



Review article

Great diversity in the utilization and reporting of latent growth modeling approaches in type 2 diabetes: A literature review

Sarah O'Connor^{a,b}, Claudia Blais^{b,e}, Miceline Mésidor^{c,d}, Denis Talbot^{c,d}, Paul Poirier^{a,b}, Jacinthe Leclerc^{a,b,f,*}^a Research Centre, Institut universitaire de Cardiologie et Pneumologie de Québec-Université Laval (IUCPQ-UL), 2725 Ch. Ste-Foy, Québec City, Québec, G1V 4G5, Canada^b Faculty of Pharmacy, Université Laval, Ferdinand Vandry Pavillon, 1050 de La Médecine Avenue, Québec City, Québec, G1V 0A6, Canada^c Department of Social and Preventive Medicine, Faculty of Medicine, Université Laval, Ferdinand Vandry Pavillon, 1050 de La Médecine Avenue, Québec City, Québec, G1V 0A6, Canada^d Research Centre, CHU de Québec – Université Laval, 2400 D'Estimauville Avenue, Québec, QC, G1E 6W2, Canada^e Bureau D'information et D'études en Santé des Populations, Institut National de Santé Publique Du Québec, 945, Wolfe Avenue, Québec City, Québec, G1V 5B3, Canada^f Department of Nursing, Université Du Québec à Trois-Rivières, 3351 des Forges Boulevard, Trois-Rivières, Québec, G8Z 4M3, Canada

HIGHLIGHTS

- There is a growing body of literature on trajectory modeling in type 2 diabetes.
- Latent class growth analysis can be used in many different contexts.
- The current reporting of methods used should be improved.

ARTICLE INFO

Keywords:

Group-based trajectory modeling
 Latent class growth modeling
 Epidemiology
 Care trajectory
 Health-care utilization
 Diabetes mellitus

ABSTRACT

Introduction: The progression of complications of type 2 diabetes (T2D) is unique to each patient and can be depicted through individual temporal trajectories. Latent growth modeling approaches (latent growth mixture models [LGMM] or latent class growth analysis [LCGA]) can be used to classify similar individual trajectories in *a priori* non-observed groups (latent groups), sharing common characteristics. Although increasingly used in the field of T2D, many questions remain regarding the utilization of these methods.

Objective: To review the literature of longitudinal studies using latent growth modeling approaches to study T2D.

Methods: MEDLINE (Ovid), EMBASE, CINAHL and Wb of Science were searched through August 25th, 2021. Data was collected on the type of latent growth modeling approaches (LGMM or LCGA), characteristics of studies and quality of reporting using the GROLTS-Checklist and presented as frequencies.

Results: From the 4,694 citations screened, a total of 38 studies were included. The studies were published between 2011 and 2021 and the length of follow-up ranged from 8 weeks to 14 years. Six studies used LGMM, while 32 studies used LCGA. The fields of research varied from clinical research, psychological science, healthcare utilization research and drug usage/pharmaco-epidemiology. Data sources included primary data (clinical trials, prospective/retrospective cohorts, surveys), or secondary data (health records/registries, medico-administrative). Fifty percent of studies evaluated trajectory groups as exposures for a subsequent clinical outcome, while 24% used predictive models of group membership and 5% used both. Regarding the quality of reporting, trajectory groups were adequately presented, however many studies failed to report important decisions made for the trajectory group identification.

Conclusion: Although LCGA were preferred, the contexts of utilization were diverse and unrelated to the type of methods. We recommend future authors to clearly report the decisions made regarding trajectory groups identification.

* Corresponding author.

E-mail address: jacinthe.leclerc@criucpq.ulaval.ca (J. Leclerc).

1. Introduction

Type 2 diabetes (T2D) is a major public health threat worldwide, with approximately 1 person out of 11 living with the disease in 2019 [1]. To prevent complications, the management of T2D requires a long-term control of glycemia, healthy lifestyle behaviors, adequate medication usage as well as regular medical examinations [1]. However, in a real-life context, the management of T2D can differ between patients due to individual and environmental factors, but may also change in time (e.g. changes in treatment guidelines, preferences, physicians, living contexts, etc.). This heterogeneity may be reflected through different longitudinal trajectories in adherence patterns, glycemic control, lifestyle behaviors or healthcare utilization, unique to each patient.

Latent growth modeling approaches are statistical methods that consider individuals being part of a heterogeneous population, composed of unobserved groups of individual trajectories sharing similar characteristics [2, 3]. The goal of latent growth modeling approaches is to estimate a given number of *a priori* unobserved latent groups within a population, based on the probability of membership of individuals to a specific trajectory group [2]. Two distinct methods of latent growth modeling approaches can be identified, namely Latent Growth Mixture Modeling (LGMM) and Latent Class Growth Analysis (LCGA) [LCGA is also known as Latent Class Growth Modeling (LCGM) or Group-Based Trajectory Modeling (GBTM)] [4]. LGMM allows the modelization of within-class heterogeneity [5]. In contrast, LCGA assumes that no within-group heterogeneity remains within the trajectory groups [2]. Latent growth modeling approaches, which have been initially developed in the fields of social sciences, psychopathology or criminology [3] are now increasingly used in the field of T2D. These approaches are well suited for the study of T2D given the repeated follow-ups in time, and also encourage the consideration of the overall temporal trajectories of care for clinical decisions [6, 7]. The latent groups may help identifying subgroups of the population with different needs and developing adapted care strategies [7]. The possibility of using these approaches with multiple types of data, long follow-up lengths and unstructured datasets makes

these methods highly appealing in the study of T2D [3]. However, latent growth modeling approaches are complex to use; many decisions must be made and should be reported correctly in order to ensure comparability between studies [4]. Accordingly, the *Guidelines for Reporting on Latent Trajectory Studies* (GRoLTS) checklist was built up for maximised transparency and to improve comparison between studies [4].

Although there is a growing number of studies using latent growth modeling approaches in the field of T2D, the variety of methods used, their context of utilization and quality of reporting remain unknown. Hence, by identifying longitudinal studies using latent growth modeling approaches among individuals with T2D, we sought (1) to identify the latent growth modeling approaches used, (2) to describe the context of utilization of the different methods, and (3) to assess the quality of reporting, using the GRoLTS checklist framework [4]. Of note, this review did not seek to describe the methods *per se*, which have been done previously [2, 5, 8]. To our knowledge, this review is the first to comprehensively describe the latent growth modeling approaches used in the field of T2D as of December 2021.

2. Materials and methods

2.1. Eligibility criteria

This study is a review of longitudinal studies using latent growth modeling approaches among human cases with T2D, as defined by the authors. Studies needed to have at least 50% of participants with T2D to be included. Eligible studies must have used a method for discriminating latent groups of individual trajectories, no matter the duration of the follow-up. Studies determining groups using standard/arbitrary thresholds (e.g. body mass index categories, age groups, etc.) or sample stratifications (quartiles, percentiles, medians, interquartile ranges) were excluded. Only published peer-reviewed studies written in English or French were included. Case-control, cross-sectional and qualitative studies as well as reviews, grey literature and conference abstracts were excluded.

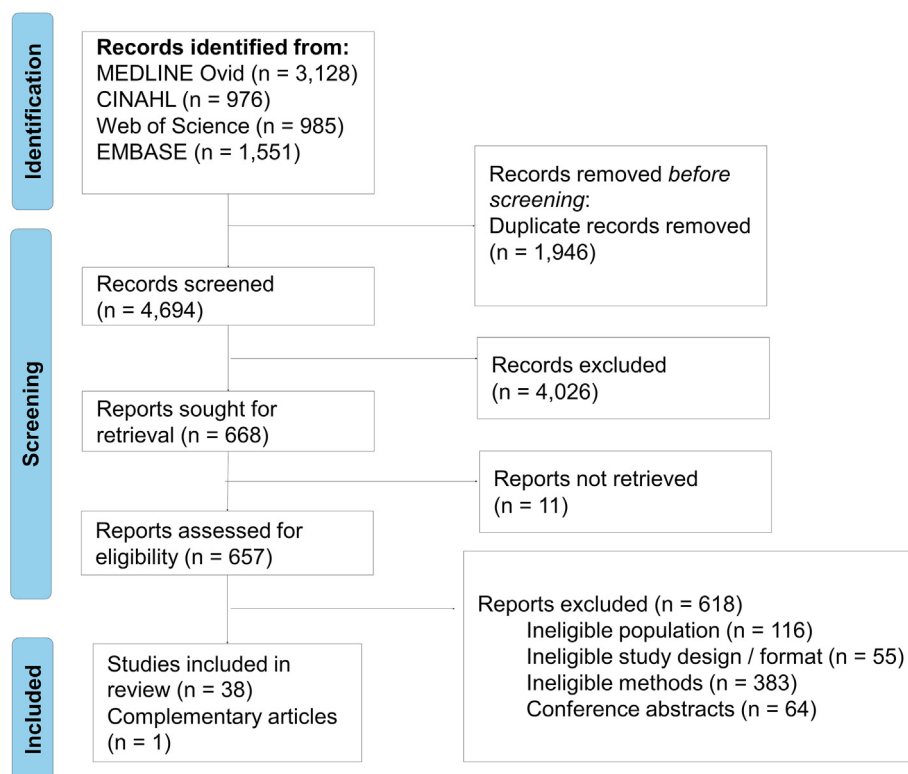


Figure 1. PRISMA 2020 Flow diagram of included studies [10].

Table 1. Characteristics of studies using latent growth modeling approaches among patients with type 2 diabetes.

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
Latent growth mixture modeling									
de Vries McClintock 2016 [17,18]	Cohort: Brief Intervention to Improve Adherence through Integrated Management of type 2 Diabetes Mellitus and Depression Treatment (2010–2011) Type of data source: Primary data (clinical trial)	To understand: 1- The course of oral hypoglycemic agent adherence patterns over 12 weeks among T2D patients in primary care. 2-Whether such patterns are related to patient characteristics. 3-Whether patterns predict glycemic control at 12 weeks.	Drug utilization research/pharmaco-epidemiology	Prevalent cases (n = 182)	Integrated Care intervention Tool: Educational intervention by integrated care managers Explicative model: Adherence trajectories	Adherence trajectories Tool: Medication Event Monitoring System caps Explicative model: HbA1c at 12 weeks after trajectory analysis Tool: in2it A1C Analyzer	Comparison between groups: X ² and ANOVA Explicative model: multivariate logistic regression	Time-point studied: 1 week Trajectory identification: 12 weeks Outcome assessment: at week 12 Follow-up period: 12 weeks	General growth curve mixture modeling [21, 22, 23, 24, 25, 26]
Hertroijs 2018 [14]	Cohort: Zwolle Outpatient Diabetes Project Integrating Available Care (2006–2013); Primary care group ZIO (validation cohort) (2009–2014) Type of data source: Secondary data (electronic health records)	1-To identify, predict and validate distinct glycemic trajectories among patients with newly diagnosed T2D treated in primary care.	Clinical research	Incident cases (n = 10,528)	Predictors of trajectories Tool: Sociodemographic factors, clinical/laboratory measures, comorbidities, family medical history	HbA1c trajectories Tool: Not reported	Comparison between groups: ANOVA, X ² Validation or predictive model: multinomial logistic backward regression	Time-point studied: 1 year (±3 months) Follow-up period: 4–5 years	Latent growth mixture modelling [2, 3, 27]
Laiteerapong 2017 [16]	Cohort: Kaiser Permanente Northern California Diabetes Registry (1997–2012) Type of data source: Secondary data (electronic health records, medico-administrative)	1-To classify trajectories of long term HbA1c values after diagnosis of T2D. 2- To examine each trajectory's associations with subsequent microvascular and macrovascular events and mortality.	Clinical research	Incident diabetes (n = 28,016) (n = 25,732 for outcomes)	HbA1c trajectories Tool: Not reported	Diabetes complications Tool: ICD-9 and procedures codes	Explicative model: Multivariate Cox proportional hazard regression model	Time-point studied: 1 year Trajectory identification: 10 years Outcome assessment: 4 years Total follow-up period: 14 years	Latent growth mixture modeling [24, 28]
Wang 2011 [15]	Cohort: San Antonio Longitudinal Study of Aging (1992–1996 to 2000–2005) Type of data source: Primary data (prospective cohort)	1-To examine whether better glycemic control improves the maintenance of lower-extremity physical function over a 36-month period among participants with diabetes.	Clinical research	Prevalent cases (n = 119)	HbA1c trajectories Tool: Not reported	Lower-extremity function Tool: Short physical performance battery	Comparison between classes: Pseudo-class estimation technique Explicative model: Mixture path modeling with propensity score	Time-point studied: 6 months Outcome assessment: 18 and 36 months Total follow-up period: 36 months	Latent growth mixture modeling [22, 29]
Whitworth 2017 [12]	Cohort: Fremantle Diabetes Study Phase II (2008–2011, follow-up until 2016) Type of data source:	1-To describe the long-term trajectories of depression symptom severity in people with T2D.	Psychological science	Prevalent cases (n = 1,201)	Predictors of trajectories: Clinical, biochemical, demographic,	Trajectories of depression symptoms Tool: Patient Health Questionnaire	Comparison between groups: One-Way ANOVA, X ² , Mann-Whitney U	Time-point studied: 2 years Trajectory identification follow-	Latent class growth analysis [30]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
	Primary data (prospective cohort)	2-To identify predictors and associates of these trajectories.			depression, psychosocial data		test, Kruskal-Wallis test Predictive model: Binary logistic regression model	up: 4 years Total follow-up: 4 years	
Whitworth 2020 [13]	Cohort: Fremantle Diabetes Study Phase II (2008–2011, follow-up until 2016) Type of data source: Primary data (prospective cohort)	1-To identify distinct trajectories of anxiety symptoms in individuals with T2D over time. 2-To identify demographic, self-management and clinical predictors of anxiety trajectory membership. 3-To assess whether having a lifetime history of anxiety or depression predicts anxiety trajectory membership beyond significant demographic predictors. 4-To examine whether anxiety trajectory membership is associated with important self-management and clinical outcomes, after controlling for these variables at baseline.	Psychological science	Prevalent cases (n = 1,549)	Obj 1–3: Predictors of trajectories: Clinical, biochemical, demographic, depression, psychosocial data Obj 4: Trajectories of anxiety symptom severity Tool: Generalized Anxiety Disorder Scale	Obj 1–3: Trajectories of anxiety symptom severity Tool: Generalized Anxiety Disorder Scale Obj 4: Self-management and clinical variables at year 4 Tool: Frequency of self-monitoring of blood glucose	Comparison between groups: X^2 , one-way ANOVA, Kruskal-Wallis Predictive model: Bivariate logistic regression model (unadjusted and adjusted) Explicative model: Linear regression model, adjustment for confounding factors	Time-point studied: 2 years Trajectory identification follow-up: 4 years Total follow-up: 4 years	Latent growth Mixture modeling [30]
Latent class growth analyses									
Bayliss 2011 [31]	Cohort: Health maintenance organization cancer registry (1998–2008) Type of data source: Secondary data (Medico-administrative)	1-To assess the effect of incident stage 0, 1, 2 or 3 breast, colon or prostate cancer; incident depression; or an exacerbation of COPD on control of T2D.	Clinical research	Prevalent cases of T2D and incident cancer/depression or COPD exacerbation (n = 5,883)	Months since diagnostic of cancer Tool: ICD-9, prescriptions, hospitalizations	HbA1c trajectories Tool: Not reported	Descriptive analyses for comparing trajectory groups (graphs)	Time-point studied: 3 months Trajectory identification: 18 months Total follow-up: 84 months	Latent class growth modeling [32]
Bocquier 2019 [33]	Cohort: Permanent Sample of Beneficiaries (2006–2015) Type of data source: Secondary data (Medico-administrative)	1-To identify temporal trajectories of seasonal influenza vaccination uptake 2- To describe their clinical characteristics.	Healthcare utilization research	Prevalent cases (n = 15,766)	Predictors of trajectories Tools: Demographic, clinical, and healthcare utilization factors	Trajectories of seasonal influenza vaccination Tool: binary variable for each influenza season	Comparison between groups: ANOVA, X^2 Predictive model: multivariate	Time-point studied: 1 year Follow-up period: 10 years	Group-based trajectory modeling [3, 34, 35, 36]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
							logistic regression		
Botvin Moshe 2020 [37]	Cohort: Israel Diabetes and Cognitive Decline study (from 1998, years not reported) Type of data source: Primary data (prospective cohort, registry)	1-To investigate the associations of long-term measurements of body mass index with indices of carotid stiffness and atherosclerosis among non-demented diabetes patients.	Clinical research	Prevalent cases (n = 471)	Body mass index trajectories Tool: body mass index equation	Carotid intima-media thickness, distensibility coefficient and elastography strain ratio, carotid plaque volume Tool: Carotid ultrasound Doppler	Comparison between groups: t-tests, X ² , ANOVA, Wilcoxon Explicative model: linear regression model or logistic regression model	Time-point studied: 1 month Trajectory identification follow-up: 120 months prior baseline Outcome assessment: 36 months after baseline	Multinomial modeling strategy [34]
Chen 2016 [38]	Cohort: National Health insurance claims data (2002–2008) Type of data source: Secondary data (Medico-administrative)	1-To identify medication adherence trajectories among patients with newly diagnosed diabetes 2-To examine the association of continuity of care and medication adherence among various adherence trajectories.	Healthcare utilization research	Incident case (n = 12,123)	Continuity of care Tool: Continuity of care index	Adherence trajectories Tool: Medication possession ratio	Explicative model: Multivariate logistic regression model	Time-point studied: 1 year Follow-up period: 6 years	Group-based trajectory modeling [32, 34, 35, 39]
Chiu 2013 [19]	Cohort: Health and Retirement Study (1992–2002) Type of data source: Primary data (retrospective cohort)	1-To identify the main patterns of weight and disability trajectories experienced by middle aged and older adults with diabetes. 2-To identify the proportion of each trajectory in the population. 3-To examine the association between weight trajectories and disability trajectories later in life, as well as baseline sociodemographic, clinical, behavioral, and diabetes-related factors.	Clinical research	Prevalent cases (self reported) (n = 1,064)	Body mass index trajectories Tool: body mass index equation	Disability trajectories Tool: Activities of daily life score; Instrumental activities of daily life	Comparison between classes: ANOVA and X ² Dual trajectory modeling: conditional probability of membership across trajectory groups.	Time-point studied: 1.5–2 years Follow-up period: 10 years	Group-based semi-parametric mixture modeling approach and dual trajectory modeling [34, 34, 40, 41]
Chiu 2017 [42]	Cohort: Taiwan Longitudinal Study on Aging (1996–2007) Type of data sources: Primary data (survey)	1-To identify distinct trajectories of depressive symptoms after diagnosis of diabetes in middle-aged and older adults. 2-To ascertain the proportion of adults in each trajectory.	Psychological science	Incident cases, no diabetes in 1996, but diabetes in 1999 (self-reported) (n = 487)	Predictive model: Predictors of trajectories: laboratory and clinical measures, socio-demographic measures, mobility score, self-rated health, comorbidities, lifestyle behaviors	Predictive model: Depression symptoms trajectories Tool: Center of Epidemiological Studies – Depression scale Explicative model: future disability	Comparison between groups: X ² and ANOVA Predictive model: Multinomial logistic regression	Time-point studied: 1 year Trajectory identification: 8 years Outcome assessment: 2007	Latent class growth modeling [32, 35, 39]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
		3-To examine antecedent and subsequent factors associated with each trajectory.			Explicative model: trajectory groups	Tools: Activities of daily life score; Instrumental activities of daily life Nagi's mobility item	Explicative model: Multiple regression		
Cooke 2020 [43]	Cohort: SMARTER randomized controlled trial (2012–2016) Type of data source: Primary data (clinical trial)	1-To examine indicators of trajectory membership of both steps/day and changes from baseline steps/day over the 1-year intervention.	Clinical research	Overweight/obese adults with T2D and/or hypertension (n = 118)	Predictors of trajectory group membership Tools: Baseline sociodemographic, T2D, time of intervention, anthropometric, clinical data	Trajectories of mean steps/day Tool: Average steps/day on a 30-day period	Predictive model: Cumulative logistic regression	Time-point studied: 30 days Follow-up period: 1 year	Group-based trajectory modeling [3, 34]
Davis 2016 [44]	Cohorts: Fremantle Diabetes Study Phase 1 (1993–1996) and Western Australia data linkage system (1998–2001) Type of data source: Primary data (prospective cohort)	1-To determine whether there was a mortality benefit of tight glycemic control beyond the period in which it was implemented in recently diagnosed patients; a neutral or increased risk of death in those with long-duration diabetes.	Clinical research	Prevalent cases (n = 531) Subgroups according to diabetes duration	HbA1c trajectories Tool: Not reported	Mortality at 5 years Tool: Not reported	Explicative model: Kaplan-Meier and multivariate Cox proportional model	Time-point studied: 1 year Outcome assessment: at year 5 Total follow-up period: 5 years	Semi-parametric group-based modeling strategy [34, 45, 46]
Davis 2016 [47]	Cohorts: Fremantle Diabetes Study Phase 1 (1993–1996) and Western Australia data linkage system (1998–2001) Type of data source: Primary data (prospective cohort)	1-To investigate the association between estimated GFR and all-cause mortality, including the contribution of temporal eGFR changes.	Clinical research	Prevalent cases (n = 1,296)	eGFR trajectories Tool: Laboratory measures	Mortality at year 5 Tool: Linkage to Western Australia Data Linkage System, all death registration and hospitalisations	Comparison between groups: Fisher exact test Explicative model: Kaplan-Meier and multivariate Cox proportional model	Time-point studied: 1 year Outcome assessment: at year 5 Total follow-up period: 5 years	Semi-parametric group-based modeling [45, 46]
Deschènes 2018 [48]	Cohort: Evaluation of Diabetes Treatment study (2011–2014) Type of data source: Primary data (surveys)	1-To examine latent longitudinal trajectories of anxiety symptoms in adults with T2D and their associations with incident cardiovascular disease.	Psychological science	Prevalent cases (n = 832)	Trajectories of anxiety symptoms Tool: Generalized Anxiety Disorder Scale	Cardiovascular disease at 24 months, 36 months, 48 months. Tool: Diabetes Complications Index (DCI)	Comparison between groups: X ² Explicative model: Univariate and multivariate logistic regression model	Time-point studied: 1 year Trajectory identification follow-up: 2 years Outcome assessment: 3 years Total follow-up period: 5 years	Semi-parametric latent class growth modeling [34, 39]
Goh 2015 [49]	Cohort: Not reported (primary care clinic northeastern part of Singapore) (2013–2014) Types of cohorts: Primary data (clinical trial)	1-To identify and describe short-term trajectories of use of the <i>Interactive Diet and Activity Tracker</i> app in a primary care setting 2-To identify patient characteristics	Clinical research	Prevalent cases (n = 84)	Predictors of trajectories: Laboratory and clinical measures, sociodemographic measures, lifestyle behaviors	Trajectories of utilization of a telephone application Tool: monitored weekly	Comparison between groups and predictive model: Univariate/multivariate stepwise polytomous	Time-point studied: 1 week Follow-up period: 8 weeks	Latent class growth modeling [3, 34, 35, 50]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
		associated with each trajectory.					logistic regression		
Lee 2018 [51]	Cohort: A medical centre in Taiwan (2009–2012, until 2014) Type of data source: Primary and secondary data (retrospective cohort, hospital records)	1-To investigate the effect of changes in fasting plasma glucose variability, as assessed by 2-year trajectories of fasting plasma glucose variability, on mortality risk in patients with T2D.	Clinical research	Prevalent cases (n = 3,569)	Glucose variability Tool: mean glucose and coefficient of variation of visit-to-visit of fasting glucose level	Mortality Tool: Hospital records	Comparison between groups: X ² and ANOVA Explicative model: Kaplan-Maier & log-rank test Cox proportional hazard for cardiovascular mortality	Time-point studied: 3 months Trajectory identification: 2 years Outcome assessment: 4 years Follow-up period: 6 years	Latent class growth modeling [3, 50, 52]
Li 2018 [53]	Cohort: The China Kailuan study (2006–2014) Type of data source: Primary data (retrospective cohort)	1-To investigates the effect of long-term systolic blood pressure trajectory on kidney damage in the diabetic population.	Clinical research	Prevalent diabetes (n = 4,556)	Blood pressure trajectories Tools: corrected desktop mercury sphygmomanometer & electronic sphygmomanometer	Onset of kidney damage in 2014 Tool: eGFR, proteinuria	Comparison between groups: ANOVA, Least Significant Difference test and Dunnett test, X ² Explicative model: Multivariate logistic regression	Time-point: 2 years Trajectory identification: 8 years Outcome assessment: in 2014 only Total follow-up: 8 years	Trajectory model [3, 34, 54, 55]
Li 2021 [56]	Cohort: Taiwan's National Health Insurance Research Database (2002–2003) Type of data source: Secondary data (medico-administrative, electronic health records)	1-To investigate associations between exposure to various trajectories of severe hypoglycemic events and risk of dementia in patient with T2D.	Healthcare utilization research	Prevalent cases (n = 677,618)	Hypoglycemic episodes trajectories Tool: ICD-9-CM (clinical modification) codes	Dementia Tool: ICD-9-CM codes	Explicative model: Multivariate Cox proportional hazard regression model, with sub-distribution Hazard Ratio	Time-point studied: 6 months. Trajectory identification: 3 years Outcome assessment: average: 6 years, max 7 years	Group-based trajectory modeling [34, 39, 54]
Lipscombe 2015 [57]	Cohort: Evaluation of Diabetes Treatment study (2011–2014) Type of data source: Primary data (surveys)	1-To identify and describe a set of distinct longitudinal trajectories of diabetes distress over 4 years of follow-up time.	Psychological science	Prevalent cases (n = 1,135)	Trajectories of diabetes distress Tool: Diabetes Distress Scale	Characteristics: Socio-demographic, mental health, diabetes-related and lifestyle factors	Comparison between groups: ANOVA, X ² , Bonferroni correction	Time-point studied: 1 year Follow-up period: 3 years	Latent class growth modeling [32]
Lo-Ciganic 2016 [58]	Cohort: Pennsylvania Medicaid administrative claims data (2007–2011) Type of data source: Secondary data (medico-administrative)	1-To examine the association between adherence trajectories for oral hypoglycemics and subsequent hospitalizations among diabetic patients.	Pharmacoepidemiology/ Drug utilization research	New users of oral hypoglycemic agents (n = 16,256)	Adherence trajectories to oral hypoglycemics Tools: number of days covered by at least one hypoglycemic medication/30 days	Hospitalisation related to diabetes/all-cause hospitalisation Tools: ICD-9 codes, Current Procedural Terminology procedure codes	Comparison between groups: X ² and other tests (not mentioned) Explicative model: Multivariate Cox proportional model	Time-point studied: 1 month Trajectory identification: 1 year Outcome assessment: 1 year Total follow-up period: 2 years	Group-based trajectory modeling [3, 8, 39, 45]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
Low 2019 [11]	Cohort: Name of the cohort not reported (2002–2017) Type of data source: Primary data (prospective cohort)	1-To characterize HbA1c trajectories and examine their associations with chronic kidney disease progression.	Clinical research	Prevalent cases (n = 770)	HbA1c trajectories Tool: Standard laboratory measures	Chronic kidney disease progression over time Tool: Decline of eGFR category	Comparison between groups: X ² , one way ANOVA, Mann-Whitney test, t-test Explicative model: Multivariate Cox proportional hazards regression	Time-point studied: 1 year Trajectory identification follow-up: 8 years Outcome assessment: 8 years Total follow-up period: 8 years	Group-based trajectory modeling [39, 45]
Luo 2017 [59]	Cohort: Singapore Consortium of Cohort Studies Diabetes Cohort (2004–2010) Type of data sources: Primary and secondary data (prospective cohort, medico-administrative, electronic health records)	1-To examine longitudinal trends in HbA1c in a multi-ethnic Asian cohort of diabetes patients, 2-To examine the associations of these trends with future risk of acute myocardial infarction, stroke, end stage renal failure and all-cause mortality.	Clinical research	Prevalent cases (n = 6,079)	HbA1c trajectories after catheterization Tool: Not reported	Trends of change in serum lipids Tool: NR Time-to-event: death, acute myocardial infarction, stroke, end stage renal failure, serum lipids, all-cause mortality Tool: Linkage to National Registry of Disease Office	Comparison between groups: Kruskal-Wallis, X ² Explicative model: Multivariate generalized estimating equation, multivariate Cox proportional hazards regression	Time-point studied: 12 months Trajectory identification: median of 4.1 years prior recruitment Outcome assessment: median of 7 years–8.3 years, post-recruitment depending on the outcome	Latent class growth modeling [3, 8, 60]
Luo 2019 [61]	Cohorts: Singapore Population Health Studies Diabetic Cohort (2004–2010) and National Healthcare Group Polyclinics (2011–2016) Types of cohorts: primary and secondary data (prospective cohort, electronic health records)	1-To analyze diabetes treatment and treatment changes in association with long-term glycemic patterns in an Asian population with diabetes.	Clinical research/ healthcare utilization research	Prevalent cases (n = 6,218)	HbA1c trajectories Tool: Not reported	Annual treatment plans Tool: Antidiabetic medication, alone or in combination	Comparison between groups: ANOVA, Kruskal-Wallis, Pearson X ² Comparison between groups in time: Cochrane Armitage	Time-point studied: 1 year Total follow-up period: 6 years	Latent class growth analysis [3, 8, 60]
Niaz 2021 [62]	Cohort: Alberta Health Databases (2008–2018) Type of data source: Secondary data (medico-administrative)	1-To characterize adherence to oral antihyperglycemic medications in the year before a depressive episode 2-To examine the association between depression and medication adherence.	Pharmaco-epidemiology	New users of metformin (n = 165,056)	No exposition for trajectory modeling	Trajectories of adherence Tool: Proportion of days covered to oral antihyperglycemic medications (<80% as threshold)	Descriptive analysis: Presentation of trajectories of proportion of days covered prior depression diagnosis	Time-point studied: 30 days Follow-up period: 1 year	Group-based trajectory modeling [55]
Obura 2020 [63]	Cohort: Medicines Initiative-Diabetes Research on Patient Stratification (IMI-DIRECT) Study Type of data source:	1-To examine the association between subgroups based on their glucose curves during a five-point mixed-meal tolerance	Clinical research	Incident and prevalent cases (within 6–24 months) (n = 789)	Glucose curves subgroups following a mixed-meal tolerance test	Metabolic traits and glucose deterioration at 18 months	Comparison between groups: general linear model Explicative model:	Time-point studied: 30 min Trajectory identification follow-up: 120 min Outcome assessment: At	Latent class trajectory analysis (reference not reported)

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
	Primary data (prospective cohort)	test and metabolic traits at baseline and glycemic deterioration in individuals with T2D.					multivariate general linear model & Bonferroni correction	18 months Total follow-up: 18 months	
Raghavan 2020 [6]	Cohort: United States veteran healthcare system (2005–2006) Type of data source: Secondary data (medico-administrative, electronic health records)	1-To identify glycemic control trajectories. 2-To describe the characteristics of patients with distinct HbA1c trajectories 3- To compare the associations of HbA1c trajectories and single HbA1c measurements with short-term mortality.	Clinical research	Prevalent cases (<2 years since T2D diagnosis) (n = 7,780)	HbA1c trajectories Tool: Not reported	Mortality Tool: vital status data	Explicative: Joint latent class survival model	Time-point studied: days/1 month Follow-up period: 2 years	Joint Latent class growth analysis [20, 21, 22, 64]
Rathmann 2019 [65]	Cohort: Diabetes Patienten Verlaufsdocumentation (years not reported) Type of data source: Secondary data (registry)	1-To identify groups of heterogeneous HbA1c trajectories over time in newly diagnosed T2D.	Clinical research	Incident cases (n = 6,355)	Comparison between groups in time: HbA1c trajectories Predictive model: Predictors of trajectories Tool: Diabetes treatment, microvascular complications, sociodemographic factors	Comparison between groups in time: Diabetes medication trends Predictive model: HbA1c Trajectories Tool: standardized using the multiple of the mean method according to the Diabetes Control and Complications trial	Comparison between groups: X^2 , Kruskal-Wallis Comparison between groups in time: descriptive analysis Predictive model: multinomial logistic regression, Bonferroni-Holm correction	Time-point studied: 1 year Follow-up period: 5 years	Latent class growth modeling [3, 66]
Schmitz 2013 [67]	Cohort: Montreal Diabetes Health and Well Being Study (2008–2011) Type of data source: Primary data (surveys)	1-To identify and describe longitudinal trajectories of self-rated health status in people with diabetes.	Clinical research	Prevalent self-reported diabetes (n = 1,288)	Trajectories of self-reported health Tool: Likert-scale, with scores from 0 to 100	Global functioning Tool: World Health Organization Disability Assessment Schedule II	Comparison between groups: general linear model, X^2 Explicative model: multivariate logistic regression model	Time-point studied: 1 year Trajectory identification follow-up: 3 years Outcome assessment: 1 year Total follow-up period: 4 years	Semi-parametric trajectory modeling [34, 35]
Sidorenkov 2018 [68]	Cohort: Groningen Initiative to Analyze Type 2 Diabetes Treatment database (2007–2013) Type of data source: Secondary data (medico-administrative, electronic health records)	1-To identify subgroups of patients with T2D following distinct trajectories of HbA1c after insulin initiation. 2-To explore underlying differences in clinical characteristics.	Clinical research	Prevalent and incident cases with insulin initiation (n = 1,459)	HbA1c trajectories Tool: Standard laboratory measures	Baseline characteristics: Laboratory and clinical measures, socio-demographic measures, history of diabetes-related comorbidities, diabetes duration	Comparison between groups: one way ANOVA, Kruskal-Wallis, X^2 , Tukey Honestly Significant Difference test	Time-point studied: 6 months Total follow-up period: 4 years	Latent class growth modeling [50]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
Tsai 2019 [7]	Cohort: Longitudinal Health Insurance Database 2000 (2000–2010) Type of data source: Secondary data (medico-administrative)	1-To explore the longitudinal care seeking patterns of diabetic patients. 2-To identify baseline characteristics associated with trajectories of care seeking behaviors.	Healthcare utilization research	Incident cases (n = 3,987)	Predictors of trajectories: Sociodemographic factors, severity of diabetes at diagnosis, diabetes complications, severity index	Trajectories of seeking patterns Tool: Regularity of visits to specialized/generalized providers (intervals <90 days for regularity)	Comparison between groups: X ² and ANOVA Predictive model: multinomial logistic regression	Time-point studied: 1 year Total follow-up period: 11 years	Group-based trajectory [34, 39, 40, 69]
Tsai 2019 [70]	Cohort: Longitudinal Health Insurance Database 2000 (2000–2010) Type of data source: Secondary data (medico-administrative)	1-To investigate diabetes outcomes by long-term trajectories of patients care settings among diabetes patients with regular follow-up.	Healthcare utilization research	Incident cases (n = 1,268)	Trajectories of care settings Tool: At least 1 visit/year	Diabetes complications Tool: Invasive procedures using intervention and ICD-9-CM codes	Explicative model: Cox proportional hazards with cumulative incidence function	Time-point studied: 1 year Trajectory identification: 5 years Outcome assessment: 5 years Follow-up period: 10 years	Group-based trajectory modeling [32, 34, 39]
Vistisen 2019 [71]	Cohort: Not reported (Steno Diabetes Center Copenhagen, 2001–2017) Type of data source: Secondary data (registry)	1-Assessing potential heterogeneity in eGFR development among persons with diabetes and normo-albuminuria after entering stage 3 chronic kidney disease.	Clinical research	Prevalent cases (subset of individuals with normo-albuminuria) and Type 1 diabetes (n = 935) or T2D (1,984)	eGFR trajectories Tool: Chronic Kidney Disease Epidemiology Collaboration standard equation	Baseline characteristics: Laboratory and clinical measures, socio-demographic measures, duration of diabetes, lifestyle behaviors, medication, retinopathy status	Comparison between groups: t-tests, X ²	Time-point studied: 1 year Follow-up period: 10 years	Latent class trajectory modeling [60, 72, 73]
Walraven 2015 [74]	Cohort: Diabetes Care System West-Friesland (1998–2001) Type of data source: Primary and secondary data (prospective cohort, electronic health records)	1-To identify subgroups of T2D patients with distinct HbA1c trajectories. 2-To investigate the prevalence of microvascular complications over time.	Clinical research	Prevalent cases (n = 5,423)	Comparison between groups in time: HbA1c trajectories Tool: Standard laboratory measures Predictive model: predictors of trajectories Tools: clinical, laboratory measures, sociodemographic factors, lifestyle behaviors	Comparison between groups in time: diabetes complications over time, medication use over time Tools: Urinary albumin-creatinine ratio; EURODIAB classification Predictive model: HbA1c trajectories Tool: Standard laboratory measure	Comparison between groups: ANOVA, X ² Comparison between groups in time: Binomial mixed modeling approach Predictive model: multinomial logistic backward regression	Time-point studied: 1 year Total follow-up period: 9 years	Latent class growth modeling [8]
Walraven 2015 [75]	Cohort: Diabetes Care System West-Friesland (1998–2010) Type of data source: Secondary data (electronic health records)	1-To identify subgroups of T2D patients with distinct trajectories of systolic blood pressure levels.	Clinical research	Prevalent cases (n = 5,711)	Predictive model: predictors of trajectories Tools: clinical, laboratory measures, sociodemographic factors, lifestyle behaviors Explicative model: trajectories of blood	Predictive model: trajectories of blood pressure control Explicative model: prevalence of retinopathy & microalbuminuria Tools: fundus photography &	Comparison between groups: ANOVA, X ² Predictive model: multinomial logistic backward	Time-point studied: 1 year Follow-up period: max 9 years (mean follow-up of 5.7 years)	Latent class growth modeling [3, 76, 77]

(continued on next page)

Table 1 (continued)

Study	Data source(s) and type	Objectives	Field/context of research	T2D population	Exposures evaluated and tools	Outcome evaluated and tools	Methods used for research question	Time scale	Method used*
					pressure control Tool: sphygmomanometer/ oscillometer device after 2003	EURODIAB classification, urinary albumin-creatinine ratio Tool: mortality in death registry (cause of death with ICD-9 codes)	regression Explicative model: Binomial mixed modeling approach, multivariate Cox proportional model		
Wang 2019 [78]	Cohort: Five diabetic clinics in Taiwan (2014–2016) Type of data sources: Primary and secondary data (prospective cohort, medical records)	1-To identify quality of life trajectory patterns and the determinants in patients with T2D.	Clinical research	Prevalent cases (n = 466)	Predictors of trajectories: laboratory and clinical measures, socio- demographic measures, psychosocial factors, lifestyle behaviors, complication of diabetes, insulin use	Trajectories of quality of life Tool: Diabetes- Specific Quality of Life scale	Comparison between groups: X ² and ANOVA Predictive model: multinomial logistic regression	Time-point studied: 6 months. Follow-up period: 2 years	Latent class growth analysis [21, 39, 40]
Zavrelova 2011 [79]	Cohort: Diabetes Care System West-Friesland (1998–2006) Type of data source: Secondary data (electronic health records)	1-To identify distinct developmental patterns of diabetic retinopathy 2-To assess the risk factor levels of patients in these clusters.	Clinical research	Prevalent cases (n = 3,343)	Diabetic retinopathy developmental patterns (trajectories) Tool: Fundus photography, EURODIAB grading system	Baseline characteristics: laboratory and clinical measures, socio-demographic measures	Comparison between groups: ANOVA with post-hoc Bonferroni or X ² tests, Kruskal- Wallis	Time-point studied: 1 year Follow-up period: 6 years	Latent class growth modeling [2, 21]

COPD: Chronic obstructive pulmonary disease, eGFR: Estimated glomerular filtration rate, EURODIAB: European Community funded Concerted Action Programme into the epidemiology and prevention of diabetes; HbA1c: Glycated hemoglobin A1c, ICD = International classification of diseases; T2D: Type 2 diabetes.

* as mentioned and cited in text.

Table 2. Characteristics of studies and context of utilization of latent growth modeling approaches in the field of type 2 diabetes.

	LGMM, 6 studies		LCGA, 32 studies		Total, 38 studies	
	Number	Proportion, %	Number	Proportion, %	Number	Proportion, %
Field/context of research						
Clinical research	3	50	22	69	25	66
Pharmaco-epidemiology/drug utilization research	1	17	2	6	3	8
Research in healthcare utilization	0	0	5	16	5	13
Psychological science	2	33	3	9	5	13
Utilization of trajectory groups						
Exposure: Comparison of baseline data only	0	0	4	13	4	11
Exposure: Comparison between groups over time	0	0	2	6	2	5
Exposure: Association to an outcome (explanatory model)	3	50	16	50	19	50
Logistic regression model	1	17	4	13	5	13
Linear regression model	0	0	1	3	1	3
Cox proportional model or survival analysis	1	17	8	25	9	24
Logistic or linear regression model	0	0	1	3	1	3
Other/multiple models	1	17	2	6	3	8
Outcome: Descriptive only/trends	0	0	2	6	2	5
Outcome: Predictive model	2	33	7	22	9	24
Both explanatory and predictive models	1	17	1	3	2	5
Data sources						
Medico-administrative databases	0	0	7	22	7	18
Trials or prospective/retrospective cohorts	4	67	8	25	12	32
Surveys	0	0	4	13	4	11
Health records/registry	1	17	4	13	5	13
Medico-administrative database and health records	1	17	3	9	4	11
Clinical studies & health records	0	0	5	16	5	13
Mixed data sources	0	0	1	3	1	3
Population						
Incident cases of T2D	2	33	6	19	8	21
Recruitment max 2 years after T2D diagnosis	0	0	2	6	2	5
Prevalent cases of T2D	4	67	22	69	26	68
Both, prevalent and incident cases of T2D	0	0	2	6	2	5

LCGA: Latent class growth analysis, LGMM: Latent growth mixture modeling, T2D: type 2 diabetes.

2.2. Search of the studies

MEDLINE (Ovid), EMBASE, CINAHL and Web of Science were systematically searched through August 25th, 2021, using a prespecified search strategy available in the **Supplemental material**. Citations were combined and duplicates were removed in Endnote 20 software (Clarivate, London, United Kingdom). Selection and full-text eligibility were performed by one reviewer (SO) in Covidence software (Melbourne, Australia). Since some published articles shared the same population/data sources but had different primary outcomes and methods, they were considered as two separated studies. However, any article using the same data source with duplicate methods and identical objectives were considered as only one study.

2.3. Data extraction

Data extraction was performed by one reviewer (SO) using a pre-piloted form in Excel software (Microsoft, Redmond, Washington, USA). Extracted data included: (1) **Type of method**: Method as identified in the studies and associated references, (2) **Context of utilization**: Study design, year of publication, data source(s), objective(s), field of research, population (incident or prevalent cases of T2D), number of subjects, exposure(s)/outcome(s) of interest, time-points, follow-up periods, other statistical method(s) (descriptive analysis, trends, predictive models, explanatory models, trajectory groups as exposure or outcomes) and, (3) **Quality of reporting** using the GRoLTS checklist [4]. The

GRoLTS checklist is composed of 16 items, covered by 21 questions [4]. In addition, we collected information about actions taken for preventing indication bias and immortal time bias. We tabulated the results in summary tables inspired by Jandoc et al. [9].

3. Results

3.1. Characteristics of included studies

From the 4,694 citations screened, a total of 38 publications were included in this review (Figure 1). The characteristics of studies are summarized in Table 1 and Table 2. The studies were published between 2011 and 2021, from which 76% were published between 2016 and 2021. Sample sizes varied from 84 to 677,618 participants and total follow-up ranged from 8 weeks to 14 years. However, 95% of studies had a total follow-up of ≥ 1 year and 57%, ≥ 5 years. The variables used for modeling trajectory groups varied from: (1) biomarkers/clinical tests (glycated hemoglobin A1c (HbA1c) [n = 12], blood glucose levels [n = 2], blood pressure [n = 2], estimated glomerular filtration rate (eGFR) [n = 2], body mass index [n = 2], fundus photography [n = 1]), (2) drug usage outcomes (medication event monitoring system caps [n = 1], proportion of days covered [n = 2], medication possession ratio [n = 1]), (3) healthcare utilization indicators (vaccination [n = 1], medical visits [n = 2]), (4) patient reported outcomes (depressive symptoms [n = 2], anxiety symptoms [n = 2], diabetes distress symptoms [n = 1], self-reported health [n = 1], quality of life [n = 1]), along with other types

of variables (step counts, utilization of cellphone application, hypoglycemic episodes). Eight studies included only incident cases of T2D, while 2 studies recruited participants with a diagnosis of T2D in the last 2 years. The other studies included participants with various lengths of T2D duration (prevalent cases).

The objectives of the studies evaluating trajectory groups could be separated in three main categories. (1) **The description/comparison of trajectory groups:** Twenty-nine (76%) of studies compared baseline characteristics of groups using comparison tests, such as t-tests, chi-square, ANOVA, Wilcoxon or Kruskal-Wallis tests. Three studies evaluated the trends of a specific outcome according to the trajectory groups, either descriptive analyses, Cochran Armitage test or binomial mixed models. One study depicted trajectory groups in graphs, without further analysis. (2) **Predictive model of group membership:** Eleven studies used a predictive model (multivariate, bivariate, polytomous or cumulative logistic regression models) to identify predictors of belonging to the different trajectory groups. (3) **Trajectory groups as exposition to an outcome of interest:** Twenty-one studies used a latent growth modeling approach as a first step to identify exposure groups, then used regression models to test the association between trajectory groups and outcome(s) of interest (univariate and/or multivariate linear or logistic regression models, Kaplan-Meier, Cox proportional hazard model, joint survival analysis). From these studies, two evaluated the exposition trajectory groups and outcomes simultaneously [6, 11], while the rest (n = 19) identified trajectory groups during a given period, then evaluated outcomes during a subsequent period following the end of the trajectory assessment.

A total of 6 studies considered within-class heterogeneity by using LGMM [12, 13, 14, 15, 16] or a General Growth Class Mixture Modeling (GGCMM) for randomized intervention [17, 18]. Most studies (n = 32) used LCGA. Although citing the same references, the terminology used by the investigators varied between LCGM, GBTM, LCGA, "Latent class trajectory analyses", "group-based trajectory", "group-based semi-parametric mixture modeling approach", "semi-parametric group-based modeling strategy" or "multinomial modeling strategy". One study used a dual trajectory modeling, which consisted of identifying trajectory groups for two distinct variables (body mass index and disability trajectories) [19]. One study used a joint LCGA to assess the association between HbA1c trajectories and mortality, in which survival analysis and trajectory group identification was performed simultaneously [20].

3.2. Quality of reporting using the GROLTS checklist

The summary of the GROLTS checklist items across studies is presented in Table 3. Well-reported items included the metric of time used, the statistical software used, the characteristics and graphical depiction of the final model and confirmation of group membership of the final solution using average posterior probability and/or entropy. Ninety-two percent of studies reported the tools used for trajectory identification, from which 94% used the Bayesian information criterion (BIC). The number of latent groups identified varied between 2 and 7, where most studies identified either 3 (n = 12), 4 (n = 15) or 5 trajectory groups (n = 7). Items partially reported included missing data mechanisms and management, the total number of fitted models tested or the distribution of observed data and the functional forms of trajectories in the final model. The items poorly reported included the consideration of mean and variance of time as well as within-class heterogeneity, the alternative across-class variance-covariance matrices considered, the number of random start values, the number of cases or graphical depiction of other models tested and availability of syntax files. From the 21 questions covered by the GROLST checklist, the total of elements reported ranged from 3 to 17 within studies while most studies adequately reported 7 to 10 elements. The details of the items for individual studies are available in Supplemental Material.

4. Discussion

4.1. Major findings

This review depicted the utilization of latent growth modeling approaches in longitudinal studies conducted among individuals with T2D. The heterogeneity in environmental, genetic, socio-economic factors influencing the outcomes of patients with T2D throughout time justifies the relevance of using latent growth modeling approaches [1, 2]. Our results brought evidence on the wide diversity of contexts of utilization, follow-up periods, type of data sources and research questions considered using these methods. In this first literature review on the topic to our knowledge, we observed an overall preference in the choice of LCGA over LGMM. In contrast, we did not observe any context of utilization specific to a method or another. Finally, this review identified some recurrent issues in reporting, notably on consideration of within-group heterogeneity, how the decisions were made or the management of missing data, as examples.

4.2. Current applications and limits of latent growth modeling approaches

The research questions covered in the included studies were diverse. Indeed, only a handful of studies limited their analysis to simply identify trajectory groups; most studies used the identified groups to answer subsequent questions. The utilization of trajectory groups as exposures to a clinical outcome was used in half of included studies. However, some issues can be documented with the utilization of trajectory groups as exposures.

First, most authors used a two-step approach, by identifying trajectory groups during an initial follow-up period, then using a regression model during a subsequent follow-up period for outcome assessment [6]. This design imposes individuals to survive up to the beginning of the second period in order to be considered in the analysis. Depending on the length of the follow-up, this 2-step approach may introduce an immortal time bias, notably because individuals must be exempt from the outcome of interest (and above all: survive) during the initial period of group identification [6]. Unfortunately, no study presented the potential impact of this immortal time bias on their results. Furthermore, given the broken temporality between group trajectory identification and outcome assessment, the clinical applicability of results is likely reduced [6]. In alternative, two studies evaluated trajectories and clinical outcomes simultaneously. This approach tackles the issue of immortal time bias yet might bring other issues about whether the temporality between the exposure and the outcome is adequately respected. Another issue of using latent classes in regression models is that the statistical uncertainty associated with finding the latent classes is not taken into account in the subsequent analysis. [80] While this may be expected to lead to underestimating the uncertainty of the regression parameters, it has been argued that this process leads to adequate large sample inferences if the data on the outcome are not used for the latent growth modeling step [80]. However, in case of poor classification (i.e.: low average posterior probabilities and/or low entropy), the interpretation of regression parameters using uncertain classes should be done with caution.

Second, another issue with the utilization of trajectory groups as exposures is the adjustment of covariates. Given the long follow-up expected in T2D progression, the advent of concurrent events or life changes might affect the occurrence of the outcome of interest [3]. In the context of a long-term follow-up, as the trajectories can vary in time, some covariates may be time-dependent and thus, may act as confounding and/or intermediate factors [80]. In other words, covariates may have an impact on both the exposure and the outcome of interest at some point in time; yet these might be influenced by the trajectory itself and become a mediator in the association. Both ignoring or adjusting for these changing covariates using standard approaches could lead to invalid estimates, especially when trajectory groups are modeled as exposures [80]. Only a few studies clearly reported having considered

Table 3. Assessing the quality of reporting of latent growth modeling approaches using the GROLTS checklist.

Items	Answer	N	%	Comments
1. Is the metric of time used in the statistical model reported?	Yes	38	100	The metric of time used was reported either in text, graph or both. Most studies used years/months or weeks from baseline, one study used the age of participants and two studies used time prior/after an index date. The spacing between points were also adequately reported.
	No	0	0	
	Unclear	0	0	
2. Is information presented about the mean and variance of time within a wave?	Yes	2	5	Most studies did not mention if variation across individuals' time intervals were present, or if data were analysed as time-unstructured or time-structured. From the 2 studies who considered the variance of time, details on consideration of time variance remains sparse; one study modelled time with both fixed and random effects. One study mention that the function of time was freed across groups.
	No	35	92	
	Unclear	1	3	
3a. Is the missing data mechanism reported?	Yes	8	21	Eight studies clearly identified missing data mechanisms. Five studies mentioned the causes of missing data, such as loss in follow-up (2 studies), censoring (1 study) or exclusion from dataset (2 studies). Most studies did not report the mechanism of missing data.
	No	25	66	
	Unclear	5	13	
3b. Is a description provided of what variables are related to attrition/missing data?	Yes	14	37	Fourteen studies presented a clear comparison between the characteristics of included/excluded individuals or complete/incomplete datasets. Two studies performed sensitivity analyses comparing the models with/without individuals with missing data. The other studies did not describe the variables with missing data.
	No	24	63	
3c. Is a description provided of how missing data in the analyses were dealt with?	Yes	20	53	Eighteen studies used exclusion of individuals with missing data from the dataset or exclusion of the last follow-up time-points. Two studies used full information maximum likelihood (FIML) estimation and one study used both exclusion and FIML. One study used multiple imputation and two studies explored the mechanism of missing data using model comparison. The other studies did not report how missing data was dealt with.
	No	18	47	
4. Is information about the distribution of the observed variables included?	Yes	14	37	From the studies who reported the distribution of observed variables, 13 reported a censored normal distribution and one study used a logit distribution.
	No	24	63	
5. Is the software mentioned?	Yes	38	100	All studies reported the statistical software used for trajectory modeling, 33 of which reported the version used. MPlus software was used in 9 studies, R was used in 5 studies (package lcmm reported in 3 studies); 16 studies used SAS (13 studies reported using the PROC TRAJ procedure), 7 studies used the "traj" plugin in STATA, 1 study used MLwin.
	No	0	0	
6a. Are alternative specifications of within-class heterogeneity considered (e.g., LCGA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	7	18	Six studies considered within-class heterogeneity using either LGMM or GGCMM. Two studies with unclear reporting added within-class confidence intervals on time intervals, although reporting using LCGA [43, 67] and 1 study mentioned considering within-class heterogeneity in text, although reporting using LCGA.
	No	28	74	
	Unclear	3	8	
6b. Were alternative specifications of the across-class variance-covariance matrix structure considered? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	0	0	Three studies mentioned the matrix structure considered; 2 studies assumed an auto-regression correlation and one study assumed a constant variance-covariance structure, without precisising the matrix chosen. No study considered alternative variance-covariance structure or justify the utilization of the chosen matrices. The other studies did not report information about across-class variance-covariance matrix structure. To note, LCGA assumes conditional independence of individuals at each point in time.
	No	38	100	
7. Are alternative shape/functional forms of the trajectories described?	Yes	19	50	From the 19 studies that reported considering alternative shapes of trajectories, linear, quadratic and cubic shapes where the shapes commonly evaluated for trajectory modeling. The shape of trajectories varied from linear functional forms (n = 7), quadratic (n = 4), cubic (n = 4) or mixed shapes (linear and/or quadratic and/or cubic) (n = 6). One study mentioned comparing alternative shapes in the method section, the results of which were not presented. From the remaining 18 studies, 9 studies only reported the trajectory shape(s) of the final solution, and nine studies did not report the final shape of the trajectories (although the shapes could sometimes be guessed from the graphs).
	No	18	47	
	Unclear	1	3	
8. If covariates have been used, can analyses still be replicated?	Yes	4	11	Thirty-one studies did not report using covariates for the identification of trajectory groups. Seven studies used covariates to predict the growth parameters/class membership. From these, one studies included both fixed and time-varying covariates (either in trajectory modeling or subsequent explanatory modeling), while 5 studies only considered fixed covariates. From our analysis, 4 studies gave sufficient details for replication.
	No	2	5	
	Unclear	1	3	
	No covariates used	31	82	

Table 3 (continued)

Items	Answer	N	%	Comments
9. Is information reported about the number of random start values and final iterations included?	Yes	1	3	One study reported rerunning the programs with multiple iterations of random start values. One study considered mentioned using a gridsearch function within the method in order to take in account convergence towards local maxima, yet did not mention iterations with random start value. The remaining studies did not report this information.
	No	37	97	
10. Are the model comparison (and selection) tools described from a statistical perspective?	Yes	35	92	From the 35 studies using model comparison tools, 33 used the BIC, either alone (21 studies) or in combination with other tools (13 studies). The other tools used were AIC, the Bayes factor, Vuong-lo-mendell-rubin likelihood ratio test, Nagin's criteria for adequacy along with other considerations such as posterior-probability, entropy, clinical relevance or interpretability. A total of 16 studies observed a steadily decrease in BIC, 7 studies observed some disagreement between the tools. Thirteen studies considered a minimal sample size threshold for each trajectory, going from >5% of the study population (5 studies); >1% (2 studies); >3% (1 study), >10 individuals per class (1 study). Three studies mentioned their intention to impose a minimum percentage of individuals, without precision regarding the threshold. From the 35 studies, 25 studies were lacking details on the interpretation of tool and how the decision of the number of trajectories was made, 15 of which did not report their decision.
	No	2	5	
	Insufficient	1	3	
11. Are the total number of fitted models reported, including a one-class solution?	Yes	16	42	A total of 22 studies reported the number of trajectories in models compared, from which 16 included a one-class solution. The maximum number of trajectories tested went from 3 to 8. Six studies used the one trajectory model for comparison.
	No one-class solution	6	16	
	No	16	42	
12. Are the number of cases per class reported for each model (absolute sample size, or proportion)?	Yes	7	18	Despite not reporting for all the models tested, all studies reported the number or the proportion of participants in each trajectory for the final model.
	No	31	82	
13. If classification of cases in a trajectory is the goal, is entropy or the number of misclassifications reported?	Yes	33	87	All studies had the goal of classifying individuals in specific trajectories, 7 studies reported calculating entropy of the chosen model and 27 studies reported average posterior probability of class membership, from which 14 studies set a minimal threshold going from >0.5 to >0.8. Five studies used entropy to choose between models.
	No	5	13	
14a. Is a plot included with the estimated mean trajectories of the final solution?	Yes	37	97	The majority of studies presented the trajectory groups for the final solution, while 4 studies presented graphically each model tested (usually presented in supplemental material). For studies using LGMM, confidence intervals or other dispersion measures were not presented. Individual trajectories were presented in one study.
	No	1	3	
14b. Are plots included with the estimated mean trajectories for each model?	Yes	4	11	
	No	34	89	
14c. Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?	Yes	1	3	
	No	37	97	
15. Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?	Yes	32	84	Most studies presented the baseline characteristics of the different trajectory groups identified.
	No	6	16	
16. Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?	Yes	1	3	No study made their syntax publicly available; one study reported the possibility of sharing syntax files by contacting the authors. In the supplemental materials, 3 studies gave additional details on their decisions for trajectory modeling or presented complementary methodological content.
	No	37	97	

GROLTS checklist from van de Schoot et al. (2016) [4].

AIC: Akaike information criterion, BIC: Bayesian information criterion, CI: Confidence intervals, FIML: Full information maximum likelihood, GBTM: group-based trajectory modeling, GGCMM: general Growth Class Mixture Modeling, LCGA: Latent Class Growth Analysis, LCGM: Latent Class Growth Modeling; LGMM: Latent Growth Mixture Modeling, SD: Standard deviation, SE: Standard error.

covariates to identify trajectory groups, from which a minority integrated time-varying covariates in the models.

Third, many studies included prevalent cases of T2D with differences in the duration of the disease. The non-consideration of T2D duration, for instance, might create some systematic distortions in the association between trajectory groups and a given outcome. Many strategies can be used for tackling this confounder, notably by restricting to incident cases, which was done by some researchers [16, 58, 65, 70]. Even though improving internal validity, restricting the selection of cases could affect generalizability of results. One study used propensity score matching [15], while other studies adjusted explanatory models with T2D duration or severity of T2D [37, 58, 59, 63, 67].

In sum, although using trajectory groups as exposure to an outcome was the most preferred approach in the included studies, one should recognize its complexity and should be aware of the limits inherent to its utilization.

4.3. Needs for improvement in the quality of reporting of latent growth modeling approaches

We identified many recurring issues in the quality of reporting using the GROLTS checklist. Among the most important issues noted stands the vast utilization of LCGA, without considering potential residual within-group heterogeneity, which could lead to important consequences on

the interpretability and inference of results. In accordance, we observed a certain confusion of the “within-group heterogeneity” concept; Some investigators mentioned using models such as LGMM without reporting how within-group heterogeneity was estimated, or on the opposite, mentioned information about within-group heterogeneity while using LCGA. In most cases, adequate reporting on the matter was lacking. Another important aspect in LGMM and LCGA is the utilization of tools/criteria to identify the optimal number of trajectory groups. The BIC was by far the most used tool for model identification, in accordance with what is usually recommended in the literature [4]. Other tools have been proven efficient to support the BIC, such as the Bayes Factor, the Lo-Mendell-Rubin Likelihood ratio test, the bootstrap likelihood ratio test, the Akaike information criterion and other subjective assessments such as clinical relevance, parsimony and interpretability [3, 81]. Although the criteria for model selection were globally well reported, we observed some reporting issues regarding how these tools were actually used for group identification. Indeed, information was often missing on the results of the tools between the models tested (statistical or subjective criterion), how inconsistencies were managed and how the final solution was identified.

Another element often missing in the included studies is the reporting of the number of different models tested, including the one-class model. The GRoLST-Checklist recommend testing the one-class model, which could fit the data best in comparison with multiple-classes models. Although adequation criteria (e.g. BIC) are usually used to compare models, these criteria could be misleading and favorize more complex models [72]. The final decision should thus combine adequation criteria with clinical plausibility of one-class versus multiple-classes models. Another strategy would be to visualize individual trajectories from raw data and evaluate if these individual patterns are concordant with the average trajectory identified [72]. This issue highlights all the complexity and subjectivity in the selection of the optimal trajectory class model and further supports the importance of rigorous reporting of the decision process within the scientific literature. We observed that two-third of the included studies used the average posterior probability as the favourite tool to assess model adequacy, while only 5 studies reported entropy. Usually, model adequacy is confirmed by ensuring all average posterior probability exceed a minimal threshold which is recommended at 0.7 [3]. In the included studies, however, 14 studies reported a clear threshold (between 0.5 and 0.8), while 13 studies mentioned considering average posterior probability without detailing how the tool was used. Finally, we observed a major issue with the treatment of missing data. Latent growth modeling approaches are renowned for their flexibility with irregular datasets, where it remains possible to identify trajectories even if some data are missing, but only if these data are missing at random or completely at random [3, 82]. However, a minority of studies clearly reported the nature of missing data [83].

4.4. The non-standardized terminology associated with latent growth modeling approaches

In a global perspective, we felt important to emphasize the great diversity in the terminology related to latent growth modeling approaches. During the process of this review, we also encountered the term “trajectory” in a wide variety of contexts, sometimes as statistical methods, as indicators of processes of care, as a description of a prospective phenomenon or sometimes purely arbitrary. This non-standardized terminology illustrates the importance of consistency in trajectory evaluation, in a way of clarifying and improving their use in future studies. Overall, we observed some issues relating to transparency and quality of report of latent growth modeling approaches, which unfortunately limit the comparability between studies [4]. This lack of information might reflect some confusion in the utilization of these methods, which further justifies the presence of experienced biostatisticians while planning latent growth modeling approaches, no matter the method used.

4.5. Strengths and limitations

This review has strengths, notably the utilization of a rigorous and objective methodology inspired by the Cochrane handbook for systematic reviews of interventions [84]. However, the review has limitations, notably the absence of two reviewers for abstracts and studies selections as well as data collection/extraction. Regarding the review process, we restricted to peer-reviewed published studies only. However, given the novelty of latent growth modeling approaches in the field of type 2 diabetes, some studies reported in conference abstracts or grey literature may have been missed. The review may thus be subject to publication bias. The generalizability of the findings is limited to T2D, even though similarities could be expected with other cardiometabolic/chronic diseases such as hypertension or dyslipidemia. Finally, we decided to exclude studies on type 1 diabetes due to the differences of its inherent pathophysiology and clinical management compared to T2D.

5. Conclusions

In conclusion, this literature review reports the relevance, the application and the issues related to latent growth modeling approaches in the study of T2D. We encourage future investigators interested into latent growth modeling approaches to consider the following suggestions: (1) to use caution when trajectory groups are treated as exposure in explanatory models, (2) to carefully choose the terminology used to characterize trajectory models in light of current literature, (3) to acknowledge the complexity behind latent growth modeling approaches by involving an experienced statistician in the process and finally, (4) to ensure that methods are reported transparently and comprehensively.

Declarations

Author contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interest's statement

The authors declare the following conflict of interests:

Sarah O'Connor had received a prize award (unrelated to this work) from the Faculty of Pharmacy, Université Laval, the funds were provided by Pfizer Canada (June 2021). Dr Poirier was a member of the Clinical Practice Guidelines of the Canadian Diabetes Association (macrovascular complications). Although not relevant to this work, Dr Paul Poirier declared having received fees for CME/consultants from Abbott, Amgen, Astrazeneca, Bayer, Bausch Health, Boehringer Ingelheim, Eli Lilly, HLS Therapeutics Inc, Janssen, Novartis, NovoNordisk, Sanofi and Servier. Jacinthe Leclerc is a professor of nursing at Université du Québec à Trois-Rivières. Within her role of professor, she provides (1) Continuous Medical Education sessions for health care professionals, accredited by the Fédération des médecins omnipraticiens du Québec and its local affiliates and (2) statistical expertise on Data Safety Monitoring Board Committees managed by JSS Medical Research (both unrelated to this work). Other authors declare no conflict of interests.

Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2022.e10493>.

Acknowledgements

We thank Chantal Beauregard, librarian at Université Laval, for her help in the search strategy design and development.

References

- [1] International Diabetes Federation, IDF Diabetes Atlas, ninth ed., 2019.
- [2] T. Jung, K.A.S. Wickrama, An introduction to latent class growth analysis and growth mixture modeling, *Soc Personal Psychol Compass* 2 (2008) 302–317.
- [3] D.S. Nagin, C.L. Odgers, Group-based trajectory modeling in clinical research, *Annu. Rev. Clin. Psychol.* 6 (2010) 109–138.
- [4] R. van de Schoot, M. Sijbrandij, S.D. Winter, S. Depaoli, J.K. Vermunt, The GROLTS-checklist: guidelines for reporting on latent trajectory studies, *Struct Equ Model Multidiscip J* 24 (2017) 451–467.
- [5] H.L. Nguena Nguéfac, M.G. Pagé, J. Katz, M. Choinière, A. Vanasse, M. Dorais, et al., Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches, *Clin. Epidemiol.* 12 (2020) 1205–1222.
- [6] S. Raghavan, W.G. Liu, S.A. Berkowitz, A.E. Barón, M.E. Plomondon, T.M. Maddox, et al., Association of glycemic control trajectory with short-term mortality in diabetes patients with high cardiovascular risk: a joint latent class modeling study, *J. Gen. Intern. Med.* 35 (2020) 2266–2273.
- [7] T.-H. Tsai, N. Huang, I.-F. Lin, Y.-J. Chou, Variation in the 11-year trajectories of medical care seeking behaviors in diabetes patients under a single payer system: persisting gaps to be filled, *BMC Health Serv. Res.* 19 (2019) 580.
- [8] J. Twisk, T. Hoekstra, Classifying developmental trajectories over time should be done with great caution: a comparison between methods, *J. Clin. Epidemiol.* 65 (2012) 1078–1087.
- [9] R. Jandoc, A.M. Burden, M. Mamdani, L.E. Lévesque, S.M. Cadarette, Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations, *J. Clin. Epidemiol.* 68 (2015) 950–956.
- [10] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *BMJ* 339 (2009) b2535.
- [11] S. Low, X. Zhang, J. Wang, L.Y. Yeoh, Y.L. Liu, S.F. Ang, et al., Impact of haemoglobin A1c trajectories on chronic kidney disease progression in type 2 diabetes, *Nephrology* 24 (2019) 1026–1032.
- [12] S.R. Whitworth, D.G. Bruce, S.E. Starkstein, W.A. Davis, T.M.E. Davis, T.C. Skinner, et al., Depression symptoms are persistent in Type 2 diabetes: risk factors and outcomes of 5-year depression trajectories using latent class growth analysis, *Diabet Med J Br Diabet Assoc* 34 (2017) 1108–1115.
- [13] S.R. Whitworth, D.G. Bruce, S.E. Starkstein, T.M.E. Davis, T.C. Skinner, W.A. Davis, et al., Risk factors and outcomes of anxiety symptom trajectories in type 2 diabetes: the Fremantle Diabetes Study Phase II, *Diabet. Med.* 37 (2020) 1688–1695.
- [14] D.F.L. Hertroijs, A.M.J. Elissen, M.C.G.J. Brouwers, N.C. Schaper, S. Köhler, M.C. Popa, et al., A risk score including body mass index, glycated haemoglobin and triglycerides predicts future glycaemic control in people with type 2 diabetes, *Diabetes Obes. Metabol.* 20 (2018) 681–688.
- [15] C.P. Wang, H.P. Hazuda, Better glycemic control is associated with maintenance of lower-extremity function over time in Mexican American and European American older adults with diabetes, *Diabetes Care* 34 (2011) 268–273.
- [16] N. Laiteerapong, A.J. Karter, H.H. Moffet, J.M. Cooper, R.D. Gibbons, J.Y. Liu, et al., Ten-year hemoglobin A1c trajectories and outcomes in type 2 diabetes mellitus: the Diabetes & Aging Study, *J. Diabet. Complicat.* 31 (2017) 94–100.
- [17] H.F. de Vries McClintock, K.H. Morales, D.S. Small, H.R. Bogner, Patterns of adherence to oral hypoglycemic agents and glucose control among primary care patients with type 2 diabetes, *Behav Med Wash DC* 42 (2016) 63–71.
- [18] H.F. de Vries McClintock, K.H. Morales, D.S. Small, H.R. Bogner, A brief adherence intervention that improved glycemic control: mediation by patterns of adherence, *J. Behav. Med.* 38 (2015) 39–47.
- [19] C.-J. Chiu, L.A. Wray, F.-H. Lu, E.A. Beverly, BMI change patterns and disability development of middle-aged adults with diabetes: a dual trajectory modeling approach, *J. Gen. Intern. Med.* 28 (2013) 1150–1156.
- [20] C. Proust-Lima, M. Sène, J.M.G. Taylor, H. Jacqmin-Gadda, Joint latent class models for longitudinal and time-to-event data: a review, *Stat. Methods Med. Res.* 23 (2014) 74–90.
- [21] B. Muthén, L.K. Muthén, Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes, *Alcohol Clin. Exp. Res.* 24 (2000) 882–891.
- [22] B. Muthén, C.H. Brown, K. Masyn, B. Jo, S.-T. Khoo, C.-C. Yang, et al., General growth mixture modeling for randomized preventive interventions, *Biostat Oxf Engl* 3 (2002) 459–475.
- [23] B. Muthén, in: G.A. Marcoulides, R.E. Schumacker (Eds.), *Latent Variable Mixture Modeling*, New Dev. Tech. Struct. Equ. Model, Lawrence Erlbaum Associates, 2001, pp. 1–33.
- [24] B. Muthén, K. Shedden, Finite mixture modeling with mixture outcomes using the EM algorithm, *Biometrics* 55 (1999) 463–469.
- [25] M.R. Elliott, J.J. Gallo, T.R. Ten Have, H.R. Bogner, I.R. Katz, Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction, *Biostat Oxf Engl* 6 (2005) 119–143.
- [26] B. Jo, B.O. Muthén, Modeling of Intervention Effects with Noncompliance: A Latent Variable Approach for Randomized Trials. New Dev. Tech. Struct. Equ. Model., Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, 2001, pp. 57–87.
- [27] B. Muthén, D. Kaplan, M. Hollis, On structural equation modeling with data that are not missing completely at random, *Psychometrika* 52 (1987) 431–462.
- [28] D. Tofghi, C.K. Enders, Identifying the correct number of classes in growth mixture models, *Inf. Age* (2007) 317–341.
- [29] C.-P. Wang, C. Hendricks Brown, K. Bandeen-Roche, Residual diagnostics for growth mixture models, *J. Am. Stat. Assoc.* 100 (2005) 1054–1076.
- [30] N. Ram, K.J. Grimm, Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups, *Int. J. Behav. Dev.* 33 (2009) 565–576.
- [31] E.A. Bayliss, P.J. Blatchford, S.R. Newcomer, J.F. Steiner, D.L. Fairclough, The effect of missed cancer, depression and pulmonary disease exacerbations on type 2 diabetes control, *J. Gen. Intern. Med.* 26 (2011) 575–581.
- [32] D.S. Nagin, Analyzing developmental trajectories: a semiparametric, group-based approach, *Psychol. Methods* 4 (1999) 139–157.
- [33] A. Bocquier, S. Cortaredona, L. Fressard, P. Loulégue, J. Raude, A. Sultan, et al., Trajectories of seasonal influenza vaccine uptake among French people with diabetes: a nationwide retrospective cohort study, 2006–2015, *BMC Publ. Health* 19 (2019). N.PAG-N.PAG.
- [34] B.L. Jones, D.S. Nagin, K. Roeder, A SAS procedure based on mixture models for estimating developmental trajectories, *Socio. Methods Res.* 29 (2001) 374–393.
- [35] B.L. Jones, D.S. Nagin, Advances in group-based trajectory modeling and an SAS procedure for estimating them, *Socio. Methods Res.* 35 (2007) 542–571.
- [36] A.M. Haviland, B.L. Jones, D.S. Nagin, Group-based trajectory modeling extended to account for nonrandom participant attrition, *Socio. Methods Res.* 40 (2011) 367–390.
- [37] C. Botvin Moshe, S. Haratz, R. Ravona-Springer, A. Heymann, L. Hung-Mo, M. Schnaider Beerli, et al., Long-term trajectories of BMI predict carotid stiffness and plaque volume in type 2 diabetes older adults: a cohort study, *Cardiovasc. Diabetol.* 19 (2020).
- [38] C.-C. Chen, S.-H. Cheng, Continuity of care and changes in medication adherence among patients with newly diagnosed diabetes, *Am. J. Manag. Care* 22 (2016) 136–142.
- [39] D.S. Nagin, *Group-based Modeling of Development*, Harvard University Press, Cambridge, Massachusetts, 2005.
- [40] D.S. Nagin, R.E. Tremblay, Analyzing developmental trajectories of distinct but related behaviors: a group-based method, *Psychol. Methods* 6 (2001) 18–34.
- [41] A.V. D'Unger, K.C. Land, P.L. McCall, D.S. Nagin, How many latent classes of delinquent/criminal careers? Results from mixed Poisson regression analyses, *Am. J. Sociol.* 103 (1998) 1593–1630.
- [42] C.-J. Chiu, Y.-H. Tseng, Y.-C. Hsu, S.-T. Wu, L.A. Wray, Depressive symptom trajectories in the first 10 years of diabetes diagnosis: antecedent factors and link with future disability in Taiwan, *Soc. Psychiatr. Psychiatr. Epidemiol.* 52 (2017) 829–836.
- [43] A.B. Cooke, E. Rahme, A.K. Defo, D. Chan, S.S. Daskalopoulou, K. Dasgupta, A trajectory analysis of daily step counts during a physician-delivered intervention, *J. Sci. Med. Sport* 23 (2020) 962–967.
- [44] T.M.E. Davis, S.A.P. Chubb, D.G. Bruce, W.A. Davis, Metabolic memory and all-cause death in community-based patients with type 2 diabetes: the Fremantle Diabetes Study, *Diabetes Obes. Metabol.* 18 (2016) 598–606.
- [45] B.L. Jones, D.S. Nagin, A note on a stata plugin for estimating group-based trajectory models, *Socio. Methods Res.* 42 (2013) 608–613.
- [46] K.S. Berlin, G.R. Parra, N.A. Williams, An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models, *J. Pediatr. Psychol.* 39 (2014) 188–203.
- [47] T.M.E. Davis, S.A.P. Chubb, W.A. Davis, The relationship between estimated glomerular filtration rate trajectory and all-cause mortality in type 2 diabetes: the Fremantle Diabetes Study, *Eur. J. Endocrinol.* 175 (2016) 273–285.
- [48] S.S. Deschênes, R.J. Burns, N. Schmitz, Trajectories of anxiety symptoms and associations with incident cardiovascular disease in adults with type 2 diabetes, *J. Psychosom. Res.* 104 (2018) 95–100.
- [49] G. Goh, N.C. Tan, R. Malhotra, U. Padmanabhan, S. Barbier, J.C. Allen, et al., Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting, *J. Med. Internet Res.* 17 (2015) e33.
- [50] H. Andruff, N. Carraro, A. Thompson, P. Gaudreau, B. Louvet, *Latent Class Growth Modelling*, A Tutorial, 2009.
- [51] C.-L. Lee, W.H.-H. Sheu, I.-T. Lee, S.-Y. Lin, W.-M. Liang, J.-S. Wang, et al., Trajectories of fasting plasma glucose variability and mortality in type 2 diabetes, *Diabetes Metab.* 44 (2018) 121–128.
- [52] A. Haviland, D.S. Nagin, P.R. Rosenbaum, Combining propensity score matching and group-based trajectory analysis in an observational study, *Psychol. Methods* 12 (2007) 247–267.
- [53] J.-C. Li, J. Tian, S.-L. Wu, Z.-J. Wang, X.-F. Zhang, D. Jia, et al., Effect of long-term systemic blood pressure trajectory on kidney damage in the diabetic population: a prospective study in a community-based Chinese cohort, *Chin. Med. J.* 131 (2018) 1199–1205.
- [54] D.S. Nagin, C.L. Odgers, Group-based trajectory modeling (nearly) two decades later, *J. Quant. Criminol.* 26 (2010) 445–453.

- [55] D.S. Nagin, Group-based trajectory modeling: an overview, *Ann. Nutr. Metab.* 65 (2014) 205–210.
- [56] C.Y. Li, C.L. Kuo, Y.H. Chang, C.L. Lu, S. Martini, W.H. Hou, Association between trajectory of severe hypoglycemia and dementia in patients with type 2 diabetes: a population-based study, *J. Epidemiol.* (2021).
- [57] C. Lipscombe, R.J. Burns, N. Schmitz, Exploring trajectories of diabetes distress in adults with type 2 diabetes; a latent class growth modeling approach, *J. Affect. Disord.* 188 (2015) 160–166.
- [58] W.-H. Lo-Ciganic, J.M. Donohue, B.L. Jones, S. Perera, J.M. Thorpe, C.T. Thorpe, et al., Trajectories of diabetes medication adherence and hospitalization risk: a retrospective cohort study in a large state medicaid program, *J. Gen. Intern. Med.* 31 (2016) 1052–1060.
- [59] M. Luo, W.Y. Lim, C.S. Tan, Y. Ning, K.S. Chia, R.M. van Dam, et al., Longitudinal trends in HbA1c and associations with comorbidity and all-cause mortality in Asian patients with type 2 diabetes: a cohort study, *Diabetes Res. Clin. Pract.* 133 (2017) 69–77.
- [60] C. Proust-Lima, V. Philipps, B. Liquet, Estimation of extended mixed models using latent classes and latent processes: the R package lmm, *J. Stat. Software* 78 (2015).
- [61] M. Luo, C.S. Tan, W.Y. Lim, K.S. Chia, W.E. Tang, E.S. Tai, et al., Association of diabetes treatment with long-term glycemic patterns in patients with type 2 diabetes mellitus: a prospective cohort study, *Diabetes Metab Res Rev* 35 (2019) e3122.
- [62] D. Niaz, C. Neczyk, S.H. Simpson, Depression and antecedent medication adherence in a cohort of new metformin users, *Diabet. Med.* 38 (2021).
- [63] M. Obura, J.W.J. Beulens, R. Slieker, A.D.M. Koopman, T. Hoekstra, G. Nijpels, et al., Post-load glucose subgroups and associated metabolic traits in individuals with type 2 diabetes: an IMI-DIRECT study, *PLoS One* 15 (2020), e0242360.
- [64] J. Sun, J.D. Herazo-Maya, P.L. Molyneaux, T.M. Maher, N. Kaminski, H. Zhao, Regularized latent class model for joint analysis of high-dimensional longitudinal biomarkers and a time-to-event outcome, *Biometrics* 75 (2019) 69–77.
- [65] W. Rathmann, A. Schwandt, J.M. Hermann, O. Kuss, M. Roden, K. Laubner, et al., Distinct trajectories of HbA1c in newly diagnosed Type 2 diabetes from the DPV registry using a longitudinal group-based modelling approach, *Diabet. Med.* 36 (2019) 1468–1477.
- [66] D.S. Nagin, B.L. Jones, V.L. Passos, R.E. Tremblay, Group-based multi-trajectory modeling, *Stat. Methods Med. Res.* 27 (2018) 2015–23.
- [67] N. Schmitz, G. Gariépy, K.J. Smith, A. Malla, R. Boyer, I. Strychar, et al., Trajectories of self-rated health in people with diabetes: associations with functioning in a prospective community sample, *PLoS One* 8 (2013).
- [68] G. Sidorenkov, J.F.M. van Boven, T. Hoekstra, G. Nijpels, K. Hoogenberg, P. Denig, HbA1c response after insulin initiation in patients with type 2 diabetes mellitus in real life practice: identifying distinct subgroups, *Diabetes Obes. Metabol.* 20 (2018) 1957, 64.
- [69] J.M. Franklin, W.H. Shrank, J. Pakes, G. Sanfélix-Gimeno, O.S. Matlin, T.A. Brennan, et al., Group-based trajectory models: a new approach to classifying and predicting long-term medication adherence, *Med. Care* 51 (2013) 789–796.
- [70] T.H. Tsai, N. Huang, I.F. Lin, Y.J. Chou, Severe diabetes complications among patients with diabetes with regular follow-up: does care setting matter? *J. Eval. Clin. Pract.* (2019).
- [71] D. Vistisen, G.S. Andersen, A. Hulman, F. Persson, P. Rossing, M.E. Jørgensen, Progressive decline in estimated glomerular filtration rate in patients with diabetes after moderate loss in kidney function even without albuminuria, *Diabetes Care* 42 (2019) 1886–1894.
- [72] H. Lennon, S. Kelly, M. Sperrin, I. Buchan, A.J. Cross, M. Leitzmann, et al., Framework to construct and interpret latent class trajectory modelling, *BMJ Open* 8 (2018), e020683.
- [73] C. Proust-Lima, L. Letenneur, H. Jacqmin-Gadda, A nonlinear latent class model for joint analysis of multivariate longitudinal data and a binary outcome, *Stat. Med.* 26 (2007) 2229–2245.
- [74] I. Walraven, M.R. Mast, T. Hoekstra, A.P.D. Jansen, A.A.W.A. van der Heijden, S.P. Rauh, et al., Distinct HbA1c trajectories in a type 2 diabetes cohort, *Acta Diabetol.* 52 (2015) 267–275.
- [75] I. Walraven, M.R. Mast, T. Hoekstra, A.P.D. Jansen, S.P. Rauh, F.R. Rutters, et al., Real-world evidence of suboptimal blood pressure control in patients with type 2 diabetes, *J. Hypertens.* 33 (2015) 2091–2098.
- [76] B. Muthén, The potential of growth mixture modeling, *Infant Child Dev.* 15 (2006) 623–625.
- [77] D.M. Blei, Build, compute, critique, repeat: data analysis with latent variable models, *Annu Rev Stat Its Appl* 1 (2014) 203–232.
- [78] R.-H. Wang, K.-C. Lin, H.-C. Hsu, Y.-J. Lee, S.-J. Shin, Determinants for quality of life trajectory patterns in patients with type 2 diabetes, *Qual. Life Res.* 28 (2019) 481–490.
- [79] H. Zavrelova, T. Hoekstra, M. Alssema, L.M.C. Welschen, G. Nijpels, A.C. Moll, et al., Progression and regression: distinct developmental patterns of diabetic retinopathy in patients with type 2 diabetes treated in the Diabetes Care System West-Friesland, The Netherlands, *Diabetes Care* 34 (2011) 867–872.
- [80] A. Diop, C. Sirois, J.R. Guertin, D. Talbot, Marginal structural models with latent class growth modeling of treatment trajectories, *Stat* (2021).
- [81] N.H. Shah, A.E. Hipwell, S.D. Stepp, C.-C.H. Chang, Measures of discrimination for latent group-based trajectory models, *J. Appl. Stat.* 42 (2015) 1–11.
- [82] V. Dupere, E. Lacourse, F. Vitaro, R. Tremblay, Méthodes d'analyse du changement fondés sur les trajectoires de développement individuel, *Bull. Méthodol. Sociol.* 95 (2007).
- [83] C. Genolini, B. Falissard, KML: k-means for longitudinal data, *Comput. Stat.* 25 (2010) 317–328.
- [84] J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, et al., *Cochrane Handbook for Systematic Reviews of Interventions*, 2021 version 6.2 (updated February 2021). Cochrane.