred to exposure in the prior number of months indicated. These time-varying predictors tested negative for multi-collinearity and were therefore left in the model. Results cannot be interpreted on numerical values alone for time-invariant factors with time interaction. Please see appendix for graphical depiction of effects on damage trajectory.
Chapter 5

From Childhood to Adulthood: Latent Classes of Disease Trajectories in Childhood-onset Systemic Lupus Erythematosus Patients

Arthritis Care & Research 2017 Jul 13. DOI: 10.1002/acr.23319

LSH Lim¹,²,³, MBBS, MRCPCH, FRCPC, PhD(c)

E Pullenayegum²,³, PhD

BM Feldman²,³, MD, MSc, FRCPC

L Lim⁴, MD

DD Gladman⁵, MD, FRCPC

ED Silverman⁴, MD, FRCPC

¹ Children’s Research Institute of Manitoba, University of Manitoba

² Children’s Health Evaluative Sciences program, SickKids Research Institute

³ Institute of Health Policy Management and Evaluation, University of Toronto

⁴ The Hospital for Sick Children, Toronto, Canada

⁵ Centre for Prognosis Studies, Toronto Western Hospital Research Institute, Toronto

Corresponding author:

Earl Silverman

Address: 555 University Ave (Suite 8253)

Toronto, Ontario M5G1X8

Fax: 4168134989    Tel: 4168137500

Total word count: 2804
ABSTRACT

No previous study has studied the longitudinal disease course of childhood-onset SLE (cSLE).

**Objectives:** 1) To assess distinguishable differences in disease activity trajectories in cSLE patients; 2) To determine baseline factors predictive of disease trajectory membership and 3) To assess if the different disease activity trajectories are associated with different damage trajectories.

**Methods:** This is a retrospective, longitudinal inception cohort of cSLE patients. Patients were followed from diagnosis as children, until they were adults. SLE disease activity (DisAct) was modeled, as a latent characteristic, jointly using the SLE disease activity index 2000 and prednisone in a Bayesian growth mixture model. Baseline factors were tested for membership prediction of the latent classes of disease trajectories. Differences in damage trajectories by DisAct classes were tested using a mixed model.

**Results:** Four hundred and seventy-three patients (82% females) with median age at diagnosis of 14.1 years were studied. We studied 11992 visits (2666 patient years). We identified 5 classes of DisAct trajectories. Baseline major organ involvement, number of ACR criteria and age at diagnosis predicted memberships into different classes. A higher proportion of Asians was in class 2 compared to class 5. Class 1 was associated with the most accrual of damage while class 5 with no significant damage accrual even after 10 years.

**Conclusions:** There are five distinct latent classes of disease trajectory in patients with cSLE. Membership within disease trajectories is predicted by baseline clinical and demographic factors. Membership in different disease activity trajectory classes is associated with different damage trajectories.

(250 words)
Significance and Innovations

1) This is the first time that longitudinal disease activity of cSLE patients has been studied from diagnosis in childhood and extending into adulthood.

2) This is the first time that an objective method has been applied to delineate latent classes of disease activity trajectory in SLE.

3) Baseline demographics and clinical factors can be used to predict individuals’ memberships in different subclasses of disease activity trajectory in cSLE.

4) Different patterns of disease activity trajectory are significantly associated with significantly different rates of longitudinal damage accrual trajectories.
The longitudinal disease course of patients with childhood-onset Systemic Lupus Erythematosus (cSLE) is unknown. In adult-onset SLE (aSLE), SLE disease activity persists for up to 10 years after diagnosis. In cSLE, patients continue to have disease activity and require steroids in adulthood. There appears to be increased mortality, more persistent disease activity and more significant morbidity in adults with cSLE patients compared to adult onset patients. However, as only a single occasion of outcome ascertainment was reported in these studies, they were not designed to answer questions about the longitudinal disease course of cSLE.

Others have attempted to classify the disease patterns of aSLE. Based on predefined rules, SLE has often been classified into 5 patterns: persistently active, remitting and relapsing, serologically active but clinically quiescent (SACQ), clinically active but serologically quiescent or remission. These patterns have been mostly defined using relatively short windows of time, e.g. 2–5 years. As these categories have been based on arbitrary time intervals, individual patients might classify into different categories should the time intervals in the definitions or the windows of observation change. This is a very real possibility as most of the studies used prevalent cohorts, therefore depending on when a patient enters the study cohort, their disease patterns would have been classified differently at different times. For instance, an individual might not have not any clinical or serological activity for 4 years and then developed serological activity without clinical activity in the past year. She would have been classified as being in remission according to one definition that sets the period of disease quiescence as 1 year and unclassifiable according to the another definition that requires 5 years (different time intervals). If the window of observation is extended to 2 years and she remains in the same state, she would have been classified as SACQ (different window). If she enters a new clinic
cohort for a year during the last 6 months of her quiescent disease and in the first 6 months of her serologically active but clinically quiescent period, she would have been unclassifiable (different window). Most importantly, fixed time intervals do not reflect the continually evolving nature of a patient’s disease trajectory.

With knowledge about the longitudinal disease trajectory for a given patient, physicians might tailor their management strategies. For instance, those who are destined to have disease refractory to current standards of therapy might be considered for testing newer therapies; for those who will have low-grade disease activity, attention should be paid to further reduce treatment toxicities. Accurate predictions about future disease courses will therefore help refine patient management.

We designed this longitudinal study to assess whether there are identifiable subgroups of latent disease activity (DisAct) trajectory— as measured by the SLE disease activity index 2000 (SLEDAI2K) and prednisone exposure— in a population of cSLE patients followed throughout childhood and into adulthood. We determined factors that predicted membership in different DisAct disease trajectories. Finally, we assessed whether different DisAct disease trajectories were associated with different damage trajectories.

PATIENTS AND METHODS

Study population and data collection

This was a single center longitudinal inception cohort study. All patients were diagnosed and followed at The Hospital for Sick Children cSLE clinic from between 1st January 1985 to 30th September 2011. We follow about 60-70% of all cSLE patients in Ontario (unpublished data). Patient inclusion criteria: 1) Satisfied at least 4 out of 11 of the American College of Rheumatology (ACR) or the SLE International Collaborating Clinics (SLICC) classification
criteria for SLE\textsuperscript{135,136}, 2) Age less than 18 years at diagnosis, 3) Diagnosed in our clinic or transferred to our clinic within 6 months of diagnosis with full treatment information available (i.e. inception cohort), 4) Seen at our cSLE clinic for at least 3 visits (according to the definition of longitudinal design definition)\textsuperscript{71}, 5) Not treated with continuous corticosteroids for more than 3 months (e.g. as another autoimmune disease) prior to SLE diagnosis. Patients were excluded if any of the above criteria was not met.

The following patient data were obtained at every clinic visit: disease manifestations, disease activity and laboratory results to score SLEDAI\textsubscript{2K} and SLICC/ACR damage index (SDI), as well as medications. While patients were followed at SickKids, information was obtained from the prospectively collected information in the cSLE database. This database was established in 1986 and has prospectively collected clinical information, laboratory results and all medications at each visit. Data for SLEDAI and the SDI for 1985 had been retrospectively entered. Of the 473 patients, 301 had transferred to adult care. Following transfer to an adult rheumatologist, data was obtained in one of 2 ways: clinical research database or chart review. The Toronto Lupus database provided data for 66\% of the transferred subjects\textsuperscript{18}. The rest of the transferred subjects were followed by another 17 rheumatologists; data were extracted from patients’ charts. For information about handling of missing data in this cohort, please see appendix.

**Outcome measures**

Disease activity in SLE is a latent construct as it is not directly measurable (unlike height, a directly measurable trait). Although many tools have been devised to measure disease activity in SLE, none measures the full spectrum and nuances of disease activity in SLE\textsuperscript{137}. Rare manifestations of SLE are usually not included, but these are still important as they may require
treatment and may lead to significant morbidity. Therefore, using only a disease activity tool as the sole measure of the disease activity trait is insufficient. Since prednisone is the most frequently used medication in SLE and its use signals the degree of disease activity, we used prednisone dose as another measure of the disease activity.

We therefore conceptualized the latent SLE disease activity (DisAct) as represented jointly by 2 measures: the SLEDAI(2K) and prednisone$^{138}$. SLEDAI or SLEDAI(2K) (as appropriate for date of visit) at each clinic visit was used to construct individual SLEDAI(2K) trajectories; they had demonstrated strong correlations to each other in previous work$^{138}$. Prednisone dose at every clinic visit was used to construct the individual prednisone trajectories.

**Prognostic factors**

In this study, we sought to identify baseline factors that predict membership into different classes of DisAct. We tested the following: sex, (self-reported) ethnicity, age at diagnosis, pubertal status at diagnosis (yes/no), time period of diagnosis, number of ACR classification criteria at diagnosis$^{135}$ and life-threatening major organ involvement. Life-threatening major organ involvement was defined a priori as including any of: lupus nephritis classes III–V, pulmonary hemorrhage, myocarditis, psychosis, cerebrovascular disease or major organ vasculitis.

**Statistical analysis**

We plotted individual patients’ trajectories for both outcomes (Appendix). We then determined the optimal shapes of the SLEDAI2K and prednisone population trajectories using fractional polynomials (FP) separately$^{139}$. See appendix for details.

*Modelling*
There are two main approaches to identifying latent class trajectories within a heterogeneous population: group based trajectory model (GBTM) or growth mixture model (GMM)\(^{140}\). These two approaches have been used to study juvenile arthritis, rheumatoid arthritis and osteoarthritis\(^{141-144}\). We choose GMM as the concept of this model is consistent with our clinical observation that there are clustering of different disease courses. GMM, in contrast to GBTM, allows incorporation of individual heterogeneity (random effects) into class trajectories, so that individuals’ trajectories within a class can vary from the mean class trajectory. This makes intuitive sense clinically. In GBTM, individuals within a group must all follow the group trajectory, which is likely not the case in a heterogeneous disease like SLE\(^{140}\).

We applied a GMM to identify the number of possible classes for the individual components of the DisAct trajectories\(^ {145}\). This was performed in R (lcmm package)\(^ {146,147}\). As the number of patients with longer disease durations decreased over time, we limited the disease course duration to the first 10 years.

We then jointly modelled the DisAct (dual outcomes) trajectories starting with 2 classes, increasing the number of class by 1 iteratively (joint GMM). Each model has 2 outcomes trajectories moving together. The solution of each class has 2 outcome trajectories. We stopped testing at 6 classes as the number of patients in the smallest class was very small (1% of patients). As 5 classes was deemed clinically plausible to the disease experts (EDS, DDG), we chose this as ideal. See appendix for more information about model choice.

We fitted the joint GMM using OpenBUGS (Bayesian analysis)\(^ {148}\). We modelled the 5-class joint outcomes model using a single chain. We ordered the classes from the smallest (class 1) to the largest (class 5). The 5-class model did not demonstrate non-convergence on the Gweke, Heidelberger and Raferty tests\(^ {149}\). The posterior probability of membership in each class
was calculated for each individual. Individuals were assigned class membership according to their highest posterior probability.

We then tested baseline covariates for their abilities to predict class membership in each DisAct class. We tested the baseline covariates in a multinomial logistic regression model using the assigned membership class for each individual as the outcome (SAS Inc, Cary, NC, USA). For this part, we only included individuals with >80% probability of belonging in a particular class. The candidate predictors that were significant at p<0.2 were included in a multivariable model.

**Damage association with DisAct**

We plotted the mean population longitudinal damage trajectories by class using the loess plot \cite{150} in R. We fitted the damage trajectory of each class using fractional polynomials and tested the difference between damage trajectories by comparisons of the slopes of all the trajectories in a mixed model (see appendix).

**RESULTS**

**Study population**

There were 568 patients in our cSLE database. Of these, 95 were excluded: 33 were either not diagnosed in our institution, or treated as another autoimmune disease before the final diagnosis of SLE; 62 patients were from outside the study period or did not have at least 3 visits by the study end date. Four hundred and seventy-three patients were finally included; 14% were eventually lost to follow-up; data were included up to the time of loss. Thirty-eight patients had ≥ 20 years, 67 patients had ≥ 15 years, 130 patients had ≥ 10 years and 213 patients had ≥ 5 years of follow-up. There were 11992 visits (2666 patient years). The median number of visits
per patient was 23 (25th-75th percentile: 14-35, maximum 97). The median duration of follow-up was 5.44 years (25th-75th percentile: 2.98-9.21 years).

The demographics of this population have been reported. Briefly, 82% were females; 38% were Asian, 32% Caucasian, 13% Afro-Caribbean, and the remaining were of mixed or other race. Median age at diagnosis was 14.1 (25th-75th percentiles: 11.5-15.8) years. The median age at last follow-up was 19.4 (25th-75th percentiles: 16.9-22.2) years.

**Joint Outcomes**: DisAct

We identified 5 latent classes of DisAct with distinct class mean trajectories. Thirty patients had their highest probability of membership in class 1 (6%), 57 patients in class 2 (12%), 79 patients in class 3 (17%), 92 patients in class 4 (19%) and 215 patients in class 5 (45%). The mean posterior probabilities of class membership for each class were 0.94 (each for classes 1, 2, 3), 0.90 (class 4) and 0.96 (class 5).

Loess plots of the joint outcomes of DisAct are shown in Figure 1. Both the individuals’ as well as the mean population-joint trajectories were used in interpretation (Appendix). Class 1 represents a relapsing or disease-transforming group. Class 2 represents high early disease activity but long-term remission group. Class 3 represents moderate early disease activity with long-term low-grade disease. Class 4 represents a late relapsing group. Class 5 represents a low-grade chronic disease group.

**Patient characteristics by latent class membership**

Table 1 shows the demographics and baseline characteristics of the patients by latent class. There is a substantially higher proportion of patients with baseline major organ involvement in class 2; Class 5 has the lowest proportion. Asians are represented most highly in
class 2 and least in class 5. Afro-Caribbeans are represented most highly in class 1 and least in class 3. Caucasians are most often in class 5 and least in class 2.

**Predictors of membership in each class**

*Univariable analysis*

For this part, we studied 415 out of the total of 473 patients (those with a class membership posterior probability >0.80, see appendix). Table 2 shows the global effects of membership predictors in predicting class memberships referenced to class 5. Table 3 shows the specific effects of each subcategory of the membership predictors (where applicable) and when the reference class is changed, to distinguish probability of membership between different classes. Only the significant results are shown. For a table of more comparisons, please refer to the appendix.

*Multivariable model*

Table 4 describes the multivariable model. Age at diagnosis, number of ACR criteria at diagnosis and baseline major organ involvement all predict membership in different subclasses. Even though Asian heritage does not have a global effect in differentiating membership, it specifically predicts significantly higher likelihood of membership in class 2 compared to class 5.

**Effects on damage trajectory according to membership in different classes**

Figure 2 shows the different SLICC/ACR damage trajectories by disease activity class. No significant damage (SDI <1) was accrued for the average patient in class 5 despite 10 years of disease. Although class 2 had the steepest damage slope initially, the average patient in this group stopped accruing more damage after the first 5 years. The average patient in class 1 with frequent relapsing disease accrued the most damage after 10 years. The average patient in class 4
with late relapsing disease activity added damage more rapidly later (after 5 years) compared to their earlier course.

**DISCUSSION**

In this longitudinal cSLE inception cohort, we identify 5 distinct latent classes of disease activity: Relapsing/transforming (class 1), high initial activity with long-term remission (class 2), moderate initial activity with long-term low-grade disease (class 3), late relapsing (class 4) and chronic low-grade disease (class 5). Baseline factors—major organ involvement, the number of ACR criteria at diagnosis, age at diagnosis and Asian heritage—predicted class membership. Those who followed a relapsing/transforming disease trajectory accumulated the most damage and at the fastest rate. Those who followed a chronic low-grade disease activity accrued little or no damage, on average, even after long durations of disease.

We have approached the concept of SLE disease activity differently from previous studies. Previous studies have used modified SLEDAI (M-SLEDAI)\textsuperscript{152-154} or physician global activity (PGA) scores as measures of disease activity\textsuperscript{152}. However, it is important to remember that the underlying DisAct cannot be directly measured. Any disease measure is only an imperfect surrogate. Prednisone use represents an indirect marker of DisAct that is not directly measured in SLEDAI(2K). It adds information about the underlying DisAct that is unmeasured or unmeasurable. Therefore, by jointly modelling both outcomes, we gained a more complete picture of the DisAct than either.

Although our study was not designed to address the question of remission specifically, our results support recent work on remission\textsuperscript{133,155-158}. Several groups have attempted to define a low or no disease state, though somewhat differently; it is important to note that most of these studies used single occasion outcome ascertainment and are not directly comparable to our design. It is
clear that only a minority of patients attained remission, whether it be clinical only or clinical and serological, with or without treatment. Only about 1-7% of SLE patients maintained remission for 5 years\textsuperscript{133,156-158}. Patients who attain this state accrue less damage on average. Our work corroborates with these observations. We showed that even those with moderate to high activity initially improve to low disease states over time, as evidenced by mean SLEDAI2K of \(<4\) (see appendix)\textsuperscript{155}. Prednisone (mean 10 mg/day) is required to maintain low disease states in most of the latent classes (2 to 5). It is also a common theme in the remission literature that those with less severe disease are more likely to attain the low or no disease state and to have less damage\textsuperscript{133,155-158}. Patients in class 5 (chronic low grade disease), as demonstrated in our study, are less likely to have had major organ disease at baseline and have a low average risk of damage.

Our approach using an inception longitudinal cohort has refined our understanding of DisAct evolution in SLE by giving a more complete view of this process. Tailored care, based on predicted longitudinal class trajectories, may potentially attain the best balance of disease control and adverse effects for patients, following validation of our work.

We identify 3 predictors of SLE DisAct trajectories in this study. Chronic low-grade disease activity patients have fewer baseline ACR criteria than the other classes. Similarly, the LUMINA group has found that a higher number of baseline ACR criteria to be inversely associated with decline to low disease activity in adult SLE\textsuperscript{159}. We found that the long-term remission group contains a greater proportion of Asians than the low-grade chronic disease group. Other Canadian SLE cohorts also noted Asians have a more severe disease phenotype\textsuperscript{160,161}. Those in the low-grade disease subclass are more likely to be younger at disease diagnosis compared to those in other classes. While there have been several studies showing that cSLE has
more severe phenotypes and greater propensity for damage than aSLE\textsuperscript{90,162,163}, there have been few studies that compared differences within the pediatric age group\textsuperscript{164,165}. We are unable to compare our findings directly with the results of those studies as they focused primarily on clinical characteristics at diagnosis and not on longitudinal disease course\textsuperscript{164,165}.

We have shown that different DisAct trajectories are associated with different damage trajectories. In the high early DisAct group (class 2) where the majority of patients have major organ involvement at baseline, the damage accrual trajectory is steep in the first few years and then plateaued. This may be a result of the initial high dose immunosuppression that these patients are exposed to early on in their disease. Patients with relapsing disease (class 1 and 4) have increasing damage trajectories over time, consistent with a recent study that demonstrated flares predicted increased damage accrual\textsuperscript{166}. Most interestingly, patients with low-grade chronic disease have a good prognosis, as the average patient will not have accrued any significant damage after 10 years of disease. This is the first time that differences in longitudinal outcomes have been associated with different disease activity trajectories. As there are significant differences in the damage trajectories associated with the different latent DisAct groups, this validates the existence of these identified groups.

Our study conclusions must be interpreted in the light of potential limitations. We have only focused on the first 10 years after diagnosis. We were unable to study the effects of time-varying predictors as the model does not allow this. Preliminary work has been published about the use of time-varying predictors that influence disease course in latent class trajectory analysis\textsuperscript{167}. In the future, this may help us identify disease course changing time-varying predictors.

However, we identified early clinical factors that can predict membership in different subclasses of disease activity and damage trajectory. Fourteen percent of our cohort was lost to follow-up.
However, this is a relatively low percentage considering the long duration of study and the inclusion of patients transitioning away from our hospital to adult care; previous studies including transitioning patients have attrition rates of up to 55%\textsuperscript{168}. We could not study effects of adherence in influencing the observed disease activity as this was not measured routinely in our cohort. For the future, investigators should consider measuring adherence routinely in chronic disease cohorts to help interpret the observed differences in disease courses, and to discern whether variations in disease course are due to true disease variations or undertreated disease due to patient non-adherence. We did not study additional therapeutic trajectories other than prednisone. This is because other medications, like the immunosuppressants, do not change as frequently or as much (in dose variations) as prednisone and the changes might be due to intolerance and not related to disease control. Furthermore, the computations for the joint growth mixture model with two outcome trajectories are already very complex and time consuming. As such, we chose to focus on just the dual outcomes model for this study.

Our study will benefit from validation in another cohort. We would also like to extend this work into longer disease duration to further improve our understanding. For future work, inclusion of molecular markers may help refine class membership predictions.

In conclusion, we identify 5 latent classes of SLE DisAct in cSLE. Age, major organ involvement, number of ACR criteria at baseline and Asian heritage predicted membership in different classes. A relapsing course is more likely associated with increasing long-term damage accrual. These results can be used to inform patients and their families, and to personalize management for new cSLE patients.
Table 5.1: Demographics and baseline clinical characteristics of patients in the most probable latent classes of SLE activity (DisAct)

<table>
<thead>
<tr>
<th></th>
<th>Total cohort N=473</th>
<th>Class 1 N=30</th>
<th>Class 2 N=57</th>
<th>Class 3 N=79</th>
<th>Class 4 N=92</th>
<th>Class 5 N=215</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females (%)</strong></td>
<td>386 (82)</td>
<td>26 (87)</td>
<td>43 (75)</td>
<td>66 (84)</td>
<td>77 (84)</td>
<td>174 (81)</td>
</tr>
<tr>
<td><strong>Ethnicities (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasians</td>
<td>153 (32)</td>
<td>7 (23)</td>
<td>12 (21)</td>
<td>28 (35)</td>
<td>24 (26)</td>
<td>82 (38)</td>
</tr>
<tr>
<td>Asians</td>
<td>181 (38)</td>
<td>12 (40)</td>
<td>31 (54)</td>
<td>34 (43)</td>
<td>37 (40)</td>
<td>67 (31)</td>
</tr>
<tr>
<td>Blacks</td>
<td>60 (13)</td>
<td>7 (23)</td>
<td>8 (14)</td>
<td>7 (9)</td>
<td>12 (13)</td>
<td>26 (12)</td>
</tr>
<tr>
<td>Others</td>
<td>78 (16)</td>
<td>4 (14)</td>
<td>6 (11)</td>
<td>10 (13)</td>
<td>19 (21)</td>
<td>39 (18)</td>
</tr>
<tr>
<td><strong>Median age at diagnosis (25-75th percentile)</strong></td>
<td>14.0 (11.3-15.7)</td>
<td>14.5 (11.1-15.2)</td>
<td>14.8 (11.9-16.4)</td>
<td>14.6 (12.6-16.1)</td>
<td>13.7 (11.8-15.6)</td>
<td>13.5 (11.0-15.4)</td>
</tr>
<tr>
<td><strong>Baseline major organ involvement (%)</strong>#</td>
<td>202 (44)</td>
<td>15 (54)</td>
<td>47 (84)</td>
<td>40 (51)</td>
<td>52 (57)</td>
<td>48 (23)</td>
</tr>
<tr>
<td><strong>Median baseline ACR (25th-75th percentile)</strong></td>
<td>5 (4-6)</td>
<td>6 (5-6)</td>
<td>6 (5-7)</td>
<td>5 (4-6)</td>
<td>6 (5-6)</td>
<td>5 (4-6)</td>
</tr>
<tr>
<td><strong>Pubertal status (%)</strong></td>
<td>334 (71)</td>
<td>25 (83)</td>
<td>39 (68)</td>
<td>64 (81)</td>
<td>68 (74)</td>
<td>138 (64)</td>
</tr>
</tbody>
</table>

*The ethnicity of 1 patient was unknown. Asians in our population included South Asians. #The baseline major organ involvement status was unclear for 10 patients. The denominators have been modified accordingly within subgroups.
Table 5.2: Results of univariable screen for predictors of membership in latent SLE DisAct classes (class 5 as reference class)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effects</th>
<th>p (Global effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time period of entry into cohort (ref= 1985-1989)</td>
<td>No differences across the time periods for the different classes</td>
<td>0.5458</td>
</tr>
<tr>
<td>Ethnicity (ref= whites)</td>
<td>Asians more likely in class 2 than 5, Blacks more likely in class 1 than 5</td>
<td>0.1107</td>
</tr>
<tr>
<td>Asian Ethnicity</td>
<td>Asians more likely in classes 2, 3 &amp; 4 than 5</td>
<td>0.0268</td>
</tr>
<tr>
<td>Pubertal status at baseline</td>
<td>More likely to have entered into puberty in class 3 than 5.</td>
<td>0.0320</td>
</tr>
<tr>
<td>Age at baseline</td>
<td>Older age at diagnosis in classes 2 &amp; 3 compared to 5</td>
<td>0.0430</td>
</tr>
<tr>
<td>Major organ involvement at baseline</td>
<td>Major organ involvement more likely in all classes compared to 5</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Number of ACR criteria at diagnosis</td>
<td>Higher numbers of ACR criteria in classes 1, 2 and 4 compared to 5</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
Table 5.3: Specific effects of univariable membership predictors of SLE DisAct trajectory classes (variable class of reference)

<table>
<thead>
<tr>
<th>Membership predictors</th>
<th>Class#</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asians vs Caucasion</td>
<td>2 vs 5</td>
<td>3.08</td>
<td>1.42-6.68</td>
</tr>
<tr>
<td>Asians vs Other minorities</td>
<td>2 vs 5</td>
<td>2.71</td>
<td>1.02-7.16</td>
</tr>
<tr>
<td>Asians vs Afro-Caribbean</td>
<td>3 vs 5</td>
<td>2.02</td>
<td>0.75-5.44</td>
</tr>
<tr>
<td>Asians vs Other minorities</td>
<td>3 vs 5</td>
<td>2.18</td>
<td>0.90-5.26</td>
</tr>
<tr>
<td>Afro-Caribbean vs Caucasion</td>
<td>1 vs 5</td>
<td>3.60</td>
<td>1.01-12.82</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 vs 5</td>
<td>1.20</td>
<td>1.00-1.25</td>
</tr>
<tr>
<td></td>
<td>3 vs 5</td>
<td>1.15</td>
<td>1.04-1.27</td>
</tr>
<tr>
<td>Baseline major organ involvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 vs 5</td>
<td>4.77</td>
<td>1.98-11.49</td>
</tr>
<tr>
<td></td>
<td>2 vs 5</td>
<td>25.00</td>
<td>10.00-62.52</td>
</tr>
<tr>
<td></td>
<td>3 vs 5</td>
<td>3.84</td>
<td>2.14-6.87</td>
</tr>
<tr>
<td></td>
<td>4 vs 5</td>
<td>4.82</td>
<td>2.69-8.63</td>
</tr>
<tr>
<td></td>
<td>2 vs 1</td>
<td>5.24</td>
<td>1.61-17.00</td>
</tr>
<tr>
<td></td>
<td>2 vs 3</td>
<td>6.52</td>
<td>2.45-17.31</td>
</tr>
<tr>
<td></td>
<td>2 vs 4</td>
<td>5.19</td>
<td>1.95-13.77</td>
</tr>
<tr>
<td>Number of ACR criteria at diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 vs 5</td>
<td>1.50</td>
<td>1.08-2.10</td>
</tr>
<tr>
<td></td>
<td>2 vs 5</td>
<td>1.88</td>
<td>1.45-2.43</td>
</tr>
<tr>
<td></td>
<td>4 vs 5</td>
<td>1.56</td>
<td>1.34-1.96</td>
</tr>
<tr>
<td></td>
<td>1 vs 3</td>
<td>1.50</td>
<td>1.03-2.18</td>
</tr>
<tr>
<td></td>
<td>2 vs 3</td>
<td>1.88</td>
<td>1.38-2.55</td>
</tr>
<tr>
<td></td>
<td>4 vs 3</td>
<td>1.56</td>
<td>1.17-2.07</td>
</tr>
</tbody>
</table>
Table 5.4: Results of multivariable model for membership predictors in latent SLE (DisAct) trajectory classes (Class 5 is the reference class)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Class</th>
<th>OR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major organ</td>
<td>1</td>
<td>4.144</td>
<td>1.637-10.487</td>
<td>0.0027</td>
</tr>
<tr>
<td>Major organ</td>
<td>2</td>
<td>21.713</td>
<td>8.368-56.342</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Major organ</td>
<td>3</td>
<td>4.776</td>
<td>2.532-9.008</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Major organ</td>
<td>4</td>
<td>4.115</td>
<td>2.216-7.639</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Asians#</td>
<td>1</td>
<td>1.761</td>
<td>0.732-4.24</td>
<td>0.2066</td>
</tr>
<tr>
<td>Asians</td>
<td>2</td>
<td>2.444</td>
<td>1.199-4.982</td>
<td>0.0139</td>
</tr>
<tr>
<td>Asians</td>
<td>3</td>
<td>1.638</td>
<td>0.906-2.962</td>
<td>0.1027</td>
</tr>
<tr>
<td>Asians</td>
<td>4</td>
<td>1.475</td>
<td>0.815-2.67</td>
<td>0.1988</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>1</td>
<td>1.108</td>
<td>0.95-1.292</td>
<td>0.1905</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>2</td>
<td>1.180</td>
<td>1.037-1.343</td>
<td>0.0121</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>3</td>
<td>1.159</td>
<td>1.044-1.287</td>
<td>0.0056</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>4</td>
<td>1.075</td>
<td>0.976-1.184</td>
<td>0.1437</td>
</tr>
<tr>
<td>ACR criteria°</td>
<td>1</td>
<td>1.305</td>
<td>0.91-1.87</td>
<td>0.1477</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>2</td>
<td>1.378</td>
<td>1.027-1.849</td>
<td>0.0325</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>3</td>
<td>0.827</td>
<td>0.633-1.079</td>
<td>0.1613</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>4</td>
<td>1.299</td>
<td>1.018-1.657</td>
<td>0.0355</td>
</tr>
</tbody>
</table>

^Major organ involvement at baseline # Asians compared to all other ethnicities. °Number of ACR criteria at baseline.
Figure 5.1: Plots of Latent SLE Disease Activity (DisAct) Trajectories in cSLE by Subclasses

Figure 5a) Loess plots of mean population SLEDAI2K trajectory (with 95% confidence intervals).

Figure 5b) Loess plots of mean population prednisone trajectory (with 95% confidence intervals).

Each latent class mean population trajectory is represented by a different colour. The same colour represents the joint outcomes trajectories constituting DisAct.
Figure 5.2: Damage trajectories of cSLE patients by disease activity classes

Legend: Damage plots start from 6 months after diagnosis. 415 patients with posterior membership probability >0.80 were used for this plot.
Chapter 6

Synthesis

6.1 What has been achieved?

With this thesis, I have achieved my overarching aims:

1) I have found that prognosis studies (in SLE, as an example of a rare disease) are often poor quality. In the majority of studies, important domains of study design—confounding, attrition and study population—which directly contribute to the validity of study results and conclusions, are found to be at high risk of bias.

2) I have shown that the longitudinal study design has been rarely used in prognosis (using SLE as an example of a rare disease). Even when used, appropriate longitudinal analytic methods are often not used to analyze the data.

3) I have shown that the longitudinal study design can be applied to a retrospective rare disease clinic-based cohort with an irregular visit schedule, to study disease evolution and to identify predictors of the disease trajectory.

4) I have demonstrated the shape of the population average trajectory of morbidity accrual over time in a rare disease population. The population of cSLE patients have, on average, a continual and steady gain in morbidity without prolonged remission from further damage.

5) I have successfully used Bayesian latent class statistical methods to identify latent subclasses of patients with more homogeneous disease activity trajectories within a heterogeneous rare disease.
I have achieved my clinical aims:

1) I have shown that the damage trajectory of cSLE patients did not plateau but instead continued to increase steadily over time, suggesting that these patients remain at persistent risk of gaining more damage over time.

2) I have shown that demographics and baseline clinical factors predict the evolution of organ damage trajectory in cSLE patients. Afro-Caribbean patients have a higher level of damage over time, and that major organ involvement predicts a more rapid initial damage accrual.

3) I have identified time-varying prognostic factors – clinical features and therapeutics – that predict changes in the subsequent evolution of the population damage trajectory, within precise timeframes.

4) I have identified 5 latent classes of disease activity trajectories, demonstrating that there are more homogeneous clusters of patients with more similar disease activity trajectories within the heterogeneous population of cSLE patients.

5) I have found baseline prognostic factors – age at diagnosis, major organ involvement, number of ACR criteria and Asian race – that predict different probabilities of membership in the different latent classes of cSLE disease activity trajectories.

6) I have demonstrated that different latent classes of disease activity trajectories are associated with significantly different trajectories of damage.

6.2 What have I learned?

Modern longitudinal analytic methods can be applied to clinic-based rare disease cohorts with irregular, individualized visit schedules. Currently, most cohorts (especially multi-centre studies) using the longitudinal design have the same standard visit schedule for all members of
the cohort. This schedule is always a compromise between cost (including burden on the study participants) and possible gain of information. However, a “standardized visit schedule” design might not be as helpful in providing precise information about disease activity variations. In clinical practice, patients are seen as required by their disease activity; they may also return for complications or flares of disease in between scheduled visits. With a standardized visit schedule, there is a tendency towards observing the mean disease experience as the peaks and troughs of disease changes (that often occur in between scheduled visits) are not captured if the intervals between visits are long; unless alternative and flexible options for recording important events (e.g., flares of disease) that occur between visits are available. Valuable information about some aspects of disease evolution might thus be lost with a standardized visit schedule.

Especially in the setting of a rare disease, the gain of information about disease course evolution from an individualized, irregular visit schedule might be considerable as it is so much harder to study these diseases due to their rare occurrence. In this thesis, I have shown that longitudinal methods can be applied successfully – with due considerations of the irregular schedule – to clinic-based cohort data with an irregular schedule to obtain precise information about prognosis.

Modern longitudinal analytic methods can be used to address different kinds of prognosis questions in rare diseases. Depending on the underlying assumptions and the clinical questions of interest, different longitudinal analytic methods can be applied to address prognosis questions, even in rare diseases.

When the question of interest pertains to how the disease outcome of a population of patients evolves longitudinally, marginal models such as the weighted generalized estimating equation (WGEE), can be used to study the average population outcome trajectory\textsuperscript{169}. I used the WGEE to study the population average damage trajectory in a rare disease cohort (cSLE
patients). The weighting was important to satisfy the underlying assumption of data missing at
random (MAR) in the dataset\textsuperscript{104}.

Given the heterogeneity of the underlying disease (cSLE), I asked if there are different
and more homogeneous subgroups within the total population even though these are not directly
identifiable (i.e., latent)\textsuperscript{48}. This was answered using \textit{latent class growth mixture analysis to
identify statistically distinct and clinically meaningful subgroups}– with distinct mean subgroup
disease trajectories– \textit{within a rare disease (cSLE) cohort}. Importantly with this approach, we
predict distinct disease courses; this is more informative about possible patients’ disease
experience than single-occasion outcomes. The latent class growth mixture model can be applied
to any other heterogeneous rare disease, with the caveat that a population size of 300 is necessary
for stable identification of the subgroups, when implemented in the frequentist setting\textsuperscript{170}.
Bayesian implementation of this method can be considered – and had been successfully applied –
in very rare diseases with much smaller population sizes (<50 patients) than the population used
in this thesis\textsuperscript{171}.

When the outcome to be studied is a complex trait, which might not be adequately
represented or measured by one tool, multiple outcomes measures can be studied together to
reflect the underlying latent construct using a joint (or multivariate) model. In this thesis, \textit{I
studied the complex, unmeasurable underlying (SLE) disease activity trait using 2 measures,
with joint (multivariate) modelling}\textsuperscript{172,173}. The implementation of this approach can rapidly get
very complex with more outcomes being jointly modelled\textsuperscript{173}. In my work, the joint model was
complex due to the long duration of the data, and the changing relationships of the two outcome
trajectories over time (which violated the assumptions of the frequentist software I was using in
analysis initially)\textsuperscript{174}. When model specification gets complex, with high integral equations,
frequentist methods often fail to converge. However, *Bayesian statistics* provide greater flexibility (so that we would not be limited by ill-fitting assumptions) and the ability to deal with highly complex equations. I performed the joint (latent class growth mixture) model analysis using Bayesian statistical software\textsuperscript{175,176}. This joint modelling approach was especially exciting as we used this for the first time in the setting of a rare disease. However, this approach can potentially be used in any other disease with complex outcome traits, even if the disease is not rare\textsuperscript{177-179}.

Longitudinal analytic methods provide a more complete picture of prognosis by addressing clinical questions in rare diseases using the entire disease course information for both outcomes and covariates. *By studying time-varying factors throughout the disease course* using a longitudinal analytic method, researchers can identify important prognostic factors that *appear during the disease course*, not only those present at baseline, and additionally target these factors to modify prognosis. This approach also *allows the effects of such time-varying factors to be more precisely delineated*. In the example of rare disease (cSLE) damage trajectory in this thesis, I lagged time-varying factors at different time intervals in relationship to a subsequent occasion of outcome measurement, to achieve a more precise prediction of the effects of time-varying factors on the damage trajectory. The use of whole disease course information, including time-varying covariates, potentially adds tremendous gain to the quality of information that can be acquired in rare disease prognosis studies. Of course, such an approach can equally benefit common diseases as well.

**6.3 Limitations and Future Directions**

Attrition could potentially bias study results. It is well known that it is hard to retain patients for sustained collection of longitudinal and very long-term data. Patients who are
available for study in long-term research cohorts may be very different from those who are lost to follow-up. For example, they might be sicker (and therefore require continued tertiary care) compared to those lost to follow-up. If the direction of bias is unknown, it will be hard to know how to interpret and use the information from available patients. Considering this, the rare disease cohort that I worked with had only 14% attrition, a very low rate for the length of follow-up. In my clinical studies, I built in strategies to explore the direction of bias that might result from attrition, such as attempts to obtain the most recent outcome information (after loss to follow-up) of the group of patients lost to follow-up and exploring the trajectory of outcomes up to attrition. My attempts suggested that those who have more morbidities (accrued damage) were more likely to continue in the cohort. This information gave a hint as to the possible direction of bias – i.e., sicker patients remained in this cohort – and therefore may help readers assess the generalizability of my study results.

It was very difficult to track down patients lost to follow-up. Tracking down patients with rare diseases has additional unique challenges. Seventy-five percent of rare diseases have their onset in childhood\textsuperscript{180}. Often little is known about the adult outcomes of these patients. \textit{The transition of care from the pediatric to the adult setting is known to be associated with significant attrition}; for example, in a complex congenital heart disease cohort, only about half go to the adult cardiology appointment after transfer, and a quarter stop any follow-up after age 18\textsuperscript{181}. \textit{Young adults are notoriously difficult to study as they are highly mobile and difficult to locate as they leave parental homes for school or work, sometimes in different cities or countries}\textsuperscript{168}. The physician who started this rare disease cohort had the foresight to implement several measures to help keep track of these patients, including: partnering with an adult cohort with standardized data collection (protocol-guided prospective data collection), transferring
patients only to a handful of highly experienced physicians in academic settings (easy tracking while ensuring quality care), and maintaining good rapport with the physicians who took over care of these patients (resulting in clinical status updates, or at least a point of contact for future studies). Despite this, it was still very challenging when I tried to find some of the patients from earlier periods in the cohort who did not go to our partner institution for follow-up, as some of these patients’ last contacts were more than 10 years prior to my study. I had designed multiple strategies to reconnect with these patients, e.g., through advertisements in local chapters of Lupus Canada across the country, but these did not yield any contact from our targeted group. I found more patients through contacts with rheumatologists than through my attempts at trying to contact them directly.

While a prospective, multi-centre funded cohort of a rare disease aimed at answering focused questions would be ideal, it is very unlikely that such a cohort would be funded indefinitely. Therefore, physician-initiated, clinic-based cohorts will likely continue to provide the bulk of information about longitudinal and very long-term outcomes in rare diseases. Measures should be implemented— at the outset— to ensure sustainability of data collection for rare disease cohorts whether in a funded or unfunded (clinic-based) setting; some measures are of course more applicable to a funded setting. Other than the measures above, other strategies that have proven to be effective in retaining patients in a longitudinal study or in studying challenging populations include: collection of multiple pieces of contact information (e.g., phone and email information for patients, their friends, their families, healthcare providers, social services agencies where applicable, and health insurance or medical chart numbers), regular updates of contact information, flexibility in assessment locations (e.g., allowing patients to go to a local specialist for some visits rather than travelling a long distance to a study centre for every
visit) or mode of data collection (e.g., phone interviews, electronic or paper forms), hiring staff with good interpersonal skills to build positive relationships with participants, keeping in contact with birthday cards and/or study newsletters and monetary incentives\textsuperscript{168,182-184}. In addition, technology is now available that can be leveraged towards this end: for instance, through the use of a secure electronic health records website or smartphone app data collection from either participants or local physicians who are sharing in the care of such patients; this could save monetary and time costs for all involved in the study\textsuperscript{185,186}. These newer technology aids to data collection would help patients continue to participate even if they were to move away from the centre where they were recruited. \textit{I would certainly ensure that in any future prospective cohort study I design, whether clinic-based or multi-centre network based, multiple strategies are built in at the outset.}

Even though there have been many advances in strategies – both design and analysis – to handle missing data in longitudinal studies, there are limited strategies to tackle missing data in a retrospective study setting with an irregular visit schedule\textsuperscript{187}.

In terms of study design, it has been suggested that \textit{auxiliary variables} – measures of the main dependent variable that are not in the equation and that can be used to predict missing values – should be collected and used to help mitigate bias\textsuperscript{188,189}. Given the retrospective nature of my studies, this was not a viable strategy for me. I managed to collect data from 9\% of patients lost to follow-up in my clinical studies as a strategy suggested to explore mechanisms of missingness\textsuperscript{190}. Exploration with appropriate disease duration adjustments, showed that missingness was predictable by observed data (therefore, at least MAR)\textsuperscript{104}. Arguably, this small group that provided data despite being lost-to-follow-up might not be a truly random sample but this was deemed to still be helpful, a decision supported by findings from previous missing data
methods research\textsuperscript{190}. The possible mechanism of missingness was considered in my decisions for analytic strategies in this thesis. In our study on analyzing longitudinal data with irregular visits, we have made some recommendations to help make the assumptions of visits missing at random (and therefore missing data being more likely MAR) more plausible and to allow more robust analysis of such data\textsuperscript{191}. A protocol for follow-up should be specified, but in contrast to the standard visit schedule, sicker patients can have a different visit schedule compared to those less sick. Recommended visit intervals and the reasons for deviation from protocol specified visit schedules should be recorded\textsuperscript{191}. These strategies should be implemented in clinic-based cohorts to allow better discernment of the nature of missing data/visits (whether greater extent of MAR or MNAR). Ultimately, these measures will allow a more appropriate and accurate choice of analytic methods.

Most of the current strategies in mitigating bias from missing data \textit{during data analysis} are based on a prospective study design, the use of a standardized and finite measurement schedule and variables built into the study design to enable subsequent exploration of the missing data mechanism\textsuperscript{107}. Missing data indicator methods and missing data diagnostics could not be used in my irregular schedule study data, as each patient has an individualized measurement schedule and there is no finite data end point\textsuperscript{107,192}. Multiple imputation is also not an option due to the irregular schedule. I incorporated elements in my study design to allow me to explore missing data mechanisms (as above). This is helpful as the correct choice of model is important for analyzing the data to obtain unbiased estimates of prognostic factors; for instance, wrongly using MNAR models can result in worse performance than MAR models\textsuperscript{193}. By incorporating the design strategies that we have suggested above, \textit{I will be better positioned to deal with missing data in my future, clinic-based, rare disease cohorts with an irregular visit schedule}. 
Longitudinal prospective cohorts take a long time to yield results. However, the data derived from such a design can potentially transform our understanding of disease evolution and outcomes. If there is a developmental process underpinning the outcome trajectory of interest, then a more time-efficient longitudinal design, such as an accelerated cohort design, can be used. An accelerated cohort design is a structured multi-cohort sequential design. Instead of following a cohort for the entire period of interest, multiple overlapping-age cohorts can be recruited and each followed for a shorter length of time, to generate population-average trajectory information. The caveat here is that the process studied must have age-related effects. This design allows disentangling of age, period and cohort effects. In this case, this method has been shown to result in a similar population trajectory as that derived from the single longitudinal cohort design. This approach has been used successfully to generate information about the marginal hearing trajectory of children who are hard of hearing in the first 10 years of their lives, with children in each cohort followed for just 3 years. I am interested in the young adulthood (between 18-30 years old) outcomes of patients with childhood-onset chronic rheumatic diseases, specifically their employment outcomes. Young adulthood is a period associated with important age-related developmental changes as individuals transition through several important stages in life – e.g., leaving home, graduating from school, entering employment – with increasing maturity and independence. I believe an accelerated cohort would be helpful to delineate employment trajectory, and to identify factors associated with adverse outcomes that can become targets of future intervention studies. Concerns about immortality bias have been raised about this design when used in a disease cohort. To address this concern, I will document the total eligible population at the point of transition (from pediatric rheumatology to adult rheumatology care) as well as important disease characteristics, such as severity, duration
of illness and irreversible damage (e.g., in joints or organs) of the eligible study population. I will then compare the characteristics of the eligible individuals to those of individuals subsequently recruited into the sequential cohorts. These measures will allow the information from such a study to be interpreted in context, with better understanding of possible biases. There will be missing data in the periods before entry into the various study cohorts but, as this is planned missingness, this is considered benign\textsuperscript{197}. This kind of missing data may reduce power but will not introduce bias\textsuperscript{195}. The accelerated cohort design also requires further work in extensions of current available analytic methods\textsuperscript{187}. These are the two areas that I will be working on in the near future.

### 6.4 Conclusions

With this body of work, I have advanced the field of rare disease prognosis studies. Using a rare disease (SLE) model, I identified areas of deficit in the design of rare disease prognosis studies that can become targets of improvement. I have demonstrated the use of multiple modern longitudinal analytic methods to answer questions of prognosis in a rare disease. Importantly, I have applied the methods in an observational setting where the visit schedule is irregular and likely informative of the underlying disease outcomes of interest, showing that longitudinal analytic methods can be used successfully in this manner. I have used Bayesian analysis to solve the problem of highly complex prognostic models in latent class trajectory analysis of dual outcomes trajectories. With this body of work, I have demonstrated how the application of these techniques in the setting of a rare disease clarifies our understanding of prognosis. The methods from this body of work can be easily applied to study other rare (or common) chronic, heterogeneous diseases with fluctuating disease courses to clarify questions of prognosis and potentially personalize management of patients according to predicted disease courses.
7 Appendices

7.1: Applying Modern Longitudinal Analytic Methods to an Irregular Visit Schedule Observational Cohort

This thesis used the longitudinal study design to examine disease trajectory in a cohort of cSLE patients. Longitudinal studies require special analytic methods, which have advanced tremendously in recent years. Traditional longitudinal analytic methods have restrictive assumptions (e.g. no missing data) and strict requirements of study design (e.g. balanced measurement schedule) that are not practical for observational disease cohort studies. Modern longitudinal analytic methods, such as mixed random effects models (MRM) and generalized estimating equations (GEE), have much more flexibility to deal with the realities of longitudinal observational cohort data. They allow for missing data (with different degrees of stringencies and assumptions of etiology), irregular measurement schedules, and complex covariance structures for the repeated measures.

Concomitant with the advancement of the longitudinal analytic methods, latent class models (considered a subset of structural equation modeling) have also been advanced. With the latent class models, the assumption is that there are different subclasses within a heterogeneous population that are not directly observable or measurable. Longitudinal latent class trajectory models were developed using modern longitudinal models along with latent class models to allow identification of clusters of similar longitudinal trajectories.

I have prepared a review and demonstration of the uses of modern longitudinal analytic methods to address different kinds of research questions as applied to an observational cohort of Juvenile Dermatomyositis (JDM) patients with an irregular, personalized visit schedule (see
paper following).
Comprehensive Project

Methods for Analyzing Observational Longitudinal Prognosis Studies for Rheumatic Diseases: A Review & Worked Example

Using a Clinic-Based Cohort of Juvenile Dermatomyositis Patients

Open Access, Pediatric Rheumatology, BioMed Central

Published 29th March 2017

Lily Siok Hoon Lim, Children’s Hospital Research Institute of Manitoba, University of Manitoba; Department of Pediatrics, University of Manitoba; Institute of Health Policy Management & Evaluation, University of Toronto. Email: lim@chrim.ca.

Eleanor Pullenayegum, The Child Health Evaluative Sciences, SickKids Research Institute; Institute of Health Policy Management & Evaluation, University of Toronto. Email: eleanor.pullenayegum@sickkids.ca.

Rahim Moineddin, Department of Family and Community Medicine, University of Toronto. Email: rahim.moineddin@utoronto.ca.

Dafna D Gladman, Department of Medicine, University of Toronto; Toronto Western Research Institute; Deputy director, Centre for Prognosis Studies, University Health Network, Toronto Western Hospital. Email: dafna.gladman@utoronto.ca.

Earl D Silverman, Division of Rheumatology, SickKids; Physiology & Experimental Medicine Program, SickKids Research Institute; Department of Pediatrics & Immunology, University of Toronto. Email: earl.silverman@sickkids.ca.

Brian M Feldman, Division of Rheumatology, SickKids; Child Health Evaluative Sciences, SickKids Research Institute, Department of Pediatrics, University of Toronto; Institute of Health
Policy Management & Evaluation and the Dalla Lana School of Public Health, University of Toronto. Email: brian.feldman@sickkids.ca.

Corresponding author:

Lily SH Lim

Address: RR149– 800 Sherbrook St, Winnipeg, Manitoba R3A1M4, Canada

Email: llim@chrim.ca

ABSTRACT

Most outcome studies of rheumatic diseases report outcomes ascertained on a single occasion. While single assessments are sufficient for terminal or irreversible outcomes, they may not be sufficiently informative if outcomes change or fluctuate over time. Consequently, longitudinal studies that measure non-terminal outcomes repeatedly afford a better understanding of disease evolution.

Longitudinal studies require special analytic methods. Newer longitudinal analytic methods have evolved tremendously to deal with common challenges in longitudinal observational studies. In recent years, an increasing number of studies have used longitudinal design. This review aims to help readers understand and apply the findings from longitudinal studies. Using a cohort of children with juvenile dermatomyositis (JDM), we illustrate how to study evolution of disease activity in JDM using longitudinal methods. 3 (120 words)

Keywords: epidemiology, longitudinal study, childhood-onset dermatomyositis, biostatistics

Total word count: 3695
BACKGROUND

Outcome studies in rheumatic diseases have generally focused on single-occasion outcome assessments, e.g., 5-year outcome. When the outcome of interest is not a terminal event (e.g., death), but a dynamic one (e.g., disease activity, functional status), a cross-sectional view is usually not the best way to look at the data. Two patients may have similar outcomes at a point in time but how they arrived at their outcomes may have been very different. To understand the disease course, information about how outcomes change over time is necessary. By measuring patients’ outcomes repeatedly (by definition, on ≥ 3 occasions), a longitudinal study provides information about the shape of outcome trajectory, e.g., whether the disease goes into remission, waxes and wanes or remains persistently active.¹

Longitudinal studies require special longitudinal statistical analysis. Although some of these methods have been available for many years, they are not commonly used in the literature. These complex methods are harder to understand and use. This paper aims to: 1) Provide a review of common methods used to analyze longitudinal trajectory data; and 2) Demonstrate how to interpret results from longitudinal trajectory analysis.¹⁹⁸⁻²⁰⁰ We will focus on application of these methods to a real-life clinic-based rheumatic disease cohort.

Questions that can be addressed by a longitudinal study

A longitudinal study can answer questions about the sources of variability in observed outcomes. In studies where outcomes are assessed once, the differences in outcomes are usually attributed to differences between individuals. The effects of within-individual differences cannot be differentiated from that of between-individual differences in cross-sectional studies. In contrast, by measuring the outcomes repeatedly over time— for each individual— the longitudinal
design captures important prognostic information about within-individual differences and allows these effects to be distinguished from between-individual differences.

Prognostic factors that fluctuate or emerge later during the course of the disease can only be understood using the longitudinal design. Examples of prognostic factors that often vary over time may include a biomarker of disease activity, or alterations in treatment. Using a longitudinal study design, we can repeatedly and simultaneously measure both the time-varying prognostic factors and the outcome(s) of interest, allowing direct relationships to be established.

Special considerations in analyzing the longitudinal study

Methods to analyze longitudinal data have been available since the 1960s. Traditional longitudinal analyses include repeated measures analysis of variances (RANOVA) and multivariate analysis of variance (MANOVA). Newer methods of longitudinal analysis include the generalized estimating equation (GEE), mixed effects regression model (MRM), latent class trajectory analyses (LCTA), joint modeling, and multi-state modeling. For illustrating longitudinal disease trajectory in this review, we will focus on the first four of these newer models. In this section, we will also compare the traditional with the modern methods of longitudinal analysis.

In a longitudinal study, each individual contributes at least three observations (by definition). As observations originating from the same individual are less variable than those originating from different individuals, longitudinal analysis needs to account for this relationship. If analyzed without consideration for within-individual correlations, the conclusion will be inaccurate. However, traditional methods like RANOVA have highly restrictive assumptions, such as the assumption that the correlation between two measurements is constant, i.e., the correlation between measurements is similar, whether the measurements have been two days or
Appendix 1

Applying LAMs

two years apart. In contrast, newer methods attempt to account for the fact that within-individual correlations likely vary over time \(^{210}\) (Appendix).

In a longitudinal study, patients may be followed for differing lengths of time, resulting in a different number of visits for each patient and different visit schedules among patients. While traditional methods require an equal number of visits and/or the same schedule of visits, newer methods can accommodate an unequal number of visits and irregular measurement schedules \(^{210}\).

Missing data is inevitable in observational longitudinal studies. As patients are followed over long periods, there will be times when patients may leave a cohort and then return, or be lost to follow-up. Traditional longitudinal methods have a requirement of no missing data, which is impractical in longitudinal observational studies \(^{211}\). In contrast, newer methods are able to handle missing data with varying degrees of flexibility \(^{203,204,212}\) (Appendix).

We will now apply four of the longitudinal methods—GEE, MRM, LCTA and Joint Modelling— to illustrate the use of these methods in an observational cohort. We have chosen these four models, as they all provide a view of the shape of the longitudinal outcome trajectory and form the basic models from which more complex models can be developed.

**JUVENILE DERMATOMYOSITIS AS A DISEASE MODEL FOR LONGITUDINAL ANALYSIS**

We will use Juvenile Dermatomyositis (JDM) as a convenient disease model to show how longitudinal design and analysis can be used in rheumatic diseases. Multiple cross-sectional studies have determined that, when assessed 2-3 years after diagnosis, there are three disease course patterns in JDM: monocyclic, polycyclic or chronic \(^{213-216}\). Although previous studies have shown that a substantial proportion of patients have active disease many years after the
Appendix 1

Applying LAMs

diagnosis \(^{217,218}\), these studies could not differentiate patients with active disease throughout their entire disease course from those with a polycyclic course. A longitudinal study can help to clarify this question.

The study population was 95 JDM patients followed at The Hospital for Sick Children, Toronto, Canada. Information about this cohort has been previously reported \(^{219,220}\). We used clinical data from the first four years of follow-up to demonstrate the application of longitudinal analytic methods. The frequency of patients’ visits was based on the severity of their disease, i.e., the visit schedule was irregular across the population. See Appendix for the baseline characteristics of this cohort.

The primary outcome was the modified Disease Activity Score (DASm) \(^{220,221}\). The skin component of modified DAS (SDAS) scores up to 4 points and the musculoskeletal component (MDAS) up to 7 points.

We used: 1) The DASm at diagnosis (bDAS) as an example of a time-invariant (unchanging) prognostic factor; and 2) The steroid dose (in mg/kg) or methotrexate use (yes or no) from visits before each DASm measurement as examples of time-varying (changing) prognostic factors. As we used bDAS as a predictor, we excluded the first visit DASm (bDAS) in this dataset. We tested 3 different lag times (of 3, 6 and 12 months) when the time-varying predictor (treatment) was measured (e.g., methotrexate use 3 months before, 6 months before or 12 months before DASm measurement).

**QUESTIONS AND ANSWERS**

*Question 1: What is the disease activity course for a population of JDM patients?*

The GEE, which determines the mean population disease activity trajectory, is frequently used to answer this kind of question \(^{203}\). The GEE calculates the average DASm of the whole
population at each visit. These population averages are then joined to make a “trajectory” of DASm for the whole population over time. The GEE assumes that the measurement schedule is unrelated to the outcome, i.e., both sick and well patients are presumed to have the same visit schedule on average. As sicker patients were likely seen more frequently than those who were well, this assumption was probably not supported in our clinic-based cohort. However, there have been extensions of the GEE (the weighted GEE) that allow GEE to be used even when the measurement schedule is related to the outcome.

The GEE is the most popular method used for longitudinal analysis for several reasons. With the GEE, the mean population response at each occasion is modelled as the result of only the prognostic factors of interest and not of previous responses or random effects (individual heterogeneity). If the researcher’s interest is in the average population prognosis, then the GEE provides a simple interpretation.

We will briefly discuss the results when we used the GEE to analyze our study cohort. From the smoothed local regression plot of Figure 1, we can see that the average population DASm decreases rapidly and then plateaued to a lower degree of disease activity. This is clinically sensible as there are effective treatments for lowering disease activity in JDM, yet many patients remain chronically active.

In summary, the GEE is a good choice for analyzing studies where the overall population prognosis is the primary interest, e.g., when addressing questions about population trends or healthcare utilization.

**Question 2a: How do we determine the disease activity course of an individual with JDM?**

The best way to address this question is to use the MRM, a so-called subject-specific model, that allows for potential inference at the level of the individual. For a continuous
Appendix 1

Applying LAMs

outcome like DASm, the MRM also provides us with the average population trajectory. The smoothed plot (Figure 2) shows the DASm trajectories of both the individual patients (gray lines) and the mean population trajectory (bold black line). As the outcome is continuous, fitting the MRM resulted in a similar longitudinal trajectory to that derived from the GEE. However, if the outcome is not continuous, calculating the mean population trajectory from the MRM is very complex.

In the MRM, each individual’s trajectory is different from the mean population trajectory because of “random effects.” Random effects are a combination of unmeasured prognostic factors, confounders, environmental factors and genetic factors, which account for heterogeneity within the population. Any difference between an individual’s trajectory and the average population trajectory is the result of that individual’s random effects; this forms the basis for individual level inference from the MRM. Random effects may be applied at intercepts, which represent between-individual differences. When random effects are applied to slopes, this allows individuals’ time trend to deviate from the average population trajectory. Although the random effects are taken into account in modeling, the results reported are that of the mean population trajectory in this case.

Although the results of the MRM may appear much like the GEE in continuous outcomes models, the interpretations are very different. The MRM models the average of individuals’ trajectories (subject specific) but the GEE models the trajectory from DASm averages at each time point (population average). Practically, in the context of the continuous outcomes models, the forms and the results from the 2 models are very similar. The differences between MRM and GEE are obvious when the outcomes are non-continuous, e.g., binary or ordinal. Taking the example of a random intercepts (logistic) MRM, where the outcome could be disease
remission (yes/no), and a time-invariant covariate could be calcinosis at baseline (yes/no), the odds of remission would correspond to the effect of calcinosis plus an individual’s random effects. The effect of calcinosis therefore changes with different individuals. The odds ratios estimated by the logistic MRM are subject specific as these additionally adjust for heterogeneity between individuals. In the binary outcome GEE, the odds ratio corresponds to odds of an event among those with baseline calcinosis divided by the odds of an event in those without. As these are ratios of subpopulation risks, the GEE estimates are termed population average. From this, it should also be obvious that the results of non-continuous MRM will always be greater that than of non-continuous GEE due to inclusion of the random effects. The choice of model depends on the purposes of the study. If the mean prevalence of disease remissio
ty predicted by baseline calcinosis is of interest, the GEE is suitable. If the investigators want to study individuals’ risk factors for personal predictions, then the MRM is the model of choice. As the estimates of the MRM are adjusted for random effects (unobserved individual characteristics), the results reflect the effect of baseline calcinosis status of an individual with a specific random effect (or individuals with the same random effect).

*Question 2b: How do we predict the disease activity course of an individual with JDM?*

i) *Time-invariant prognostic factor*

The MRM can be used to determine the individual level predictive effect of a time-invariant prognostic factor, such as the baseline DASm (bDAS), on the trajectory of disease activity. From Figure 3, we can see that the average trajectory for individuals with a higher bDAS had a faster initial improvement (up to month 10) followed by persistently higher activity, compared to the average trajectory of individuals with a lower bDAS (p<0.0001). Baseline state therefore predicted the rate of change of the slope of the trajectory. We can therefore tell patients
that their disease activity will likely improve within 1 year (with treatment). Those who are more active initially (higher bDAS) will likely see a relatively faster improvement in their symptoms compared to those who are less weak.

ii) Time-varying prognostic factor

Using the MRM, we tested the individual level effects of time-varying prognostic factors, including corticosteroid and methotrexate exposures from 3, 6 and 12 months before each visit. In this case, the time-varying predictors tested did not influence the DASm trajectory. Had these time-varying predictors been significant, their effects would have been to shift the trajectory up or down, corresponding to slower or faster resolution of disease activity.

In summary, if the primary interest of a study is to understand how covariates impact disease course, based on the individual patient’s prognosis trajectory, the MRM is the best choice. The MRM is also more suitable when there is significant heterogeneity in individuals’ disease courses in the population.

Question 3a: Are there different patterns (or subgroups) of disease activity course among individuals with JDM?

The models used in the preceding sections assumed that the shape of patient trajectories were homogenous, i.e., all patients follow the same disease course. If the investigator suspects that his study population contains heterogeneous patient trajectories, then latent class trajectory analysis (LCTA) (Appendix) can be used to identify distinct subgroups with different disease trajectories (those with a clinically distinct prognosis) 145,229.

When we applied LCTA to our cohort, we found three distinct subclasses. Each of these three subclasses has a class-specific average population disease activity trajectory, derived using individuals’ trajectories (with the MRM as in question 2a). From Figure 4, we can see that Class
2 patients have high disease activity at diagnosis, and then rapidly improve. Class 1 patients have moderate disease activity at diagnosis that improves gradually to low disease activity over time. Class 3 patients have high disease activity at diagnosis that improves very gradually. For this kind of model, we determine a probability of belonging in each of the three classes but for each individual, but the probability of any individual of belonging to each class varies. When classified according to the highest probability of class membership for each individual, 42% are in class 1, 55% in class 2 and 3% in class 3. None of the patients was classified into 2 classes with similar probabilities, i.e., was ambiguous in the class membership probability (see Appendix).

The LCTA is very flexible (Appendix): 1) More than one and more than one kind of outcome trajectory can be modelled at a time (see next section); and 2) Trajectories of different subclasses can take on different shapes (reflecting different patterns of outcome evolutions)\(^\text{171,229,230}\).

**Question 3b:** What factors predict an individual’s membership in the different subgroups of disease activity?

The LCTA allows us to study the effects of time-invariant baseline factors, such as bDAS, in predicting membership in distinct subgroups. We use the bDAS to predict membership in the three subclasses. The higher the bDAS, the less likely is a patient to belong in class 1 (OR 0.11, 95% CI 0.02-0.63) or 2 (OR 0.27, 95% CI 0.05-1.41) compared to class 3. Furthermore, the higher the bDAS, the less likely is a patient to belong to class 1 (OR 0.41, 95% CI 0.25-0.68) compared to class 2. This means that those with highest bDAS are most likely to belong to class 3, with chronic high-grade disease activity. High bDAS patients are less likely to improve substantially or go into low disease activity states unlike those whose initial bDAS are low to
moderate. This may appear different from the results of MRM where high bDAS predicted quicker resolution. In MRM, the whole population was studied together although there was significant heterogeneity within the population. LCTA allows us to group patients into more homogeneous subgroups and then more precisely clarify the effect of potential baseline membership predictors.

Time-varying covariates can also be studied in LCTA and can be formulated in different ways depending on the underlying question. As the computations are very complex and interpretations challenging, we chose to leave out time-varying covariates for this review.

In summary, if a researcher suspects that there are several prognostic subgroups within the study population, the LCTA can be used to identify these subpopulations. Baseline factors may identify patients’ memberships in different prognostic subpopulations, thus allowing more individualized management. Time-varying covariates can also be studied but interpretations are more complex (Appendix).

Question 4a: What are the separate disease activity courses for the skin and musculoskeletal components of JDM and what is the relationship between these two disease components?

In this case, we have two potential outcomes: the modified skin (SDAS) and musculoskeletal (MDAS) components of JDM. We are interested in how the disease courses evolve over time and the correlation of disease activity in the two components of JDM over time. The best way to study how two different outcomes evolve simultaneously over time is a joint model as shown in Figure 5 (Appendix). The joint model refers to the concurrent modeling of two outcome trajectories (SDAS and MDAS) in one model. The basic model used here to model the trajectories is also the MRM.
The two disease components of JDM follow different trajectories. The MDAS improves more rapidly (within the first 10 months after diagnosis) to minimal activity. In contrast, the SDAS improves less rapidly and persists at a low activity level for almost 20 months (Figure 5). The two disease components therefore follow discordant courses, as confirmed by the low correlation between the two trajectories ($r = 0.32$)\(^{172}\). This relationship between two or more outcome trajectories can only be studied using the joint model and not by using separate models for the individual components.

In our example above, both outcomes were treated as continuous outcomes. However, outcomes of different natures can be jointly modeled. We can model a continuous outcome (e.g., DASm), with a time-to-event outcome (e.g., calcinosis), a binary outcome (e.g., nailfold capillary abnormality) or a count (e.g., active joint count). Furthermore, as alluded to previously, more than one outcome can be jointly studied in the LCTA; in this case, members within the same subclass of joint outcomes would have similar trajectories or risk of events\(^ {171,230,234}\). This way of examining outcomes more closely resembles the real world where several clinically relevant outcomes may be considered equally important in clinical decision making.

*Question 4b*: What factor(s) predicts the joint disease activity courses of the skin and musculoskeletal components of disease in an individual with JDM?

We evaluated the predictive effects of bDAS (time-invariant) on the joint MDAS and SDAS trajectories (Table 1). In this joint model, individuals with higher bDAS have a faster decline in the slope of the MDASm trajectory; but bDAS does not significantly affect the slope of the SDAS trajectory. Prior treatment with methotrexate or steroid (time-varying factors) does not significantly influence the MDAS and SDAS trajectories, perhaps because everyone received
Applying LAMs

the same protocol of treatment. Thus, joint modeling can distinguish the prognostic effects of predictors for the major components of JDM disease activity.

Table 7.1.1: Predictors identified from the joint multivariate model of MDAS and SDAS

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Outcome</th>
<th>Predictor Estimate</th>
<th>Standard Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>bDAS</td>
<td>MDAS</td>
<td>0.0363</td>
<td>0.0316</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>SDAS</td>
<td>0.0214</td>
<td>0.0342</td>
<td>0.53</td>
</tr>
<tr>
<td>Time₁*bDAS</td>
<td>MDAS</td>
<td>1.0140</td>
<td>0.1321</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>SDAS</td>
<td>0.0362</td>
<td>0.1261</td>
<td>0.77</td>
</tr>
<tr>
<td>Steroid²</td>
<td>MDAS</td>
<td>-0.0919</td>
<td>0.0753</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>SDAS</td>
<td>0.1211</td>
<td>0.0688</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Baseline DASm measurement, bDAS.
1The shapes of the MDAS and SDAS models are defined using 2 time terms each (fractional polynomials). MDAS and SDAS crossed with their common time form (p= −1.,) of their respective fractional polynomials (see Appendix).
2Steroid treatment from 3 months before each occasion of DASm measurement, i.e., it is a time-varying predictor. Significant results have been bolded.

In summary, if a researcher is interested in more than one (or more than one kind of) outcome and wants to evaluate the relationship (i.e., correlation) among these outcomes over time, the joint model can be used. The joint model can be used in the LCTA context so that latent classes of the different outcomes of interest can be delineated. The joint model also allows direct comparisons of the differing effects of the same predictors on multiple outcomes simultaneously.
CONCLUSIONS

In this paper, we have illustrated the use of longitudinal design and analysis in studying prognosis. Longitudinal studies are superior to cross-sectional studies as they use all available data, thereby giving a more complete view of patients’ outcomes.

We have presented four different ways of analyzing longitudinal observational data depending on the question(s) of interest. The GEE is best used if a population-level view is preferred and the visit schedule is not related to the outcome studied. This population view may be more relevant to health services researchers addressing population level questions, e.g., healthcare utilization. The MRM should be used if an individual level view is of interest. The LCTA should be used when the researcher wants to identify subgroups of patients within a heterogeneous cohort, with different outcome trajectories, and identify the factors predicting their membership in these subgroups. Joint modeling is best used to study the evolution and correlation among multiple outcome trajectories and the differing effects of predictors on simultaneous multiple outcome trajectories. We have summarized salient points about the four models in an overview table (Table 2). These four methods can also be combined with other methods, e.g., propensity scoring\textsuperscript{235} and marginal structural modeling\textsuperscript{236}, to answer other kinds of questions (e.g., therapeutic) in a longitudinal study\textsuperscript{237}.

While the methods presented in this paper have the potential to transform our understanding of prognosis, we acknowledge that these methods could be challenging to use without necessary expertise. We therefore recommend consulting with biostatisticians knowledgeable in these methods to help design and analyze longitudinal studies. Our review is not meant to be an exhaustive review of all available longitudinal analytic methods. For example, Markov multistate models can also be used to determine patients’ transition between disease
states longitudinally. In the interest of simplicity, we tested a minimum number and kind of predictors in these models. In practice, more predictors can be tested in these models, with a far greater confidence in predicting individuals’ disease trajectories.

The methods outlined in this paper will allow for a more complete understanding of longitudinal outcomes and a more precise understanding of the effects of predictors. Combination of these methods with molecular information holds great potential to transform clinical practice towards the ultimate goal of precision medicine.
## Table 7.1.2: Overview of the 4 modern longitudinal analytic methods

<table>
<thead>
<tr>
<th>Model</th>
<th>Questions</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEE</td>
<td>What is the averaged outcome trajectory for the population? (Trajectory of averages)</td>
<td>Parameter estimates robust to misspecification of the covariance structure. Both time-invariant and time-varying predictors can be studied.</td>
<td>No individual level inference Assumes missing data to be missing completely at random (MCAR), which may not be true for many longitudinal studies.</td>
</tr>
<tr>
<td>MRM</td>
<td>What is the outcome trajectory of the individual? What is the average outcome trajectory for the population? (Average of trajectories)</td>
<td>Individual level inference possible with the incorporation of random effects. Both time-invariant and time-varying predictors can be studied. Assumes missing data to be missing at random (MAR), which is more likely in longitudinal studies.</td>
<td>Misspecification of covariance structure may bias parameter estimates.</td>
</tr>
<tr>
<td>LCTA*</td>
<td>Are there distinct subgroups within the study population? What are the trajectories of the identifiable subgroups within the population?</td>
<td>Objectively identifies latent distinct subgroups within a heterogenous population. Able to use time-invariant factors to predict group membership. Able to study effects of time-varying covariates in different ways (depending on question and underlying theoretical framework)</td>
<td>Complex and time-consuming computing procedures. Interpretation of time-varying covariates can be challenging depending on the formulation.</td>
</tr>
</tbody>
</table>

---

238
<table>
<thead>
<tr>
<th><strong>Joint Model#</strong></th>
<th>What are the trajectories of (multiple) outcomes of interest?</th>
<th>Multiple outcome trajectories of disparate nature (e.g., continuous with binary, binary-Poisson, continuous-survival) can be studied simultaneously.</th>
<th>Objective determination of the longitudinal correlation of the trajectories. Joint model with time-to-dropout may be used as a means to adjust for data missing not at random (MNAR).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What is the correlation between the outcome trajectories of interest (i.e., are the trajectories concordant or discordant)?</td>
<td></td>
<td>Modeling procedures can be complex with increasing number and kinds of outcomes modeled jointly.</td>
</tr>
</tbody>
</table>

*Usually modeled with MRM as the base model.

#MRM may be used as the base model for continuous, binary and count data. Proportional hazard is used for time-to-event outcomes.
Figure 7.1.1: Population-averaged modified DAS trajectory in JDM patients

(representation of GEE)

Disease activity score, DAS. Total population= 95.
Figure 7.1.2: Plot of all JDM patients’ individual and the population-averaged modified DAS trajectories (representation of MRM)

Disease activity score, DAS. Total population= 95.
Figure 7.1.3: Effect of baseline (time-invariant) modified DAS on the modified DAS trajectory in JDM patients (MRM)

Activity Score, DAS; Baseline modified DAS (bDAS).
Figure 7.1.4: Latent classes of JDM disease activity trajectories (LCTA)

Disease Activity Score, DAS. Class 1, n=49; Class 2, n=43; Class 3, n=3.
Figure 7.1.5: Musculoskeletal and skin disease activity trajectories in JDM (joint multivariate modeling)

Disease Activity Scores, DAS; Musculoskeletal DAS, MDAS; Skin DAS, SDAS. MDAS and SDAS are components of the modified DAS.

The left y axis (black) represented the MDAS (maximum score=7) and the right y-axis (grey) represented the SDAS (maximum score=4). The MDAS and SDAS curves are colour coded to match their respective axes.
List of abbreviations

RANOVA, repeated measures analysis of variance; MANOVA, multivariate analysis of variance; GEE, generalized estimating equation; MRM, mixed effects regression model; LCTA, latent class trajectory analysis; JDM, juvenile dermatomyositis; DASm, modified disease activity score; SDAS, skin-disease activity score; MDAS, musculoskeletal disease activity score; bDAS, baseline disease activity score.

Declarations

Ethics approval

Ethics approval was obtained for the use of JDM data in this review at SickKids Toronto (1000026088). Consent was waived for this study due to the nature of the study.

Consent for publication

Not applicable.

Availability of data and material

The data is not available for public use as our research ethics board did not permit public sharing of data.

Competing interests

None of the authors has declared competing interests.

Funding

The Canadian Institute of Health Research (CIHR) and the SickKids clinician-scientist training program funded Dr Lim’s salary and research for this work.

Authors’ contributions
Lily Lim, Rahim Moineddin, Eleanor Pullenayegum, Brian Feldman conceptualized the design of the study. Lily Lim and Brian Feldman acquired the clinical data. Lily Lim, Rahim Moineddin and Eleanor Pullenayegum were involved in the analysis of data. All authors were involved in the interpretation of data, drafting and revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Authors’ information

1Children’s Hospital Research Institute of Manitoba, University of Manitoba; 2Department of Pediatrics, University of Manitoba; 3Institute of Health Policy Management & Evaluation, University of Toronto; 4The Child Health Evaluative Sciences, SickKids Research Institute; 5Department of Family and Community Medicine, University of Toronto; 6Department of Medicine, University of Toronto; Toronto Western Research Institute; 7Centre for Prognosis Studies, University Health Network, Toronto Western Hospital; 8Division of Rheumatology, SickKids; 9Physiology & Experimental Medicine Program, SickKids Research Institute; 10Department of Pediatrics, University of Toronto; 11The Dalla Lana School of Public Health, University of Toronto.
7.2: Supplementary Material for Appendix 1

Appendix accompanying “Methods for Analyzing Observational Longitudinal Prognosis Studies for Rheumatic Diseases: A Review & Worked Example Using a Clinic- Based Cohort of Juvenile Dermatomyositis Patients”

This supplementary document has been prepared to include further information about the study cohort used for illustration and to provide interested readers a more in-depth understanding of the technical aspects of newer longitudinal analytic methods.

STUDY COHORT

We used a cohort of Juvenile Dermatomyositis (JDM) patients followed at our centre for illustration of the application of modern longitudinal analytic method. Patients were diagnosed and followed between 1st January 1991 and 31st December 2010. Table A1 shows the baseline characteristics and follow-up information of this cohort of patients.

Table 7.2.1: Baseline characteristics and follow-up information of study patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: Males (%)</td>
<td>33 (35)</td>
</tr>
<tr>
<td>Median age at diagnosis, years (25%-75% percentile)</td>
<td>7.8 (4.9-12.1)</td>
</tr>
<tr>
<td>Time-period of diagnosis, n (%)</td>
<td></td>
</tr>
<tr>
<td>1991-1999</td>
<td>37 (39)</td>
</tr>
<tr>
<td>2000-2010</td>
<td>58 (62)</td>
</tr>
<tr>
<td>Median baseline DAS (modified) at first visit (25%-75% percentile)*</td>
<td>7 (6-9)</td>
</tr>
<tr>
<td>Median duration of follow-up, years (25%-75% percentile)</td>
<td>3.5 (2.4-3.8)</td>
</tr>
<tr>
<td>Median number of visits per patient (25%-75% percentile)</td>
<td>14 (10-16)</td>
</tr>
<tr>
<td>Total follow-up of the cohort, patient years</td>
<td>285.64</td>
</tr>
</tbody>
</table>

* Disease Activity Score, DAS.

DATA ANALYSIS

We first examined the modified disease activity score (DASm) trajectories of all patients within this cohort (Fig 1) for possible trajectory shapes. From perusal of the individual
Appendix 2: Supplementary Material for Appendix 1

trajectories, it was clear that the relationship of DASm with time was non-linear and did not clearly fit with any conventional low order polynomials.
Appendix 2: Supplementary Material for Appendix 1

Figure 7.2.1: Plot of all JDM patients’ modified DAS trajectories

Disease Activity Score, DAS. This referred to the modified version (DASm).

In order to model the evolution of the DASm trajectories more accurately, we tested various combinations of fractional polynomials. We tested first order (m=1) and then second order (m=2) fractional polynomials. For the second order fractional polynomials, we tested the powers (p) that have been suggested by Royston to fit most biomedical data, i.e. p= -2, -1, -0.5, 0, 0.5, 1, 2, 3. The DASm curve was best fitted by a 2nd order fractional polynomial (m=2,p= (-2,-1)), with maximal gain.

The analyses to be presented in the following sections were performed using SAS (Cary, North Carolina, USA) or R (http://www.r-project.org) where appropriate.

LONGITUDINAL ANALYTIC METHODS

1) GENERALIZED ESTIMATING EQUATION (GEE)
Appendix 2: Supplementary Material for Appendix 1

Currently, the GEE is probably the most commonly used analytic method for longitudinal studies\textsuperscript{203,240}. In the simplest form, this is like a linear regression, plus a correction factor for the within-person correlation of measurements. It is a marginal model, meaning it presents a population averaged view of the repeatedly measured outcomes at each time point.

How does the GEE answer our question about the modified disease activity score (DASm) trajectory in JDM patients? The GEE equation fits a trajectory connecting mean population DASm at each occasion of measurement, giving an overview of the population’s average DASm evolution over time. This has two implications. First, it is important to note that the marginal trajectory is a curve of averages (at the occasions of measurements), not an average of curves: we therefore cannot necessarily infer an individual’s DASm trajectory from this equation, especially if the outcome is non-continuous. Second, to accurately understand the population averaged DASm at each measurement, it is implied that there is a similar schedule of visits for the cohort; although some degree of irregularity can be tolerated. Furthermore, there is an assumption that this schedule of visits is not related to the outcome in any way. For cohorts where there are no regular schedules of assessments or where patients may visit more often if they are sicker (i.e. the visit schedule is informative), the GEE is not a good choice of analytic method\textsuperscript{241}. In such cases, an extension of the GEE, weighted GEE (WGEE), which uses inverse probability weightings of the visits, will be more appropriate\textsuperscript{169}.

In the GEE, within-subject correlation is accounted for by including a covariance structure in the model. As GEE uses quasi-likelihood, this covariance matrix need not be completely accurate for the estimates to be consistent\textsuperscript{203}. Misspecifying the covariance matrix will still however negatively impact on the efficiency of the model.
The GEE does allow missing data. The assumption here is data are missing completely at random (MCAR)\textsuperscript{77,242}. This means that the missing data must be neither related to the missing outcome nor be predictable by covariates or the history of outcomes. An example of this would be a child who stops attending our JDM clinic as his family has relocated to another province for his father’s new job. This has nothing to do with his disease condition and is not predictable based on his intrinsic characteristics or his disease outcomes up to that point. The MCAR assumption may be difficult to fulfill in real life cohorts. It is important to explore the mechanisms of missing data as having missing data that are not MCAR will bias the inference derived from the GEE. If the interest is in population inference and there is evidence of data missing at random, the WGEE can be used instead as it tolerates missing at random (MAR) mechanism for missing data. MAR missing data are not related to the missing outcome but may be predictable from covariates (e.g. patient characteristics) or outcomes up to the point of attrition. An example of this would be the observation that JDM patients who live further away (patient characteristic) tend to not return for visits when they are better (history of outcomes up to attrition).

Both time-invariant and time-varying predictors can be evaluated using GEE. The time-invariant predictor associates with the intercept and predicts the slope of the trajectory. To interpret the meaning of a time-varying predictor in GEE however, one needs to consider where the source of variation is, whether it is between or within-individual. In this case, as most subjects were started on a standard dose of prednisone (i.e. minimal between-individual variation), most of the variation occurred within the individual over time. Conversely, if a time-varying covariate varies more between individuals and fluctuates less within an individual, most of the variation would be between-individual and such a variable would predict differences
between individuals. One can easily see that if a time-varying covariate varies between individuals and over time, interpretation becomes difficult. If the investigator wants to use GEE to study time-varying predictors and discern the effects more clearly, he/she will have to separately specify the effects, both between-individual and within-individual, using additional variables.

In summary, the GEE is a population view model that is good for studying population level mean outcome trajectory over time. It does not offer the possibility of individual level inference. The GEE also has a potentially restrictive assumption regarding missing data that needs to be taken into consideration when choosing this method. This model should be used when the population view is desired, e.g. healthcare utilization.

2) MIXED EFFECTS REGRESSION MODEL (MRM)

The MRM has many synonyms: random effects, multilevel model, hierarchical-linear-model. It has been used extensively in fields such as psychology, behavioural sciences and criminology. In contrast to the GEE, the MRM is subject-specific and models each individual’s outcome trajectory.

In MRM, each subject has a subject-specific outcome trajectory. All subjects’ trajectories are summarized into an average population trajectory. The equation below illustrates a simple case of a linear trajectory for an individual:

\[ Y_{it} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{it} + \varepsilon_{it} \]

Outcome Y for person i at time t
Intercept for person i
Slope for person i
Random error for person i at time t
Appendix 2: Supplementary Material for Appendix 1

In the equation above, $\beta_0$ represents the mean population intercept and $\beta_1$ the mean population slope. Each individual i’s trajectory is related to the population intercept and slope (or whichever forms the time function takes) through their random effects $b_{0i}$ and $b_{1i}$ respectively. The random effects are unique person-specific characteristics that contribute to variability in outcome trajectories within a cohort. The random effect represents a combination of measured and unmeasured factors, e.g. genetic factors, environmental factors, prognostic factors and confounders. The final term in the equation $\varepsilon_{it}$ represents the random error on each occasion of measurement, such that the actual values of the outcome measure would vary around each individual’s mean outcome trajectory.

Both the GEE and the MRM produce a population trajectory. While the mean trajectory in GEE is a curve of averages, the mean trajectory in MRM is an average of curves. As alluded to in the previous section, the mean outcome trajectory in GEE cannot necessarily be inferred to resemble an individual’s trajectory, especially in the cases of non-continuous outcomes. In the case of MRM, the individual’s outcome trajectory can be inferred from the population average outcome trajectory and is individualized by that individual’s random ($b_{0i}, b_{1i}$) and fixed effects.

The MRM allows data to be MAR. This is a much less restrictive assumption than the MCAR assumption for GEE. This kind of missing data is probably also more common in real-life clinic-based cohorts and therefore a more realistic assumption.

Both time-invariant and time-varying predictors can also be evaluated in the MRM. The time-invariant first encounter baseline DASm (bDAS) is an example of a between-individual prognostic factor that serves to distinguish patients’ responses. Time-invariant predictors associate with the intercept and predict the slope of disease trajectory. This kind of baseline factor may potentially be used to select different management strategies for individual patients.
who have been predicted to follow certain trajectories. A time-varying factor is an example of a within-individual prognostic factor that predicts within-individual change over time. This kind of factor may be targeted to improve the trend of the outcome trajectory within an individual over time.

As mentioned earlier, within-individual observations are usually correlated. Failure to adjust or misspecifying the covariance structure may lead to misleading conclusions from this model. In addition to the common covariance structures available in GEE, the MRM also allows easy implementation of spatial covariance structures. These covariance structures were initially developed for geospatial modeling. Spatial covariance structures are suitable when visits were not the result of a pre-planned schedule. When used in longitudinal models, the “distances” in spatial covariance structures are treated as time intervals.

In summary, the MRM is a subject view model that is potentially helpful in individual level inference of disease trajectory over time. One can say that an individual with a certain prognostic factor may follow a trajectory like so. In addition, it also offers a population view through the average population trajectory. The MRM can accommodate irregular and individualized visit schedules. This model should be used when the more refined individual view is preferred, such as the physician trying to refine management strategy according to patient characteristics or seeking to inform patients regarding prognosis.

**Additional model information in applying the GEE & MRM**

For the GEE model, we modeled within-individual correlations using a first order autoregressive structure (AR(1)). For the MRM, we modeled within-individual correlations with a spatial exponential covariance structure to reflect the individualized visit schedules in this
Appendix 2: Supplementary Material for Appendix 1

observational cohort \(^{244}\). For the MRM, our model posits natural heterogeneity i.e. random effects, in the intercept and the slope.

3) LATENT CLASS TRAJECTORY ANALYSIS (LCTA)

As rheumatologists, we have often observed that patients with similar diagnoses follow different courses and have different outcomes. Within the same disease cohort, patients may cluster into smaller subclasses in terms of disease evolution. Researchers have traditionally applied arbitrary criteria to classify patients into groups, e.g. observed disease activity at certain defined times\(^{216,246}\). This approach runs the risk of being a self-fulfilling prophesy: there can only be as many groups as defined by the researchers and the outcome groups can only behave as defined by the investigators. It is more scientific to use a data-driven, statistical approach to identify these groups of patients with different prognoses. Furthermore, the use of a formal statistical structure also helps to distinguish random variations between individuals from real and systematically different variations.

Studies in social sciences and psychology have used latent class analysis to identify unobserved (i.e. latent) subgroups within a population, e.g. identifying latent classes of high-risk behaviors among adolescents\(^ {247}\). In much the same way, outcome trajectories of patients may cluster into a few more homogenous classes over time. Latent class trajectory analysis (LCTA) is the application of latent class analysis technique to longitudinal data, such that identifiable latent classes of (longitudinal) trajectories can be distinguished within the study population \(^{248}\). Combined with baseline or early prognostic factors that predict an individual’s membership in a certain class of outcome trajectory, we can potentially use this information to refine patient management.
Appendix 2: Supplementary Material for Appendix 1

There are two main methods within the group of latent class trajectory analyses: group-based trajectory modeling \(^{170,206}\) and growth mixture modeling (GMM) \(^{145}\). The GMM has gained great usage in literature. GMM identifies subclasses with distinct mean trajectories. Individuals’ trajectories within a subclass cluster around the mean class trajectory, deviating from it by their random effects, i.e. each class trajectory is defined by a class-specific MRM. GMM has been used to study diverse topics such as the development of bladder control in children, the prediction of treatment response among patients with interstitial cystitis, patterns of alcohol use, the evolution of aggressive behavior, the evolution of attention deficit/hyperactivity disorder, patterns of post myocardial infarction depression and patterns of fetal growth \(^{171,230,249-252}\). A few recent papers in osteoarthritis have also used GMM to predict disease course \(^{198,253}\). As such, we chose to demonstrate the application of GMM in this paper.

How can we use this class of model? We can ask if there are identifiable subclasses of outcome trajectories within a heterogeneous population of patients. The GMM is very flexible. Multiple outcome trajectories of disparate nature can be studied simultaneously (like the joint model in the previous section) \(^{171,230,250,254}\). Furthermore, subclasses of an outcome can also take on trajectories of different shapes, e.g., some may have a linear shape while others have a quadratic shape.

In the MRM, we modeled subject specific trajectories and averaged them into an average population trajectory. The assumption in MRM is that all patients belong in the same population. In LCTA, we recognize that there may be more than one subgroup within the population and more than 1 kind of disease trajectory. Using LCTA, we can split a heterogeneous population of patients into several subgroups, each with its own subgroup specific average population trajectory. Each subgroup’s average population trajectory is an average of all individuals’
trajectories within that subgroup. This is akin to fitting multiple MRMs in one analysis. The subgroups can have different shapes of trajectories, thus informing on how one disease evolves differently in the different subgroups. The LCTA therefore has similar potential for individual inference as MRM.

After we identify distinct subclasses, we will then proceed to identify the prognostic factor(s) that predicts membership in the subclasses, i.e., membership predictors. When subclasses are robustly identified, the membership predictors will be able to clearly assign individuals into specific subclasses (i.e. with clear high probability in one of the subclasses).

Membership predictors are often time-invariant predictors. There are several ways of dealing with time-varying covariates in the context of growth mixture modeling. They can be studied as modifiers of disease outcome trajectories, with the assumption of a common effect across all subgroups or within subgroups of trajectories (i.e. different effects between subgroups). They can be studied in a joint model with the outcome trajectory of interest. Recent extension of the method now allows the effects of time-varying covariates to be studied in a piecewise fashion across the entire observation period with dynamic class changing between periods of observation. The specific formulation for time-varying covariates depends on the underlying research question and theoretical framework of the area of study. The interpretation of time-varying covariates in some of these formulations can be potentially difficult and complex. This is an area awaiting more work.

As the underlying model of LCTA is MRM, the assumption of missingness follows that of MRM, i.e. MAR. Furthermore, as in joint modeling, the LCTA has also been studied as a means of exploring for the effects of data missing not at random (MNAR).
In summary, researchers can use the LCTA to ask if there are more homogenous subgroups in a heterogeneous population of patients. They can ask about the shapes of multiple outcome longitudinal trajectories, e.g. SDASm and MDASm, which may together form a more global construct of outcome. Different kinds of outcomes e.g. continuous and binary, can be studied simultaneously, like in joint modeling. Identified membership predictors can classify individuals into more homogenous subgroups. Physicians can then use such predictors to individualize patient management according to their predicted prognostic trajectories.

**LCTA application to JDM cohort**

In the past, the LCTA (GMM) had to be performed using special software (MPlus) specifically developed for this analysis (http://www.statmodel.com). Now, common statistical software such as R, SAS (proc Traj) and Stata will also perform such an analysis 255. Another alternative is to cast this analysis in a Bayesian framework and then perform the analysis in a Bayesian software, e.g. WinBUGS (www.mrc-bsu.cam.ac.uk/software/bugs). We performed this analysis using the package lcmm (v1.6.3) in R 255.

We modeled trajectories using fractional polynomials with random intercepts and random slopes. We tested 2 to 5 latent classes for fit, using Bayesian Information Criteria (BIC) for comparisons. The 3-class, random slopes model fitted the data the best and most parsimoniously. We then went on to model the correlation between measurements, testing the 2 covariance structures available in the package: Brownian motion (BM) and Autoregressive (AR). The initial models using number of months since diagnosis as the basic time metric had difficulties converging, therefore we switched to using number of years (since diagnosis) as the basic time metric. The AR (time in years) models fitted significantly better than the BM models. We tested the AR model using multiple sets of starting values (>10), taking care to stay away
Appendix 2: Supplementary Material for Appendix 1

from parameter boundaries, to ensure true estimate of the global maximum in parameter estimates.

From our analyses, 3 subclasses (AR corrected) appeared to be the best fit of our cohort. All the patients were classified clearly into 1 of the 3 classes, i.e. no patient was classified into 2 classes with similar probabilities. The 3 classes did not overlap significantly (see table 1). The proportions in the 3 classes were: 42%, 55% 3%. The last class only contained 3 patients. Future studies in another cohort of JDM patients will be helpful to validate the number of latent classes identified in this analysis.

Table 7.2.2: Mean posterior probabilities of belonging in the 3 classes

<table>
<thead>
<tr>
<th></th>
<th>Probability in class 1</th>
<th>Probability in class 2</th>
<th>Probability in class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>0.9292</td>
<td>0.0689</td>
<td>0.0002</td>
</tr>
<tr>
<td>Class 2</td>
<td>0.0776</td>
<td>0.9214</td>
<td>0.0010</td>
</tr>
<tr>
<td>Class 3</td>
<td>0.0001</td>
<td>0.0243</td>
<td>0.9756</td>
</tr>
</tbody>
</table>

Those in each of the 3 classes had correspondingly much lower probabilities of belonging in other classes, suggesting good separation of classes.

4) JOINT MODELING

In monitoring a patient’s disease course, it is common to have more than one outcome measure of interest. The researcher may find interest in studying the shape of outcome trajectories, i.e. how outcomes evolve over time. We can ask whether and how two or more outcomes are related in their evolutions over time, e.g., lupus disease activity (SLEDAI) and depression scores. Joint modeling allows us to test for concordance or discordance in the ways that the outcome trajectories evolve over time. The joint outcomes of interest do not have to be
of a similar nature; they can be of any combination, e.g. a continuous with a binary outcome, a continuous with a time-to-event outcome and any number of outcomes. Joint modeling has also been used as a means of reducing bias when there is informative attrition (e.g. when patients drop out because they are very well or very sick). Joint modeling of the outcome of interest with the time-to-dropout has been used to address this problem of informative attrition. Joint modeling is thus a potentially powerful tool that can expand our understanding of how outcomes trajectories evolve over time.

As we have demonstrated in the preceding sections, longitudinal data can be studied at a marginal level (e.g. GEE) or an individual level (e.g. MRM), so can joint models be formulated in either form, depending on the underlying research questions. Due to the assumptions required of joint marginal models, these are not very flexible. The MRM formulation is more flexible and can be easily extended to outcomes of disparate nature (e.g. continuous, binary or count). As such, the MRM formulation is more commonly used. For the MRM formulation, it is assumed that the outcomes are independent, conditional on the random effects, i.e. the underlying characteristics of the individual (random effects) govern the outcomes. The longitudinal process can also be jointly modeled with a time-to-event outcome using a proportional hazards model. Newer literature does not even require that random effects be shared between outcomes now. Instead, it is assumed that the outcome trajectories are manifestations of a common latent process.

Missing data is allowed in the joint model. The mechanism of missing data depends on the underlying model used: marginal (GEE) or subject-specific (MRM), used in specifying the joint model. If the investigator chooses to use the GEE, then missing data is assumed to be MCAR. If the investigator chooses the more common choice of MRM, then missing data is
Appendix 2: Supplementary Material for Appendix 1

assumed to be MAR. A special form of the joint model— the joint model of random effects— where the longitudinal trajectory of interest is jointly modeled with the time-to-dropout, has been proposed as a means of exploring the impact of MNAR, i.e., informative attrition\textsuperscript{84,179,242}. When data is MNAR, the missing outcome is related to the missing process\textsuperscript{84}. An example of this would be the sicker patients who no longer return for follow-up as they are too sick to leave their home. This kind of missing data can potentially bias any inference made based on patients remaining in the cohort. The joint model can be used to explore and adjust for this kind of bias\textsuperscript{179,233}. This does not mean that all the bias of MNAR can be resolved, as such modeling tends to lean on many assumptions, some of which may not be possible to verify.

The joint model allows for testing of both time-invariant and time-varying predictors with similar interpretations as in the MRM. Within the same model, one can study the relative effects (in both the significance and the effect size) of different predictors on each of the outcome trajectories modeled.

These models have been used in a few areas of medicine. Multiple continuous outcomes, e.g., hearing at various frequencies in both ears, have been jointly modeled to evaluate the change in hearing ability as patients age and to identify factors influencing those changes\textsuperscript{173}. A continuous (sequential organ failure assessment score) outcome has been jointly modeled with discharge status from intensive care unit (ICU) to evaluate the effect of treatment on outcome trajectory, adjusted by the dropout process, through either death or discharge from the ICU\textsuperscript{178}. This class of models has not yet been commonly used in the rheumatology literature. We believe that it has a lot of potential in answering real-life clinical questions.

In summary, the researcher should consider joint modeling when seeking to study multiple outcome trajectories concurrently. The multivariate analyses may be more accurate
Appendix 2: Supplementary Material for Appendix 1

reflections of the clinical decision making process, whereby several outcomes are followed and all factored into the clinician’s global perception of a patient’s outcome. The researcher can also investigate the degree of concordance among outcome trajectories. In addition, the researcher can potentially use this kind of model to explore for the biasing effects of missing data on observed outcomes.

**Joint Modeling applied to JDM outcomes**

This joint model was specified using the MRM. We used the spatial exponential covariance structure for modeling the MDASm and SDASm trajectories, similar to the section on MRM. The MDAS was modeled with a 2\textsuperscript{nd} order fractional polynomial (m=2, p=−1,−2). The SDAS was also modeled with a 2\textsuperscript{nd} order fractional polynomial but of a different form (m=2, p=−1,−1). Both forms were selected based on the maximal gain in fit\textsuperscript{139}. When modelling for predictors, we crossed the fractional polynomial time forms of the MDAS and SDAS curves with bDAS to test the effects of bDAS on the slopes. Only the one time term of MDAS (p=−1, which is in common with that of the SDAS) was significant, so this was retained in the model. We left the product of the bDAS crossed with first time term of SDAS in the equation for easier comparison but this product was not significant, ie. bDAS did not predict the slope evolution of SDAS.

**Summary**

Modern longitudinal analytic methods are now more flexible and powerful. Observational studies, with challenges from irregular visit schedules and missing data, can now be analyzed using these longitudinal methods. The choice of method will depend on the researcher’s question. This review is not meant to be exhaustive. Markov multistate models\textsuperscript{208} or
latent transition regression\textsuperscript{258} are considered longitudinal as well but as they do not deal with "trajectories" but a change in state, we did not include them in this review.
### Appendix 3: SLE Measures

#### 7.3 SLE Measures

**7.3.1 The American College of Rheumatology/ SLE International Collaborating Clinics SLE Damage Index (SDI)**

<table>
<thead>
<tr>
<th>SYSTEMIC LUPUS INTERNATIONAL COLLABORATING CLINICS</th>
<th><strong>Score</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Name:</strong></td>
<td><strong>Study No.</strong></td>
</tr>
<tr>
<td><strong>Assessment Data:</strong></td>
<td>L / L</td>
</tr>
<tr>
<td>Damage occurring since diagnosis of lupus, ascertained by clinical assessment and present for at least 6 months unless otherwise stated. Repeat episodes mean at least 6 months apart to score 2. The same lesion cannot be scored twice.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Item</strong></th>
<th><strong>Score</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OCULAR</strong> (Eye by clinical assessment)</td>
<td></td>
</tr>
<tr>
<td>Any cataract present</td>
<td>0</td>
</tr>
<tr>
<td>Papilledema / Optic atrophy</td>
<td>0</td>
</tr>
<tr>
<td><strong>NEUROPSYCHIATRIC</strong></td>
<td></td>
</tr>
<tr>
<td>Cognitive impairment (eg. memory deficit, difficulty with calculations, poor concentration, difficulty in spoken or written language, impaired performance level)</td>
<td>0</td>
</tr>
<tr>
<td>OR Major psychoses</td>
<td>0</td>
</tr>
<tr>
<td>Seizures requiring therapy for 6 months</td>
<td>0</td>
</tr>
<tr>
<td>Cortical venous sinus thrombosis (Score 3 if &gt; 2) and for malignancy</td>
<td>0</td>
</tr>
<tr>
<td>Cranial or peripheral neuropathy (excluding optic)</td>
<td>0</td>
</tr>
<tr>
<td>Transient myelitis</td>
<td>0</td>
</tr>
<tr>
<td><strong>RENAL</strong></td>
<td></td>
</tr>
<tr>
<td>Estimated or measured GFR &lt; 50%</td>
<td>0</td>
</tr>
<tr>
<td>Proteinuria 24 h, g 3.3 g</td>
<td>0</td>
</tr>
<tr>
<td>OR Most severe renal disease (regardless of dialysis or transplantation)</td>
<td>0</td>
</tr>
<tr>
<td><strong>PULMONARY</strong></td>
<td></td>
</tr>
<tr>
<td>Pulmonary hypertension (right ventricular prominence, or initial PE)</td>
<td>0</td>
</tr>
<tr>
<td>Pulmonary edema (physiologic and X-ray)</td>
<td>0</td>
</tr>
<tr>
<td>Shunting lung (X-ray)</td>
<td>0</td>
</tr>
<tr>
<td>Pleural effusion (X-ray)</td>
<td>0</td>
</tr>
<tr>
<td>Pulmonary infection (X-ray) OR resection not for malignancy</td>
<td>0</td>
</tr>
<tr>
<td><strong>CARDIOVASCULAR</strong></td>
<td></td>
</tr>
<tr>
<td>Angina OR Coronary artery bypass</td>
<td>0</td>
</tr>
<tr>
<td>Myocardial infarction (Score 2 if &gt; 1)</td>
<td>0</td>
</tr>
<tr>
<td>Cardiomyopathy (ventricular dysfunction)</td>
<td>1</td>
</tr>
<tr>
<td>Valvular disease (includes murmur or a systolic murmur)</td>
<td>1</td>
</tr>
<tr>
<td>Pancarditis &gt; 6 months or pericarditly</td>
<td>1</td>
</tr>
<tr>
<td><strong>PERIPHERAL VASCULAR</strong></td>
<td></td>
</tr>
<tr>
<td>Claudication &gt; 6 months</td>
<td>0</td>
</tr>
<tr>
<td>Major tissue loss (P&amp;L space)</td>
<td>0</td>
</tr>
<tr>
<td>Significant tissue loss (e.g. loss of digit or limb, resection) (Score 2 if &gt; 1)</td>
<td>0</td>
</tr>
<tr>
<td>Venous thrombosis with swelling, ulceration, or venous stasis</td>
<td>0</td>
</tr>
<tr>
<td><strong>GASTROINTESTINAL</strong></td>
<td></td>
</tr>
<tr>
<td>Intestinal resection of bowel (below duodenum), spleen, liver or gall bladder (Score 2 if &gt; 1)</td>
<td>0</td>
</tr>
<tr>
<td>Mesenteric ischemia</td>
<td>0</td>
</tr>
<tr>
<td>Chronic pancreatitis</td>
<td>0</td>
</tr>
<tr>
<td>Sigmoid OR upper gastrointestinal tract surgery</td>
<td>0</td>
</tr>
<tr>
<td>Pancreatic insufficiency requiring enzyme replacement or with pseudocyst</td>
<td>0</td>
</tr>
<tr>
<td><strong>MUSCULOSKELETAL</strong></td>
<td></td>
</tr>
<tr>
<td>Arthritis or weakness, hand loss</td>
<td>0</td>
</tr>
<tr>
<td>Deforming or severe arthritis (including reducible deformities, excluding avascular necrosis)</td>
<td>0</td>
</tr>
<tr>
<td>Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis)</td>
<td>0</td>
</tr>
<tr>
<td>Avascular necrosis (Score 2 if &gt; 1)</td>
<td>0</td>
</tr>
<tr>
<td>Osteosclerosis</td>
<td>0</td>
</tr>
<tr>
<td>Ruptured tendon</td>
<td>0</td>
</tr>
<tr>
<td><strong>SKIN</strong></td>
<td></td>
</tr>
<tr>
<td>Alopecia</td>
<td>0</td>
</tr>
<tr>
<td>Actinic scarring or peau d'orange aside lesions and scalp disease</td>
<td>0</td>
</tr>
<tr>
<td>Skin ulceration (excluding those with), for more than 6 months.</td>
<td>0</td>
</tr>
<tr>
<td><strong>PREMATURE GONADAL FAILURE</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DIABETES</strong></td>
<td></td>
</tr>
<tr>
<td>(Complications of diabetes)</td>
<td></td>
</tr>
<tr>
<td><strong>MULTIPLE MYELOMA</strong> (Exclude dyspnea)</td>
<td>0</td>
</tr>
</tbody>
</table>

# Appendix 3: SLE Measures

## 7.3.2 The SLE Disease Activity Index 2000 (SLEDAI2K)

<table>
<thead>
<tr>
<th>Weight</th>
<th>SLEDAI Score</th>
<th>Descriptor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3</td>
<td>Seizure</td>
<td>Recent seizure, exclude metabolic, infectious or drug cause.</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>Psychosis</td>
<td>Altered ability to function in normal activity due to severe disturbance in the perception of realty. Include hallucinations, delusions, marked mood changes, impoverished thought content, marked irregular thinking, libido, disorientation, or bizarre behavior. Exclude seizures and drug causes.</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>Cerebral vascular disease</td>
<td>Altered mental function with impaired orientation, memory, or other intellectual function, with rapid onset and fluctuating clinical features, inability to maintain attention to environment, plus at least 2 of the following: perceptual disturbances, incoherent speech, inattention or daydreaming, or increased or decreased psychomotor activity. Exclude metabolic, infectious, or drug causes.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Visual disturbance</td>
<td>Retinal changes of SLE. Include retinal hemorrhages, papillitis, retinal infarction, or papilledema in the absence of optic neuritis. Exclude exogenous, infectious, or drug causes.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Cranial nerve disorder</td>
<td>New onset of sensory or motor neuropathy involving cranial nerve.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Lupus headache</td>
<td>Severe, persistent headache may be migraine, but must be nonresponsive to narcotic analgesics.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>CVA</td>
<td>New onset of cerebrovascular accident(s). Exclude aneurysmal.</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Vasculitis</td>
<td>Ulceration, panniculitis, tender, larger nodules, purpuric infarct, splinter hemorrhages, or biopsy or angiogram proof of vasculitis.</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Arthritis</td>
<td>≥2 joints with pain and signs of inflammation (i.e. tenderness, swelling or effusion).</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Myositis</td>
<td>Proximal muscle aching/weakness, associated with elevated creatine phosphokinase, aldolase, or electromyogram changes or a biopsy showing myositis.</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Ciliary casts</td>
<td>Hematocrit or red blood cell casts.</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Hematuria</td>
<td>&gt;5 red blood cell/high-power field. Exclude stone, infection, or other cause.</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Prokainuria</td>
<td>≥0.5 gram/24 hours.</td>
</tr>
<tr>
<td>≤1</td>
<td>1</td>
<td>Pyrexia</td>
<td>≥3 white blood cell/high power field. Exclude infection.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Rash</td>
<td>Infarct clot or red blood cell casts.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Alopecia</td>
<td>Ablated, patchy or diffuse loss of hair.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Minimal ulcers</td>
<td>Oral or nasal ulcerations.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Pneumonitis</td>
<td>比利时. Close to paraclinical, but not the following: rub or effusion, or echocardiogram or edecardiosgram confirmation.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Pericarditis</td>
<td>Positive, not at the following: rub, effusion, or echocardiogram or edecardiosgram confirmation.</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Decreased complement C3 or C4 below the lower limit of normal for being laboratory</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Increased DNA binding</td>
<td>Increased DNA binding by Farr assay above normal range for being laboratory.</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Fever</td>
<td>&gt;38°C, exclude infectious cause.</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Thrombocytopenia</td>
<td>&lt;100,000 platelets/μL, exclude drug causes.</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Leukopenia</td>
<td>&lt;4,000 white blood cells/μL, exclude drug causes.</td>
</tr>
</tbody>
</table>

**TOTAL SLEDAI SCORE**


7.4: Supplementary Material for Chapter 3

Appendix accompanying “A Systematic Review of the Quality of Prognosis Studies in Systemic Lupus Erythematosus”

Handling those journals with less than a full 12 years of impact factors

Among the general medicine journals, PLOS Medicine’s impact factor grading only started in 2005 and would have ranked within the top 5 general medical journals using available gradings (from 2005-2011). As there is a trend for impact factors to stay stable or increase over time in high ranking journals, we imputed the 2005 impact factor grading for PLOS Medicine for 2000-2004. This approach would likely overestimate PLOS Medicine’s grading (had it been available in the missing years). With the imputed gradings, PLOS Medicine did not rank in the top 5.

Among the top rheumatology journals, Arthritis Research & Therapy (and its antecedent, Arthritis Research) did not have impact factor grading in 2000-2001. The ranking of this journal remained stable within the top 5 journals when the same approach as above was used to impute impact factors.

Preparation for use of QUIPS to assess study quality

As observational studies vary widely in design and clarity of presentation, we tested the QUIPS instrument for agreement between the 2 reviewers before embarking on the study. Our aim was to have an inter-rater correlation coefficient (ICC2,1) of 0.9, but a priori we allowed for a minimal agreement of 0.7. We calculated that we would need 17 articles to achieve this. We tested 17 random rheumatic disease (not SLE) prognosis articles from between 2006 to 2011. After the first round of articles, we did not reach the minimum agreement for all the domains. After further discussion and standardizing our assessments, we repeated the test using an additional 10 articles. After the second round, we reached our target ICC of 0.9 for the domains.
The ICC (2,1) was calculated for assessment of inter-rater agreement in the pre-study preparation for using QUIPS.

**Table 7.4.1: Basic characteristics of reviewed and non-reviewed studies**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Reviewed studies Number (%)</th>
<th>Non-reviewed studies Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study periods</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990-2000</td>
<td>53 (36)</td>
<td>213 (29)</td>
</tr>
<tr>
<td>2000-2011</td>
<td>97 (64)</td>
<td>525 (71)</td>
</tr>
<tr>
<td><strong>Types of journals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top-ranking medical journals</td>
<td>2 (1)</td>
<td>11 (1)</td>
</tr>
<tr>
<td>Other medical journals</td>
<td>45 (30)</td>
<td>219 (30)</td>
</tr>
<tr>
<td>Top-ranking rheumatology journals</td>
<td>39 (26)</td>
<td>172 (23)</td>
</tr>
<tr>
<td>Other rheumatology journals</td>
<td>64 (43)</td>
<td>336 (46)</td>
</tr>
</tbody>
</table>
7.5: Supplementary material for Chapter 4

Appendix accompanying “From Childhood to Adulthood: The Longitudinal Trajectory of Damage in Childhood-onset Systemic Lupus Erythematosus Patients”

Methods

Dealing with missing data

Missing items were handled according to pre-specified rules. Urine parameters were imputed within 6 months from each occasion when there was data; thereafter, the parameters were resolved as missing. If a patient never had renal disease up to the end of the follow-up period, then the urine parameters were imputed as normal. Anti-DNA antibody levels were imputed within 12 months from a visit with missing data. Complement levels were imputed within 6 months from an occasion with missing data. Multiple imputation was not performed due to the highly individualized visit schedule (different schedule and frequency of visits for individuals).

We defined not being lost to follow-up as having data up to within 18 months before the study end date or time of death (whichever was earlier). This period was chosen in compliance with the definition that the Toronto Lupus cohort used to define a patient as being lost to follow-up. To explore the nature of loss-to-follow-up, we modelled the time-to-loss-to-follow-up using a Cox proportional hazards model, regressing onto the (time-varying) damage score. If damage predicted lost to follow-up, this would be an indication of possibly data Missing At Random (MAR) or Missing Not At Random (MNAR). The data would be MAR if loss to follow-up was predicted by observed parameters. In MNAR, the missing outcome could be the cause for loss to follow up, or the missing outcome might be systematically different for those lost to follow-up compared to those who remained in the study, which may then result in a biased
Appendix 5: Supplementary material for Chapter 4

Conclusion. MNAR is always a concern but cannot be definitely proven unless the missing outcome data is available. We tried to collect the most recent damage data from those lost to follow-up to assess the nature of loss-to-follow-up\textsuperscript{82}. We sent out 34 packages to the most recent address known to us (through our clinical database, last address in chart, contacting family physicians, contacting rheumatologists) but only received 5 questionnaires replies. The majority of packages were returned unopened due to obsolete addresses. As we were only able to obtain recent (close to study end date) damage scores from 6 patients (9\% of those lost to follow-up), we did not attempt to analyse them further. Of the 6, 5 patients reported their damage using the Lupus Damage Index Questionnaire\textsuperscript{260}; 1 rheumatologist scored the remaining patient. Of these, 2 patients had no damage, 1 patient each scored at 1,2,4 and 10.

Identification of predictors using WGEE

Baseline predictors were entered one at a time into a basic model that only contained the time terms, first alone (i.e., affecting initial damage) and then interacting with the time form (i.e., affecting the rate of change of the trajectory). Candidate predictors were chosen based on clinical understanding of the disease and from review of current literature for predictors of damage. Predictors were retained for testing in the multivariable model if $p<0.1$ on initial screening. Time-varying predictors were entered directly into the model (without interactions with the time form). In the final multivariable model (other than the time form parameters), predictors were kept if $p<0.05$.

The inverse-intensity-weighted Generalized Estimating Equation (WGEE) model with an independent working covariance matrix was used to determine the overall damage trajectory\textsuperscript{241}. The weights were calculated based on variables deemed to likely affect both the damage trajectory and visit frequency (see Table 8.4.1 for the list of variables).
Figure 7.5.1: Plot of damage trajectories in cSLE patients

All grey lines represented individual patients. The black line is a LOESS plot of the average population damage trajectory over time.

Table 7.5.1: List of items considered in deriving weights for the weighted GEE analysis

<table>
<thead>
<tr>
<th>Items considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
</tr>
<tr>
<td>Ethnicity</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Time period of diagnosis</td>
</tr>
<tr>
<td>Number of ACR- SLE criteria at diagnosis</td>
</tr>
<tr>
<td>Major organ involvement</td>
</tr>
</tbody>
</table>
Appendix 5: Supplementary material for Chapter 4

Prednisone treatment at every visit
Prednisone treatment 3 months, 6 months, 9 months, 12 months, before each visit
Antimalarial treatment
Cyclophosphamide treatment
Significant immunosuppression each treatment

Table 7.5.2: Results of univariable analysis of individual SLEDAI2K items for the damage trajectory

<table>
<thead>
<tr>
<th>Lagged SLEDAI2K individual items</th>
<th>Parameter estimates (standard errors)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6 months</td>
</tr>
<tr>
<td>Psychosis</td>
<td>0.320(0.112)*</td>
</tr>
<tr>
<td>Acute confusion state</td>
<td>0.353(0.155)*</td>
</tr>
<tr>
<td>Visual changes</td>
<td>0.342(0.357)</td>
</tr>
<tr>
<td>Lupus headache</td>
<td>0.746(0.353)*</td>
</tr>
<tr>
<td>Vasculitis</td>
<td>0.078(0.107)</td>
</tr>
<tr>
<td>Arthritis</td>
<td>0.049(0.059)</td>
</tr>
<tr>
<td>Myositis</td>
<td>-0.033(0.090)</td>
</tr>
<tr>
<td>Urine casts</td>
<td>0.205(0.254)</td>
</tr>
<tr>
<td>Haematuria</td>
<td>0.206(0.166)</td>
</tr>
<tr>
<td>Proteinuria</td>
<td>0.192(0.123)</td>
</tr>
<tr>
<td>Pyuria</td>
<td>0.273(0.349)</td>
</tr>
<tr>
<td>Rash</td>
<td>-0.152(0.058)*</td>
</tr>
<tr>
<td>Alopecia</td>
<td>0.077(0.089)</td>
</tr>
<tr>
<td>Mucous membrane</td>
<td>-0.170(0.093)*</td>
</tr>
<tr>
<td>Pleurisy</td>
<td>0.129(0.187)</td>
</tr>
<tr>
<td>Pericarditis</td>
<td>-0.161(0.097)*</td>
</tr>
<tr>
<td>Hypocomplementemia</td>
<td>-0.123(0.076)</td>
</tr>
<tr>
<td>Elevated anti-dsDNA titres</td>
<td>-0.083(0.073)</td>
</tr>
<tr>
<td>Fever</td>
<td>0.144(0.110)</td>
</tr>
</tbody>
</table>
Appendix 5: Supplementary material for Chapter 4

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrombocytopenia</td>
<td>-0.038</td>
<td>0.005</td>
<td>0.021</td>
<td>-0.099</td>
</tr>
<tr>
<td>Leucopenia</td>
<td>-0.050</td>
<td>-0.079</td>
<td>-0.067</td>
<td>-0.049</td>
</tr>
</tbody>
</table>

*Indicates p<0.1

---

**Figure 7.5.2: Effects of time-invariant predictors in predicting the population-averaged trajectory in childhood-onset SLE patients**

**Figure 7.5.2a: Effect of Race In Individuals with Major Organ Involvement at Baseline, Diagnosed in the 2000-2011 Period of the Cohort***

[Graph showing the effect of race over time]

**Legend:** Black patients with had persistently higher damage trajectory compared to white patients when all other time-invariant predictors were held constant and all time-varying predictors set to zero. *Plotted from 6 months as most damage items can only be scored after 6 months. Plot stopped at 120 months (~10 years) as the proportion of patients with continued data became significantly reduced after that.
Appendix 5: Supplementary material for Chapter 4

Figure 7.5.2b: Effects of Major Organ Involvement In Black Patients Diagnosed in the 2000-2011 Period of the Cohort*

![Graph showing effects of major organ involvement in black patients](image)

Legend: Patients with major organ involvement (red lines) showed a more rapid damage accrual initially but the trajectory of increase slowed down later in the course compared to those without major organ involvement (blue lines). Black patients have greater damage trajectory compared to white patients. All other time-invariant factor and time-varying factors were held constant or set to zero. *Plotted from 6 months as most damage items can only be scored after 6 months. Plot stopped at 120 months (~10 years) as the proportion of patients with continued data became significantly reduced after that.

Figure 7.5.2c: Effect of Major Organ Involvement In White Patients Diagnosed in the 2000-2011 Period of the Cohort*

![Graph showing effects of major organ involvement in white patients](image)
7.6: Supplementary Material for Chapter 5

Appendix accompanying “From Childhood to Adulthood: Latent Classes of Disease Trajectories in Childhood-onset Systemic Lupus Erythematosus”

Methods

Missing data

Any incomplete scoring of SLEDAI2K in the pediatric database was re-scored by crosschecking with the charts. For the Toronto Lupus Cohort, the values of the missing items were imputed if there were data within 6-12 months before or after the missing visit and the value of the item in the visit had not changed from the previous visit where there was a value. In addition, for patients who never had any renal disease up to the end of the study, any missing renal item in the SLEDAI was set to zero. The same approach was adopted when data were extracted from the patients' charts from rheumatologists' offices.

Statistical analysis

Fractional polynomials

We calculated gain and picked the fractional polynomials with the maximal gain as the final forms for time for the 2 outcomes. The best form for SLEDAI2K was a 2nd order FP (-2,0) and that for prednisone was also similar (-2,-1).

After separation into latent classes, we calculated the best fractional polynomials for the damage (SLICC/ACR SDI) trajectories by class following the same approach as above. All damage trajectories were best defined using 2nd order FPs. The FPs for the 5 classes were: (−1,0.5), (1,1), (0.5,2), (0.5,3) and (3,3).

Model Choice
Appendix 6: Supplementary Material for Chapter 5

From preliminary testing of single outcome trajectories, we found that there were 2-6 latent subgroups within each of the outcomes. This gave us an idea of the possible number of subgroups for the joint (dual) trajectories.

We tested the joint trajectories sequentially, starting from 2 subgroups to 6. It is important to note that the GMM tested the SLEDAI2K and prednisone trajectories together. The output solution has 1 mean SLEDAI2K and 1 mean prednisone trajectory within each subgroup.

We planned to use the nadir of the Bayesian information criterion (BIC) to help guide the choice for the best number of groups. However, the BIC was still decreasing at 6 classes. Other authors have cautioned against over-extraction such that the smallest group is < 5-10% of the total population. The smallest group of the 6-class solution had only 1% of the population. The most important consideration guiding model choice in the end should be guided by theory or clinical experience. Upon consulting the Lupus experts in the group, we chose the 5-class solution to be more consistent with our clinical experience. The 4-class solution had bigger groups but class 1 of the 5-class solution (relapsing disease) was not identified by the 4-class solution, though relapsing disease is a clinically observed group. Therefore, we chose the 5-class solution over the 4-class solution.

**Predictors**

After the basic (no membership covariate) model was obtained, we used proc genmod (SAS) to screen for significant membership covariates by regressing onto class membership. We tested these candidate baseline covariates: age of diagnosis, ethnicity, gender, pubertal status, number of ACR criteria at diagnosis, time-period of diagnosis and life-threatening organ manifestations.
Figure 7.6.1: Individuals’ and mean population trajectories of latent SLE disease activity (DisAct) by latent classes

Figure 7.6.1A) Loess plots of individuals’ SLEDAI2K trajectories by latent classes.

Figure 7.6.1B) Loess plots of mean population SLEDAI2K trajectories by latent classes (with 95% confidence intervals in gray).

Figure 7.6.1C) Loess plots of individuals’ prednisone (mg) trajectories by latent classes.

Figure 7.6.1D) Loess plots of mean population prednisone (mg) trajectories by latent classes (with 95% confidence intervals in gray).

Table 7.6.1: Ethnicities of cSLE patients by latent classes

<table>
<thead>
<tr>
<th>Ethnicities n=472*</th>
<th>Class 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total (row)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whites</td>
<td>7(5)</td>
<td>12(8)</td>
<td>28(18)</td>
<td>24(16)</td>
<td>82(54)</td>
<td>153</td>
</tr>
<tr>
<td>Asians</td>
<td>12(7)</td>
<td>31(17)</td>
<td>34(19)</td>
<td>37(20)</td>
<td>67(37)</td>
<td>181</td>
</tr>
<tr>
<td>Blacks</td>
<td>7(12)</td>
<td>8(13)</td>
<td>7(12)</td>
<td>12(20)</td>
<td>26(43)</td>
<td>60</td>
</tr>
</tbody>
</table>
#179 patients had major organ involvement at diagnosis

Table 7.6.2: Results of univariable screen for baseline predictors of SLE DisAct trajectory

different reference classes

<table>
<thead>
<tr>
<th>Membership predictors</th>
<th>Class</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethnicities</strong>#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asians vs Caucasians</td>
<td>2 vs 5</td>
<td>3.08</td>
<td>1.42-6.68</td>
</tr>
<tr>
<td>Asians vs Other minorities</td>
<td>2 vs 5</td>
<td>2.71</td>
<td>1.02-7.16</td>
</tr>
<tr>
<td>Asians vs Afro-Caribbeans</td>
<td>3 vs 5</td>
<td>2.02</td>
<td>0.75-5.44</td>
</tr>
<tr>
<td>Asians vs Other minorities</td>
<td>3 vs 5</td>
<td>2.18</td>
<td>0.90-5.26</td>
</tr>
<tr>
<td>Afro-Caribbeans vs Caucasians</td>
<td>1 vs 5</td>
<td>3.60</td>
<td>1.01-12.82</td>
</tr>
<tr>
<td><strong>Age at diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 vs 5</td>
<td>1.20</td>
<td></td>
<td>1.00-1.25</td>
</tr>
<tr>
<td>3 vs 5</td>
<td>1.15</td>
<td></td>
<td>1.04-1.27</td>
</tr>
<tr>
<td>5 vs 4</td>
<td>0.952</td>
<td></td>
<td>0.871-1.04</td>
</tr>
<tr>
<td>1 vs 5</td>
<td>1.07</td>
<td></td>
<td>0.933-1.228</td>
</tr>
<tr>
<td>1 vs 2</td>
<td>0.956</td>
<td></td>
<td>0.812-1.126</td>
</tr>
<tr>
<td>3 vs 2</td>
<td>1.025</td>
<td></td>
<td>0.897-1.17</td>
</tr>
<tr>
<td>1 vs 3</td>
<td>0.933</td>
<td></td>
<td>0.797-1.092</td>
</tr>
<tr>
<td>3 vs 4</td>
<td>1.092</td>
<td></td>
<td>0.97-1.228</td>
</tr>
<tr>
<td>2 vs 4</td>
<td>1.065</td>
<td></td>
<td>0.939-1.209</td>
</tr>
<tr>
<td>1 vs 4</td>
<td>1.019</td>
<td></td>
<td>0.875-1.185</td>
</tr>
<tr>
<td><strong>Baseline major organ involvement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 vs 5</td>
<td>4.77</td>
<td></td>
<td>1.98-11.49</td>
</tr>
<tr>
<td>2 vs 5</td>
<td>25.00</td>
<td></td>
<td>10.00-62.52</td>
</tr>
<tr>
<td>3 vs 5</td>
<td>3.84</td>
<td></td>
<td>2.14-6.87</td>
</tr>
<tr>
<td>4 vs 5</td>
<td>4.82</td>
<td></td>
<td>2.69-8.63</td>
</tr>
<tr>
<td>2 vs 1</td>
<td>5.24</td>
<td></td>
<td>1.61-17.00</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>6.52</td>
<td></td>
<td>2.45-17.31</td>
</tr>
<tr>
<td>2 vs 4</td>
<td>5.19</td>
<td></td>
<td>1.95-13.77</td>
</tr>
<tr>
<td>1 vs 3</td>
<td>1.244</td>
<td></td>
<td>0.486-3.189</td>
</tr>
<tr>
<td>3 vs 4</td>
<td>0.796</td>
<td></td>
<td>0.406-1.56</td>
</tr>
<tr>
<td>1 vs 4</td>
<td>0.99</td>
<td></td>
<td>0.387-2.536</td>
</tr>
<tr>
<td><strong>Number of ACR criteria at diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 vs 5</td>
<td>1.50</td>
<td></td>
<td>1.08-2.10</td>
</tr>
<tr>
<td>2 vs 5</td>
<td>1.88</td>
<td></td>
<td>1.45-2.43</td>
</tr>
<tr>
<td>4 vs 5</td>
<td>1.56</td>
<td></td>
<td>1.34-1.96</td>
</tr>
<tr>
<td>1 vs 3</td>
<td>1.50</td>
<td></td>
<td>1.03-2.18</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>1.88</td>
<td></td>
<td>1.38-2.55</td>
</tr>
<tr>
<td>4 vs 3</td>
<td>1.56</td>
<td></td>
<td>1.17-2.07</td>
</tr>
<tr>
<td>5 vs 3</td>
<td>0.998</td>
<td></td>
<td>0.784-1.271</td>
</tr>
<tr>
<td>1 vs 2</td>
<td>0.798</td>
<td></td>
<td>0.549-1.160</td>
</tr>
<tr>
<td>4 vs 2</td>
<td>0.829</td>
<td></td>
<td>0.624-1.100</td>
</tr>
<tr>
<td>1 vs 4</td>
<td>0.963</td>
<td></td>
<td>0.674-1.378</td>
</tr>
</tbody>
</table>
# Due to the number of ethnic groups and the number of different latent classes, there are many possible permutations. The results of the multiple permutations of race and class can be available upon contact of the corresponding author. Bolded results are statistically significant.

**Table 7.6.3: Results of multivariable model predicting class membership in latent SLE**

**DisAct trajectory classes (class 5 as reference class)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimates (overall)</th>
<th>P (Global effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at baseline</strong></td>
<td>Older age at diagnosis in classes 2 &amp; 3 compared to 5</td>
<td>0.0166</td>
</tr>
<tr>
<td><strong>Asian ethnicity</strong></td>
<td><em>OR of membership in class 2 = 2.444,</em> 95% <em>CI 1.199-4.982</em></td>
<td>0.1274</td>
</tr>
<tr>
<td><strong>Major organ involvement at baseline</strong></td>
<td>Major organ involvement more likely in all other classes compared to 5</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Number of ACR criteria at diagnosis</strong></td>
<td>Higher numbers of ACR criteria in classes 2 and 4 compared to 5</td>
<td>0.0064</td>
</tr>
</tbody>
</table>
8 References for all chapters


Complete References


Complete References

Complete References


Complete References


244. Repeated Measures Modeling with PROC MIXED. SAS Users Global International (SUGI); 2004; Montreal, Canada. SAS.


Complete References


