2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

Incidence and risks for hypoglycemic episodes in stable type 2 diabetic patients

Promotor :
Prof.dr. Christel FAES

Mohammed A. Ibrahim

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt | Maastricht University

2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

Incidence and risks for hypoglycemic episodes in stable type 2 diabetic patients<u></u>

Promotor :
Prof.dr. Christel FAES

Mohammed A. Ibrahim
Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

universiteit hasselt | Maastricht University

Certification

This is to certify that this report was written by Mohammed A. Ibrahim under my supervision.

Mohammed A. Ibrahim	Prof. Dr. Christel FAES
Student	*Supervisor*
.....	.....
Signature	Signature
.....	.....
Date	Date

Acknowledgment

Next to almighty God for his unlimited blessing in my entire career, my special thanks goes to my supervisor Prof. Dr. Christel FAES for her constructive comments and remarks, valuable guidance and directives, and enlightening discussions. The completion of this thesis could not have been possible to be appearing in its current form without support of several people and organizations. Firstly, my sincere gratitude goes to the Belgian Government through its Vlaamse InterUniversitaire Raad (VLIR) scholarship program, for supporting me finantially and giving me an opprtunity to study my MSc in Biostatistics from talented and well experiencd professors at University of Hasselt. I would like to thank all my professors for passing on their knowledge. Their names deserve to be highly recognized in this humble acknowledgment.

Secondly, my gratitude extends to my friends and classmates for sharing ideas, who in one way or another extended their help with open arms. Special thanks to all my friends who wish my success, especially Abduljewad Wele, Yimer Wasihun, and Belay Belete for their invaluable comments and sharing their life experience. Thanks to my group-mates in my second year master program, Lazaro Mwandigha, Jedelyn Cabrieto, and Kevin Watjou. Thanks to peaceful Belgian people who make me live with them my whole two years like in my own country, starting from Mr. Luc Van Herck.

Finally, I would like to express my deepest gratitude to my parents for their unconditional love, pray and unending support. Thanks to my lovely brothers, sisters, and all my relatives.

Mohammed

September 2014

Genk, Belgium

Contents

List of Tables

List of Figures

List of Abbreviations

BB	Beta-Binomial model
BMI	Body Mass Index
CI	Confidence Intervals
CM	Combined overdispersed and normal random effects Model
COMMM	Combined Overdispersed and Marginalized Multilevel Model
CPG	Capillary Blood Glucose
GLMM	Generalized Linear Mixed Model
LRT	Likelihood Ratio Test
MAR	Missing At Random
MMM	Marginalized Multilevel Model
OAD	Oral Antidiabetic Drugs/agents
OR	Odds Ratios
SL	Simple Logistic model
SMBG	Self Monitoring of Blood Glucose

Abstract

Hypoglycemia and hyperglycemia are serious health problems for those with diabetes. Hypoglycemia appears when blood glucose level is below normal, whereas hyperglycemia occurs for blood glucose level above normal. A prospective observational study was conducted of 52 patients with type 2 diabetes mellitus who were on insulin and metabolically stable. They were followed for 12 weeks, performing two daily capillary blood glucose measurements before and 2 hours after breakfast, lunch, and dinner in turn. A profile for 6 capillary blood glucose measurements performed before and after meals was also prepared every four days. The aim of this study was to investigate the optimal frequency and timing of self monitoring of blood glucose to detect hypoglycemia and hyperglycemia. In addition, it was of interest to investigate the clinical characteristics associated with episodes of hypoglycemia or hyperglycemia.

The simple logistic model, beta-binomial model, logistic-normal model, and combined model were fitted to the data and compared. Several covariates were included in the model: moment, age, insulin, oral antidiabetic agents, body mass index (BMI), HbA1c level, increment in HbA1c level from start to end of study (HbA1c_{inc}), time from diagnosis, and gender. The combined model did a better job, taking into account both overdispersion and correlation of observations within a patient. Moment and BMI with their interaction were found significantly associated with the risk of hypoglycemia. For hyperglycemia, the moment, HbA1c, HbA1c_{inc}, and interaction of moment with both HbA1c and HbA1c_{inc} were found significant.

In general, the highest risk of hypoglycemia was found before breakfast and before lunch, with a decreasing risk over BMI. On the other hand, measurements at pre-dinner had the highest risk of hyperglycemia, with an increasing risk over HbA1c or HbA1c_{inc}.

Keywords: Self monitoring of blood glucose; Type 2 diabetes mellitus; Hypoglycemia; Hyperglycemia; Bernoulli model; Beta-binomial model; Logistic-normal model.

1 Introduction

Diabetes mellitus is a chronic illness that requires continuing medical care and ongoing patient self-management education and support to prevent acute complications and to reduce the risk of long-term complications. The prevalence of diabetes for all age-groups worldwide was estimated to be 2.8% in 2000 and 4.4% in 2030 [13]. The most important demographic change to diabetes prevalence across the world appears to be the increase in the proportion of people > 65 years of age. The overall prevalence of diabetes mellitus in Spain is about 13.8% [14], where the prevalence of diabetes and impaired glucose regulation increased significantly with age, and was higher in men than in women. Diabetes care is complex and requires multifactorial risk reduction strategies beyond glycemic control. A large body of evidence exists that supports a range of interventions to improve diabetes outcomes.

Hypoglycemia and hyperglycemia are serious health problems for those with diabetes. Hypoglycemia appears when blood glucose level is below normal and the body tries to compensate (by increasing blood glucose level via increasing the glucagon hormone and gluconeogenesis, forming glucose), but somehow that mechanism is failed to increase the blood glucose level to normal again. Ross *et al.* [11] reported a reduced risk of hypoglycemia for patients with age > 65 years or with a body mass index (BMI) > 25 kg/m^2 and an increased risk with previous history of microvascular disease, hypoglycemia, increased number of oral hypoglycemic agents (OHAs), OHA intensification, or use of glinides. Further, Seufert *et al.* [12] investigated hypoglycemic episodes in patients with poorly controlled type-2 diabetes treated with a single daily bolus of insulin on top of insulin glargine and oral antidiabetic drugs (Basal-plus regimen). They found an increased risk of hypoglycemia for patients with HbA1c $< 7.0\%$, females, or patients with diabetes duration > 10 years, and a reduced risk for patients with higher BMI, or treated with basal bolus.

Hyperglycemia develops when there is too much sugar or glucose in the blood. When a person with diabetes has hyperglycemia frequently or for long periods of time, damage to nerves, blood vessels, and other body organs can occur. Hyperglycemia can also lead to more serious conditions, including ketoacidosis – mostly in people with type 1 diabetes – and hyperglycemic hyperosmolar nonketotic syndrome (HHNS) in people with type 2 diabetes or in people at risk for type 2 diabetes [16]. The most common precipitating factor in the development of hyperosmolar hyperglycemia is infection. Other precipitating factors include cerebrovascular accident, alcohol abuse, pancreatitis, myocardial

infarction, trauma, and drugs. Elderly individuals with new-onset diabetes (particularly residents of chronic care facilities) or individuals with known diabetes who become hyperglycemic and are unaware of it or are unable to take fluids when necessary are at risk for hyperosmolar hyperglycemic state (HHS) [15]. Diabetic patients with very high blood concentrations of glucose have from 2 to 3 times more HbA1c than normal individuals [17].

An expert panel of diabetes specialists that met in Tampa, FL, March 28 – 29, 2012 [3], selected 70 – 180 mg/dL as the default target range of capillary plasma glucose (CPG). They selected < 70 mg/dL as the criterion for "reportable hypoglycemia", and > 180 mg/dL as a criterion for hyperglycemia. This was consistent with the 2013 American Diabetes Association (ADA) standards of medical care [14], which state in the summary of glycemic recommendations that the preprandial capillary plasma glucose and peak postprandial capillary plasma glucose goals are 70 – 130 mg/dL and 70 – 180 mg/dL, with room for individualization, respectively.

Self monitoring of blood glucose (SMBG) provides real time information that helps patients to manage their diabete. The purpose of SMBG is to facilitate appropriate clinical interventions to achieve and maintain blood glucose within an acceptable range by detecting hypoglycemia and hyperglycemia, thus improving metabolic control and patient satisfaction [7]. International guidelines recommend the frequency and timing to be determined by the needs and goals of each patient. Hence, it is not uniform, and more frequent monitoring may be required for poorly controlled patients than those with stable blood glucose level. In this report, the risk of hypoglycemia and hyperglycemia in well-controlled diabetes will be investigated, and differences interms of timing during the day will be looked at. In addition, the clinical characteristics which are associated with episodes of hypoglycemia or hyperglycemia will be examined.

This paper is organized as follows. In Section 2, the dataset is described and explored. Statistical methodologies used in this study is the topic of Section 3. Results are discussed in Section 4. Finally, Section 5 presents the discussion of the findings, conclusions, and recommendations.

2 Case Study

2.1 Data Description

A prospective observational study of risk factor control, in stable-insulin-treated patients with type 2 diabetes, was conducted. Patients were eligible for this study if they met the following criteria: a) older than 40 years, b) body mass index (BMI) < 40 kg/m^2, c) type 2 diabetes mellitus of at least 1 year duration, d) on insulin therapy for at least 3 months before recruitment, e) ability to perform self-monitoring, and f) stable metabolic situation, defined as having no need to add new treatments or make changes in insulin dose $> 10\%$. The patients were recruited when they came to their routine check at the University Hospital of Santiago de Compostela, Spain.

Subjects were instructed to measure their blood sugars before and after breakfast, lunch, and dinner, for a period of 12 weeks, alternating throughout the week following the scheme as presented in Table A1 (Day 1: pre-breakfast and 2 hours post-breakfast, Day 2: pre-lunch and 2 hours post-lunch, Day 3: pre-dinner and 2 hours post-dinner, and Day 4: pre- and post-prandial 6-point profile). Blood glucose levels were recorded using a blood glucose meter (Accu-Check Complete, Roche Diagnostics, Montclair, NJ) with the capacity for storing 1,000 readings with the dates and times of testing. Plasma HbA1c was determined at the start of the study and end of week 12.

According to previously introduced reasoning in Section 1, a binary variable for hypoglycemia and hyperglycemia were created. This was done to simplify the information in the blood glucose levels, and further inhance a simple interpretation for hypoglycemia as well as hyperglycemia. It was defined as follows:

$$Hypoglycemia_{ij} = \begin{cases} 1 & \text{if CBG}_{ij} < 70 \text{ mg/dL} \\ 0 & \text{otherwise} \end{cases}$$

$$Hyperglycemia_{ij} = \begin{cases} 1 & \text{if Moment}_{ij} = \text{Pre-prandial and CBG}_{ij} > 130 \text{ mg/dL} \\ 1 & \text{if Moment}_{ij} = \text{Post-prandial and CBG}_{ij} > 180 \text{ mg/dL} \\ 0 & \text{otherwise} \end{cases}$$

where CBG_{ij} is Capillary Blood Glucose of patient i at measurement j.

Table 1: *Covariates and their descriptions*

Variable	Label.	Name	Description
ID			Patient identification number
Moment	Mom1.	Pre-breakfast	Timing in a day when blood
	Mom2.	Post-breakfast	sugar measured
	Mom3.	Pre-lunch	
	Mom4.	Post-lunch	
	Mom5.	Pre-dinner	
	Mom6.	Post-dinner	
Sex		Men	Gender of the patient
		Women	
Insulin	Treat1.	Basal insulin	Insulin treatment
	Treat2.	Mixtures of standard insulins	
		(Basal + Quick) (30/70, 50/50)	
	Treat3.	Basal insulin + 3 rapid insulin	
OAD	OAD1.	Insulin alone	Oral Antidiabetic Drugs
	OAD2.	Insulin + oral monotherapy	
	OAD3.	Insulin + OAD combination	
HbA1c			Glycated haemoglobin level at baseline
HbA1c_2			Glycated haemoglobin level at end-of-study
HbA1c_{inc}			The increment in HbA1c from start to
			end of study ($\text{HbA1c}_2 - \text{HbA1c}$)
BMI			Body mass index of the patient in kg/m^2
Age			Age of the patient in years
Timediag			Duration of diabetes in years

Glycated haemoglobin level -(percentage of haemoglobin which has been glycated)

Several baseline variables, as well as the moment at which CBG measurements taken, were considered as potential covariates in this study. Table 1 gives a decription of the covariates in the dataset.

2.2 Exploratory Data Analysis (EDA)

The dataset contained measurements for 52 patients. Table 2 shows the distribution of the patients over insulin and oral antidiabetic drugs (OAD) by sex. There were 21 men and 31 women in the study. In general, the majority of patients or 20 (38.5%) have taken an insulin "basal insulin + 3 rapid insulin". Among those, 11 were women. Considering OAD, 21 (40.4%) of patients have taken "Insulin + oral monotherapy".

As observed in Figure A1, of the 144 capillary blood glucose measurements scheduled for each patient

within the 12 weeks, not all patients have performed all the measurements. The median number of measurements taken was 141 with a range of 70 – 144. Only 15(28.8%) of the patients had a complete profile.

Figure A2 shows the distribution of HbA1c at baseline. The median HbA1c level was 7.5 with a range of 5.6 – 11.3. Further, the distribution of BMI is presented in Figure A3. The median BMI was 31.25 kg/m^2 with a range of 23.0 – 39.1 kg/m^2. Furthermore, the distribution for Age, HbA1c$_{inc}$, and Timediag can also be consulted in Figures A4 to A6. The median age for the patients was 61 years with a range of 49 – 77 years. The median HbA1c$_{inc}$ was -0.09 with a range of -1.6 – 0.7. Finally, the median time from diagnosis was 11 years with a range 1 – 28 years (see also Table 3).

Figure 1 and 2 show the proportion of hypoglycemia and hyperglycemia by patient, respectively. For hypoglycemia, many patients have proportion of zero. For hyperglycemia, a larger variation of the observed proportions between patients was observed. The distribution for proportion of hypoglycemia and hyperglycemia per each patient by moment are shown in Figure 3 and 4, respectively. On both hypoglycemia and hyperglycemia, there are fluctuations in the proportions over the moments. For hypoglycemia, the higher proportions are observed before lunch (moment 3), whereas for hyperglycemia, the proportions were higher for all pre-prandial measurements (moment 1, 3 and 5).

Table 2: *Distribution of patients over insulin and OAD by sex*

		Sex		
Treatment type	Treatment levels	Men 21 (40.4%)	Women 31 (59.6%)	Total
Insulin	Basal	7 (13.5%)	10 (19.2%)	17 (32.7%)
	Basal+Quick	5 (9.6%)	10 (19.2%)	15 (28.8%)
	Basal+3rapid	9 (17.3%)	11 (21.2%)	20 (38.5%)
OAD	Insulin alone	6 (11.5%)	12 (23.1%)	18 (34.6%)
	Insulin+oral	10 (19.2%)	11 (21.2%)	21 (40.4%)
	Insulin+OAD comb	5 (9.6%)	8 (15.4%)	13 (25.0%)

Table 3: *Distribution of the continuous covariates*

Covariate	Mean	Median	Standard deviation	Range
Timediag	11.48	11.00	7.33	1.0 – 28.0
Age	61.60	61.00	7.52	49.0 – 77.0
BMI	31.70	31.25	3.99	23.0 – 39.1
HbA1c	7.66	7.50	1.39	5.6 – 11.3
HbA1c_{inc}	-0.09	0.00	0.46	-1.6 – 0.7

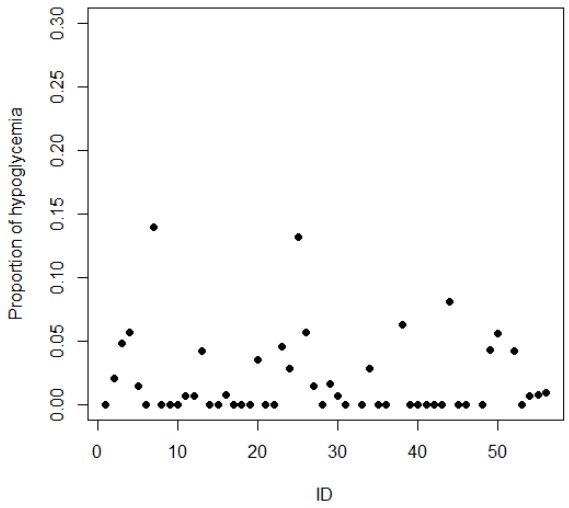

Figure 1: *Proportion of Hypoglycemia by Patient*

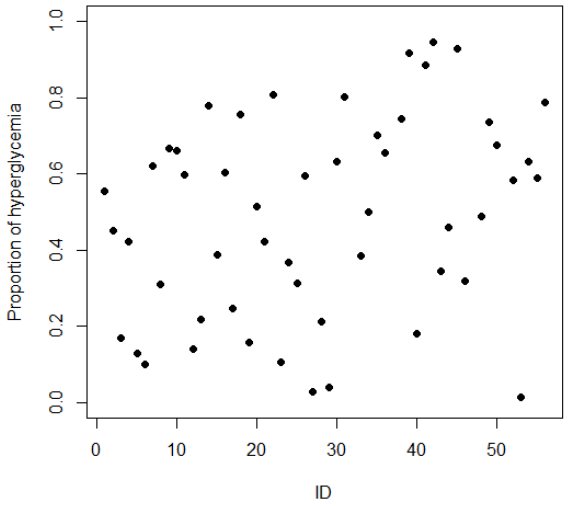

Figure 2: *Proportion of Hyperglycemia by Patient*

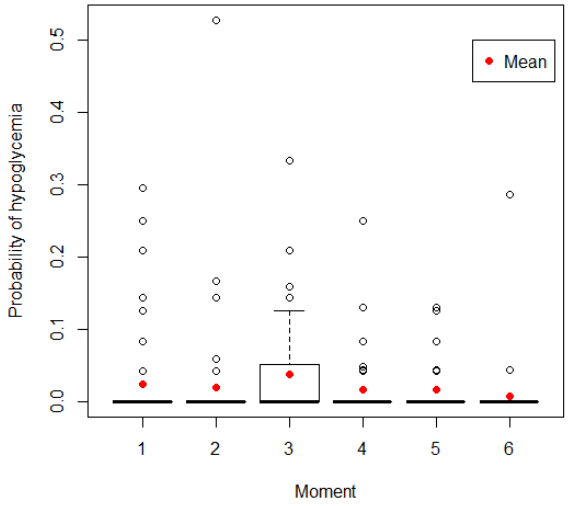

Figure 3: *Proportion of Hypoglycemia by Moment*

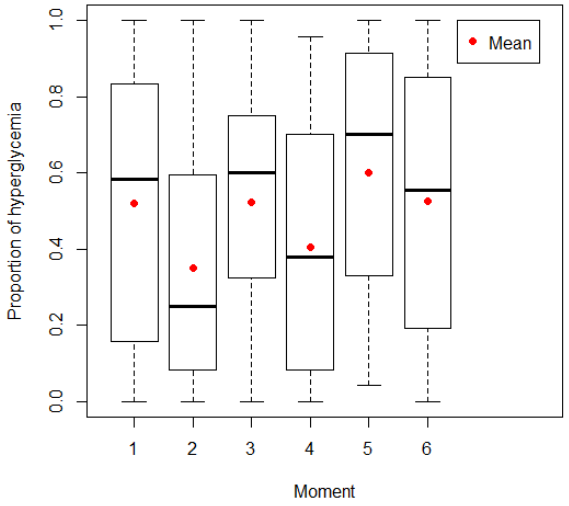

Figure 4: *Proportion of Hyperglycemia by Moment*

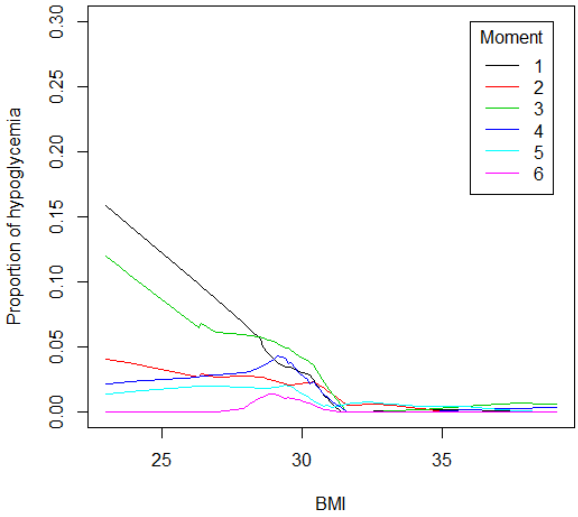

Figure 5: *Proportion of Hypoglycemia by Moment and BMI*

Further, it was of interest to investigate the relationship between the proportions of hypoglycemia and different covariates. In order to visualize the relationship of the continuous covariates with hypoglycemia or hyperglycemia, smoothed proportions were used. As observed in Figure 5, there is a decreasing risk with increasing BMI. Figure A7 shows that there is no clear difference on the risk of hypoglycemia for males versus females. Further exploration shows that, no clear difference was observed in the risk of hypoglycemia between the levels of Oral antidiabetic drugs (Figure A8). Furthermore, there is no much deviation between the treatment levels (Figure A9). The comparison of proportions by other covariates (Age, Timediag, HbA1c, and $HbA1c_{inc}$) were also performed and presented in Figures A10 to A13. No visible trend for the aformensioned covariates were observed.

Similar exploration was conducted for the outcome hyperglycemia. As can be observed from Figure 6 and 7, there is an increasing trend over the HbA1c levels and HbA1c_{inc}. For age, and time from diagnosis, there is a slight increasing trend (Figure A14 and A15, respectively). Further, no clear trend for the risk of hyperglycemia for covariate BMI was observed (Figure A16). It seems male patients has higher risk of hyperglycemia compared to their female counterparts (Figure A17). Further exploration shows that, patients treated with the mixture of standard insulins (Ins2) tend to have higher risk of hyperglycemia (Figure A18). Furthermore, patients treated with combination of insulin and oral monotherapy (OAD2) has lower risk (Figure A19).

The above discussed views are just graphical observations. A formal assessement will be discussed

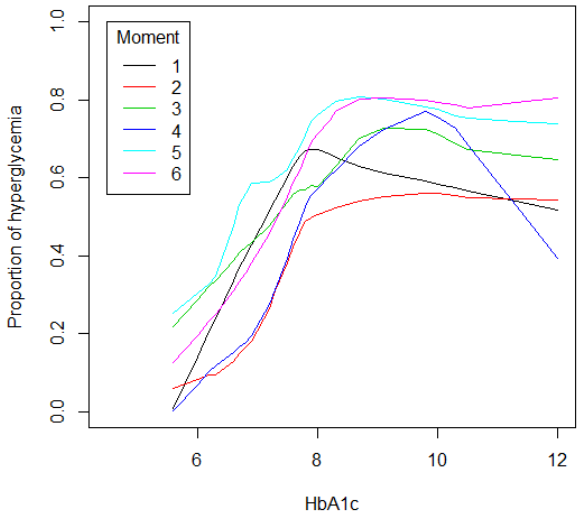

Figure 6: *Proportion of Hyperglycemia by Moment and HbA1c*

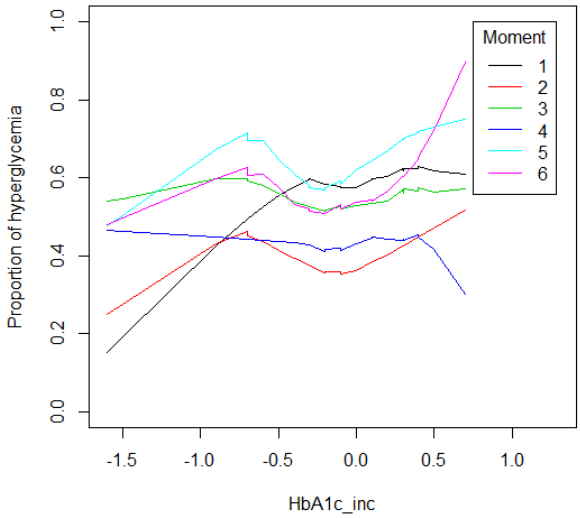

Figure 7: *Proportion of Hyperglycemia by Moment and* $HbA1c_{inc}$

in the following sections.

3 Methodology

This study deals with binary correlated data, where repeated measurements collected for a patient over the six moments for 12 weeks. From the previous section, it was indicated that there is a variation of the observed proportions between patients (Figure 1 and 2). Four different types of models will be considered to analyze this data: (1) The obvious choice for fitting binary data, ignoring both overdisperssion and correlation, is Simple Logistic model (**SL**); (2) Generalized Linear Mixed Model (**GLMM**) are commoly used models to take the variation into account [8]; (3) The other thing one should consider is overdispersion, which may occur due to various reasons, like omission of important covariates, subject heterogeneity, misspecification of link function, and other data complexities not well understood [2]. Overdispersion happens when count observations exhibit variability exceeding that of predicted by the binomial model. This can be taken into account by a Beta-Binomial model (**BB**); (4) Furthermore, one might be interested in modeling both overdispersion and correlation at the same time, combining beta and normal random effects. This can be undertaken by the model so-called Combined Model (**CM**). Hence, in this paper the aforementioned models were considered for the purpose of comparing and selecting a better model.

The comparison of the four models, as well as finally variable selection based on the better model were carried out using Akaike's information criterion (AIC) and Likelihood Ratio Test (LRT). In selecting a model, one can not find the "correct" model. Any model is a simplification of reality. Hence, a simple model that fits adequately has the advantage of model parsimony. The well known AIC judges a model how close its fitted value tend to be to the true expected values [1]. This is the model that minimizes:

AIC = -2(loglikelihood) + 2(number of parameters in the model)

The test statistic for LRT is given by:

$$G^2 = -2log\Lambda = -2log(l_0/l_1) = -2(L_0 - L_1), \tag{3.1}$$

where L_0 and L_1 denote the maximized log-likelihood function under reduced and full model, respectively. G^2 has a limiting null chi-squared distribution, as $n \to \infty$. The degree of freedom is given by the difference of the parameters under full and reduced model [1].

In both modeling for hypoglycemia and hyperglycemia, several continuous covariates (BMI, age, HbA1c level, increment in HbA1c level from start to end of study, and time from diagnosis) and categorical covariates (moment, insulin, oral antidiabetic drugs, and gender) were considered. The continuous covariates (X) were standardized in order to obtain unitless parameter estimates, permit comparisons of the estimated regression coefficients in common units, and to control roundoff errors in normal equation calculations,

$$x_i^s = \frac{x_i - \bar{x}}{s_x},$$

where x_i^s and x_i are the standardized and original X-values for patient i, respectively, $\bar{x}$ and s_x are the mean and standard deviation of X values (given in Table 3).

3.1 Simple Logistic Model

Logistic regression is widely used in medical studies to investigate the relationship between a binary response variable Y and a set of potential predictors X. This model takes into account neither overdispersion nor correlation of observations in the data. It is the simplest model that one may consider for binary data, which is defined as:

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ logit(\pi_{ij}) &= log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \mathbf{x_{ij}'}\beta \end{aligned} \tag{3.2}$$

where Y_{ij} is the measurement for hypoglycemia or hyperglycemia for patient i at measurement j, $\mathbf{x_{ij}}$ and β are p-dimensional vectors of known covariate values and unknown fixed regression coefficients, respectively. The logit function, which is a natural link, was used to link the random part (Y_{ij}) of 0/1 values to the systematic part ($\mathbf{x_{ij}}\beta$) which spans from $-\infty$ to ∞. This link will help also to have an odds ratio interpretation for the fixed effects (β).

3.2 Generalized Linear Mixed Model

This model is an extension of Generalized Linear Models (GLM) (Appendix C), incorporating patient-specific effects in the model. A random-effects model is used to describe the risk of hypoglycemia and hyperglycemia for each patient, by the inclusion of random (patient-specific) effects. This model was found important to take into account the variation of the observed proportions between patients (Figure 1 and 2) and to take into account the possible association of measurements taken on same patient. The model further enjoys the advantages of utilizing likelihood-based estimation, expressing the full probability distribution of the response. Hence, producing valid inference under the assumption of Missing At Random (MAR). MAR refers where there is missing on the response and this missingness does not depend on the unobserved response, but may depend on the observed ones. Working with this assumption was important in this study, since the dataset exhibits missingness, as observed in Figure A1. The model is given as:

$$\begin{aligned} y_{ij}|\mathbf{b_i} &\sim \text{Bernoulli}(\pi_{ij}) \\ \mathbf{b_i} &\sim N(\mathbf{0}, D) \\ logit(\pi_{ij}) = log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) &= \mathbf{x'_{ij}}\beta + \mathbf{z'_{ij}}\mathbf{b_i} \end{aligned} \tag{3.3}$$

where y_{ij} is the measurement for hypoglycemia or hyperglycemia for patient i at measurement j, $\mathbf{x_{ij}}$ and $\mathbf{z_{ij}}$ are the p-dimensional and q-dimensional design matrices for fixed and random effects, and β and $\mathbf{b_i}$ are p-dimensional unknown fixed effects and q-dimensional unknown patient-specific effects, respectively. The random effects ($\mathbf{b_i}$) account for the varying between patient variation and possible correlation of measurements within a patient. They are drawn from a normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix D.

This model describes the likelihood of the data conditioned on the fixed (β) and random effects $\mathbf{b_i}$,

$$f_i(\mathbf{y_i}|\mathbf{b_i}, \beta, \phi) = \prod_{j=1}^{n_i} f_{ij}(y_{ij}|\mathbf{b_i}, \beta, \phi) \tag{3.4}$$

where ϕ is the overdispersion parameter.

The marginal distribution can also be obtained when using this procedure, namely by integrating over

the random effects [8]. The likelihood contribution for patient i could be presented as:

$$f_i(\mathbf{y_i}|\beta, D, \phi) = \int f_i(\mathbf{y_i}|\mathbf{b_i}, \beta, \phi) f(\mathbf{b_i}|D) d\mathbf{b_i} \tag{3.5}$$

When one wants to determine the likelihood for the entire dataset, the following integral has to be solved for β, D and ϕ:

$$L(\beta, D, \phi) = \prod_{i=1}^{N} \int f_i(\mathbf{y_i}|\mathbf{b_i}, \beta, \phi) f(\mathbf{b_i}|D) d\mathbf{b_i} \tag{3.6}$$

However, only for special cases does there exist a closed form for this expression, because of the presence of N integrals over the random effects b_i. For binary response with logit link this closed form expression does not exist, and the likelihood can be maximized by applying numerical tools such as approximating either integrand, data, or integral. The integral method was used in this paper, since it is the most reliable for binary response. The likelihood contribution of each individual could be represented by the following sum:

$$\int f(z)\phi(z)dz \approx \sum_{q=1}^{Q} w_q f(z_q), \tag{3.7}$$

where Q is the order of the approximation. The higher Q, the more accurate the approximation will be. Further, w_q are the weights and the nodes z_q are the solutions to the Q^{th} order Hermite polynomial. When using *Gaussian Quadrature*, the nodes are chosen based on $\phi(z)$ (the density of the multivariate standard normal distribution), independent of $f(z)$ in the integrand. With *Adaptive Gaussian Quadrature*, the nodes and weights depend on the unknown parameters β, D and ϕ, which are updated at each iteration step. This implies that z_q and w_q need to be updated as well. In this paper, *Adaptive Gaussian Quadrature* was used, since it needs only few quadratue points to have closure approximation (convergence). To assure the convergence of the estimates obtained, models with various quadrature points were fitted, until similar results upto the fourth decimal places obtained.

The parameter estimates in this model has a patient-specific interpretation. To obtain the population averaged (marginal) parameter estimates a Marginalized Multilevel Model (**MMM**) [4] was used. This

can be expressed as:

$$
\begin{aligned}
g(\pi_{ij}^m) &= \mathbf{x}'_{\mathbf{ij}}\zeta^{\mathbf{m}}, \\
h(\pi_{ij}^c) &= \Delta_{ij} + \mathbf{z}'_{\mathbf{ij}}\mathbf{b_i}, \\
\mathbf{b_i} &\sim N(\mathbf{0}, D), \\
Y_{ij}|\mathbf{b_i} &\sim Bernoulli(\pi_{ij}^c)
\end{aligned} \tag{3.8}
$$

where π_{ij}^m and π_{ij}^c are marginal and conditional probabilities, respectively, $\mathbf{x_{ij}}$ and $\mathbf{z_{ij}}$ are p-dimensional and q-dimensional vectors of known covariates, g(.) and h(.) are the link functions, ζ^m and $\mathbf{b_i}$ are the p-dimensional vector of fixed effects and q-dimensional vector of random effects, respectively.

Now, the marginal mean $\pi_{ij}^m = E(Y_{ij})$ only depends on the linear predictor $\mathbf{x_{ij}}$ through a link function g(.), obtaining parameters estimates $\zeta^{\mathbf{m}}$ which have marginal (population averaged) interpretation. The logit link was used for g(.) in order to obtain an odds ratio interpretations for the parameter estimates. The conditional mean (π_{ij}^c) was used to take into account the patient-specific random effects $\mathbf{b_i}$ through the link function h(.). For the purpose of computational simplicity, probit link (ϕ^{-1}) was used, hence logistic-probit-normal model formulated. Δ_{ij} links the marginal mean π_{ij}^m and conditional mean π_{ij}^c. It is obtained from the solution to the integral equation:

$$
\mu_{ij}^m = expit(\mathbf{x}'_{\mathbf{ij}}\zeta^{\mathbf{m}}) = \int_b \phi(\Delta_{ij} + z'_{ij}\mathbf{b_i})\varphi(\mathbf{b_i}|\mathbf{0}, \mathbf{D})db_i \tag{3.9}
$$

and its solution is given by

$$
\Delta_{ij} = \left(\sqrt{1 + z'_{ij}Dz_{ij}}\right)\phi^{-1}\{expit(\mathbf{x}'_{\mathbf{ij}}\zeta^{\mathbf{m}})\}. \tag{3.10}
$$

3.3 Beta-binomial Model

When overdispersion is present in the data, the variability in the data may not be adequately described by the models which often exhibit a prescribed mean-variance link. Hence, it was of interest to take this into account in the model. This was done by combining Bernoulli distribution to that of its conjugate Beta distribution as a random effect, giving the so-called beta-binomial model. Conjugacy refers when hierarchical (Y_{ij}) and random-effects (θ_{ij}) densities have similar algebraic forms [9] (see

Appendix C). This is important to get a general and closed form solution for the marginal model. The model is defined as follows,

$$\begin{aligned} Y_{ij}|\theta_{ij} &\sim \text{Bernoulli}(\theta_{ij}) \\ \theta_{ij} &\sim Beta(\alpha,\beta), \quad (\alpha,\beta) \geqslant 0 \\ logit(\theta_{ij}) &= log\left(\frac{\theta_{ij}}{1-\theta_{ij}}\right) = \mathbf{x'_{ij}}\beta \end{aligned} \tag{3.11}$$

where Y_{ij} is the measurement for hypoglycemia or hyperglycemia for patient i at measurement j, $\mathbf{x_{ij}}$ and β are the p-dimensional design vector and p-dimensional vector of unknown fixed effects, respectively. It is possible to allow α and β vary with i and/or j. However, they are taken constant in this paper to avoid overspecifying the model. These parameters are not always jointly identified and hence it is important to impose a constraint such that only one parameter is assigned. In this paper, the contraint c $= \beta/\alpha$ was used. This model may fail short to appropriately estimate the parameters in the model, as explained in Appendix C, for binary data hierarchies need to be present. Hence, a model combining GLMM and BB, which was proposed by Molenberghs *et al.* [9], will be discussed next.

3.4 Combined Model

It is a model which combines the conjugate (Beta) random effects and Normal random effects, taking into account the occurence of overdispersion as well as hierarchical structure in the data. The normal random effects, at the individual level, accomodates the hierarchical structure and captures some overdispersion, whereas the Beta random effects, at the observation level, handles overdispersion not accounted for by the normal random effects [2]. It is defined as:

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij} = \theta_{ij}\kappa_{ij}) \\ \theta_{ij} &\sim Beta(\alpha,\beta), \quad (\alpha,\beta) \geqslant 0 \\ logit(\kappa_{ij}) &= log\left(\frac{\kappa_{ij}}{1-\kappa_{ij}}\right) = \mathbf{x'_{ij}}\beta + \mathbf{z'_{ij}}\mathbf{b_i} \\ \mathbf{b_i} &\sim N(0,D) \end{aligned} \tag{3.12}$$

where Y_{ij} is the measurement for hypoglycemia or hyperglycemia for patient i at measurement j, $\mathbf{x_{ij}}$ and $\mathbf{z_{ij}}$ are the design matrices for fixed effects and random effects, and β and $\mathbf{b_i}$ are vectors of fixed and random effects, respectively. The partially marginalized density over the random effects θ_{ij}, derived analytically by Molenberghs *et al.* [9], takes the form:

$$f(y_{ij}|b_i) = \frac{1}{\alpha+\beta}(\kappa_{ij}\alpha)^{y_{ij}}[(1-\kappa_{ij})\alpha+\beta]^{1-y_{ij}}, \tag{3.13}$$

Using the constraint $c = \beta/\alpha$ for the overdispersed parameters, the log-likelihood is given as:

$$ll = -log(1+c) + y_{ij}log(\eta_{ij}) - y_{ij}log(1+\eta_{ij}) + (1-y_{ij})log[(1-\kappa_{ij})+c] \tag{3.14}$$

where $\eta_{ij} = exp(\mathbf{x'_{ij}}\zeta + \mathbf{z'_{ij}}\mathbf{b_i})$.

When logit link and normal random effects used, there is no closed form solutions for neither the mean nor the variance, because the beta distribution does not allow for the multiplicative invariance or no strong conjugacy. However, approximations of the means, and variance-covariance elements were derived by Molenberghs *et al.* [9]. Strong conjugacy refers when hierarchical $Y_{ij}|b_i$ and random effects θ_{ij} have similar algebraic forms (see Appendix C). It is needed for an easy integration over the normal random effects.

As that of GLMM, CM has an appealing feature of being able to provide patient-specific predictions when they are needed by making use of Emperical Bayes estimates. However, in this study the interest was to obtain marginal population average predictions. Hence, a flexible framework that combines the MMM, discussed in Section 3.2, and CM were used to form the combined overdispersed and marginalized multilevel model (**COMMM**) [6]. This can be expressed as:

$$\begin{aligned}
g(\pi^m_{ij}) &= \mathbf{x'_{ij}}\zeta^{\mathbf{m}} \\
h(\kappa_{ij}) &= \Delta_{ij} + \mathbf{z'_{ij}}\mathbf{b_i} \\
\pi^c_{ij} &= \theta_{ij}\kappa_{ij} \\
\theta_{ij} &\sim Beta(\alpha,\beta), \quad (\alpha,\beta) \geqslant 0 \\
\mathbf{b_i} &\sim N(0,D) \\
Y_{ij}|\mathbf{b_i} &\sim Bernoulli(\pi^c_{ij})
\end{aligned} \tag{3.15}$$

In a similar fashion as that of GLMM, a logit link was used for the function g(.) to have an odds ratio interpretation, and a probit link for h(.) used for computational simplicity, generating logistic-probit-beta-normal model. Δ_{ij} is an expression that links the marginal mean π_{ij}^{m} and conditional mean π_{ij}^{c}. In contrast to MMM case, here Δ_{ij} involves two integrations, for θ_{ij} and $\mathbf{b_i}$. It is given as:

$$\begin{aligned}\pi_{ij}^{m} = expit(\mathbf{x'_{ij}}\boldsymbol{\zeta}^{\mathbf{m}}) &= \int_b \int_\theta \pi_{ij}^{c} d\Theta_\theta dN_b \\ &= \int_b \int_\theta \theta_{ij}\phi(\Delta_{ij} + \mathbf{z'_{ij}b_i}) d\Theta_\theta dN_b \\ &= \int_b E(\theta_{ij})\phi(\Delta_{ij} + \mathbf{z'_{ij}b_i}) dN_b\end{aligned}$$

such that

$$\Delta_{ij} = \left(\sqrt{1 + \mathbf{z_{ij}Dz_{ij}}}\right).\phi^{-1}\{(1+c).expit(\mathbf{x'_{ij}}\boldsymbol{\zeta}^{\mathbf{m}})\}, \tag{3.16}$$

with $c = \beta/\alpha$ is the constraint used in this paper to avoid overparametrizing the model.

SAS (version 9.3) procedures Genmod, GLIMMIX, and NLMIXED were used in this study. For hypothesis testing a significance level of 5% was employed. Further, odds ratios (OR) and their 95% confidence intervals (CI) were calculated for the fixed effects. An effect is statistically significant if the CI does not contain 1.

4 Results

The process of buiding a model for each response (hypoglycemia or hyperglycemia) was as follows. First, the four models (SL, BB, GLMM, and CM) were fitted with all main effects included in the model. After comparing the four models, the best model was selected for further analysis. Then, using the best model, step-up procedure were employed for variable selection. Finally, if the best model is either GLMM or CM, marginalization was applied using MMM or COMMM, respectively.

4.1 Modeling Hypoglycemia

The initial models for SL, BB, GLMM, and CM, representing parameter estimates with and without beta-binomial component on one hand, and with and without normal random effects on the other hand are presented in Table 4. Comparison of the parameter estimates in the four models simultaneously is not possible, since they have different interpretation. Simple Logistic and Beta-binomial models have marginal interpretation, whereas GLMM and CM have patient-specific interpretation. Comparing SL and BB, the standard errors of the parameter estimates are higher for BB, indicating that their was overdispersion that was taken into account in the BB model. It was also observed that GLMM generated higher standard errors than SL. Further, it was observed that the standard errors of the parameter estimates in CM were higher comparing to the rest three models, indicating that this model took into account both overdispersion and correlation of measurements within a patient.

Furthermore, looking at the AICs of the four models, AIC of BB was similar to that of SL. Comparing GLMM and SL, GLMM had a way lower AIC. However, further decrease in AIC was observed under CM as compared to GLMM. As was anticipated from AIC, the overdispersion parameter in BB was not significant, but it became significant under CM. This was expected, as was explained in section 3.3, that for binary data, BB is hard to fit when normal random effects are not present in the model. In contrast to the overdispersion parameter, the correlation parameter was significant on both GLMM and CM. Hence, it was concluded that, even though the normal random effects (Logistic-normal or GLMM) improved significantly over SL, there was further improvement upon adding the overdispersion parameter (CM). This implies that, the data exhibit at the same time overdispersion and correlation. Hence, it was found important to consider the combined model for further analysis

Table 4: *Parameter estimates and standard errors of the initial models for Hypoglycemia*

		SL	BB	GLMM	CM
Effect	Par.	Estimate(s.e.)	Estimate(s.e.)	Estimate(s.e.)	Estimate(s.e.)
Intercept	β_0	-4.3326(0.4934)*	-2.2746(0.8063)*	-4.7877(0.9688)*	-2.8468(1.8379)
Mom2	β_2	-0.3462(0.2969)	-0.2117(0.5063)	-0.3355(0.3017)	-1.7586(0.7709)*
Mom3	β_3	0.4099(0.2488)	0.8699(0.6021)	0.4343(0.2545)	0.7238(0.6793)
Mom4	β_4	-0.4541(0.3062)	-0.3623(0.4910)	-0.459(0.3109)	-0.8091(0.752)
Mom5	β_5	-0.5323(0.3113)	-0.3384(0.5953)	-0.548(0.3158)	-0.3127(0.7752)
Mom6	β_6	-1.3954(0.4255)*	-1.6007(0.5533)*	-1.4075(0.4290)*	-4.4741(1.3310)*
Age^s	β_7	-0.3709(0.1100)*	-0.4923(0.2117)*	-0.2791(0.2669)	-0.6550(0.5352)
Treat2	β_8	0.6338(0.3702)	0.9785(0.5912)	0.6530(0.8494)	1.6066(1.6549)
Treat3	β_9	0.5639(0.3671)	0.7673(0.5094)	0.1921(0.8444)	1.1476(1.6657)
BMI^s	β_{10}	-0.8741(0.1076)*	-1.3065(0.5237)*	-1.1794(0.2758)*	-2.0524(0.6007)*
HbA1c^s	β_{11}	-0.0348(0.1209)	-0.1403(0.2279)	-0.1997(0.3060)	-0.5839(0.6186)
HbA1c$^s_{inc}$	β_{12}	-0.0659(0.1120)	-0.0899(0.1562)	-0.0652(0.2879)	-0.0815(0.5754)
Timediag^s	β_{13}	0.0899(0.1043)	0.0844(0.1603)	0.1246(0.2634)	0.1583(0.5304)
Women	β_{14}	-0.3769(0.2237)*	-0.6404(0.4873)	-0.5492(0.5444)	-1.3086(1.0971)
OAD2	β_{15}	-0.0436(0.2460)	-0.1631(0.3664)	-0.3937(0.5965)	-0.6845(1.1727)
OAD3	β_{16}	-0.0005(0.4840)	0.3758(0.7909)	0.2844(1.0090)	0.9051(1.9174)
c	β/α	–	7.8517(6.2859)	–	8.5054(1.5181)*
Std. dev. of RE	$\sqrt{d}$	–	–	1.2192(0.2443)*	2.6144(0.5745)*
AIC		1200.3	1200.9	1131.5	1128.9

** -significantly different from 0, Par. -Parameter, s.e. -standard error, s -standardardized variable*

to have valid inference.

The fixed effects were assessed using LRT and AIC. Only moment and BMI were identified as significant covariates with minimum AIC, whereas the others had G^2 of 6.9 with p-value 0.6475. Further, the interaction of moment and BMI was found significant, generating G^2 of 15.1 with p-value 0.0099. Since the parameter estimates of the final model have patient-specific interpretations, the marginalized model (COMMM) was fitted and presented in Table 5, under COMMM(1). After marginalizing the parameter estimates, there was a gain in precision of all parameters. This had a consequence on the significance of two marginal parameters β_3 and β_9. This implies that, although the two parameter estimates are not significantly different from zero at the patient-specific level, they are significant on the population average level.

Finally, the parameter estimates under COMMM(1) can be used to calculate the odds, predicted probability, and odds ratios of hypoglycemia. The odds of hypoglycemia for patients with average BMI (31.7 kq/m^2) at pre-breakfast was $\exp^{\beta_0} = exp^{-5.3355} = 0.0048$. Whereas at post-breakfast, for

Table 5: *Parameter estimates and standard errors of the final models for Hypoglycemia.*

		CM	COMMM(1)	COMMM(2)		
Effect	Par.	Estimate(s.e.)	Estimate(s.e.)	Effect	OR	95% CI [L : U]
Intercept	β_0	-3.751(0.975)*	-5.3355(0.4739)*	Mom1	1.00	Reference
Mom2	β_2	-0.7485(1.1063)	0.2519(0.5500)	Mom2	1.29	[0.43 : 3.88]
Mom3	β_3	1.5221(0.9208)	1.2065(0.4680)*	Mom3	3.34	[1.31 : 8.55]
Mom4	β_4	0.2049(0.9703)	0.4507(0.5205)	Mom4	1.57	[0.55 : 4.46]
Mom5	β_5	0.6212(0.9306)	0.5763(0.5022)	Mom5	1.78	[0.65 : 4.88]
Mom6	β_6	-2.8448(1.6629)	-0.4935(0.6529)	Mom6	0.61	[1.84: 2.26]
BMI^s	β_7	-3.7156(1.0137)*	-1.9224(0.3240)*	BMI^s*Mom1	0.15	[0.08 : 0.28]
BMI^s*Mom2	β_8	1.6981(1.0232)	0.7482(0.4119)	BMI^s*Mom2	0.31	[0.07 : 1.35]
BMI^s*Mom3	β_9	1.2029(0.9282)	0.7982(0.3299)*	BMI^s*Mom3	0.32	[0.09 : 1.21]
BMI^s*Mom4	β_{10}	2.1358(0.9672)*	0.9641(0.3995)*	BMI^s*Mom4	0.38	[0.09 : 1.64]
BMI^s*Mom5	β_{11}	2.3021(0.9462)*	1.2232(0.4112)*	BMI^s*Mom5	0.50	[0.11: 2.17]
BMI^s*Mom6	β_{12}	2.4719(1.4432)	1.0602(0.5754)	BMI^s*Mom6	0.42	[0.07 : 2.57]
c	β/α	8.1449(1.4252)*	2.2308(0.5998)*	–	–	–
Std. dev. of RE	$\sqrt{d}$	2.8468(0.6969)*	0.469(0.0826)*	–	–	–

** -significantly different from 0, s -standardardized variable*

the same patients, it was $\exp^{\beta_0+\beta_2} = \exp^{-5.3355+0.2519} = 0.0062$. Further, it was observed that for the same patients, the risk of hypoglycemia is significantly higher before lunch ($log(odds) = \beta_3 = 1.2065$), compared to that of pre-breakfast. Figure 8 shows the predicted probability of hypoglycemia over moments for patients with average BMI. It was observed that a higher risk was obtained before lunch.

The odds ratios were calculted and presented in Table 5 under COMMM(2). For an increase in standardard deviation of BMI (an increase in 3.9863 kg/m^2) at pre-breakfast, the odds of hypoglycemia is lower by 85% ($\exp^{\beta_7} = \exp^{-1.9224} = 0.15$). Whereas before dinner, for the same increase in BMI, the odds of hypoglycemia is lower by 50% ($\exp^{\beta_7+\beta_{11}} = \exp^{-1.9224+1.2232} = 0.50$). Figure 9 presents the predicted probability of hypoglycemia by moment and BMI (unstandardized). It was observed that for patients with BMI less than about 30 kg/m^2, the risk of hypoglycemia was higher at pre-breakfast and pre-lunch, whereas for patients with BMI above 30 kg/m^2, the risk was almost zero for all moments. Further, it was observed that for all moments there is a decreasing trend over BMI.

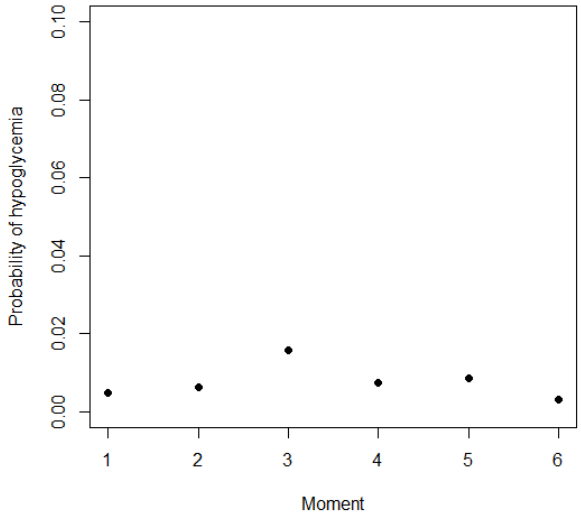

Figure 8: *Predicted probability of Hypoglycemia by Moment for patients with average BMI*

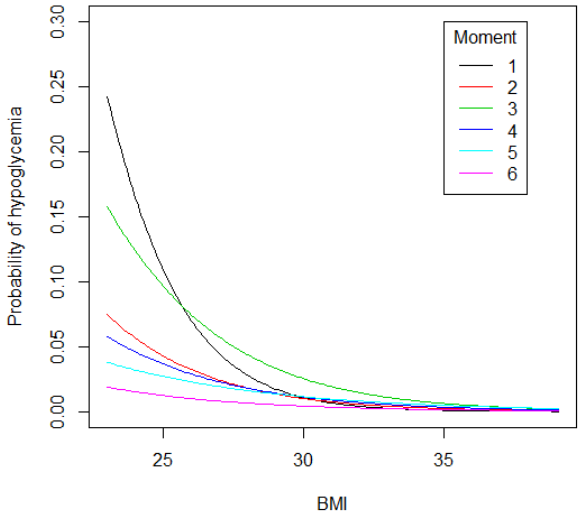

Figure 9: *Predicted probability of Hypoglycemia by Moment and Body mass index*

4.2 Modeling Hyperglycemia

In a similar fashion to that of hypoglycemia, a model was developed for hyperglycemia. Table 6 presents the parameter estimates and standard errors for the four models (SL, BB, GLMM, and CM). BB and GLMM had higher standard errors and lower AIC compared to SL. Further, CM had the higher standard errors and lower AIC among the four models. Furthermore, it was found that the overdispersion component (in BB) and normal random effects (in GLMM), each alone, was significant. In addition, when both random effects were taken into account at the same time (in CM), they were both significant. These indicate that, although each random component alone improved the model, when both were taken into account at the same time, the model further improved. Hence, the combined model was used for variable selection.

The covariates moment, HbA1c and $HbA1c_{inc}$ were found statistically significant with minimum AIC, whereas the other covariates were not significant, generating $G^2 = 9.0$ with p-value = 0.3423. Further, the pairwise interaction of moment with both HbA1c and $HbA1c_{inc}$, were found significant with $G^2 = 56.3$ and p-value < 0.0001. The parameter estimates in the CM have patient-specific interpretations. Hence, to have population average ones, the COMMM model were fitted for further inference, and presented in Table 7 under COMMM(1). The parameter estimates have higher precision in the

Table 6: *Parameter estimates and standard errors of the initial models for Hyperglycemia.*

		SL	BB	GLMM	CM
Effect	Par.	Estimate(s.e.)	Estimate(s.e.)	Estimate(s.e.)	Estimate(s.e.)
Intercept	β_0	0.4742(0.1198)*	1.9652(0.2830)*	0.5445(0.5911)	0.7334(0.6910)
Mom2	β_2	-0.8101(0.0917)*	-1.5692(0.1681)*	-0.9611(0.0997)*	-1.0725(0.1171)*
Mom3	β_3	-0.0244(0.0891)	0.01697(0.1710)	-0.0270(0.0973)	0.0101(0.1093)
Mom4	β_4	-0.5747(0.0906)*	-1.2305(0.1698)*	-0.6723(0.0986)*	-0.7380(0.1112)*
Mom5	β_5	0.3571(0.0901)*	0.4035(0.1722)*	0.4286(0.0988)*	0.4634(0.1106)*
Mom6	β_6	-0.0062(0.0899)	-0.3966(0.1660)*	-0.0135(0.0981)	-0.0159(0.1088)
Age^s	β_7	0.1536(0.0302)*	0.3898(0.0664)*	0.1123(0.1671)	0.1234(0.1926)
Treat2	β_8	0.4533(0.0973)*	0.1266(0.1778)	0.5108(0.5359)	0.7063(0.6359)
Treat3	β_9	-0.1264(0.0987)	-0.6881(0.1841)*	-0.1365(0.5361)	-0.0510(0.6234)
BMI^s	β_{10}	0.0449(0.0289)	-0.2708(0.0750)*	0.0724(0.1605)	0.1287(0.1868)
HbA1c^s	β_{11}	0.8221(0.0366)*	2.4980(0.1249)*	0.9164(0.1815)*	1.0116(0.2121)*
HbA1c$^s_{inc}$	β_{12}	0.4084(0.0344)*	0.9039(0.0858)*	0.4088(0.1809)*	0.4265(0.2109)*
Timediag^s	β_{13}	0.0460(0.0307)	0.1456(0.0817)	0.0954(0.1757)	0.0709(0.2022)
Women	β_{14}	-0.2298(0.0604)*	0.3294(0.1108)*	-0.3164(0.3303)	-0.4217(0.3837)
OAD2	β_{15}	-0.6073(0.0644)*	-0.2664(0.1210)*	-0.7280(0.3624)*	-0.7853(0.4178)
OAD3	β_{16}	-0.2737(0.1124)*	-0.5442(0.2131)*	-0.3399(0.6295)	-0.2582(0.7408)
c	β/α	–	0.3441(0.0232)*	–	0.0556(0.02036)*
Std. dev. of RE	$\sqrt{d}$	–	–	1.0427(0.1121)*	1.1975(0.1467)*
AIC		8416.7	8135.1	7490.2	7485.7

** -significantly different from 0, s -standardardized variable*

COMMM model compared to CM. One would expect that the parameter estimate and precision for the variance component d to have similar results on both CM and COMMM, because d appears on both case in the conditional part of the model. However, the connector function (3.16) exhibits a non-linear relationship between the variance component D and the fixed effect parameter ζ^m. Hence, the parameter estimate and standard error for $\sqrt{d}$ is a way lower in COMMM compared to CM.

Finally, the odds, predicted probability, and odds-ratios can be calculated using the parameter estimates under COMMM(1). The odds of hyperglycemia pre-breakfast, for patients with average HbA1c level (7.66) and an average $HbA1c_{inc}$ (-0.09), was $\exp^{\beta_0} = \exp^{0.1122} = 1.1187$. Further, the odds of hyperglycemia pre-dinner, for the same patients, was $\exp^{\beta_0+\beta_5} = \exp^{0.1122+0.3894} = 1.6514$. As shown in Figure 10, for the same patients, the predicted probability of hyperglycemia was higher before dinner.

The odds ratios for each parameters were calculated, and results are given in COMMM(2) of Table 7. For a standard deviation increase in HbA1c level (an increase in 1.39) pre-breakfast, the odds

Table 7: *Parameter estimates and standard errors of the final models for Hyperglycemia.*

		CM	COMMM(1)	COMMM(2)		
Effect	Par.	Estimate(s.e.)	Estimate(s.e.)	Effect	OR	95% CI
Intercept	β_0	0.2362(0.2102)	0.1122(0.1606)	Mom1	1.00	Reference
Mom2	β_2	-1.0823(0.1144)*	-0.9276(0.0941)*	Mom2	0.39	[0.33 : 0.48]
Mom3	β_3	-0.0176(0.1053)	-0.0195(0.0883)	Mom3	0.98	[0.82 : 1.17]
Mom4	β_4	-0.7411(0.1091)*	-0.6356(0.0911)*	Mom4	0.53	[0.44 : 0.64]
Mom5	β_5	0.4770(0.1115)*	0.3894(0.0897)*	Mom5	1.48	[1.23 : 1.77]
Mom6	β_6	0.0057(0.1094)	-0.0035(0.0910)	Mom6	0.99	[0.83 : 1.20]
HbA1c^s	β_7	0.8892(0.2178)*	0.6953(0.1694)*	HbA1c^s*M1	2.00	[1.43 : 2.82]
HbA1c^s*Mom2	β_8	0.3470(0.1253)*	0.3177(0.1011)*	HbA1c^s*M2	2.75	[1.60 : 4.74]
HbA1c^s*Mom3	β_9	-0.1173(0.1169)	-0.0857(0.0945)	HbA1c^s*M3	1.84	[1.08 : 3.13]
HbA1c^s*Mom4	β_{10}	0.3149(0.1253)*	0.2622(0.0988)*	HbA1c^s*M4	1.54	[1.52 : 4.46]
HbA1c^s*Mom5	β_{11}	0.1793(0.1219)	0.1427(0.0960)	HbA1c^s*M5	2.31	[1.36 : 3.94]
HbA1c^s*Mom6	β_{12}	0.4239(0.1243)*	0.3661(0.0987)*	HbA1c^s*M6	2.89	[1.69 : 4.95]
HbA1c$^s_{inc}$	β_{13}	0.5794(0.2139)*	0.4318(0.1724)*	HbA1c$^s_{inc}$*M1	1.54	[1.09 : 2.18]
HbA1c$^s_{inc}$*Mom2	β_{14}	-0.1554(0.1177)	-0.1042(0.0974)	HbA1c$^s_{inc}$*M2	1.39	[0.81 : 2.38]
HbA1c$^s_{inc}$*Mom3	β_{15}	-0.4215(0.1099)*	-0.3452(0.0912)*	HbA1c$^s_{inc}$*M3	1.09	[0.64 : 1.85]
HbA1c$^s_{inc}$*Mom4	β_{16}	-0.3016(0.1138)*	-0.2278(0.0944)*	HbA1c$^s_{inc}$*M4	1.23	[0.72 : 2.09]
HbA1c$^s_{inc}$*Mom5	β_{17}	-0.0148(0.1153)	-0.0133(0.0946)	HbA1c$^s_{inc}$*M5	1.52	[0.89 : 2.60]
HbA1c$^s_{inc}$*Mom6	β_{18}	-0.0238(0.1157)	-0.0096(0.0959)	HbA1c$^s_{inc}$*M6	1.53	[0.89 : 2.61]
c	β/α	0.0470(0.0205)*	0.0535(0.0207)*	–	–	–
Std. dev. of RE	$\sqrt{d}$	1.2988(0.1570)*	0.4492(0.0496)*	–	–	–

** -significantly different from 0, M1 -Mom1, s -standardardized variable*

of hyperglycemia was higher by $\exp^{\beta_7} = 2.00$ (100%). Whereas for the same increase in HbA1c level before dinner, the odds of hyperglycemia was higher by $\exp^{\beta_7+\beta_{11}} = 2.31$ (131%). Further, for a standard deviation increase in HbA1c$_{inc}$ (an increase in 0.46) pre-breakfast, the odds of hyperglycemia was higher by $\exp^{\beta_{13}} = 1.54$ (54%). Similar increase in odds of hyperglycemia was obtained for the same increase in HbA1c$_{inc}$ before dinner, $\exp^{\beta_{13}+\beta_{17}} = 1.52$ (52%). Figure 11 shows the predicted probability of hyperglycemia by moment and HbA1c level (unstandardized). It was observed that, for patients with HbA1c level less than about 10, the risk of hyperglycemia was higher before dinner. However, for patients with HbA1c level greater than 10, the risk of hyperglycemia was higher on both before and after dinner. As shown in Figure 12, for patients who had HbA1c$_{inc}$ (unstandardized) less than about -0.65, the risk of hyperglycemia was higher before lunch, whereas for patients who had an HbA1c$_{inc}$ greater than -0.65, the risk was higher before dinner.

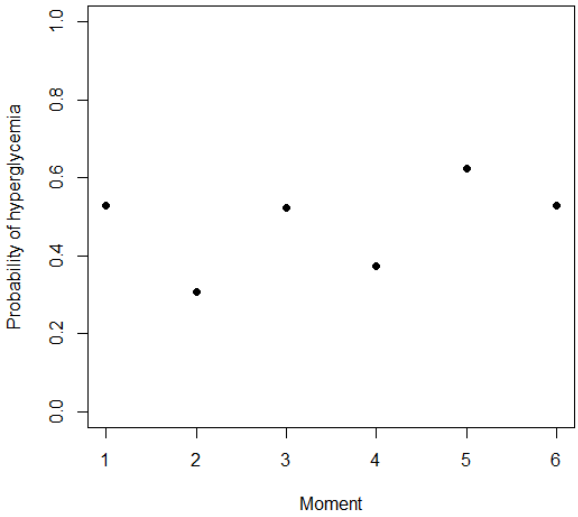

Figure 10: *Predicted probability of Hyperglycemia by Moment for patients with average HbA1c level and HbA1c$_{inc}$*

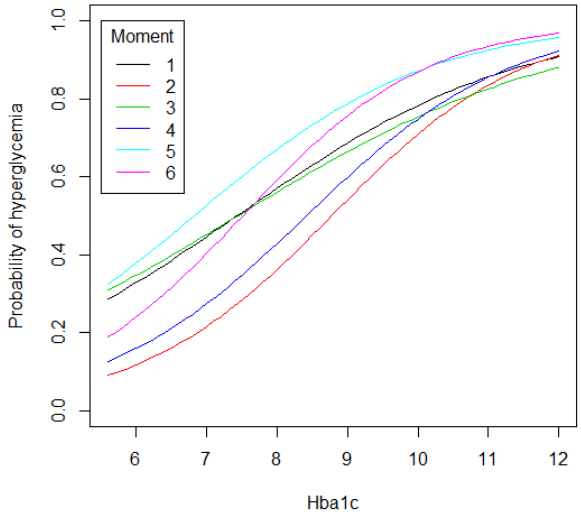

Figure 11: *Predicted probability of Hyperglycemia by Moment and HbA1c level*

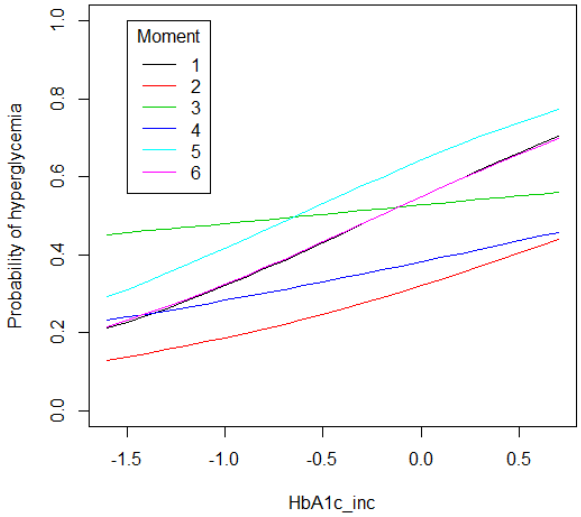

Figure 12: *Predicted probability of Hyperglycemia by Moment and HbA1c$_{inc}$*

5 Discussion, Conclusion, and Recommendation

Hypoglycemia and hyperglycemia are serious health problems for diabetic patients. Hence, it is important to monitor blood glucose of the patients to attain it in acceptible range (normal glucose level). SMBG provides real time information that helps patients to manage their diabete, whereas cost influence its use. Thus in this paper, it was important to identify the optimal frequency and timing to monitor the blood glucose level of the patients. Further, it was important to assess various clinical factors which are associated with hypoglycemia or hyperglycemia, to avoid the risks as much as possible.

GLMM is a powerful tool to deal with binary responses which have repeated measurements. This model fits the full likelihood of the data. Although binary data do not violate mean variance relationship, due to introduction of correlation in the data, overdispersion occurs. Hence, the importance of BB model arises. To deal with both overdispersion as well as correlation, a model which combines a beta random effects and a normal random effects was used in the sprit of Molenberghs *et al.* [9]. In the consequence, four different models (SL, BB, GLMM, and CM) were fitted for each response, to compare and have a better model for an appropriate inference. CM did a better job for both hypoglycemia and hyperglycemia. Since overdispersion can come from only correlated outcomes for binary data, it is difficult to identify the effect of overdispersion with reasonable precision from that of normal random effects [10]. Finally, since CM has patient-specific interpretation, to have population averaged ones, COMMM was fitted. This model had higher precision for the parameter estimates, for both hypoglycemia and hyperglycemia.

For hypoglycemia, exploration was done first for each covariate included in the study. As shown in Figure 3, it was observed that pre-lunch measurements had higher risk, post-dinner had lower risk, and the rest had medium risk. This observation was assured, checking it formally using COMMM model presented in Figure 8. Futher, from the exploration done by BMI in Figure 5, it was observed that as BMI increases the risk of hypoglycemia decreases. In addition, the risk of moment on hypoglycemia seems to depend on BMI, where for patients with low BMI, a higher risk were observed before breakfast and before lunch. Whereas for patients with high BMI, the risk of hypoglycemia was almost zero for all moments. Again, this observation was investigated by the model and presented in Figure 9. For patients with BMI less than about 30 kg/m^2, the higher risk was obtained before breakfast and

before lunch. Whereas with BMI greater than 30 kg/m^2, all moments had almost no risk. Similar result was obtained by Hoffman *et al.* [5], who reported that prebreakfast/prelunch measurements captured the largest proportion of the hypoglycemic readings. In addition, as BMI increases the risk of hypoglycemia decrease for all moments (Figure 9). This result was also in line with the one obtained by Ross *et al.* [11], who reported that a reduced risk of hypoglycemia was for patients with BMI $>$ 25 kg/m^2. Other covariates age, sex, time from diagnosis, insulin, HbA1c, HbA1c$_{inc}$, and oral antidiabetic drugs were found to be insignificant in determining the risk of hypoglycemia.

Similarly, the exploration done for hyperglycemia showed that the highest risk was seen before dinner, then medium risk for pre-breakfast, pre-lunch, and post-dinner (Figure 4). This view was confirmed by the model, showing that the higher risk was obtained before dinner (Figure 10). This result consides with the one obtained by Hoffman *et al.* [5], who reported that the predinner/bedtime measurements captured the largest proportion of hyperglycemic readings. The exploration in Figure 6 shows that, as HbA1c level at baseline increases, the risk of hyperglycemia increases. It also shows that this might depend on the moment by which the measurement was taken. This was investigated by the model and presented in Figure 11, showing that the risk increases as HbA1c level at baseline increase. This was supported by "Mayo clinic" [17], stating that patients with high blood concentration of glucose have from 2 to 3 times more HbA1c level than normal individuals. Furthermore, for patients who had HbA1c level less than about 10, the risk was higher at pre-dinner, whereas for those who had greater than 10 HbA1c level, the risk was higher at both pre-dinner and post-dinner. Figure 7 explored that, more frequent hyperglycemic episodes were present for patients with high value of HbA1c$_{inc}$, and this relationship depends on the moments. This was investigated using the model and presented in Figure 12. As was expected from the exploration, higher risk of hyperglycemia was obtained for an increase in HbA1c$_{inc}$. Pre-lunch measurements had highest risk for patients with HbA1c$_{inc}$ less than about -0.65, and pre-dinner measurements had highest risk for patients with HbA1c$_{inc}$ greater than -0.65. Other covariates age, sex, time from diagnosis, body mass index, insulin, and oral antidiabetic drugs were found to be insignificant in determining the risk of hyperglycemia.

In conclusion, generally speaking, the risk of hypoglycemia was higher before breakfast and before lunch. However, as was discussed above, it depends on BMI of the patients. Whereas for hyperglycemia, the risk was higher before dinner. This also depends on HbA1c level at baseline or on HbA1c increment from baseline to the end of study (HbA1c$_{inc}$). Further, there was a decreasing risk

of hypoglycemia for an increase in BMI. Whereas for hyperglycemia, there was an increasing risk for an increase in HbA1c or an increase in HbA1c_{inc}.

In this paper, the interest was to investigate the risk of hypoglycemia and hyperglycemia, in general. Hence, it was not investigated whether patients with hypoglycemic or hyperglycemic readings were symptomatic. For both cases, hypoglycemia and hyperglycemia, measurements taken within a patient were found to be correlated. This was taken into account by including a patient-specific intercept in the model. However, in this study there was no day indicator in the dataset, at which patients had taken the measurements. For future studies, recording day indicator is important to take into account the correlation of observations within a patient in consecutive days. Further, this study was done under the assumption of MAR. However, MAR might not be sufficient and it is important to do some sort of sensitivity analysis. But, again since the day at which measurements taken was not recorded in the dataset, it was difficult to have the pattern of missingness and do the sensitivity. Furthermore, the treatments (insulin and OAD) were not randomized over the patients due to ethical reason. This will influence to selection bias. Hence, in this study causal interpretation of the treatment will be lost.

References

[1] Agresti, A. (2002) *Categorical Data Analysis (2nd ed.).* New York: John Wiley & Sons.

[2] Aregay, M. (2013) *Flexible Statistical Modeling Approaches For Hierarchical and Overdispersed Data (Doctoral dissertation).* KU Leuven, Doctoral School in Biomedical Sciences. (ISBN. D/2013/09/27)

[3] Bergenstal, R., Ahmann, A., Bailey, T., Beck R.W., Bissen, J., Buckingham, B., Deeb, L., Dolin, R.H., Garg, S.K., Goland, R., Hirsch, I.B., Klonoff, D.C., Kruger, D.F., Matfin, G., Mazze, R.S., Olson, B.A., Parkin, C., Peters, A., Powers, M.A., Rodriguez, H., Southerland, P., Strock, E.S., Tamborlane, W., and Wesley, D.M. (2013) Recommendations for Standardizing Glucose Reporting and Analysis to Optimize Clinical Decision Making in Diabetes: The Ambulatory Glucose Profile (AGP). *Diabetes thechnology and therapeutics,* **15**(3), 198-211.

[4] Griswold, M.E. and Zeger, S.L. (2004) On marginalized multilevel models and their computation. *Johns Hopkins University, Department of Biostatistics, Working Paper 99.*

[5] Hoffman, R.M., Shah, J.H., Wendel, C.S., Duckworth, W.C., Adam, K.D., Bokhari, S.U., Dalton, C., and Murata, G.H. (2002) Evaluating once- and twice-daily self-monitored blood glucose testing strategies for stable insulin-treated patients with type 2 diabetes : the diabetes outcomes in veterans study. *Diabetes care,* **25***(10),* 1744-1748.

[6] Iddi, S. (2013) *Flexible Hierarchical and Marginalized Models for Non-Gaussian Data: Application to Life Science Studies (Doctoral dissertation).* KU Leuven, Doctoral School of Biomedical Sciences. (ISBN. D/2013/09/27)

[7] López, G.J.M., Couselo, P.M., Rodriguez, G.M., Sampedro, G.F., Santos, M.J.M., Rodriguez, J., Segade, S., and Casanueva, F.F. (2014) Optimal Time for Capillary Blood Glucose Monitoring in Patients with Type 2 Diabetes Treated with Insulin: A Cross-Sectional Observation Study. *Journal of Diabetes Metabolic Disorders and Control,* **1**(2), 00010.

[8] Molenberghs, G. and Verbeke, G. (2005) *Models for Discrete Longitudinal Data*. Springer Verlag New York, Inc.

[9] Molenberghs, G., Verbeke, G., Demétrio, C.G.B., and Vieira, A. (2010) A Family of Generalized Linear Models for Repeated Measures With Normal and Conjugate Random Effects. *Statistical Science,* **25**(3), 325-347.

[10] Molenberghs, G., Verbeke, G., Iddi, S., and Demétrio, C.G.B. (2012) A Combined Beta and Normal Random-effects Model for Repeated, Overdispersed Binary and Binomial Data. *Journal of Multivariate Analysis,* **111**, 94-109

[11] Ross, S., Dzida, G., Kaiser, M., Ligthelm, R., Meneghini, L., Nazeri, A., Orozco-Beltran, D., Pan, C., and Svendsen, A.L. (2014) Safety of once-daily insulin detemir in patients with type 2 diabetes treated with oral hypoglycemic agents in routine clinical practice. *Journal of Diabetes,* **6**(3), 243–250.

[12] Seufert, J., Brath, H., Pscherer, S,, Borck, A., Bramlage, P., and Siegmund, T. (2014) Composite Efficacy Parameters and Predictors of Hypoglycemia in Basal-Plus Insulin Therapy - A Combined Analysis of 713 Type 2 Diabetic Patients. *Diabetes Obesity and Metabolism,* **16**(3), 248-254.

[13] Wild, S., BCHIR, M., Roglic, G., Green, A., Sicree, R., and King, H. (2004) Global Prevalence of Diabetes Estimates for the year 2000 and projections for 2030. *Diabetes Care,* **27**, 1047-1053

[14] American Diabetes Association. (2013) Standards of medical care in diabetes–2013. *Diabetes Care,* **36**(1), S11-S66.

[15] American Diabetes Association. (2004) Hyperglycemic Crises in Diabetes. *Diabetes Care,* **27**(1), S94-S102.

[16] Diabetes Health Center. (2014) *Hyperglycemia and Diabetes.* Retrieved August 7, 2014, from http://www.webmd.com/diabetes/diabetes-hyperglycemia

[17] Mayo Clinic. (2014) *Hemoglobin A1c, Blood.* Retrieved August 7, 2014, from http://www.mayomedicallaboratories.com/test-catalog/Clinical+and+Interpretive/82080

Appendix

A Figures

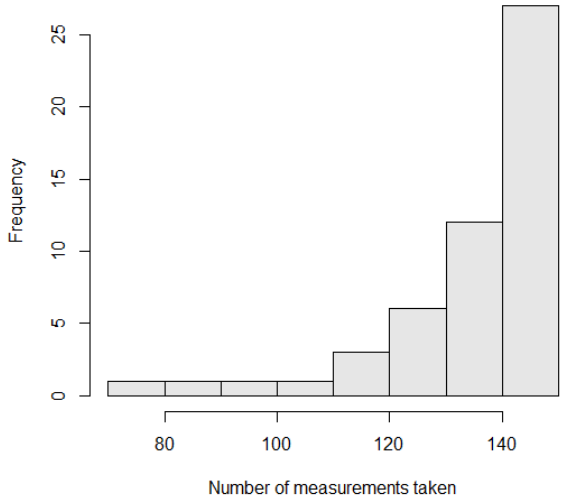

Figure A1: *The distribution for the number of measurements taken per patient*

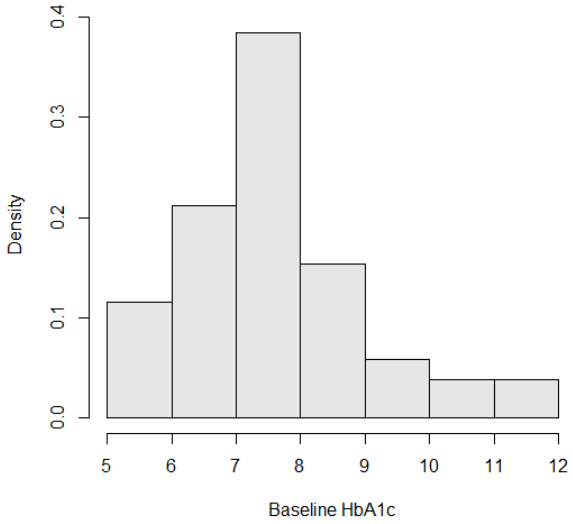

Figure A2: *Distribution of HbA1c at baseline*

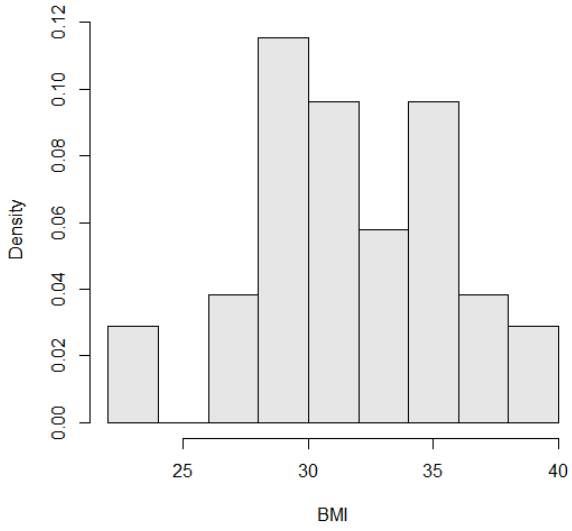

Figure A3: *Distribution of Body mass index*

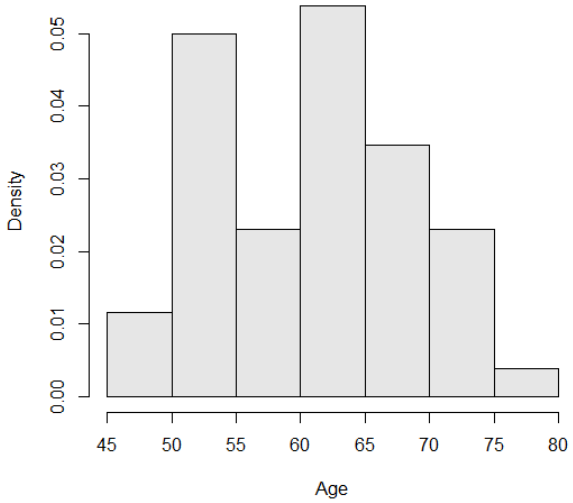

Figure A4: *Distribution of Age*

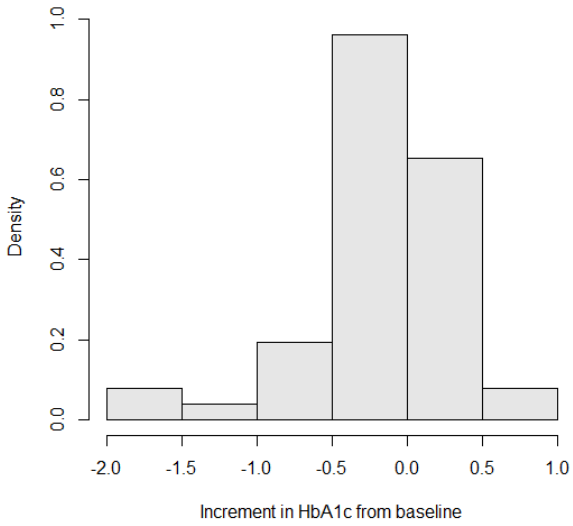

Figure A5: *Distribution of* $HbA1c_{inc}$

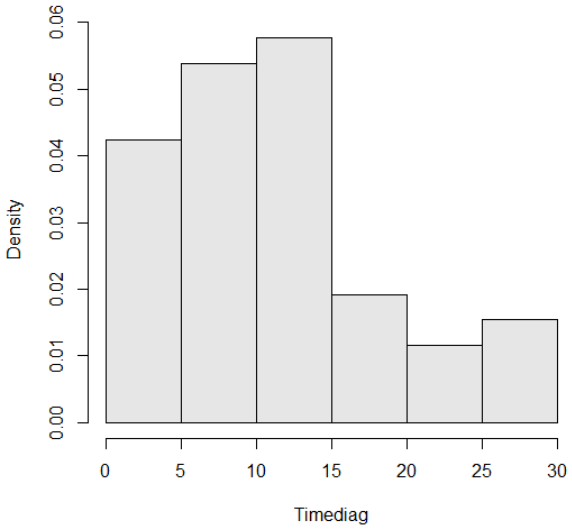

Figure A6: *Distribution of Time from diagnosis*

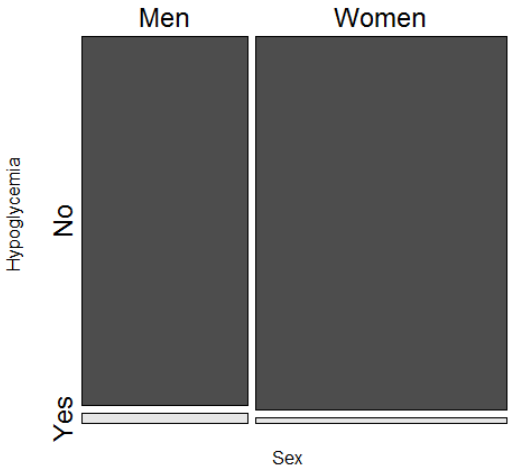

Figure A7: *Proportion of Hypoglycemia by Sex*

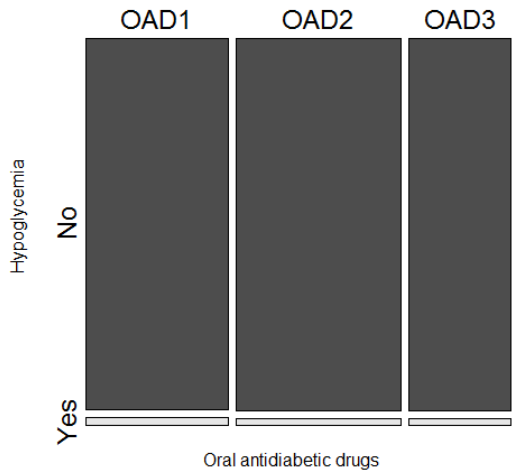

Figure A8: *Proportion of Hypoglycemia by Oral antidiabetic drugs*

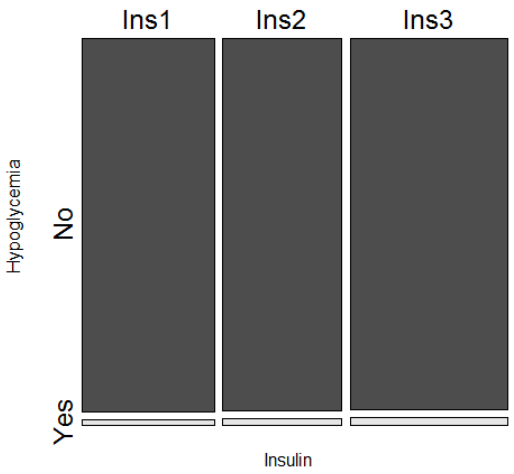

Figure A9: *Proportion of Hypoglycemia by Insulin*

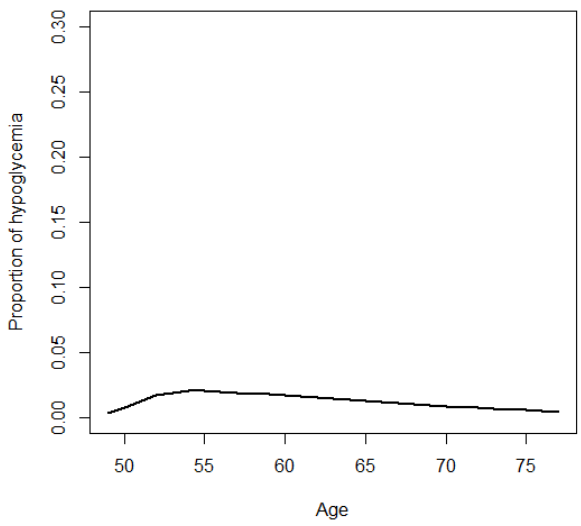

Figure A10: *Proportion of Hypoglycemia by Age*

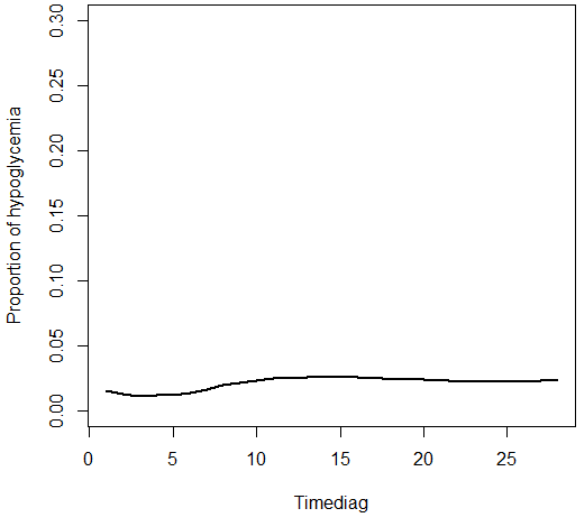

Figure A11: *Proportion of Hypoglycemia by Timediag*

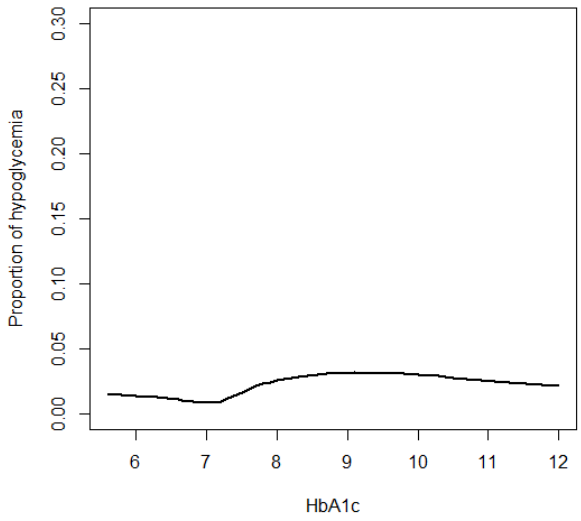

Figure A12: *Proportion of Hypoglycemia by HbA1c level*

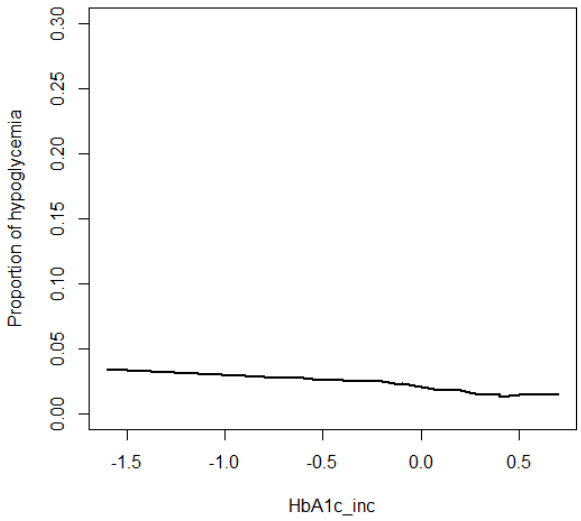

Figure A13: *Proportion of Hypoglycemia by* $HbA1c_{inc}$

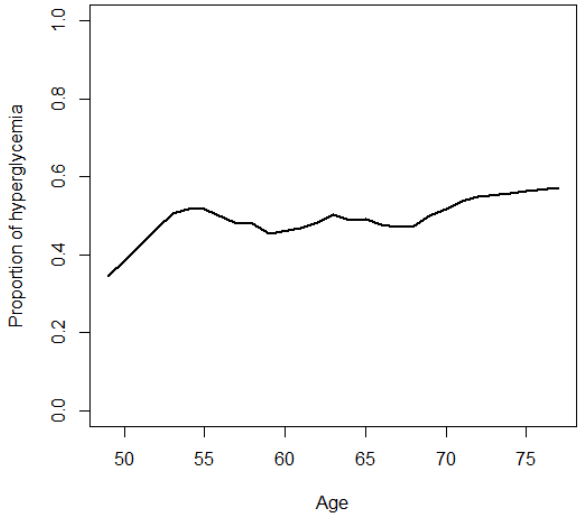

Figure A14: *Proportion of Hyperglycemia by Age*

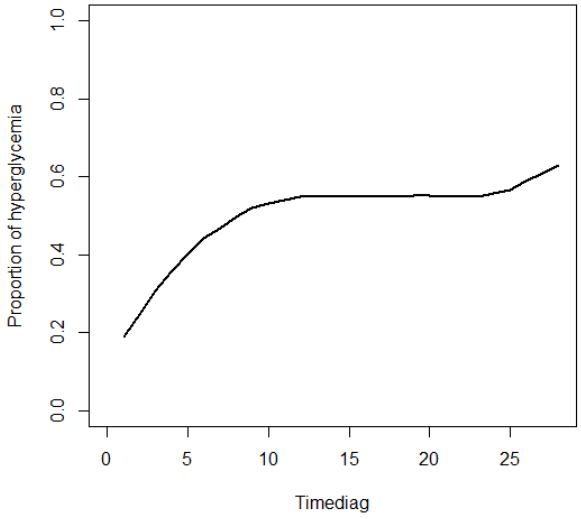

Figure A15: *Proportion of Hyperglycemia by Timediag*

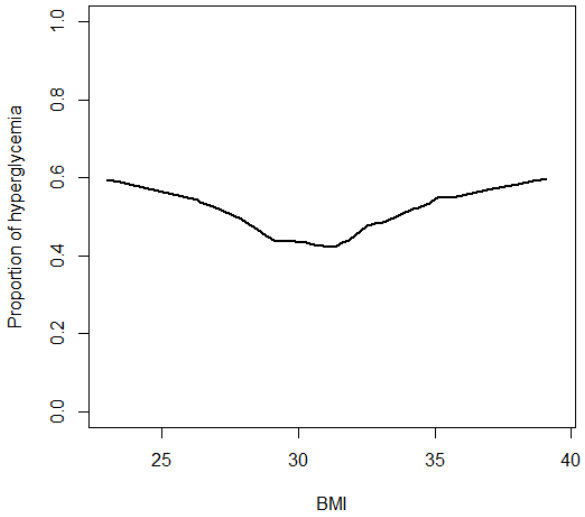

Figure A16: *Proportion of Hyperglycemia by Body mass index*

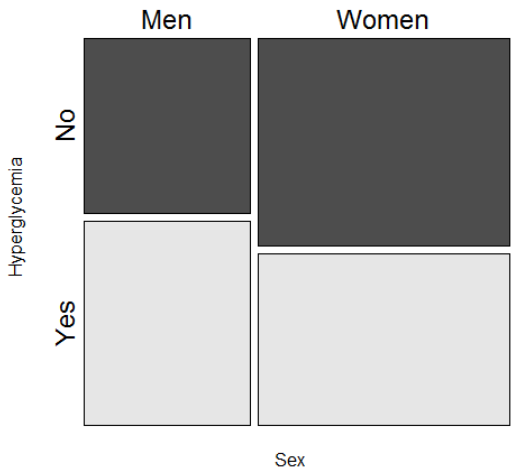

Figure A17: *Proportion of Hyperglycemia by Sex*

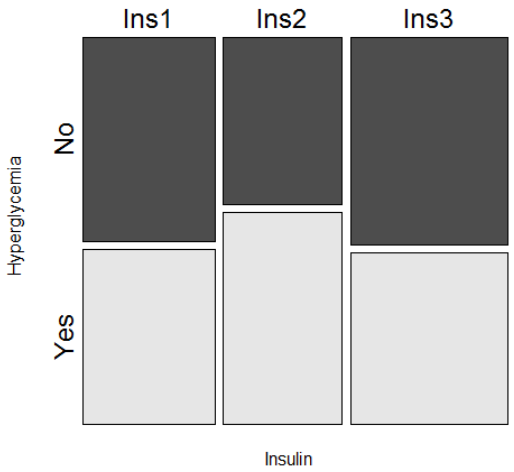

Figure A18: *Proportion of Hyperglycemia by Insulin*

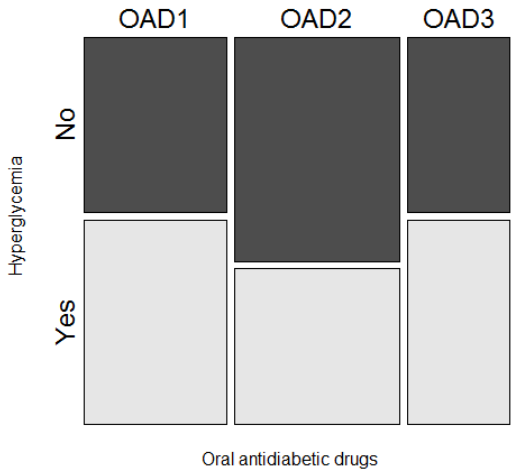

Figure A19: *Proportion of Hyperglycemia by Oral antidiabetic drugs*

B Tables

Table A1: *Schedule to perform capillary blood glucose*

Day	Pre-breakfast	Post-breakfast	Pre-lunch	Post-lunch	Pre-dinner	Post-dinner
1	X	X				
2			X	X		
3					X	X
4	X	X	X	X	X	X
5	X	X				
6			X	X		
7					X	X

C Generalized Linear Models

It is a unifying framework for exponential families [9]. A random variable follows this family if it's density is of the form:

$$f(y) \equiv f(y|\eta,\phi) = exp\{\phi^{-1}[y\eta - \psi(\eta)] + c(y,\phi)\}, \tag{C.1}$$

for a specific set of unknown parameters η and ϕ often called natural (accomodating a function of covariates, $\mathbf{x}'\zeta$) and scale parameters, respectively, and for known functions $\psi(.)$ and c(.,.). The first two moments are given by [8],

$$E(Y) = \mu = g(\eta) = \psi'(\eta), \tag{C.2}$$

$$Var(Y) = \sigma^2 = \phi\psi''(\eta) = \phi\psi''[\psi'^{-1}(\mu)] = \phi v(\mu), \tag{C.3}$$

where g is a function that links μ and η, $\psi'(.)$ and $\psi''(.)$ are the first and second derivatives, respectively, and $v(.)$ is the variance function.

For binary outcomes, the natural link function is logit(.) or $\eta = log[\pi/(1-\pi)]$, and $v(\mu) = \pi(1-\pi)$, and $\phi = 1$, with success probability $\mu = \pi$. Whereas for normal case, the mean-variance relationship is absent since $v(\mu) = 1$ and:

$$\hat{\phi} = \hat{\sigma^2} = \frac{1}{N-p}\sum_i (y_i - x_i\hat{\beta})^2 \tag{C.4}$$

This shows that the variability in binary data might not be explained properly or it is overly restrictive , since it is deterministic function of the mean, unless the observations are identically and independentily distributed. Hence, a simple approach is to use ϕ as a correction term by deviating it from 1. The two stage approach, is another, assuming $Y_i|\pi_i \sim Bernoulli(\pi_i)$ and further π_i is a random variable with $E(\pi_i) = \mu_i$ and $Var(\pi_i) = \sigma_i^2$. But, the later approach can be done if hierarchies are present in the data. This can be observed by the mean and variance of Y_i, which can be calculated using iterated

expectation,

$$\begin{aligned} E(Y_i) &= E[E(Y_i|\pi_i)] = E(\pi_i) = \mu_i, \\ Var(Y_i) &= E[Var(Y_i|\pi_i)] + Var[E(Y_i|\pi_i)] \\ &= E[\pi_i(1-\pi_i)] + Var(\pi_i) \\ &= E(\pi_i) - E(\pi_i^2) + E(\pi_i^2) - E(\pi_i)^2 \\ &= \pi_i(1-\pi_i), \end{aligned} \tag{C.5}$$

indicating that purely Bernoulli data do not capture overdispersion.

We may assume π_i to have a full distribution, the commonly used one is beta distribution, which is a conjugate. The hierarchical and random-effects densities are said to be **conjugate** if and only if they can be written in the generic form [9],

$$f(y|\theta) = exp\{\phi^{-1}[yh(\theta) - g(\theta)] + c(y,\phi)\}, \tag{C.6}$$

$$f(\theta) = exp\{\gamma[\psi h(\theta) - g(\theta)] + c^*(\gamma,\psi)\}, \tag{C.7}$$

where $g(\theta)$ and $h(\theta)$ are functions, ϕ, γ, and ψ are parameters, $c(y,\phi)$ and $c^*(\gamma,\psi)$ are normalizing constants. Then, upon constructing the joint distribution and integrating over the random effects, the marginal model resulting from (C.6) and (C.7) equals,

$$f(y) = exp\left\{c(y,\phi) + c^*(\gamma,\psi) - c^*\left(\phi^{-1} + \gamma, \frac{\phi^{-1}y + \gamma\psi}{\phi^{-1} + \gamma}\right)\right\} \tag{C.8}$$

Further, one of commonly used model for repated measures is GLMM, which captures correlation between repeated measurements and to some extent overdispersion. An alternative to the Bernoulli model with logit link is the probit model, *i.e.* $\eta = \Phi^{-1}(\pi)$, where Φ is the standard normal cumulative distribution function. This link has an appealing feature in the overdispersed and/or repeated contexts. Although probit model do not satisfy strong conjugacy, it has a nice analytical expression [9]. **Strong conjugacy** refers to conjugacy conditional upon the normally distributed random effect b_i, where the

hierarchical and random-effects densities will have generic forms [9],

$$f(y|\kappa\theta) = exp\{\phi^{-1}[yh(\kappa\theta) - g(\kappa\theta)] + c(y,\phi)\}, \tag{C.9}$$

generalizing (C.6), and retain (C.7). Upon applying the transformation theorem to (C.7) leads to

$$f(\theta|\gamma,\psi) = \kappa.f(\kappa\theta|\tilde{\gamma},\tilde{\phi}).$$

Next, requesting the parametric form (C.7):

$$f(\kappa\theta) = exp\{\gamma^*[\psi^* h(\kappa\theta) - g(\kappa\theta)] + c^{**}(\gamma^*,\psi^*)\}, \tag{C.10}$$

where the parameters γ^* and ψ^* follow $\tilde{\gamma}$ and $\tilde{\psi}$ upon absorption of κ, and $c^{**}(.,.)$ is the corresponding normalizing function. Then, the marginal model, in analogy with (C.8), equals:

$$f(y|\kappa) = exp\left\{c(y,\phi) + c^{**}(\gamma^*,\psi^*) + c^{**}\left(\phi^{-1}+\gamma^*, \frac{\phi^{-1}y + \gamma^*\psi^*}{\phi^{-1}+\gamma^*}\right)\right\}. \tag{C.11}$$

D SAS Code

The following code is for hypoglycemia.

D.1 SL

```
proc genmod data=t.fs descending;
title 'Logistic-Bernoulli GLM';
class mom(ref='1') tto(ref='1') ado(ref='0');
model hipo = mom sage tto sbmi sa1c sinc std sexo ado / dist=binomial ; run;
```

D.2 GLMM

```
proc glimmix data=t.fs method=quad(qmax=10);
class id  mom(ref='1') tto(ref='1') ado(ref='0') ;
model hipo(event='1')= mom sage tto sbmi sa1c sinc std sexo ado /
dist=binary solution;
random intercept /subject=id type=un; run;
```

D.3 BB

```
proc nlmixed data=t.fs qpoints=10;
title 'Beta-binomial';
parms Beta_0=0 Beta_2=0 Beta_3=0 Beta_4=0 Beta_5=0 Beta_6=0 Beta_7=0
 Beta_9=-0 Beta_10=0 Beta_11=0 Beta_12=0 Beta_13=0 Beta_14=0 Beta_15=0 Beta_16=0 Beta_17=0  ;
eta=Beta_0 + Beta_2*(mom=2) + Beta_3*(mom=3) + Beta_4*(mom=4) + Beta_5*(mom=5) + Beta_6*(mom=6) +
Beta_7*sage + Beta_9*(tto=2)+ Beta_10*(tto=3) +Beta_11*sbmi+ Beta_12*sa1c+ Beta_13*sinc +
Beta_14*std + Beta_15*sexo +Beta_16*(ado=1) + Beta_17*(ado=2) ;
expeta = exp(eta);
ll = -log(1+const) + hipo*eta - hipo*log(1+expeta)+ (1-hipo)*log((1-expeta/(1+expeta)) + const);
model hipo ~ general(ll);
run;
```

D.4 CM

```
*** glmm as initial;
proc nlmixed data=t.fs qpoints=10 ;
title 'CM';
parms Beta_0=-4.7877 Beta_2=-0.3355 Beta_3=0.4343 Beta_4=-0.4590 Beta_5=-0.5480 Beta_6=-1.4075
 Beta_7=-0.2791   Beta_9=0.6530 Beta_10=0.1921 Beta_11=-1.1794 Beta_12=-0.1997 Beta_13=-0.06521
  Beta_14=0.1246 Beta_15=-0.5492 Beta_16=-0.3937 Beta_17=0.2844 const=7.8517 sigma=1.2191;
eta=Beta_0 +b+ Beta_2*(mom=2) + Beta_3*(mom=3) + Beta_4*(mom=4) + Beta_5*(mom=5) +
Beta_6*(mom=6) + Beta_7*sage + Beta_9*(tto=2)+ Beta_10*(tto=3) +Beta_11*sbmi+ Beta_12*sa1c+
Beta_13*sinc + Beta_14*std + Beta_15*sexo +Beta_16*(ado=1) + Beta_17*(ado=2) ;
expeta = exp(eta);
ll = -log(1+const) + hipo*eta - hipo*log(1+expeta)+ (1-hipo)*log((1-expeta/(1+expeta)) + const);
model hipo ~ general(ll);
random b ~ normal(0,sigma*sigma) subject = id;
run;
```

D.5 COMMM

```
proc nlmixed data=t.fs qpoints=10  tech=newrap;*Final model;
title'Commm';
bounds const>0;
parms Beta_0=-5.8591 Beta_2=0.6016 Beta_3=1.3408 Beta_4=0.5525 Beta_5=0.6726 Beta_6=-0.2843
Beta_11=-2.1852 Beta_18=0.9677 Beta_19=0.9263 Beta_20=1.0704 Beta_21=1.4258 Beta_22=1.2406
const=1.5 sigma=1.6872;
eta=Beta_0+b + Beta_2*(mom=2) + Beta_3*(mom=3) + Beta_4*(mom=4) + Beta_5*(mom=5) +
Beta_6*(mom=6)+ Beta_11*sbmi +  (Beta_18*(mom=2)+ Beta_19*(mom=3)+ Beta_20*(mom=4)+
Beta_21*(mom=5)+  Beta_22*(mom=6))*sbmi;
pi_m=1/(1+exp(-eta));
delta = sqrt(1+(sigma*sigma)) * probit((1+const)*pi_m);
eta_c = delta + b;
pi_c = probnorm(eta_c);
ll =-log(1+const)+hipo*log(pi_c)+(1-hipo)*log((1-pi_c)+const);
model hipo ~general(ll);
RANDOM b ~NORMAL(0,sigma*sigma) SUBJECT=id; run;
```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Incidence and risks for hypoglycemic episodes in stable type 2 diabetic patients<u></u>

Richting: **Master of Statistics-Biostatistics**
Jaar: **2014**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Ibrahim, Mohammed A.

Datum: **10/09/2014**