DOI: https://doi.org/10.4038/sljas.v26i2.8170

Sri Lankan Journal of Applied Statistics, Vol (26-2)

Latent Growth Curve Modeling in Health Outcomes Research: An Illustrative Study

T. Oamen*

Department of Clinical Pharmacy and Pharmacy Administration,
Obafemi Awolowo University, Ife, Osun state, Nigeria

*Corresponding author: oamentheo@yahoo.com

Received: 07th February, 2025/ Revised: 22nd May, 2025/ Published: 11th June, 2025

ABSTRACT

Health outcomes research assesses and evaluates the result of health interventions, care, and therapy. Health outcomes research studies often use longitudinal data or repeated measures of health outcomes from interventions as dynamic measures. However, the traditional regression methods of analyzing such data are inherently limited in scope and flexibility. Latent growth models provide the needed flexibility and robustness for longitudinal data analysis. To illustrate the rationale and methodology of latent growth models in the context of outcomes research, the author used simulated hypothetical data of 357 hypertensive patients with controlled blood pressure attending a clinic with blood pressure measurements collected over four time points or waves (T1, T2, T3, and T4). The growth model was developed using structural equation modeling with Analysis of Moment Structures (AMOS) software. The study assumed that blood pressure control followed a linear trend. The results showed that the simulated model provided an adequate fit to the data. At the intra-individual level, the mean blood pressure measure (intercept) at the starting point was significant. The rate of change (slope or growth parameter) was positive and significant, indicating improvement of blood pressure measures from one time point to another. The correlation between the intercept and slope parameters was not substantial, implying that the initial measure did not necessarily change over time (consistency). To explain possible inter-individual differences based on assumed significant variances of the error variances of intercept and slope parameters, the effect of a time-invariant variable-sex on the conditional model was explored. The use of latent growth curve modeling in health outcomes research can enrich the scope and quality of information obtained from longitudinal outcomes data. Reporting format,

interpretation, and further implications for outcomes research are provided in the paper.

Keywords: Growth factor, Longitudinal data, Methodology, Outcomes research, Patient outcomes data, Structural equation modeling

1 Introduction

Generally, health outcomes research focuses on the assessment of the results or outcomes of a particular health intervention, care practice, or service (Normand 2008). The results, outcomes or measures depict the attainment of desired goals such as lowered blood pressure, blood sugar control, lowered morbidity or mortality, improved quality of life, and improved mobility (Jefford et al. 2003). In other words, outcomes research explores the outcomes of healthcare interventions and practices to generate scientific proof or evidence to support administrative, clinical, and operational decision-making and delivery in healthcare settings (Ting et al. 2013). Other outcomes of interest in health outcomes research are the effectiveness of healthcare delivery in areas such as cost-effectiveness, burden of disease, reduced disability, patient-reported preferences, health-related quality of life, and health status assessments (Pan et al. 2008; Tonphu et al. 2022). Health outcomes research fosters an understanding of the improvement of intermediate clinical outcomes by providing robust data and information (Ting et al. 2013).

Health outcomes research is a relatively new scientific field that draws from clinical medicine, social science, statistics, epidemiology, improvement science, and implementation science, etc., to support the improvement of the quality of future interventions and actions (Willett and Sayer 1994; Krumholz 2011; Ting et al. 2013).

Patient outcomes data [POD] are often the product of data obtained in the observational or routine setting in daily healthcare practice (Normand 2008). POD from healthcare interventions [e.g., drug therapy, physiotherapy, medical procedures, effect of health policy change, etc.], provide a plethora of useful information that may potentially enhance or improve future interventions or instigate alternative actions and/or change of approach. POD is an important source of insight for the healthcare provider for several reasons, 1] it gives information about the key outcomes of treatment, 2] provides information on the outcomes from patients treating chronic or debilitating diseases e.g. diabetic, hypertensive, cancer, stroke patients. Finally, provide insights which support informed clinical and administrative decision making by healthcare providers and practitioners to obtain better future outcomes (Ting et al. 2013).

However, POD are recorded or documented over time [they are longitudinal, panel, or repeated measures data], hence the method of data analysis should be aligned to the data type to obtain the right or correct type of information. For example, a diabetic and hypertensive patient has his or her blood pressure and glucose levels checked and recorded for every clinic visit (e.g., 4 clinic visits over 4 months). Thus, available data has been collected over

time (that is, on every clinic visit). Usually, the norm for analysis is to treat POD as cross-sectional in which linear regression methods may not be appropriate to obtain robust and precise estimations of patient health outcomes. This is because researchers need to give answers to pertinent questions such as -Are the measurement of health status changing over time? Is there a trend, pattern or trajectory of the variable (health status, parameter, or indicator) over time? What is responsible for the change? Are the patterns due to intra-individual and/or inter-individual differences in the population? The above questions are relevant due to the fact that POD collected from the individual patient from a single visit is not adequate to furnish robust and critical clinical judgment or interventions based on data. From the data analytical perspective, repeated measures or longitudinal data may produce inaccurate results if subjected to traditional data analysis methods such as linear regression, analysis of variance-ANOVA, and analysis of covariance-ANCOVA (Duncan and Duncan 2009). This is because these approaches only treat changes in mean and do not account for measurement error and non-linear growth trajectory of the variable. Moreover, when using ANOVA or ANCOVA, researchers cannot determine if the model adequately fits the data (Sparks and Alamer 2024).

To resolve this gap in analysis of POD, latent growth curve models [LGCM) based on structural equation modeling (SEM) framework are used to analyze the pattern of change from longitudinal data for a variable, construct, or phenomenon of interest (Willett and Sayer 1994; Kline 2016: Collier 2020; Alamer and Alrabai 2023). SEM is an advanced multivariate technique that is used to develop, and examine causal and complex relationships between variables (Kline 2016: Hair et al. 2019; Collier 2020). SEM is very useful for model building and to test the validity of models and theories (Kline 2016: Collier 2020). Unlike cross-sectional data analysis, where data is collected at one point in time, LGCM is precise for exploring the dynamic nature of variables or phenomena which are subject to change over time and explores the individual variations or differences that may impact the rate of the change process (Preacher et al. 2008; Duncan et al. 2013). According to Preacher et al (2008), longitudinal data compared to cross-sectional, or two-wave data provides more statistical power for hypotheses testing.

Furthermore, LGCM makes provision for the use of covariates (time-invari–ant and time-varying variables) and other predictors to explain the reasons for the pattern of change (positive or negative, linear or non-linear). That is, when explaining inter-individual differences due to variability in mean of the intercept (average values) and slope (rate of change) in outcomes where they exist. Based on the initial example given, compared to conventional analysis of variance and regression models, LGCM additionally provides answers to the following critical questions namely- 1] Are the blood pressure level of the patients increasing or decreasing over time? 2] If yes, is the rate of change due to other factors (covariates, like sex, diet change, exercise, socioeconomic status, training, adherence to medication regimen etc.)? 3] Is there a relationship between the mean blood pressure and some other variables? 4] Is there a rela-

tionship or association between the rate of change in blood pressure and other variables? 5] Are males experiencing better blood pressure control compared to females?

Furthermore, LGCM provides rich and more expansive information relating to the developmental process of change in a construct or variable of interest by using four key parameters- Intercept (or mean) and slope. LGCM provides the mean variable (intercept) of the sample population at the initial stage or time and also designates the rate of change in mean using the slope also known as latent growth factor (Byrne and Crombie 2003; Zhang 2022; Hiver and Al-Hoorie 2019).

The variances of the intercept and slope provide further evidence of the presence of within (intra-individual) and between (interindividual) subjects' variation which provides an explanation using predictors or covariates to explore these differences (Curran et al. 2010; Hiver and Al-Hoorie 2019). This implies that it addresses inter-individual factors instigating changes in the latent variable of interest. The fourth parameter is the correlation or covariance between the intercept and the slope which depicts how the initial mean relates to the rate of change.

Additionally, a unique feature of the LGCM as an SEM statistical approach is that it assesses the fit of the model to the data using model fit indices, chi-square difference test-$\chi 2$/df, absolute fit indices-RMSEA (root mean square error of approximation) and SRMR (standardized root mean square residual), and comparative indices such as Tucker Lewis and Comparative fit indices, that is TLI and CFI respectively (Kline 2016: Hair et al. 2019; Sparks and Alamer 2024).

LGCM is an advanced statistical procedure that has extensive application in many fields: psychology (Curran et al. 2010), management and marketing (Oamen 2023), clinical research (Bengtsson et al. 2015), public health (Nemoto et al. 2021), education and language research (Zhang 2022; Sparks and Alamer 2024), outcomes research (Bengtsson et al 2015; Tonphu et al 2022). In health outcomes research, the use of LGCM is not commonplace apparently because outcomes researchers are not familiar with the basics, benefits, and application of the technique, hence limiting its substantial use in conducting investigative research on longitudinal, panel, or repeated measures data. This consequentially limits the quality of analysed information available for decision-making by practitioners. Furthermore, LGCM's limited use in health outcomes research is apparently because of an insufficient understanding of the concept. Generally, regular assessment of critical health outcomes such as glucose level, blood pressure, quality of life indicators, and blood cholesterol levels is essential for improving health outcomes. The POD obtained is an essential source of information to healthcare researchers and practitioners alike. For instance, for diabetic patients attending clinic for routine medical checkup (let's say 6 times in a year), blood glucose tests are carried out on patients to assess their state of health on each visit. A lot of information can be obtained when the LGCM technique is used to analyze

the data collected from each assessment time point. They include- information on the mean value of glucose control, trajectory of blood glucose control (linear or non-linear) over the assessment time points, and the potential effect of covariates (time variant and invariant variables) on mean and trajectory of blood glucose control. The information obtained from LGCM analysis at the individual, group, and population levels are useful for healthcare researchers, providers and managers to support quality healthcare decisions. To the best of the author's knowledge, there is no known study that has illustrated the dynamics of latent growth modeling in outcomes research.

Therefore, from a data analysis standpoint, this paper seeks to introduce the rationale, methodology, application, and interpretation of LGCM to improve the quality of analysis, insights, inference, and consequent impact on decision-making when applied to health outcomes research. The author hopes that this illustrative study will stimulate the use of growth curve models by researchers in outcomes research.

This paper is restricted to considering the basic single domain LGCM (the unconditional model), the conditional LGCM (includes covariates or predictors), and the parallel or multiple domain LGCM. Additionally, the author focused on simplifying the process of developing LGCM with minimal use of structural equations.

1.1 Latent growth curve models (LGCM)

A growth curve model is a statistical technique that estimates inter-individual variability (between-individual or between-person) based on intra-individual (within-individual or within-person) trajectory of change over time (Willet and Sayer 1994; Byrne and Crombie 2003). A basic latent growth curve model is composed of three main variables or parameters namely- the main latent variable or phenomenon of interest which is measured over several time points or waves (at least 3-T1, T2, T3 etc.), the Intercept (Mean), and Slope or latent growth factor (Howard and Curran 2014). The intercept or mean refers to the average value of the key variable at the initial time. It represents the aggregate mean of the phenomenon of interest at the beginning of the study across all time points; hence, it is fixed to 1 across all time points (Collier 2020). The slope refers to the rate of change (which could be increasing or decreasing) of the key variable over each time point (Byrne and Crombie 2003).

It is noteworthy that the variances of the intercept and slope parameters also provide valuable insights into the individual differences that may or may not exist in the sample population. In other words, if the variance of intercept is significant, it suggests that there are individual differences in the mean value of the variable of interest. The same interpretation applies to statistically significant variance of slope parameter which may suggest that the rate of change is attributable to individual differences of the sample. This often warrants the use of time-invariant covariates (e.g., gender, race, and class) and time-varying covariates (e.g. perception of respondents at each time point, measurement of quality of life at each time point etc.,) variables to explain

the inter-individual differences. Time-invariant covariates are usually baseline measurements available to the individual which explain variation in baseline values and growth whereas time-varying covariates tend to be measured at each time point and operationally function as control variables measured at each time point (Byrne and Crombie 2003). Generally, LGCM describes or elaborates on intra-individual change (unconditional models) and inter-individual change (conditional) of a latent variable in a sample population (Byrne and Crombie 2003).

The basic LGCM (or single-domain growth model) is a univariate growth model that involves estimating the mean and rate of change over time of a single variable of interest. This basic model is derived from the SEM framework (see Figure 1). The basic model without the addition of a predictor or covariate is referred to as an Unconditional LGCM. However, when a covariate or predictor is included in the model, its intercept and slope values are contingent on the effects of the covariates and hence called Conditional LGCM (see Figure 2). Researchers evaluate or assess Conditional LGCM, in cases when the variances of the slope and intercept in the unconditional model are significant, indicating the presence of inter-individual differences or variations which can be explained using covariate/s-time-invariant and time-varying (see Figure 3). Additionally, conditional LGCM is also developed to test a hypothesis or confirm the theory of interest to the researcher by using the slope and intercept to predict another outcome of interest and/or use a predictor variable to examine its influence on the slope and intercept parameters of the LGCM (Willett and Sayer 1994; Byrne and Crombie 2003; Duncan et al. 2013; Sparks and Alamer 2024). Therefore, researchers in outcome research can innovatively apply conditional LGCM analysis to test hypotheses or research questions of interest.

To model individual differences over time, the basic LGCM for an individual, group, or population is represented by the structural equation below:

$$y_{it} = \eta_{0i} + \eta_{1i}\lambda_t + \varepsilon_{it}$$

Where y=Outcome measure, variable or phenomenon of interest (y1, y2, y3, y4 are the outcomes measured or assessed at time 0, 1, 2, and 3 respectively), η0=intercept value, η1=slope value (amount of vertical increase for each unit of horizontal run of the growth curve), t=ith value of time of measurement, λ=predictor variable at 0, 1, 2, 3 and 4 time points, ε=time specific error of prediction for each time point, i=value of time [from start time t=0, to 1, 2, 3]. Hence, a positive value of η0 indicates a positive or high mean value of the outcome variable at time 1 (start time equals 0; the paths of the intercept to the outcome variable in the growth model were constrained to 1 to obtain the average value for all participants or subjects). A positive value of the slope factor indicates an increasing value of the outcome variable over time. In addition, a positive covariance between η0 and η1 indicates that the initial mean value of the outcome variable tends to improve or increase over

time. A negative covariance between η0 and η1 suggests that the initial mean value of the outcome variable tends to reduce or lowered over each time point (Byrne and Crombie 2003: Duncan and Duncan 2009).

Parallel LGCM (multiple-domain growth models) are growth models that explore the developmental growth trajectories involving the simultaneous estimation of two or more phenomena or variables of interest (Muthen and Curran 1997; Byrne and Crombie 2003; Nemoto et al. 2021; Tonphu et al. 2022). Tonphu et al (2022) used a parallel LGCM to explore the relationship between metabolic syndrome disease and several health indicators based on data obtained over 6 years. Parallel LGCM are multivariate growth models that simultaneously estimate the rate of change in multiple variables over time (Preacher et al. 2008). For example, the researcher may be interested in assessing the relationship between the intercept (mean value) of blood pressure control to the intercept (mean value or initial level) of blood sugar control of a sample of diabetic hypertensive patients over time-T1, T2, T3, and T4 (see Figure 4). Similarly, he or she may desire to evaluate if the rate of change (slope parameter) in blood pressure is positively or negatively correlated to the rate of change in blood sugar levels (see Figure 4). It is noteworthy that these relationships can also be predictive, in which case regression lines connect the intercept and/or slope of blood pressure control to the intercept and/or slope of blood sugar control (see Figure 5). In this case, it answers the question of whether the mean value of blood pressure in the patients predicts or influences the mean values of blood sugar control in the sample. In addition, the initial mean value of the blood pressure can be correlated to the rate of change in blood glucose. Conversely, the correlation of initial mean value of blood glucose to the rate of change in blood pressure can be estimated.

To further illustrate the above parallel LGCM, another scenario may involve more than 2 or more variables- the LGCM of medication cost is used to predict the rate of change in LGCMs of blood pressure control and blood sugar control. Another lucid example is the influence of LGCM related to health to predict the rate of change in LGCM for blood pressure control and blood sugar control (the assumption here is that poor quality of life related to health negatively predicts blood pressure and glucose control in patients).

1.2 Statistical Software Used for LGCM

A range of structural equation modeling software is available for conducting latent growth curve analysis to varying degrees of complexity. Some of them include- Linear structural relations (LISREL®, Jöreskog and Sorbom 2006), Analysis of moment structures (AMOS®, Arbuckle 2016), Mplus® (Muthen and Muthen 2007), and JASP® (JASP Team 2024). In this illustrative study, AMOS version 24 was used.

1.3 Notations Used for LGCM

There are known conventions for building models using SEM, CIRCLES

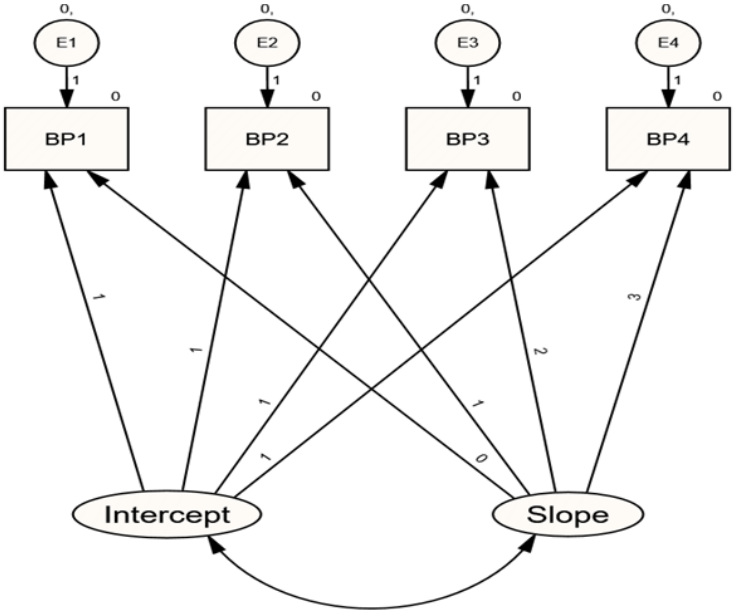
or OBLIQUE CIRCLES represent Unobserved or latent variables, RECTANGLES represent observed or manifest variables, SINGLE-HEADED ARROWS depict regression paths [impact of a variable on another], and DOUBLE-HEADED ARROWS represent covariance or correlations between two variables.

1.4 Modeling Intra-individual and inter-individual changes in LGCM

In LGCM, to achieve the objectives of the trajectory of change, modeling intra- and inter-individual change is critical. Intra-individual change, or within-person or level 1 change, typically refers to the LGC measurement model (see Figure 1). A measurement model typically shows the relationship or links between the unobserved or latent variable (intercept and slope) and the observed indicators (phenomenon or variable of research interest) measuring the latent variable. The strength of the relationship between the unobserved and the observed variables depends on the value of the regression coefficients. In essence, the intra-individual growth curve utilizes the covariance structure in which all indicator scores are computed as deviations from their means, hence, the means play no role in the specification of the intra-individual change or within-person or level 1 model. In the case of inter-individual change or between-person or level 2 changes, the analysis of mean values is associated with the slope and intercept factors. In this model, the difference in means is used to estimate the differences between persons. Based on the frame of SEM, the inter-individual aspect of the growth model, which estimates differences in means and slopes of the variable of interest, represents the Structural Model of the LGCM (see Figure 1).

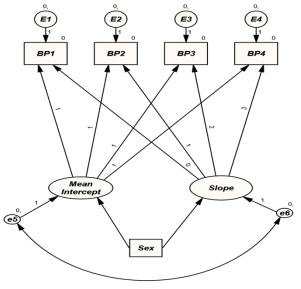
Figure 1 shows the basic LGCM with a 2-factor structure composed of two latent variables namely the intercept (referring to the average or mean blood pressure score of all the patients at the starting time, and the slope (or growth or stability factor usually fixed to 1 across) which shows rate of change in the mean values over each time point. Blood pressure measurement values (observed variables) are each time point (T1, T2, T3, T4) are depicted as BP1, BP2, BP3, and BP4, respectively. The error terms ε1, ε2, ε3, and ε4 are measurement errors (variances) that are not accounted for by the observed variables.

In Figure 2, the presence of statistically significant variances of the slope and mean intercept indicates the possible presence of inter-individual differences, which should be accounted for in the model [the illustrative model did not have significant variances of intercept and slope]. Therefore, the above explanation is for the sake of illustration only. Thus, a time-invariant covariate (also known as an individual-level predictor) is incorporated as an exogenous variable to explore the differences between individuals. In the illustrated model, the time-invariant predictor used was sex, which was operationalized as male (coded 1) or female (coded 0).



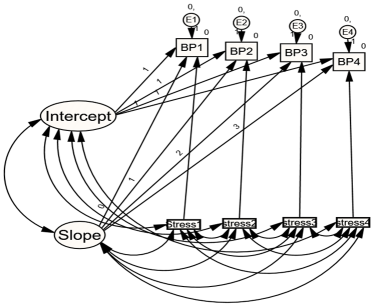
H

Fig. 1: Linear latent growth curve model (Unconditional)



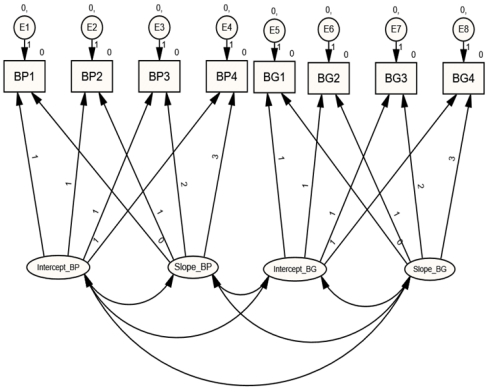
H

Fig. 2: Conditional Latent growth curve model with time-invariant covariate (Sex)



H

Fig. 3: Conditional Latent growth curve model with time-varying covariate-perception of stress

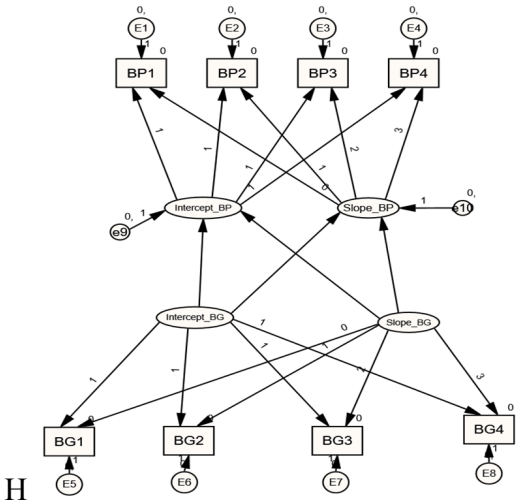


H

Fig. 4: Parallel latent growth curve model showing the correlated relationship between Blood pressure control and Blood glucose growth curves

In Figure 3, the time-varying covariate stress measured at times T1, T2, T3, and T4 was related to each of the observed (outcome) variables. Stress acts as a form of control variable such that the time-related trajectory of change [mean and slope values] of the phenomenon or variable of interest is adjusted accordingly. It is worth noting that the intercept, slope, and the covariate variable were correlated due to the fact that they are exogenous variables predicting the main dependent (outcome) variable.

Figure 4 shows the relationship between the growth curves of blood pressure (BP) and blood glucose (BG) control typifying a multiple-domain model. Here, another independent growth curve was developed using the measures of blood glucose control from the same patients over the same 4 time points. The goal is to draw parallels or connections between the two measures. The slope of both curves gives information on whether the rate of change (increasing or decreasing] of BP is associated with an increase or decrease in the rate of change in BP, 3] The association between the mean value of BP and the rate of change in BG, and/or vice versa gives information on the association between the rate of change in the slope of BP) and the mean value of BG.



H

Fig. 5: Parallel latent growth curve model with Blood glucose as the Predctor

In Figure 5, the LGCM of blood glucose control shows BG (as predictor or independent variable) used to predict the LGCM of blood pressure control, BP (Outcome or dependent variable). In other words, the LGCM of multiple domains statistically explores to what extent the mean value of the BG control directly influences or impacts the mean value of BP control (direct effect of $Slope_{BG}$ on $Intercept_{BP}$); for example, a negative regression coefficient implies that the initial level of BG control of the patient's control negatively impacts the control of the average blood pressure BP control of the patients over time.

The same analysis applies to the impact of the slope of BG on the mean value of (the direct effect of $Slope_BG$ on $Intercept_BP$). A negative value implies that the rate of change in blood glucose (positive or negative) tends to have a negative influence on the BP control values of the participants over time. In other words, a practical implication is that patients who have a good glucose control rate tend to have poor blood pressure control over time. Also, assuming a positive direct relationship between the $Slope_{BG}$ on $Intercept_{BP}$ shows that the impact of average blood glucose control on the rate of change in BP. The relationship suggests that because the average blood glucose control of the patients tends to positively influence the rate at which blood pressure levels is controlled over time. In addition, the direct effect of $Slope_{BG}$ on $Intercept_{BP}$ relationship depicts the impact of rate of change in blood glucose control on the rate of change in the blood pressure of patients over time.

A summary of the main differences between LGCM and traditional data analysis methods is presented below in Table 1.

Table 1: Comparison between LGCM and Traditional Regression Methods

Parameters	**LGCM**	**Traditional Methods**	**References**
1. Examples	Linear growth model (or unconditional), conditional, parallel, quadratic, etc.	Linear regression, multiple regression, Analysis of Variance, Analysis of Covariance, etc.	Byrne and Crombie 2003; Zhang 2022
2. Type of Data	Longitudinal	Mostly cross sectional	Duncan et al 2013
3. Model fit	Uses fit indices within SEM framework	Does not use fit indices	Sparks and Alamer 2024
4. Measurement error	Considers it as relevant	Not considered. Taken as a nuisance	Sparks and Alamer 2024
5. Growth Trajectory	Linear and nonlinear	Linear	Duncan and Duncan 2009
6. Statistical power	Higher	Lower	Muthen and Curran 1997; Preacher et al 2008; Curran et al 2010
7. Flexibility	Model can be re-specified or modified. Also, make provision for inclusion of predictors of change-time invariant, and time varying	Less flexible to re-specification or modification	Duncan et al 1999; Duncan and Duncan, 2009; Sparks and Alamer 2024
8. Difficulty level	Knowledge of SEM is required	knowledge of regression method	Sparks and Alamer 2024

2 Methodology

2.1 Study assumptions

The illustrative study had the following assumptions, namely: 1. The main outcome measure was <130/90mmHg in systolic and/or diastolic in line with the Joint National Commission on hypertension 8 guideline (Fuchs and Whelton 2020; Oamen et al 2022). The mean difference between actual blood pressure readings and cutoff values was used as an absolute number 2. Data is normally distributed. Although the structural equation modeling tool (AMOS) uses the maximum likelihood method that accounts for the non-normality of data 3. Assumes that blood pressure control follows a linear trajectory, trend, or growth pattern [occurs at a constant rate] 4. Assumed that all patients fully complied with the physician's prescription and adhered to usage 5. Patients with adequate blood pressure control were purposively selected for this illustrative study

2.2 Statistical Procedure

In the hypothesized model in Figure 1, the paths from the intercept to the observed variables at the different times were constrained to 1 because it is considered that the mean or intercept value remains constant for each individual over time (Duncan et al. 1999; Byrne and Crombie 2003). On the other hand, the slope parameter paths are assigned 0, 1, 2, and 3 depending on the number of time points. This frees up the paths for estimation of the growth trajectory or trend in the model. According to Duncan and Duncan (2009), researchers can obtain precise estimates of a growth trajectory when the number of observations or time points of outcome assessments increases. Hence, a two-wave (data collected over 2 time points) design is inadequate to define change over time because it tends to define research questions very poorly (Rogosa et al. 1982; Duncan and Duncan 2009). According to Boomsma and Hoogland (2001) and Byrne and Crombie (2003), the recommended minimum sample size for a latent growth curve study involving SEM analysis should be 200 or more. Our simulated sample was 357. The analysis was conducted using Analysis of Moment Structures Software (AMOS version 24; Arbuckle 2016). Initial analysis of data should include an estimation of the normality of the data to help identify outliers before testing the specified LCGM. Model fit estimation is also required, such as global fit index- $\chi 2$, but is no longer used because of sensitivity to increase in sample size (Jöreskog and Sörbom 1993). Chi-square difference test-$\chi 2$/df (values less than 3 to 5 is acceptable), comparative fit index (<0.90), Tucker Lewis Index (TLI>0.90), Root mean error of approximation RMSEA (<0.08 but with a limit of 0.1) and standardized root mean squared residual SRMR (<0.05) based on the recommendations of Hu and Bentler (1999) and Kock (2022).

3 Results and Discussion

3.1 Presentation of Results

Generally, researchers can apply the reporting format used by Byrne and Crombie (2003), Duncan et al. (2013), and Sparks and Alamer (2024). First, the model fit attributes of the LGCM were presented. In the hypothetical model, the fit statistics were $\chi 2/df = 4.469$, RMSEA = 0.099, with a 90% confidence interval of [0.059, 0.142], CFI = 1.00, and SRMR = 0.032, all of which are within acceptable ranges. Hence, the inference is that the estimated growth model has an adequate fit to the data.

Table 2: Descriptive Statistics and Correlation Coefficients of Observed Variables

Parameters	**BP1**	**BP2**	**BP3**	**BP4**
Mean (M)	8.34	11.42	16.49	20.00
Standard Deviation (SD)	2.26	2.32	2.81	3.04
Skewness	0.19	0.056	-0.48	-0.036
Kurtosis	-1.23	-1.27	-1.15	-1.12
BP1	1			
BP2	-0.081	1		
BP3	-0.004	0.01	1	
BP4	0.036	0.04	0.05	1

NOTE. Significance if present is indicated with asterisks; *p<0.05, **p<0.01, ***p<0.001

Thereafter, the descriptive statistics [mean, standard deviation, skewness, and kurtosis] and the correlation of the observed variables should be presented as shown in Table 2. Note that descriptive statistics of the demographic characteristics should also be reported to give readers a sense of the sample characteristics. The normality of the data was evaluated based on the absolute values between ±2 and ±7 for skewness and kurtosis, respectively, as recommended by Kline (2011) and Hair et al. (2019). The obtained values for the outcome measure-blood pressure were between -0.036 to 0.19 for skewness, and between -1.12 to -1.27 for kurtosis. Therefore, confirms that the data is normally distributed and adequate for structural equation modeling. However, the SEM software (AMOS) used for this study adopted the maximum likelihood estimator, which is robust for data that is not normally distributed.

Then, an assessment of the estimates for the key parameters of the developed model was made. As presented in Table 3, for the single-domain LGCM, the path related to the Intercept was positive and significant ($p<0.001$, $t>1.96$) with an average or mean value of 7.92 at time T1 for the study participants. The slope value was positive and statistically significant with a mean value of 3.75 ($p<0.001$, $t>1.96$). This implies that blood pressure improved from the initial mean blood pressure by an average of 3.75 from T1 to T2, from T2 to T3, and from T3 to T4. In other words, the results suggest that patients' blood pressure improved by 11.67 units on the second clinic visit, improved

by 15.42 units by the third clinic visit, and improved by 19.17 units by the fourth clinic visit. For healthcare researchers, this analysis provides important information to support the improvement of health outcomes and remedial actions when undesirable trends are observed.

Table 3: Estimates of LGCM (Unconditional Model)

Parameter	Estimates	t-value	p-value	Inference
Intercept (I)				
Mean	7.92	85.97	0.001***	significant
Variance	0.01	-0.02	0.987	not significant
Slope (G)				
Mean	3.75	72.64	0.001***	significant
Variance	0.03	0.27	0.787	not significant
Correlation				
I—->G	-0.06	-0.334	0.738	not significant

Note. Significance levels: *p<0.05. **p<0.01, ***p<0.001

Next, the variances of the intercept and slope are examined for statistical significance. In this illustration, the variances of the mean and slope parameters were not statistically significant (Intercept variance=0.01, p>0.05; slope variance=0.03, p>0.05). This indicates the absence of differences between persons or inter-individual differences. Therefore, there is no need to further investigate the possible predictors or covariates that can explain the disparities. The correlations were insignificant, implying that the mean blood pressure of the participants at the beginning was not necessarily affected by the rate of changes in the blood pressure over time. In this sense, if the covariance or correlation between the intercept and slope was found to be significant with a negative value, it suggests that patients with initial high blood pressure control tend to reduce over time. If the value is positive, it means that patients with initial high blood pressure tend to maintain or achieve positive control over time.

Generally, these variances (of intercept and slope) by design represent deviations from the average value of the intercept and slope. If significant, it gives ample indication of possible inter-individual differences in both the initial blood pressure scores at Time 1, and their change over time as study participants underwent clinical evaluation from assessment Time 1 to Time 4. Therefore, justifies the inclusion of predictor variables or covariates in follow-up (subsequent) analysis to explain the reasons for variation.

Note. Significance levels: *p<0.05. **p<0.01, p<0.001

In this illustrative study, the variances of the intercept and slope were not significant (p>0.05), hence implying the possible absence of differences between individuals. From an interpretational standpoint, such results present at least two possible explanations. first, it substantiates prior analysis presented in Table 3, which gave non-significant error variances of the intercept and slope parameters. Secondly, if the error variances were significant in the ini-

Table 4: Estimates of LGCM (Conditional Model)

Parameter	Values	t-value	p-value	Inference
Intercept (I)				
Effect of Sex	0.02	0.132	0.895	not significant
Slope (G)				
Effect of Sex	0.04	0.212	0.832	not significant
Correlation				
e5—>e6	0.439	2.05	0.01**	significant

tial unconditional model (in a case where there is no significant effect of the covariate), it means that the variations between individuals in the study population may be attributable to some other factors other than not gender.

However, for the sake of illustration, the intercept and slope variances were assumed to be significant; then, it is required to include a covariate (time-invariant or time-varying) to proffer explanations for the inter-individual differences [this is addressed by the Conditional LGCM in Figure 2].

Addressing the conditional LGCM depicted in Figure 2 and Table 4, sex was incorporated as a time-invariant predictor of the possible inter-individual differences in the intercept and slope parameters. Gender was coded as 0 for females and 1 for males. It is important to include an assessment of the model fit characteristics of the conditional model. If the model characteristics improve or worsen, an explanation is required. The illustrative conditional model showed that sex was not a predictor of inter-individual differences, which more or less substantiates the initial results of the unconditional LGCM. In other words, since the error variances of the intercept and slope parameters, it means that the variations between individuals in the study population may be attributable to some other factors, and not gender. For a health outcomes researcher, this would trigger an investigation to search for the plausible factor responsible for variations in health outcomes.

However, let's explore the interpretation of results in a case where the covariate (predictor) has a significant effect on the intercept and slope parameters. Therefore, if gender was found to be a positive and significant predictor of the intercept parameter (that is, the initial value) and the slope. This means that the average blood pressure control was higher in males than in females at time 1 (T1). However, regarding the initial blood pressure at T1, the rate of change (blood pressure control) was higher in males than in females based on the positive and significant value of the slope parameter or growth factor. Conversely, the interpretation of the slope is different if, for instance, the slope parameter was negative; it implies that females had a higher rate of blood control over time compared to males (since the time-invariant covariate is a binary predictor). By extension, the interpretation of the results will be different if the predictor is a continuous variable. In the same vein, in our illustration of the time-varying covariate [stress level measurement at T1, T2,

T3, and T4] in Figure 3, was added to the growth curve for each observed variable, respectively.

For the parallel LGCM (see Figure 4), the same process of reporting is followed, starting with good-of-fit statistics. However, the important result in this multiple-domain growth model is the covariances between the intercept of BP with the intercept for BG control and, similarly, the covariance between the slopes of BP with that of BG model/s. Also examined was the relationship between the intercept of BP with the slope of BG. Also, the association between the intercept of BP with the slope of the BG model, and/or the correlation between the slope of the BP model with the mean of BG. For more details on the reporting and interpretation of multiple domain or parallel LGCM, see or refer to Byrne and Crombie (2003), Duncan and Duncan (2009), Curran et al (2010), and Collier (2020).

This paper builds on the growing literature of LGCM in the social and behavioral sciences, which is now gaining traction in the health sciences. Data from healthcare interventions, programs, and activities are often presented as repeated measures or longitudinal data forms, which require robust and precise evaluations of trends and / or developmental trajectories of change over time. Researchers in outcomes research can apply SEM framework to develop an LGCM that helps them identify dynamic developmental trajectory in POD and to examine possible intra-individual (within-person or measurement model, or level 1, inter-individual (between-person or structural model or level 2), and differences when evaluating the growth trajectory of a variable or phenomenon of interest.

The paper presented the basic or unconditional, conditional, and parallel models of growth curve analysis using an explanatory approach to buttress the robust results and inferences that can be obtained using growth models. It is necessary to posit that these results cannot be readily obtained from the traditional methods of analysis, such as analysis of variance (ANOVA), linear and multiple regression (Duncan and Duncan 2009; Sparks and Alamer 2024). The relevance of covariates or predictors was used in enriching our understanding of the factors that can influence the trajectory of change in the model. A key advantage of using LGCM is that we can examine if the model fits the data adequately to warrant further analysis or re-specification of the model. Another unique feature is that innovative approaches can be applied to LGCM to answer hypotheses and research questions based on the theoretical leanings of the outcomes researcher (Duncan et al. 1999).

This illustrative paper provides a rationale for applying LGCM in outcomes research for longitudinal data, offering step-by-step guidance on how to develop and apply LGCM. It also explores the role of covariates (time-invariant and time-varying) in enhancing the researcher's understanding of the dynamic change process and individual differences within a target study population over time.

3.2 Implications of LGCM use for Outcomes research

This conceptual paper provides a working template to enable health outcomes researchers to apply basic and complex LGCM in analyzing longitudinal or panel data available to them. The flexibility inherent in LGCM ensures that the outcome researcher is well positioned to ask critical questions about patient outcomes data and develop growth curve models to answer them. Therefore, the measurement and structural models developed using LGCM elaborate on the dynamic nature of the explored changes in POD, which gives clarity to the researcher for better inference and robust decision making. A standout outcome of LGCM use in outcomes research is that it affords the researcher to explore and understand the trajectory of change in health outcome/s over time at the individual, group, and population level. The development of change is adequately captured using the LGCM framework to The implications of using latent growth modeling in outcomes research are numerous which include; a] outcomes researchers are equipped with the rudiments of latent growth curve modeling to improve the quality of analysis of empirical data, b] evidence provided from the analysis will justify measures taken by the healthcare management team to optimize benefits for patients, c] information obtained provides evidence of short, medium, and long term effectiveness of therapy or intervention programs on patients. The detailed analysis of repeated measures data ensures the optimal tracking of patient clinical performance outcomes taking into account the impact of time-invariant and time-varying predictors, e] the application of latent growth models will enable researchers to probe the effects of time-invariant and time-varying factors that influence treatment outcomes, f] latent growth models can be applied to support more complex relationships thus ensuring robust analysis, and finally, g] this study improves the researcher's understanding of the analytical possibilities inherent and available in latent growth curve modeling especially in outcomes research.

3.3 Limitations of the study

Although hypothetical data simulating real-life situations was used to illustrate the application of latent growth models, the study was used for illustrative purposes only. This illustrative study was limited to simple unconditional and conditional growth models. Therefore, several other variants of latent growth models can be explored in future research studies depending on the nature of the data, the phenomenon of interest, and the context of the study. These considerations inform the model specification adequate to answer the research questions and the tests of hypotheses (Muniz-Terrera et al. 2017). Finally, the paper assumes that the growth path or trajectory is linear for the paper. Therefore, readers should be informed that non-linear quadratic growth trajectories also occur or exist in LGCM (Grimm and Ram 2009; Duncan and Duncan 2009). Quadratic LGCM involves the inclusion of a third latent variable known as the Quadratic factor to account for the non-linearity or curvilinear nature of the growth trajectory. Researchers can apply this when there is evidence of the presence of non-linearity in the growth pattern.

4 Conclusion

This simplified paper outlined the relevance of LGCM to outcome research. LGCM is useful for researchers and practitioners to use observational panel and longitudinal patient outcomes data [POD] to develop models that assess the effects and trajectory of change of interventions on a health phenomenon over time at the individual, group, and population levels. Although not exhaustive, the author provided various research scenarios warranting the application of latent growth models that may best address modeling and hypothesis testing in health outcomes research. Furthermore, it supports extending the analysis to include other covariates (or predictors) that may account for changes and effects based on the theoretical basis of the study. The study addressed the need for researchers in outcome research to take advantage of the benefits of using latent growth curve modeling in the analysis of longitudinal and repeated measures data. Thereby, the researchers can avoid the causal limitations inherent in the results of analysis by using or over-relying on cross-sectional data and conventional linear regression methods. This is pertinent especially when the objective of the researcher is to explore the trajectory or pattern of change concerning a specific measure or phenomenon of interest. Finally, the study espoused the need for researchers to use this approach using the framework of structural equation modeling, so that research outcomes can be robust, precise enough to support decision making.

Acknowledgement: The author appreciates advice given by Dr. Banjo Lawal Moshood, which greatly improved the quality of the paper

References

Alamer, A., amp; Alrabai, F. (2023). The causal relationship between learner motivation and language achievement: New dynamic perspective. Applied Linguistics, 44, 148–168. Doi: 10.1093/applin/amac035

Arbuckle, J. L. (2016). IBM ® SPSS TM AMOS TM User's Guide. Amos Development Corporation.

Bengtsson, U., Kjellgren, K., Hallberg, I., Lindwall, M., amp; Taft, C. (2015) Improved blood pressure control using an interactive mobile phone support system. The Journal of Clinical Hypertension, 18(2), 101-108.

Boomsma, A., amp; Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, amp; D. Sörbom (Eds.), Structural equation modeling: A festschrift in honor of Karl Jöreskög (pp. 139–168). Lincolnwood, IL: Scientific Software.

Byrne, B. M., amp; Crombie, G. (2003). Modeling and testing change: An introduction to the latent growth curve model. Understanding Statistics, 2(3), 177-203. Lawrence Erlbaum Associates, Inc.

Collier, J. (2020). Applied structural equation modeling using AMOS: Basic to advanced techniques. New York: Routledge.

Curran, P. J., Obeidat, K., amp; Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. Journal of Cognitive Development, 11, 121–136 (2010). Doi:10.1080/15248371003699969

Duncan, T. E., amp; Duncan, S. C. (2009). The ABCs of LGM: An introductory guide to latent variable growth curve modeling. Social Personal Psychological Compass, 3, 979-991, Doi:10.111/j.1751-9004.2009.00224.x

Duncan, T. E., Duncan, S. C., amp; Strycker, L. A. (2013). An introduction to latent variable growth curve modeling: Concepts, Issues, and Application (2nd Eds.). NY, Routledge, Doi:10.4324/9780203879962.

Duncan, T. E., Duncan, S. C., Stryker, L. A., Li, F., amp; Alpert, A. (1999). An introduction to latent variable growth curve modeling. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fuchs, F. D., amp; Whelton, P. K. (2020). High blood pressure and cardiovascular disease. Hypertension, 75, 285-292.

Grimm, K. J., amp; Ram, N. (2009). Nonlinear growth models in Mplus and SAS. Structural Equation Modeling, 16(4), 676-701.

Hair, J. F., Risher, J. J., Sarstedt, M., amp; Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2-24

Hiver, P., amp; Al-Hoorie, A. H. (2019). Research Methods for Complexity Theory in Applied Linguistics. Bristol, UK: Multilingual Matters.

Howard, A. L., amp; Curran, P. J. (2014).Latent growth curve modeling. In: Michalos, A. C. (eds) Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht. Doi:10.1007/978-94-007-0753-5-1605

Hu, L., amp; Bentler. P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6. 1–55

JASP Team. (2024). JASP (Version 0.19.l3) [Computer Software]

Jefford, M., Stockler, M. R., amp; Tattersall, M. H. (2003). Outcomes research: what is it and why does it matter, Internal Medicine Journal, 33, 110-118

Jöreskog, K. G., amp; Sorbom, D. (2006). LISREL 8,54 and PRELIS 2.54. Chicago, IL: Scientific Software

Jöreskog, K., amp; Sörbom, D.(1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Chicago: Scientific Software

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). Guilford Press

Kline, R. B. (2016). Principles and practice of structural equation modeling (5th ed.). Guilford Press

Kock, N. (2022). WarpPLS User Manual (Latest version 8.0, 2022). https://www.scroiptwarp.com/warppls/#User_Manual

Krumholz, H. M. (2011). The real-world imperative of outcomes research. JAMA, 306(7), 754-755

Muniz-Terrera, G., Robitaille, K. A., Johansson, B., Hofer, S., amp; Piccinin, A. (2017). Latent growth models matched to research questions to answer questions about dynamics of change in multiple processes. Journal of Clinical Epidemiology, 82, 158-166. Doi: 10.1016/j.jclinepi.201.6.09.001

Muthen, L.K., amp; Muthen, B. O. (2007). Mplus User's Guide (Sixth edition). Los Angeles, CA.

Muthen, B. O amp; Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: a latent variable framework for analysis and power estimation. Psychological Methods, 2, 371-402

Nemoto, Y., Sakurai, R., Matsunaga, H. et al. (2021). Social Contact with Family and Non-Family Members Differentially Affects Physical Activity: A Parallel Latent Growth Curve Modeling Approach. International Journal of Environmental. Research amp; Public Health, 18, 2313.
Doi: 10.3390/ijerph18052313

Normand, S-L (2008). Some old and some new statistical tools for outcomes research. Circulation, 118(8), 872-884

Oamen, T. E., Osemene, K. P., amp; Ihekoronye, M. R. (2022). Economic analysis of hypertension management in public hospitals in southwestern Nigeria: evidence from secondary data analysis. African Journal of Health Economics, 11(2), 1-11

Oamen, T. E (2023). A longitudinal survey of job satisfaction among pharmaceutical executives: an unconditional latent growth curve study. Texila International Journal of Management, 9(2), 26-32

Pan, J. H., Song, X. Y., Lee, S. Y., et al. (2008). Longitudinal analysis of quality of life for stroke survivors using latent curve models. Stroke. 2795-2801. Doi: 10.1161/STROKEAHA.108.515460

Preacher, K. J., Wichman, A. L., Briggs, N. E., et al. (2008). Latent Growth Curve Modeling. Thousand Oaks, CA: Sage

Rogosa, D. R., Brandt, D., amp; Zimowski, M. (1982). A growth curve approach to the measure of change. Psychological Bulletin. 92,726–748 (1982)

Sparks, R., amp; Alamer. A. (2024) Language anxiety does not affect the growth of L2 reading achievement: The latent growth curve model approach. Applied Psycholinguistics, 45(3). 467–488

Ting, H. H., Xiang, M-X., amp; Wang, J-A. (2013). Outcomes research: science and action. Journal of Zhejiang University-Science B

Tonphu, S., Tanaprakob, S., Piyaraj, P., et al. (2022). Relationship between health behavior and metabolic progression: The parallel latent growth curve model. Bangkok Medical Journal, 18(1), 1-11

Willett, J. B., amp; Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. Psychological Bulletin, 116, 363–381(1994)

Zhang, F. (2022). Latent growth curve modeling for the investigation of emotional factors in L2 in longitudinal studies: A conceptual review. Frontiers in Psychology. 13:1005223. Doi: 10.3389/fpsyg.2022.1005223