Measurement Variability Related to Insulin Secretion and Sensitivity

Assessment and Implications in Epidemiological Studies

LARS BERGLUND

Dissertation presented at Uppsala University to be publicly examined in Enghoffsalen, Akademiska sjukhuset, Uppsala, Wednesday, April 29, 2009 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in Swedish.

Abstract
Berglund, L. 2009. Measurement Variability Related to Insulin Secretion and Sensitivity. Assessment and Implications in Epidemiological Studies. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 440. 67 pp. Uppsala. ISBN 978-91-554-7469-0.

There is a growing interest in random measurement variability of biological variables. In regression models, such variability of the predictors yields biased estimators of coefficients (regression dilution bias). The objectives of this thesis were to develop an efficient method to correct for such bias, to reveal the relative importance of insulin sensitivity and insulin secretion, corrected for regression dilution bias, on glucose tolerance, and to explore the seasonal nature of the variability of insulin sensitivity.

A reliability study is often designed to randomly select subjects from the main study. Our idea was to collect replicates for subjects with extreme values on their first measurement. The extreme selection design, in combination with maximum likelihood estimation, resulted in an efficient estimator of a corrected regression coefficient in a simple linear regression model. Results were presented theoretically and with an application: The relation between insulin sensitivity and fasting insulin in Uppsala Longitudinal Study of Adult Men (ULSAM) where the extreme selection design decreased the standard error of the estimated regression coefficient with 28 per cent compared with the random sampling design.

We estimated the partial longitudinal effects of the predictors insulin sensitivity and insulin secretion, corrected for regression dilution bias, on glucose tolerance in ULSAM. The effects of the predictors, when corrected, were similar.

Insulin sensitivity in ULSAM increased during summer and decreased during winter and insulin secretion exposed opposite variation keeping glucose homeostasis nearly constant. Insulin sensitivity was related to outdoor temperature.

In summary, we developed a cost-efficient reliability design for correction for regression dilution bias. Insulin sensitivity and insulin secretion had similar longitudinal effects on glucose tolerance, which implies that interventions aimed at these targets are equally important. Further, we revealed the seasonal nature of variations of insulin sensitivity and insulin secretion. This result has implications on glycaemic control in diabetic patients.

Keywords: measurement variability, insulin sensitivity, insulin secretion, reliability study, bias in regression coefficients, extreme selection, seasonal variation

Lars Berglund, Department of Public Health and Caring Sciences, Geriatrics, Box 609, Uppsala University, SE-751 85 Uppsala, Sweden

To Birgitta,
my mother Anna-Lisa,
Anders, Julia and Saga,
Emma and Tomas

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals:

I   Berglund L, Garmo H, Lindbäck J, Zethelius B. Correction for regression dilution bias using replicates from subjects with extreme first measurements. *Statistics in Medicine* 2007; **26**:2246–2257.

II Berglund L, Garmo H, Lindbäck J, Svärdsudd K, Zethelius B. Maximum likelihood estimation of correction for dilution bias in simple linear regression using replicates from subjects with extreme first measurements. *Statistics in Medicine* 2008; **27**:4397–4407.

III Berglund L, Berne C, Svärdsudd K, Garmo H and Zethelius B. Early insulin response and insulin sensitivity are equally important as predictors of glucose tolerance after correction for measurement errors. Submitted.

IV Berglund L, Berne C, Svärdsudd K, Garmo H and Zethelius B. Seasonal variations of insulin sensitivity are compensated by variations of insulin secretion and are related to outdoor temperature. Manuscript.

Contents

Abbreviations

AIR   Acute insulin response
ANOVA   Analysis of variance
BMI   Body mass index
CI   Confidence interval
CV   Coefficient of variation
CT   Clinical trial
DBP   Diastolic blood pressure
EIR   Early insulin response
FPG   Fasting plasma glucose
HbA1c   Haemoglobin A1c
HOMA-IR   Homeostasis model assessment-estimated insulin resistance
ICC   Intraclass correlation coefficient
IGT   Impaired glucose tolerance
IVGTT   Intravenous glucose tolerance test
IRI   Immunoreactive insulin
MLE   Maximum likelihood estimator
M/I   Insulin sensitivity index
OGTT   Oral glucose tolerance test
OLS   Ordinary least squares
OR   Odds ratio
RBE   Regression based estimator
SE   Standard error
SD   Standard deviation
T2DM   Type 2 diabetes mellitus
ULSAM   Uppsala longitudinal study of adult men

Notations

Y	The continuous response (dependent) variable
x	The continuous true predictor (independent) variable
μ	Mean of x
σ_{xx}	Variance of x
β_{1c}	Slope in linear regression of Y on x
$\rho_{x,Y}$	Correlation between x and Y
u	Measurement error of predictor variable
σ_{uu}	Variance of u
CV	Coefficient of variation, $CV = \frac{\sqrt{\sigma_{uu}}}{\mu}$
ρ	Reliability ratio $\rho = \frac{\sigma_{xx}}{\sigma_{xx} + \sigma_{uu}}$
X	The observed predictor variable, $X = x + u$
β_1	Slope in linear regression of Y on X, $\beta_1 = \rho \beta_{1c}$
n	Number of participants in main study
k	Number of participants in reliability study
p	Fraction of participants selected to reliability study, $p = k/n$
X_1	Measurement of X in main study
X_2	Measurement of X in reliability study
b	Estimated slope in linear regression of X_2 on X_1
$x \sim N(\mu, \sigma_{xx})$	x is normally distributed with mean μ and variance σ_{xx}

Introduction

In recent years there has been a growing interest in the implication of random measurement variability, i.e. measurement errors, of biological variables (see e.g. [1], [2], [3], [4] and [5]). This thesis deals with the presence of errors in measurement of continuous variables in biology with special reference to variables related to insulin secretion and insulin action in humans. Repeated measurements on the same individual will vary around the usual value because of measurement error. Measurement error is defined as the deviation between an observed value and a usual value. The usual value can be conceived as an individual's long-term average value. With this definition the measurement error can be divided into technical error due to an imprecise measurement method (e.g. a food frequency questionnaire) and the individuals's true biological variation over time (e.g. seasonal variation of intake of vitamins). The size of measurement error can be assessed with a validation study where observed values using an imprecise method are compared with observations from a gold-standard method without error *or* with a reliability study where observations are replicated with the same method.

The implications of measurement errors are twofold:
- (i) in, e.g. a clinical trial the required number of patients increase with the measurement error magnitude,
- (ii) measurement errors yield biased estimation of coefficients when regression models or correlations are estimated [1]

The second of these problems is a theme of these studies.

As a motivating example the relation between insulin sensitivity, measured with the expensive and labour-intensive euglycaemic insulin clamp technique [6], and fasting insulin is studied. The latter is measured with noticeable error [7] while insulin sensitivity is measured with low error ([8] and [9]).

A single fasting insulin measure is subject to random fluctuations, due partly to the measurement technique and partly to any real but temporary deviations from the usual fasting insulin level. The distribution of single measures is therefore wider than the distribution of true usual values. The term for the resulting underestimation of a predictor variable's impact on a response variable is regression dilution bias [10].

If insulin sensitivity and fasting insulin are related to each other in a regression model where fasting insulin is measured once for each study participant the regression dilution bias will yield an underestimation of the risk

for insulin resistance for a high long-term average of fasting insulin (see Figure 1). This underestimation would be smaller with two or more measurements of fasting insulin for all participants and by use of the average of these values in the regression model. A more cost-efficient approach is to select a fraction of the participants for a replicate measurement of fasting insulin and use the data from these participants to correct the regression coefficient for the measurement error in fasting insulin.

Seasonality is an important source of biological variation. In this thesis we study systematic seasonal variations of insulin sensitivity.

Reliability studies

Reliability of a measurement method of a continuous variable is the similarity of repeated measurements administered on the same individual. The amount of measurement error is the variation seen over such repeated observations. When one sample from an individual is measured repeatedly or an individual is measured repeatedly with very short time intervals the variation is denoted technical measurement error. If an individual is measured repeatedly over two or more occasions (e.g. with intervals of one week or one month) the resulting variation is the total measurement error which is the sum of technical measurement error and biological variation.

Depending upon the expected relative size of the technical measurement error and biological variation and the scope of the study a reliability study can be designed in different ways. The present studies are concerned with the total measurement error and are concentrated on the simple reliability design where a fraction of the participants in the main study are selected for a replicate measurement, e.g. a number of weeks after the first measurement. If the technical measurement error is expected to dominate the total measurement error the design is modified so that replicate measurements are analyzed of drawn samples from a fraction of the participants in the main study.

The first measurement of the variable with measurement error for all n participants in the main study is denoted X_1 and the second measurement of the k participants selected for the reliability study is denoted X_2.

If it is not feasible to re-measure all participants in the main study, the style of selection of participants to a reliability study is important. The recommendation is usually to select a random sub-sample or at least a representative sample from the main study. In this thesis another and more efficient style of selection is introduced.

Response models and measurement error models

These studies examine correction for regression dilution bias in linear regression models. The simplest regression model is the structural response
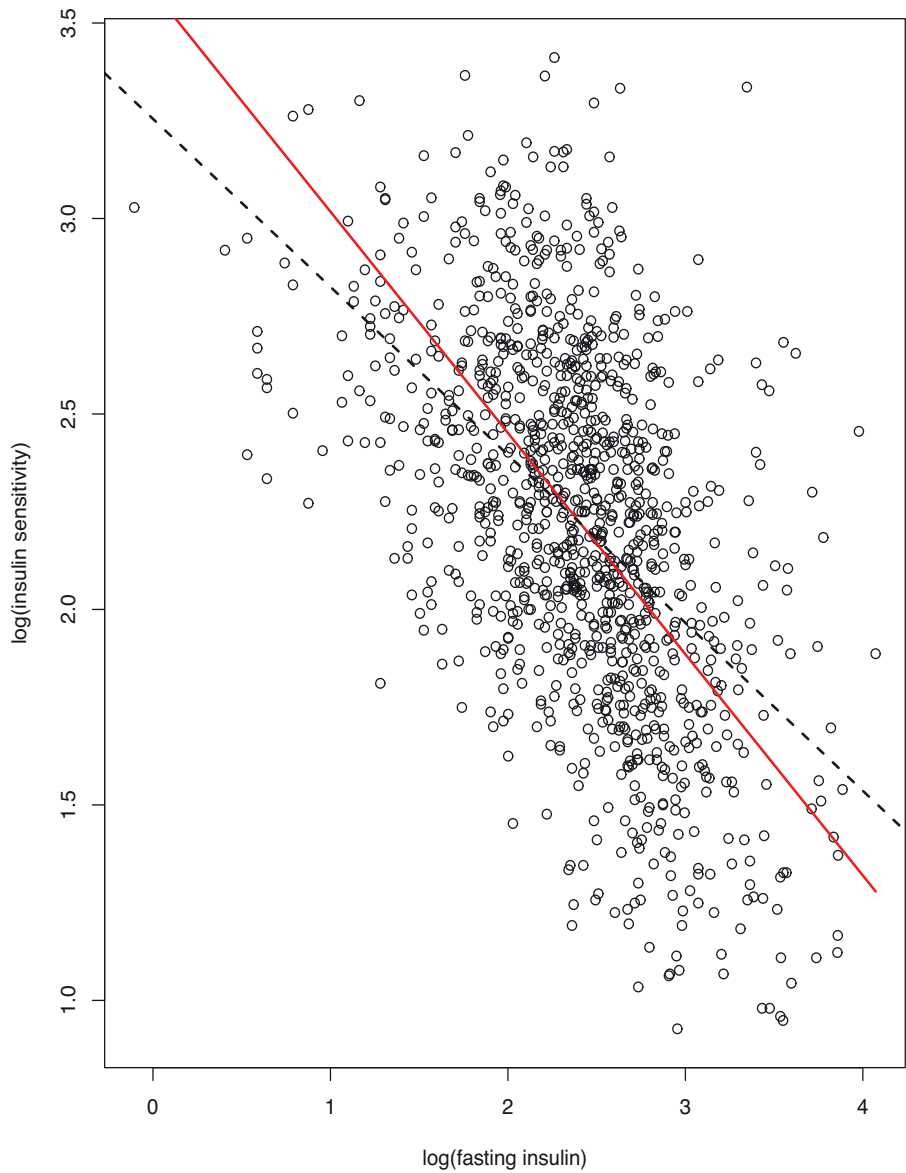
Figure 1: Log(insulin sensitivity *vs.* log(fasting insulin). Dashed black line is ordinary regression line and solid red line is regression line corrected for measurement error in measurement of fasting insulin.

model relating the response variable Y to one random predictor variable x in a linear fashion:

$$Y = \beta_{0c} + \beta_{1c}x + \delta, \tag{1}$$

where $\delta \sim N(0, \sigma_{\delta\delta})$, $x \sim N(\mu, \sigma_{xx})$ and δ and x are independent.

The model postulates that the x value is selected randomly from a normal distribution with mean μ and variance σ_{xx} and that Y conditional on x is selected randomly from a normal distribution with mean $\beta_{0c} + \beta_{1c}x$ and variance $\sigma_{\delta\delta}$. The expected effect on Y of a one unit increase of x is thus β_{1c} units. An example of this model is when Y is insulin sensitivity and x is fasting insulin. The interest is then in the change β_{1c} of insulin sensitivity for every pmol/l increase of fasting insulin concentration.

Model (1) is estimated without bias when the x values are measured without errors. When this is not the case a model for the measurement error structure must be assumed. The most common model is the classical measurement error model (see e.g. [4], p. 3):

$$X = x + u, \tag{2}$$

where $u \sim N(0, \sigma_{uu})$ and independent of δ and x.

Here, x is the unobservable true value of e.g. fasting insulin and u is a normally distributed measurement error and thus X is the predictor variable measured with error. The classical model assumes non-differential errors, i.e. the distribution of X conditional on the distribution of x gives no information on the distribution of Y.

Measures of reliability of a continuous variable

We assume that the design is such that all participants in the reliability study have two measurements of the variable with measurement error. The level of reliability is usually summarized in one measure. We will briefly discuss
- (i) the coefficient of variation (CV)
- (ii) the intra-class correlation coefficient (ICC)
- (iii) the slope b in the linear regression of the second measurement on the first

CV has dominated the presentation of such data. The CV is defined as the standard deviation of the differences of the first and second measurements divided by the mean of the mean values of the first and second measurements. The idea behind the use of CV as a measure of reliability is that the mean and the intra-individual standard deviation of a variable increases proportionally. This idea is contrary to the classical measurement error model. When a variable displays a dependence

between intra-individual variations and means, this dependence is removed with, e.g. a logarithmic transformation.

Another approach to summarize reliability data is to calculate the ICC [11]. ICC is defined as the ratio of the between-individuals variation and the total variation (which is the sum of between-individuals variation and measurement error). Hence, ICC is in the interval 0 to 1. ICC is estimated from a one-way analysis of variance (ANOVA) model where individual is the factor.

ICC is an unbiased estimator of ρ for a continuous variable when the style of selection of participants to the reliability study is random sub-sampling. If some other style of selection has been advocated, ICC will be a biased estimator. A simple example illustrates this phenomenon: suppose that the between-individuals variation and the measurement error are 1 unit each and thus ICC is $\frac{1}{2}$. If selection to the reliability study is such that subjects with extreme values are disregarded, the between-individuals variation will be underestimated (say 0.5 instead of 1). According to the classical measurement error model the measurement error variance is unchanged and the ICC will be biased downwards (in this example it would be $\frac{1}{3}$).

Yet another measure of reliability is the regression coefficient (the slope b) in the linear regression of the second measurement X_2 on the first X_1 [11]. When random sub-sampling is used the ICC and b estimates the same population quantity ρ. ICC is a slightly more efficient estimator than b especially for variables with high reliability. If selection is not random ICC is biased as illustrated above. b is an unbiased estimator regardless of selection style based on the first measurement [12].

Regression dilution bias

When Y and X are measured the model

$$Y = \beta_0 + \beta_1 X + \varepsilon, \tag{3}$$

where $\varepsilon \sim N(0, \sigma_{\varepsilon\varepsilon})$, is estimated[13].

The reliability ratio is defined as $\rho = \frac{\sigma_{xx}}{\sigma_{xx}+\sigma_{uu}}$. It is well known (see e.g. Fuller [1] p. 3) that

$$\beta_1 = \rho \beta_{1c}. \tag{4}$$

Thus, the ordinary least squares estimator of the slope is biased towards zero. The phenomenon is called regression dilution bias. In the example this bias yields an underestimation of the change β_{1c} of insulin sensitivity for every pmol/l increase of fasting insulin when the latter variable is measured with random error.

If ρ were known, according to (4), an estimate of β_{1c} would be $\hat{\beta}_{1c} = \hat{\beta}_1/\rho$. In this case the standard error of $\hat{\beta}_{1c}$, $se(\hat{\beta}_{1c})$, is $se(\hat{\beta}_1)/\rho$ [1].

When ρ is unknown it can be estimated from a reliability study. The estimator is denoted $\hat{\rho}$ which is either the intra-class correlation coefficient ICC or the regression coefficient in the linear regression of the second measurement X_2 on the first X_1 (b). The estimator of β_{1c} is then $\hat{\beta}_{1c} = \hat{\beta}_1 / \hat{\rho}$ [1]. Other estimators of β_{1c} are described in the literature [11]. The standard error of $\hat{\beta}_{1c}$ is complex and depends on the reliability design, the uncertainty of the uncorrected estimator $\hat{\beta}_1$, the uncertainty of the reliability ratio $\hat{\rho}$ and the covariance of $\hat{\beta}_1$ and $\hat{\rho}$ [14].

In paper I the slope b, in the linear regression of the second measurement X_2 on the first X_1, was used as estimator of ρ and the resulting estimator of β_{1c} was called the regression based estimator. In paper II another estimator of β_{1c} based on the maximum likelihood method was examined.

Regression calibration

Several methods are available for correction of regression dilution bias in general regression models, i.e. simple and multiple linear, logistic or proportional hazards models [4]. One of the most powerful and easily adapted general methods is the regression calibration method. The method is proposed by Armstrong [15] for generalized linear models and Rosner *et al.* [16] use the method for correction of logistic regression models and Prentice [17] applies the method to proportional hazards models. The regression calibration algorithm is suggested as a general approach by Carroll and Stefanski [18] and Gleser [19].

Regression calibration is a statistical method for adjusting point and interval estimates of effects obtained from regression models for bias due to predictor measurement errors. The method is appropriate when a gold standard is available in a validation study or when replicate measurements are available in a reliability study.

Validation or reliability data are used to obtain estimated true predictor values for all participants in the main study. These estimated true predictor values are used instead of the measurement error prone observed predictor values in ordinary regression estimation methods (e.g. OLS for linear models). Standard errors and confidence intervals and p values for tests of null hypotheses of zero effects are usually assessed with the bootstrap method [20]. The Appendix contains a description of the regression calibration method with application to the analysis in Paper III, i.e. linear and logistic regression models with three continuous predictor variables.

Diabetes mellitus

The term diabetes mellitus is used to describe a variety of metabolic disorders characterized by elevated blood glucose levels. The hormone insulin, which is produced in the pancreatic β-cells, plays a central role in diabetes

mellitus. Insulin is a peptide hormone and the main regulator of glucose uptake in muscle, liver and fat cells. An insufficient production and/or response to insulin will therefore lead to hyperglycemia. Even when treated, the disease often leads to more serious long time complications such as nephropathy, nerve damage, cardiovascular disease and retinopathy.

Diabetes can broadly be classified into two main types. Type 1 diabetes, which represents approximately 5-10 % of all cases of diabetes, is an autoimmune disease resulting in destruction of the insulin-producing β-cells located in the pancreatic islets of Langerhans, and Type 2, which is estimated to represent 90 % of all cases, is due to β-cell failure or various degrees of insulin resistance.

Type 1 diabetes usually has its onset before adulthood, whereas Type 2 diabetes most often develops in the middle aged and in the elderly.

In the early stages of Type 2 diabetes mellitus (T2DM) the muscle and fat cells become non responsive to insulin (insulin resistant), and blood glucose levels increase. The pancreas responds by making more insulin. Insulin resistant individuals have high blood levels of both insulin and glucose. Eventually, however, the insulin-producing cells in the pancreas start to malfunction, insulin secretion decreases, and frank diabetes develops. The diagnostic criteria of T2DM according to the World Health Organization is fasting plasma glucose > 7.0 mmol/l or > 11.1 mmol/l measured two hours after an oral glucose tolerance test (OGTT) [21]. Left untreated T2DM will result in severe complications due to the effect of chronic hyperglycemia. The complications include an overall increased risk for cardiovascular disease, retinopathy that can lead to blindness and nephropathy that progress until the kidneys fail completely. During the last century there has been a dramatic increase in the incidence of T2DM world wide, to the point that T2DM is referred to as an epidemic [22]. It is rapidly becoming one of the largest common diseases in the world. Today, more than 230 million people have T2DM worldwide and by the year 2025 numbers are believed to reach 350 million (http://www.idf.org 2007). The dramatic rise in T2DM incidence is mainly attributed to changes in human behavior and lifestyle leading to increased obesity [23]. The best way to deal with this epidemic is prevention and several studies have inferred that lifestyle intervention can have great success in preventing the development of T2DM in individuals with impaired glucose tolerance (IGT) which is a pre-stage to full T2DM ([24] and [25]).

This thesis concerns T2DM and the risk factors insulin resistance and β-cell dysfunction.

Insulin variability

Insulin is secreted from the β-cells, located in the pancreatic islets of Langerhans, in a pulsatile manner resulting in detection of high-frequency insulin concentration oscillations in the peripheral circulation ([26]

and [27]). These high frequency oscillations are caused by inter islet coordinated insulin secretory bursts, at a frequency of 5–15 min per pulse ([28], [29], [26], [30] and [31]). The contribution of these insulin secretory bursts to overall insulin secretion has been quantified in a canine model by direct sampling across the pancreas [31] and in a human model employing high-frequency sampling, a highly specific insulin assay, and validated deconvolution analysis [32]. In both species, the contribution of pulsatile insulin secretion is at least 70–75%.

Accordingly, variability in measurement of insulin is expected both in fasting and post-prandial states. In a study of within-individuals variation over 12 consecutive days of fasting insulin Widjaja *et al.* [33] find that the analytical CV is 6.6% with the RIA method (PhRIA100, Pharmacia Ltd, Uppsala, Sweden) and the within-individuals biological variation is 26%. Poulsen and Jensen [34] report an analytical CV of 7.5% for insulin determined with the ELISA method (Dako Cytomation, Copenhagen, Denmark).

This thesis corrects for the random variability of insulin measurements in the fasting state (Papers I and II) and the early insulin response (EIR) from an OGTT.

The EIR was defined as the ratio of the 30 min change in insulin concentration to the 30 min change in glucose concentration after oral glucose loading. The size of measurement error of EIR and insulin sensitivity contribute to the uncertainty of their relative impact on disease progression. No previous studies address the problem of measurement error and the magnitude of its implication on regression dilution bias for bivariable models with insulin sensitivity measured with the gold-standard clamp technique and EIR, when evaluating long-term effects on plasma glucose concentration or glucose tolerance (Paper III).

Seasonal variations

A variety of biological systems display fluctuations by season of the year. Humans living at high latitudes are exposed to changing patterns of diet, physical activity, light exposure, and outdoor temperature. These populations have seasonal rhythms for cerebral [35] and myocardial infarct ([35] and [36]), mood disorders [37], blood pressure ([38] and [39]), serum cholesterol [40], calcium metabolism [41], growth hormone [42], female gonodal hormone patterns [43], and thyroid hormones ([44], [45], [46], [47] and [48]).

The incidences of Type 1 and Type 2 diabetes mellitus reveal seasonal variations with peaks during the winter months ([49] and [50]). Seasonal variations of HbA1c in diabetic patients ([51], [52], [53] and [54]) and of fasting plasma glucose (FPG) in healthy individuals ([55] and [56]) are reported but it is not known if these are due to seasonal variations of insulin sensitivity. Results are inconclusive with some studies demonstrating a seasonal

effect on insulin sensitivity with an increase of sensitivity during the warm season ([57] and [58]) while other studies ([45], [59] and [60]) do not find this variation. A study by Bunout *et al.* [61] reports an opposite seasonal effect with decreased insulin sensitivity during the warm season in healthy elderly people. While most of these studies use a repeated measures design there are limitations in the number of participants. Further, the studies are confined by the use of surrogate measures of insulin sensitivity based on fasting values of insulin or on homeostasis model assessment-estimated insulin resistance (HOMA-IR). In paper IV seasonal variations of insulin sensitivity measured with the gold-standard euglycaemic insulin clamp technique and the surrogate marker (HOMA-IR) were studied.

Aims of the studies

In these studies, aspects of measurement variability, within the field of insulin secretion and insulin action, were investigated. The overall aim of the thesis was to apply methods for regression dilution bias and for description of measurement variability in the field of type 2 diabetes.

The specific aims of the studies were:

to develop a novel design for a reliability study in order to efficiently estimate corrected regression coefficients in simple linear regression models with application to the relation between insulin sensitivity and fasting insulin (Papers I and II),

to estimate the bivariate regression models between the response variables fasting glucose and HbA1c, respectively, and the predictors insulin sensitivity and insulin secretion where the measurement error of the predictors have been taken into account (Paper III),

and to explore if the biological variations of insulin sensitivity, measured with the euglycaemic insulin clamp technique, and insulin secretion are due to seasonality and/or outdoor temperature and how this affects glucose homeostasis (Paper IV).

Material and methods

Participants

Papers I-IV were based on data from the population-based Uppsala Longitudinal Study of Adult Men (ULSAM) (http://www.pubcare.uu.se/ULSAM). All 2841 men born in 1920-1924 and living in the municipality of Uppsala, Sweden, in 1970 were invited to attend a health survey. A total of 2322 men (82 % of those invited), 49 to 51 years of age, participated. The men were traced in the population register, using the individual 10-digit personal identification number given to all Swedish citizens. Men who were still alive and still living in the Uppsala region were invited for re-investigations at ages 60 (number of participants = 1860), 71 (1221), and 77 (839) years. The men who participated in the investigations at age 71 years and/or age 77 years were also invited to a fifth investigation at the age of 82 years (number of participants = 530) [62].

The present studies used data from men who attended the investigations at ages 71, 77 and 82 years.

In Papers I and II data from men who attended the investigation at age 71 years, and had measurements of fasting insulin and insulin sensitivity index from a euglycaemic insulin clamp examination (n = 1139) were used.

Paper III included data from men examined at age 71 years and with measurements of EIR, insulin sensitivity, fasting and 2-h plasma glucose (n = 1128) and the follow-up groups who also had measurements of fasting plasma glucose, HbA1c and attended the age 77 years investigation (n = 673) and the age 82 years investigation (n = 468).

In Paper IV data from men who attended the investigation at age 71 years, were examined between October 1991 and May 1995, had measurements of insulin and glucose from a 2-h OGTT and insulin sensitivity index from a euglycaemic insulin clamp examination (n = 1117) were used.

In Papers III and IV data was used from a reliability study at age 71 years where a subgroup of 20 participants was investigated twice within 4 to 6 weeks to determine the combined effects of biological variation and technical measurement error on insulin sensitivity, HOMA-IR, EIR, incremental area under the insulin curve from an OGTT, fasting and 2-h plasma glucose from an OGTT, body mass index, and waist circumference [63].

All examinations were made at the outpatient clinic for obesity and metabolic diseases at Uppsala University Hospital. The Ethics Committee

at the Faculty of Medicine, Uppsala University, Sweden, approved the study. All participants gave written informed consent.

Data management and software tools

Data were extracted from the ULSAM database (www.pubcare.uu.se/ULSAM) in SAS® format to specific SAS® analysis databases.

Software tools used were Maple® 8.00, SAS® for Windows v.9, R 2.6.2 and LaTeX.

Maple® was used for derivation of mathematical expressions. SAS® was used for data management, descriptive results, Monte Carlo simulations, bootstrap estimations and as a tool to check derived expressions. R was used to produce graphs from data generated by SAS®. This document and Papers I-II were produced with LaTeX.

Clinical measurement methods

Oral glucose tolerance test

In an OGTT at age 71 years, blood samples were drawn immediately before and 30, 60, 90 and 120 min after ingestion of 75 g anhydrous D-glucose dissolved in 300 mL water. Plasma insulin was assayed by using an enzymatic immunological assay (Enzymmun, Boehringer Mannheim, Mannheim, Germany) gauged in an ES300 automatic analyzer (Boehringer Mannheim). Plasma glucose was measured by the glucose dehydrogenase method (Gluc-DH, Merck, Darmstadt, Germany).

The EIR was defined as the ratio of the 30 minutes change in insulin concentration to the 30 minutes change in glucose concentration after oral glucose loading: $(\text{Ins}_{30} - \text{Ins}_0)/(\text{Gluc}_{30} - \text{Gluc}_0)$.

The incremental area under the curve for insulin during the OGTT was calculated with the trapezoidal method using the formula

$$\text{Ins}_{30\text{min}} + 2 * \text{Ins}_{60\text{min}} + 2 * \text{Ins}_{90\text{min}} + \text{Ins}_{120\text{min}} - 6 * \text{Ins}_{0\text{min}}$$

Insulin resistance based on the homeostasis model (HOMA-IR) was computed with the formula: fasting plasma glucose (mmol/l) times fasting serum insulin (mU/l) ([64] and [65]).

Euglycaemic insulin clamp

Insulin-mediated glucose disposal was estimated at age 71 years with a euglycaemic insulin clamp as described by DeFronzo [6], with insulin (Actrapid Human, Novo, Copenhagen, Denmark) infused at a constant rate of 56 mU/body surface area (m^2)/min during 120 minutes. This rate was estimated to suppress hepatic glucose output almost completely also in participants with type 2 diabetes. The target plasma glucose concentration

was 5.1 mmol/l. Insulin sensitivity index (M/I) was calculated as glucose disposal rate (mg glucose infused/(min x kg body weight)) divided by the mean plasma insulin concentration (mU/l), during the last 60 min of the 120 min clamp, and multiplied by 100. The unit for M/I is $100x$ mg x min^{-1}x kg^{-1}/(mU x l^{-1}).

The OGTT and the clamp procedure were separated in time by approximately one week [66].

Anthropometric measurements

At age 71 years, height was measured to the nearest whole centimeter, and body weight to the nearest 0.1 kg. The BMI was calculated as the ratio of the weight (in kilograms) to the height (in meters squared). The waist circumference was measured midway between the lowest rib and the iliac crest.

Energy intake

An optically readable, pre-coded, 7-day food record was completed by 1050 men at the age 71 years investigation, for assessment of habitual dietary intake. The design and validity of the food record used has been discussed previously [67]. The total energy intake (kcal) was calculated as the mean of the intakes over the seven days.

Measurements of temperature

The outdoor temperature in °C was recorded at the Swedish Air Force base (F16), located 4 km north of Uppsala center, using calibrated scales. Data was bought from the Swedish Meteorological and Hydrological Institute (SMHI, Norrköping, Sweden) (http://www.smhi.se) as monthly mean values for each month from August 1991 to May 1995. For each participant the mean temperature of the month for the clamp investigation and the two preceding months, representing the last quarter of the year, was used as the outdoor temperature exposure value.

Statistical methods

Paper I

In Paper I a novel design for a reliability study was developed. Using results from the field of genetic statistics ([68], [69] and [70]) the variances of estimators of a corrected regression coefficient were derived analytically under this novel design and the design of random sub-sampling. The analytical results were compared with Monte Carlo simulations which imply that data were generated according to the simple linear structural regression model

and the classical measurement error model. This was repeated 10000 times and at each time an estimate of the corrected regression coefficient was calculated for both designs. The variances from the Monte Carlo distributions were compared with the analytically derived variances. In a reality-based example on the relation between the response insulin sensitivity and the predictor fasting insulin with data from ULSAM variances from the analytical expressions were compared with variances from the bootstrap method [20].

Paper II

In Paper II an improvement of the estimation of corrected regression coefficients based on the random sub-sampling and the extreme selection design was developed. The regression based estimator [1], used in, e.g. Paper I, does not fully utilize all information in the data. The method of maximum likelihood estimation was used to take advantage of all available data by adapting a method due to Chan and Mak [71] to the design with extreme selection of a sub-group of participants from the main study [72]. Further, profile-likelihood-based confidence intervals for the true regression coefficient [73] were compared with symmetric confidence intervals based on asymptotic normality. The analytical results were compared with Monte Carlo simulations. For each simulation an estimate of the corrected regression coefficient was calculated for both designs and for the regression based estimator used in Paper I and the maximum likelihood estimator. For the maximum likelihood estimation the trust-region method [74], as implemented in SAS®, was used. In addition, in the simulations the effect of non-normal distributions of the true predictor on the estimators was highlighted. The variances from the Monte Carlo distributions were compared with the theoretical analogues. The same reality-based example as in Paper I was used to verify the derived variances with the bootstrap method.

Paper III

In Paper III, the reliability of the predictors M/I and EIR were displayed as intraclass correlation coefficients (ICC) with standard errors [11] and as coefficients of variation (CV) with standard errors. The standard errors for the CVs were calculated with the bootstrap method [20]. Reliability data were inspected in Bland-Altman plots [75] to detect if measurement errors followed a classical model, i.e. if the levels and variances of measurement errors were independent of the levels of the predictor [4]. Measurement error of the predictors M/I and EIR at age 71 years and the measurement errors of the dependent variables were assumed to be independent [76] of each other. In order to meet the assumptions of the regression models all continuous variables were transformed with a logarithmic function except for M/I for which a square root transformation was appropriate.

Associations between the predictors M/I, EIR, and their interaction [77] from the 71 years investigation, and the response variables were examined in linear regression models for continuous response variables (fasting and 2-h plasma glucose from an OGTT at age 71 years and HbA1c and fasting plasma glucose at ages 77 and 82 years), and in logistic regression models for response variables prevalent (age 71 years) and incident (from age 71 to 77 years and from age 71 to 82 years) type 2 diabetes. In the regression models, the partial regression coefficients, uncorrected and corrected with the regression calibration method ([4] and [78]) for the measurement error structure of M/I, EIR, and their interaction, were estimated.

The regression calibration method uses reliability data to obtain estimated true predictor values for all participants in the main study. These estimated true predictor values are used instead of the measurement error prone observed predictor values in ordinary regression estimation methods (ordinary least squares estimation for linear regression models and maximum likelihood estimation for logistic regression models). The regression dilution bias for the estimators of regression coefficients is then removed. With respect to standard errors for the regression calibration corrected estimates, these will be underestimated by ordinary methods as they do not take into account the variance contribution from the reliability study. Since the computation of explicit formulas for the standard errors is tedious, standard errors are typically obtained through bootstrapping [20].

The effects of predictors in models with interactions are difficult to interpret and to illustrate, because the effect of one predictor depends on the level of the other predictor(s). For the linear regression models, the effect of a predictor was estimated as the change in the dependent variable from mean levels for M/I and EIR to a decrease by one standard deviation for the predictor of interest while the other predictor was constant [79]. For the logistic regression models the prevalence or incidence of type 2 diabetes was calculated when M/I and EIR were at mean levels. The effect of a predictor was estimated as the odds ratio to be or become diabetic for one standard deviation decrease from the mean level for the predictor of interest while the other predictor was constant.

The precisions of the estimated effects of M/I, EIR, and their difference were estimated with bootstrap 95 % percentile confidence intervals [20]. P values for the null hypotheses of no differences between EIR and M/I effects were assessed with the bootstrap method [20].

For the effects on HbA1c, a pre-specified non-inferiority margin of 0.3 % ([80] and [81]) was used. Non-inferiority of the EIR versus the M/I measurement error corrected effect was declared when the upper limit of the bootstrap 95 % percentile confidence interval for the difference between measurement error corrected effects of the predictors was less than 0.3 %.

A p-value of less than 0.05 was considered a statistically significant result.

Paper IV

All continuous variables were summarized with number of observations and mean (standard deviation) for winter (October-April) and summer (May-September) season and for the whole year in Paper IV. The difference between the means of the winter and the summer seasons was expressed in % of the whole year mean.

The reliability of M/I, HOMA-IR, the incremental area under the insulin curve OGTT, fasting plasma glucose, 2 h plasma glucose OGTT, BMI, and waist circumference were displayed as intraclass correlation coefficients (ICC) with standard errors [11] and as coefficients of variation (CV).

In order to meet the assumptions of the regression models fasting plasma glucose, 2-h glucose OGTT, and BMI were transformed with a logarithmic function while M/I and the incremental area under the insulin curve were transformed with the square root function. Values from October 1991 to May 1995 for the continuous dependent variables M/I, HOMA-IR, the incremental area under the insulin curve, FPG and 2 h glucose OGTT and the predictor variables outdoor temperature and an indicator variable for winter/summer season (October-April/May-September; 0/1) were analyzed in linear regression models. The models were examined for autocorrelation of residuals with the Durbin-Watson test statistic (DW). No adjustment for autocorrelation was made when DW was between 1.5 and 2.5 [82]. All models were adjusted for age at examination.

The functional form of the relation between M/I and outdoor temperature was examined using a linear function and a sigmoid function based on the cumulative normal distribution function. The criterion for best model fit was the lowest value for the sum of the squared residuals. Three temperature intervals were defined in the relation between M/I and outdoor temperature: low temperature (LT), less than 0 °C, which corresponds to the meteorological definition of winter in Sweden (http://www.smhi.se), intermediate temperature (IT), greater than or equal to 0 and less than 10 °C, which corresponds to the meteorological definition of spring or autumn in Sweden, and high temperature (HT), greater than or equal to 10 °C, which corresponds to the meteorological definition of summer in Sweden.

A p-value of less than 0.05 was considered a statistically significant result.

Results and discussion

Paper I

Paper I [14] developed a novel design for reliability studies which makes it possible to estimate the measurement error of a variable and also to estimate the corrected regression coefficient more precise than earlier. The novelty is to use only the participants with extreme first measurement values for a replicate measurement (see Figure 2).

The estimator of β_{1c} used in Paper I is $\hat{\beta}_{1c} = \hat{\beta}_1 / \hat{\rho}$ where $\hat{\beta}_1$ is the ordinary least squares estimator of the slope in the linear regression of Y on X_1 and $\hat{\rho}$ is the slope in the linear regression of the second measurement of the predictor variable X_2 on the first measurement X_1. This estimator of the corrected regression coefficient is termed the regression based estimator.

When the main study data are collected it is possible to calculate an estimate of the relative variance gain from the extreme selection design compared with the random sampling design. Paper I is mainly theoretical but also includes Monte Carlo simulations and a reality based example from ULSAM on the relation between the response variable insulin sensitivity and the predictor variable fasting insulin to support the theory.

Results from the Monte Carlo simulations are presented in Table 1. A close agreement between expected and observed standard errors was seen over the chosen ranges of ρ and p (the fraction of participants from the main study selected to the reliability study). Extreme selection is superior to random sampling in all combinations of simulations for estimation of ρ but more importantly for estimation of β_{1c}, where the standard error is approximately halved for $p = 0.2$ and $\rho = 0.5, 0.7$ and the effect was almost as dramatic when $p = 0.3$ and/or $\rho = 0.9$. In general, the precision gain of the extreme selection design compared with the random sampling design was more pronounced when the true relation between response and predictor was strong and/or there was a large amount of error in measurement of the predictor, i.e. when ρ had a low value.

The relation between insulin sensitivity and fasting insulin, which was measured twice within one to two weeks for all participants, was explored in the age 71 years investigation of ULSAM, where the latter variable's measurement error was estimated. The study data indicate that logarithmic transformations of both variables were appropriate in order to obtain linearity as well as normality and homoscedasticity of residuals. The naive regression using only one measurement of fasting insulin resulted in
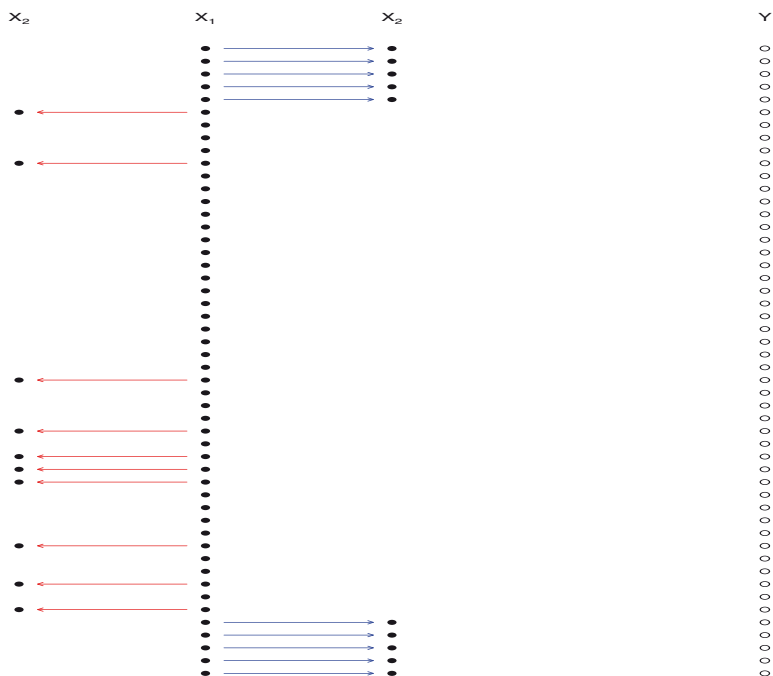
Figure 2: Schematic figure illustrating random sampling (X_2 red) and extreme selection (X_2 blue) of participants to a reliability study. The X_1 values are assumed to be sorted from lowest value to highest value.

Table 1: *Estimates (E) and theoretical (T) standard errors for $\hat{\rho}$ and $\hat{\beta}_{1c}$ for $n = 1000$ participants in the main study, $\sigma_{xx} = 1, \sigma_{\delta\delta} = 1, \beta_{1c} = 3$ and $k = np$ participants in the reliability study. Random sub-sampling (r) and extreme selection (e). Estimated results are based on 10000 Monte Carlo simulations for each combination of p and ρ*

p	ρ	E/T	$\hat{\rho}_r$	$se(\hat{\rho}_r)$	$\hat{\rho}_e$	$se(\hat{\rho}_e)$	$\hat{\beta}_{1cr}$	$se(\hat{\beta}_{1cr})$	$\hat{\beta}_{1ce}$	$se(\hat{\beta}_{1ce})$
0.2	0.5	E	0.500	0.061	0.500	0.034	3.045	0.381	3.012	0.189
		T		0.061		0.034		0.358		0.186
	0.7	E	0.700	0.051	0.700	0.028	3.016	0.217	3.003	0.109
		T		0.050		0.028		0.211		0.109
	0.9	E	0.900	0.031	0.900	0.017	3.002	0.104	2.999	0.058
		T		0.031		0.017		0.103		0.053
0.3	0.5	E	0.500	0.051	0.500	0.031	3.027	0.305	3.009	0.166
		T		0.050		0.031		0.288		0.166
	0.7	E	0.699	0.041	0.700	0.026	3.012	0.173	3.006	0.098
		T		0.041		0.026		0.170		0.098
	0.9	E	0.900	0.025	0.900	0.016	3.001	0.085	3.000	0.053
		T		0.025		0.016		0.084		0.053

Table 2: *Results for estimates of the reliability ratio ρ and corrected regression coefficient β_{1c} in the linear regression of log transformed insulin sensitivity on log transformed fasting insulin for extreme selection (e) and random (r) fasting insulin reliability sub-sampling in the ULSAM study ($p = 0.2$ and $p = 1.0$). Standard errors according to expressions and bootstrap estimation*

Selection	Expressions				Bootstrap			
	$\hat{\rho}$	$se(\hat{\rho})$	$\hat{\beta}_{1c}$	$se(\hat{\beta}_{1c})$	$\hat{\rho}$	$se(\hat{\rho})$	$\hat{\beta}_{1c}$	$se(\hat{\beta}_{1c})$
r, $p = 0.2$	0.759	0.0459	-0.567	0.0421	0.758	0.0491	-0.569	0.0442
e, $p = 0.2$	0.759	0.0239	-0.566	0.0306	0.759	0.0281	-0.567	0.0318
$p = 1.0$	0.758	0.0193	-0.566	0.0288	0.758	0.0229	-0.567	0.0294

the estimated coefficient $\hat{\beta}_1 = -0.430$ ($se_{\hat{\beta}_1} = 0.0215$). In this application standard errors from derived expressions were compared and found to agree will with standard errors from bootstrap sampling (Table 2).

The slope in the regression of insulin sensitivity on fasting insulin was strengthened 32% when corrected for measurement error in the latter variable. The importance of this finding is that when an individual has a high long-term average of fasting insulin this is an indicator of more pronounced insulin resistance than the naive regression implies (see Figure 1).

The standard error gain was 29% for the extreme selection estimate compared with the random sampling estimate of β_{1c}. The standard error for the corrected regression coefficient using replicates from all participants ($p = 1.0$) was only 8% lower than the standard error using extreme selection with $p = 0.2$, indicating how marginal the information gained about ρ was when also including the middle part of the distribution of the first measurement of fasting insulin.

An important application of extreme selection for replicates is when the aim is to relate change to initial value. For chronic diseases like T2DM or hypertension it is of interest to investigate whether the natural rate of change or the effect of an intervention is dependent on the patient's baseline level of a disease marker like fasting glucose or blood pressure. In the presence of random measurement error at baseline the estimated coefficient in the regression of change on baseline value will be biased (see, e.g. Blomqvist [83] and Edland [84]). In a setting where only a baseline visit and one follow-up visit are feasible for the participants in the main study the possibility to use extreme selection for a reliability study of the baseline data should be considered. Reanalyses of 20-30 % of the participants in the main study with extreme measurements will yield an unbiased estimator of the relation between change and initial value with a precision not far from that given by selection of all participants for a replicate baseline measurement.

Paper II

In Paper II [85] it was proved that, by adding information about the variance of the first measurement for participants that are not part of the reliability study and the information about the covariance between the second measurement of the predictor and the response variable, an estimator was obtained, based on the maximum likelihood method, that was superior to the regression based estimator. Tables 3 (normally distributed true predictor) and 4 (non-normally distributed true predictor) summarizes results of Monte Carlo simulations that compared the regression based estimator with a maximum likelihood estimator for combinations of values for the true correlation between x and Y ($\rho_{x,Y}$) and the reliability ratio ρ. In these simulations the maximum likelihood estimator was superior to the regression based estimator. This was especially true when the correlation between the true predictor x and the response Y was strong and/or ρ was low, i.e. when there was large amount of measurement error in the predictor. The success rates of the profile-likelihood-based confidence intervals were closer to the nominal level than were the symmetric confidence intervals based on asymptotic normality and the regression based estimator. Further, the latter intervals tended to have most upward misses.

A somewhat unexpected but positive finding was that the likelihood estimator was more robust to non-normal distributions and more efficient for small sample situations than the regression based estimator. The use of additional sample information seemed to play a more important role for the likelihood method's superiority than the distributional assumption did.

Other authors like Schafer and Purdy [86] and Carroll *et al.* [4] prove superiority of the likelihood approach relative to regression based estimator estimators.

Although computationally intensive, the maximum likelihood estimator should be the first choice when the distributions of the response, the predictor and the measurement errors can be carefully assessed. This is especially true when there is an anticipated strong true linear relation between response and predictor or when one awaits poor reliability in measurement of the predictor. Our application, with the response variable insulin sensitivity and the predictor fasting insulin, revealed that, in a real situation with simple transformations of data, the maximum likelihood estimator behaved as expected.

Paper III

In paper III, M/I had high (ICC = 0.95) and EIR (ICC = 0.57) had low reliability. The contribution of the measurement error to the total variance was thus 5 % for M/I and 43 % for EIR.

The uncorrected effects on fasting plasma glucose at age 77 years were larger for M/I than for EIR with a difference between effects of 0.10 mmol/l,

Table 3: *Regression based (RBE) and maximum likelihood estimates (MLE) of β_{1c} when $x \sim N(0,1)$, $p = 0.2$ and $\sigma_{\delta\delta} = 1$. Random sub-sampling (rs) and extreme selection (es). Estimated results are based on 10000 simulations for each combination. (Theoretical standard errors in parentheses.)*

n	$\rho_{x,Y}$	ρ	β_{1c}		RBE$_{rs}$	MLE$_{rs}$	RBE$_{es}$	MLE$_{es}$
200	0.35	0.9	0.374	$\hat{\beta}_{1c}$	0.378	0.375	0.374	0.374
				se($\hat{\beta}_{1c}$)	0.081 (0.079)	0.076 (0.075)	0.077 (0.076)	0.076 (0.075)
				Average width of 95% CI	0.318	0.298	0.297	0.295
				Coverage rate 95% CI	0.953	0.949	0.947	0.947
				Rate β_{1c} > upper limit	0.030	0.026	0.026	0.024
				Rate β_{1c} < lower limit	0.017	0.025	0.028	0.029
	0.85	0.5	1.614	$\hat{\beta}_{1c}$	1.802	1.665	1.640	1.633
				se($\hat{\beta}_{1c}$)	1.494 (0.439)	0.334 (0.323)	0.263 (0.239)	0.227 (0.213)
				Average width of 95% CI	4.055	1.254	1.002	0.903
				Coverage rate 95% CI	0.899	0.956	0.946	0.947
				Rate β_{1c} > upper limit	0.101	0.026	0.054	0.027
				Rate β_{1c} < lower limit	0.000	0.018	0.000	0.026
1000	0.35	0.9	0.374	$\hat{\beta}_{1c}$	0.373	0.373	0.374	0.374
				se($\hat{\beta}_{1c}$)	0.036 (0.035)	0.034 (0.034)	0.034 (0.034)	0.034 (0.033)
				Average width of 95% CI	0.140	0.132	0.133	0.131
				Coverage rate 95% CI	0.950	0.951	0.952	0.953
				Rate β_{1c} > upper limit	0.030	0.024	0.025	0.025
				Rate β_{1c} < lower limit	0.020	0.025	0.023	0.023
	0.85	0.5	1.614	$\hat{\beta}_{1c}$	1.637	1.627	1.619	1.618
				se($\hat{\beta}_{1c}$)	0.212 (0.196)	0.153 (0.144)	0.109 (0.107)	0.097 (0.095)
				Average width of 95% CI	0.810	0.603	0.424	0.384
				Coverage rate 95% CI	0.938	0.946	0.949	0.952
				Rate β_{1c} > upper limit	0.062	0.027	0.040	0.025
				Rate β_{1c} < lower limit	0.000	0.027	0.011	0.023

Table 4: *Regression based (RBE) and maximum likelihood estimates (MLE) of β_{1c} when $x \sim t(df)/(df/(df-2))^{0.5}, (df = 4)$, $p = 0.2$, and $\sigma_{\delta\delta} = 1$. Random sub-sampling (rs) and extreme selection (es). Estimated results are based on* 10000 *simulations for each combination. (Theoretical standard errors in parentheses.)*

n	$\rho_{x,Y}$	ρ	β_{1c}		RBE$_{rs}$	MLE$_{rs}$	RBE$_{es}$	MLE$_{es}$
200	0.35	0.9	0.374	$\hat{\beta}_{1c}$	0.382	0.375	0.364	0.370
				se($\hat{\beta}_{1c}$)	0.086 (0.079)	0.078 (0.075)	0.076 (0.076)	0.076 (0.075)
				Average width of 95% CI	0.329	0.303	0.295	0.297
				Coverage rate 95% CI	0.956	0.951	0.944	0.947
				Rate β_{1c} > upper limit	0.025	0.024	0.037	0.030
				Rate β_{1c} < lower limit	0.019	0.026	0.019	0.023
	0.85	0.5	1.614	$\hat{\beta}_{1c}$	2.135	1.667	1.514	1.645
				se($\hat{\beta}_{1c}$)	4.009 (0.439)	0.354 (0.323)	0.229 (0.239)	0.228 (0.213)
				Average width of 95% CI	16.215	1.308	0.886	0.917
				Coverage rate 95% CI	0.887	0.952	0.857	0.951
				Rate β_{1c} > upper limit	0.113	0.028	0.143	0.024
				Rate β_{1c} < lower limit	0.000	0.020	0.000	0.026
1000	0.35	0.9	0.374	$\hat{\beta}_{1c}$	0.376	0.374	0.364	0.370
				se($\hat{\beta}_{1c}$)	0.036 (0.035)	0.034 (0.034)	0.033 (0.034)	0.033 (0.033)
				Average width of 95% CI	0.141	0.133	0.130	0.130
				Coverage rate 95% CI	0.948	0.952	0.940	0.950
				Rate β_{1c} > upper limit	0.029	0.025	0.049	0.033
				Rate β_{1c} < lower limit	0.023	0.022	0.011	0.017
	0.85	0.5	1.614	$\hat{\beta}_{1c}$	1.671	1.633	1.485	1.628
				se($\hat{\beta}_{1c}$)	0.274 (0.196)	0.156 (0.144)	0.093 (0.107)	0.098 (0.095)
				Average width of 95% CI	0.861	0.615	0.362	0.384
				Coverage rate 95% CI	0.909	0.945	0.652	0.944
				Rate β_{1c} > upper limit	0.086	0.025	0.348	0.022
				Rate β_{1c} < lower limit	0.005	0.030	0.000	0.034

95 % CI 0.00 to 0.21, p = 0.016. There was a similar but non-significant difference for uncorrected effects on fasting plasma glucose at age 82 years (0.08 mmol/l, 95 % CI -0.09 to 0.20, p = 0.229). For HbA1c no uncorrected differences between the effects of M/I and EIR were detected at age 77 years (-0.03 %, 95 % CI -0.09 to 0.07, p = 0.51) or at age 82 years (-0.03 %, 95 % CI -0.12 to 0.06, p = 0.28) (See Figure 3).

The models corrected for measurement error had smaller estimated differences between effects of the two predictors, than in the uncorrected models, with no statistically significant differences between the measurement error corrected effects of M/I and EIR on fasting plasma glucose at age 77 years (0.02 mmol/l, 95 % CI -0.13 to 0.15, p = 0.73) or at age 82 years (0.02 mmol/l, 95 % CI -0.17 to 0.20, p = 0.85). The measurement error corrected effect of EIR on HbA1c at age 82 years was stronger than the measurement error corrected effect of M/I (-0.11 %, 95 % CI -0.28 to -0.01, p = 0.034), with a difference of the same magnitude at age 77 years, although it did not reach statistical significance (-0.10 %, 95 % CI -0.22 to 0.01, p = 0.067). The upper limits of the 95 % confidence intervals for the difference of measurement error corrected effects on HbA1c were below the non-inferiority margin 0.3 %.

The uncorrected effects on fasting plasma glucose and on 2-h glucose at age 71 years were larger for M/I than for EIR, although not statistically significant for fasting plasma glucose. The models corrected for measurement error had larger effects for EIR than for M/I on fasting and 2-h plasma glucose at age 71 years, with no statistically significant differences.

The results for response variables prevalent type 2 diabetes at age 71 years and incident type 2 diabetes from ages 71 to 77 years and from ages 71 to 82 years, were suggestive to be in the same direction as for the continuous response variables, but differences between predictor effects were not statistically significant neither for uncorrected nor for measurement error corrected models.

Corrected for measurement errors, the partial longitudinal effects of M/I and EIR, expressed per one standard deviation decrease from the mean level of each predictor, on the increase of plasma glucose concentrations and HbA1c, and the development of type 2 diabetes were of the same magnitude over a time of follow-up of 11 years. Similar cross-sectional observations were made for fasting and 2-h plasma glucose, and prevalent type 2 diabetes. The relative importance of the attenuated first-phase insulin response and insulin resistance for prediction of type 2 diabetes is under debate, with some investigators favoring impaired insulin sensitivity [87]. The results of the current study indicated that impairments of the same magnitude, i.e. one standard deviation decrease, of M/I and EIR were equally important for the elevation of glycaemia and type 2 diabetes, with the exception of HbA1c at 11 years follow-up, where a larger measurement error corrected effect of EIR than that of M/I was inferred. The use of EIR, without correction for ME underestimates the effect of an impaired first-phase insulin response with 20 to 24 %, as estimated in the present study, when ad-

Six years follow-up at age 77 years
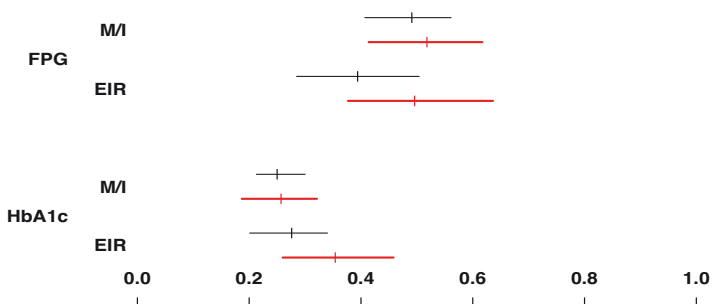
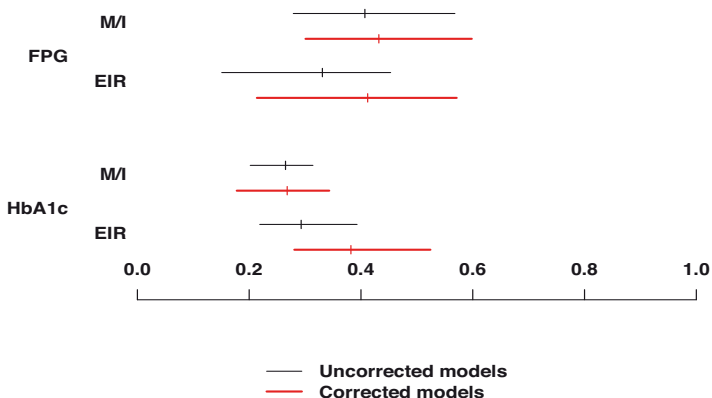**Eleven years follow-up at age 82 years**

Figure 3: The longitudinal effects (with 95 % CI) of M/I and EIR measured at baseline at age 71 years on fasting plasma glucose (FPG) (mmol/l) and HbA1c (%) as continuous response variables, uncorrected (thin black lines) and corrected (bold red lines) for measurement error, at age 77 years in upper panel (A) and at age 82 years in lower panel (B). Effects were estimated from mean levels of both predictors to mean level minus one standard deviation of each predictor while the other predictor was constant

justed for insulin sensitivity, on the plasma glucose concentrations, HbA1c and development of type 2 diabetes.

The strength of this study was that ULSAM is a large and population-based study, including 1128 investigations with the euglycaemic insulin clamp to assess insulin sensitivity. The population was homogenous for age, gender, and ethnicity. On the other hand the homogeneity of the population was also a limitation and the results need to be confirmed in women, in younger individuals, and in other ethnic groups. Most of diabetes cases develop in ages younger than examined in our study. In older ages there is less insulin resistance and more severe beta-cell function disturbances which limits the generality of our results. Incident diabetes was diagnosed on the use of diabetes medication and/or fasting plasma glucose but not on 2-h plasma glucose because there were no glucose tolerance tests at the age 77 years and the age 82 years investigations. Exclusion of diabetes cases based solely on post-prandial glucose can induce bias in the relative roles of insulin sensitivity and secretion in prediction of diabetes. However, diabetes is currently diagnosed, according to current guidelines [88], most often on fasting glucose values solely.

The present study detected a low intra-individual variation (ICC = 0.95 and CV = 13 %) of repeated measurements of insulin sensitivity index from a euglycaemic insulin clamp which is of a similar magnitude to prior studies from Soop *et al.* [89] (CV = 6 %), Bokemark *et al.* [90] (CV = 19 %) and Mather *et al.* [8] (CV = 10 %).

The high intra-individual variation in the measurement of EIR from an OGTT (ICC = 0.57 and CV = 50 %) corroborated findings by Utzschneider *et al.* [91] (AIR CV = 57 %), being higher than in the study of Cretti *et al.* [92] (AIR CV = 36 %) in which the time interval between the measurements is only 1-2 weeks.

Measurements of the acute insulin response (AIR) from an intravenous glucose tolerance test (IVGTT) has higher precision than EIR after oral glucose, as Hedstrand and Boberg [93] (CV = 20 %) and Abbate *et al.* [94] (CV = 21 %) reveal. Hanley *et al.* [95] report a strong independent effect of AIR for incident type 2 diabetes after 5.2 years of follow-up (odds ratio = 0.32 for 1 SD increase of AIR), when adjusted for insulin sensitivity and a number of other independent risk factors for type 2 diabetes. In a report from ULSAM [96] with baseline at age 50 years and with 27 years of follow-up AIR is a strong independent risk factor for incident Type 2 diabetes.

However, EIR at an OGTT is used in most epidemiological studies ([97], [98] and [99]). As uncorrected EIR underestimates the effect of first phase insulin response, due to large ME, on glucose tolerance by 20 to 24 %, we want to emphasize that to use EIR without caution, i.e. correction for ME in multivariable models including insulin sensitivity measurements, leads to wrong conclusions.

Insulin resistance and impaired insulin secretion are risk factors for type 2 diabetes [66] in the sense that each factor is manageable and a target for

the primary prevention of the disease. The results of this study imply that interventions aimed at both these targets are equally important. The role of the two risk factors for progression of type 2 diabetes can be more comparable than suggested by previous studies [87].

Paper IV

Results of the study in Paper IV revealed that, during the winter season, compared with the summer season, M/I was 12.0 % lower (4.84 *vs.* 5.44, p = 0.0003), the incremental area under the insulin curve was 14.7 % higher (1167 *vs.* 1003 mU/l, p = 0.007), fasting plasma glucose (5.80 *vs.* 5.71 mmol/l, p = 0.28) and 2-h plasma glucose OGTT (8.35 *vs.* 8.27 mmol/l, p = 0.58) were similar. Waist circumference was 1.5 % higher (95.0 *vs.* 93.7 cm, p = 0.03) during the winter season, compared with the summer season. No statistically significant differences between winter and summer seasons were detected for HOMA-IR, BMI or energy intake. All DW values were between 1.76 and 2.03.

The seasonal effect on the incremental area under the insulin curve, when adjusted for M/I, was 6.1 % higher (p = 0.36) during winter *vs.* summer season. Thus, variation of insulin sensitivity was compensated by an insulin secretion variation resulting in only a small seasonal variation of glucose concentrations. No statistically significant effects of season or of outdoor temperature were detected on HOMA-IR.

There was a statistically significant direct association (p < 0.0001) between M/I and outdoor temperature. The average increase of M/I was 0.57 units ($100x$ mg x min^{-1}x kg^{-1}/(mU x l^{-1})) (11.4 % of mean M/I) per 10 °C increase of outdoor temperature. The effect of outdoor temperature on M/I was statistically significant (p = 0.049) when adjusted for winter/summer season while the effect of the latter became non-significant (p = 0.19), i.e., insulin sensitivity variation was explained by outdoor temperature variation to a higher extent than of winter/summer season.

The best fit for the relation between M/I and outdoor temperature was a sigmoid function based on the cumulative normal distribution function (see Figure 4) with the steepest M/I rise for temperatures in the IT interval (0-10 °C) and saturations of the function for the LT (less than 0 °C) and HT (greater than or equal than 10 °C) intervals.

This study noted that seasonal variations of insulin sensitivity, with decreased sensitivity during the winter season, were compensated by inverse variations of insulin secretion resulting in only small variations of plasma glucose. Seasonal variations of insulin sensitivity were directly associated with the outdoor temperature independent of the winter/summer season effect. The seasonal effect was not detected with the surrogate measure HOMA-IR. Previous studies, using different measures of insulin sensitivity, present inconclusive results on seasonal variations of insulin sensitivity ([57], [58], [45], [59] and [60]). The compensatory seasonal variations of in-
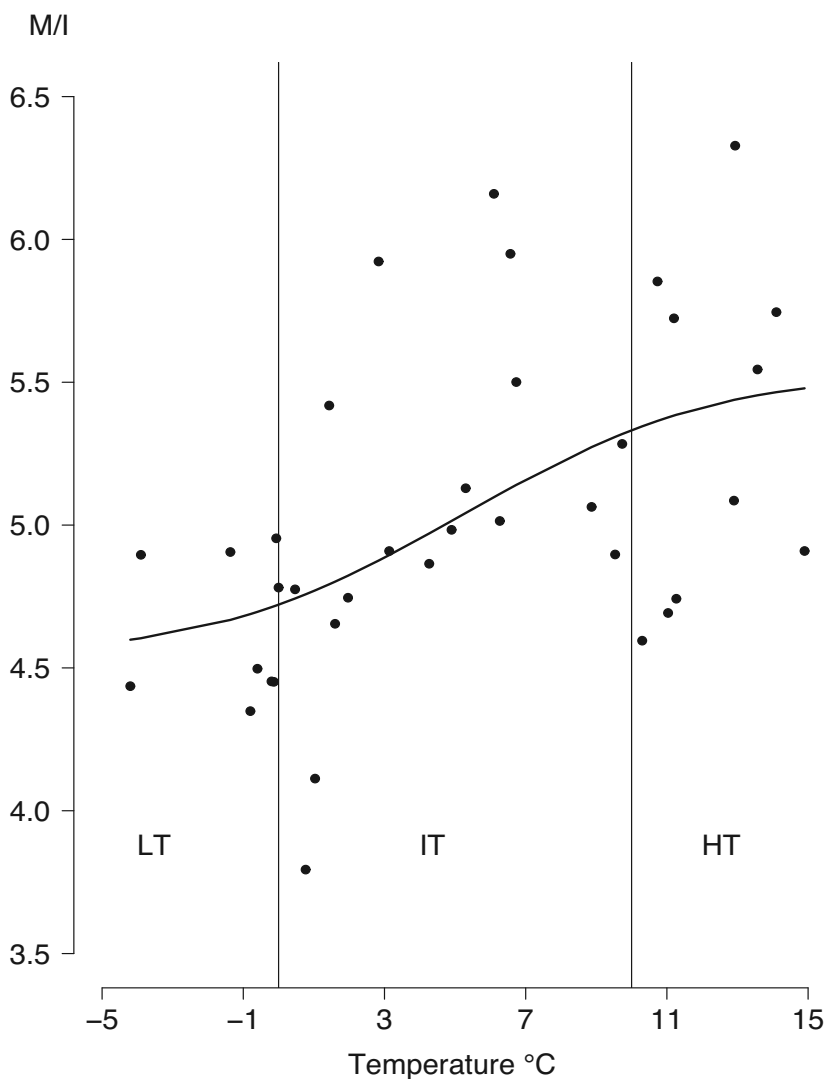
Figure 4: Monthly mean values of insulin sensitivity index M/I $(100\text{x mg x min}^{-1}\text{x kg}^{-1}/(\text{mU x l}^{-1}))$ from a euglycaemic insulin clamp *vs.* mean of outdoor temperature (°C). For each participant the mean temperature of the month for the clamp investigation and the two preceding months, representing the last quarter of the year, was used as the outdoor temperature exposure value. Monthly values from the time period October 1991 to May 1995. The solid line is predicted M/I from a sigmoid model

sulin secretion in our study corroborated findings by Fahle'n *et al.* [100] in their study, using sum of insulin from a 2-h OGTT, of one hundred 50-year-old men with myocardial infarction in Gothenburg, Sweden.

Antarctic residence (AR) exposes humans for extremely low temperatures and is associated with environmentally related thyroid alterations (AR increases the thyrotropin response to TSH-releasing hormone TRH and decreases FT4) that correlate with metabolic markers of thyroid hormone activity on hepatic and adipose tissues [101]. Patients in a tempered climate, on a fixed thyroid hormone dosage for hypothyroeosis, display seasonal variations in TSH. Further, euthyroid individuals who spend an extended number of months in polar regions prove to have a reflex TSH increase. These seasonal responses are strongly defined at high latitudes where the winter seasons are extended [102]. Furthermore, low normal FT4 levels are significantly associated with decreased insulin sensitivity [103] and TSH is negatively correlated to insulin sensitivity [104]. These associations are in line with our result that outdoor temperature had a greater impact than winter/summer season (which is a combined marker for temperature and light exposure) on insulin sensitivity fluctuations. However, also vitamin D, which is related to sun light exposure such that concentrations are higher during the summer and lower during the winter [105], is directly associated with insulin sensitivity [106]. In our study there were no data on thyroid hormones or vitamin D status.

Clinically important seasonal variations of insulin sensitivity with a 12.0 % decrease during the winter compared with the summer season in a population based sample were revealed. This effect was stated for insulin sensitivity measured with a euglycaemic insulin clamp, but not with the surrogate marker homeostasis model assessment-estimated insulin resistance (HOMA-IR). This implies that one can not adjust for the seasonal effect of insulin sensitivity using HOMA-IR. In our reliability study (n = 20) the CV for M/I was 13.9 % and for HOMA-IR the CV was 15.6 %. Sarafidis *et al.* [107] present a CV of 23.5 % (n = 78) and Matthews *et al.* [64] found a CV of 31 % (n = 18) for HOMA-IR. A partial explanation of the discrepancy between the seasonal effects on M/I and HOMA-IR is found in the different levels of reliability for these variables. Another possible interpretation is that the two variables measure different aspects of the underlying insulin sensitivity and that the temperature effect is present mainly in insulin sensitivity in the peripheral tissue skeletal muscle manifested by the euglycaemic insulin clamp [108] and less so in insulin sensitivity in the liver measured by HOMA-IR [109]. Perhaps cold climate influences peripheral tissues insulin sensitivity in skeletal muscles more than in a central organ like the liver. Such a hypothesis is supported by findings from a recent rat study [110] where weekly heat treatment of the animals during twelve weeks improves insulin sensitivity in skeletal muscle. Also in diabetic mice hyperthermia improves insulin sensitivity [111].

Our results have implications on glycaemic control in diabetic patients, who have worsened control (increased HbA1c) during winter season. Dia-

betic patients lack full ability to compensate with increased insulin secretion for worsened insulin sensitivity, i.e. resistance due to low temperature, as previously noted ([51], [52], [53] and [55]).

The results also have implications for design of clinical trials (CTs) with insulin sensitivity as an end point and on the interpretation of the mode of action of the drug tested, especially for a CT with rapid inclusion combined with short duration of follow-up, i.e. phase II studies with durations of up to 90 days. Pending on if the CT is started in the spring or the autumn different results are obtained. The seasonal effect on insulin sensitivity is also a problem in endpoint driven phase III trials where end of trial will be performed during a short time period. An example would be a trial evaluating effects of an anti-diabetic drug on cardiovascular disease where also insulin sensitivity related variables will be evaluated (see, e.g. Dormandy *et al.* [112]). The effect on the latter can be influenced by the time period when the end of trial occurs. Further, competitive inclusions give short time for inclusion and in combination with an endpoint driven trial design, where the time period for study stopping is usually as short as a few weeks, larger or smaller effects can be observed due to when in time these parts of the trial occur. Also, in phase III trials with e.g. HbA1c as primary endpoint (the usual endpoint for trials with anti-diabetic drugs) the seasonal effect can induce problems in the interpretation of results. In a trial by Bays *et al.* [80] with a treatment period of 24 weeks and in another trial by Home *et al.* [81] with treatment during 18 months, patients who start their treatment during the winter or during the summer will end their treatment on the opposite season of the year and will thus be influenced by a maximum seasonal effect. Most often the inclusions of patients are not evenly distributed over the inclusion time period. Many practical issues influence when inclusions can take place, e.g. investigators and patients usually plan together the time for inclusions to a study to avoid interference for the pre-defined visits with vacations and holidays.

A sigmoid relation between temperature and insulin sensitivity is indicated in Figure 4 where three temperature intervals are marked with different impacts of temperature on insulin sensitivity in a northern temperate or subarctic climate: In temperature group HT (greater than or equal than 10 °C) there is no or little association between temperature and M/I. This interval represents, in an evolutionary perspective, normal temperature variations to which the human species have adapted during millions of years. Temperature interval IT (greater than or equal to 0 and less than 10 °C) represents an adaption to the climate represented in the northern part of the temperate climate zone or the subarctic climate zone. Such an adaption may represent a preference for free fatty acids instead of glucose as fuel for heat production in skeletal muscle and thus resulting in lowered insulin sensitivity. This adaption represents a short time, in an evolutionary perspective, from the end of the last ice age or even shorter as the climate in Scandinavia deteriorated from the year ca 500 BC [35]. The levelling of M/I seen in group LT (less than 0 °C) may in speculation represent selection

of survivors. The humans, whose reaction on extreme cold temperatures is development of high degrees of insulin resistance, can have been abolished during evolution due to impaired reproductive ability. Thus, an adaption to northern temperate or subarctic climate zones for insulin sensitivity was observed in our study. Such hypotheses as speculated upon above can be tested in ethnic stable populations living under cold climatic circumstances for example in Siberia or in Canada or Greenland but should not be detected in newly immigrated sub-arctic populations of mixed ethnicity, e.g. in Canada.

The strength of this study was that the ULSAM study is large and population-based, including 1117 investigations with the euglycaemic insulin clamp to assess insulin sensitivity. The investigations were performed around the calendar year over almost four years. The population represents a homogenous sample considering age, gender, and ethnicity, which reveal results, not easy to examine in populations of mixed ethnicity living under similar climatic conditions. On the other hand the homogeneity of the population should be considered for extrapolation to women, younger individuals, and other ethnic groups. A putative explanation for the deterioration of insulin sensitivity during the winter season is decreased physical activity. This study could not examine the association between current physical activity and insulin sensitivity, because there are only study data on estimates of physical activity on the average over the year. Most of the study participants (71 years of age) were retired and had their professional career behind them and thus they disposed of their own time as they wished. A frequent physical activity was gardening during summer and during winter snow shovelling was frequently performed. Thus, physical activity in this age group does probably not vary much over the calendar year.

Further research, if type 2 diabetes patients are benefitted by a reinforced pharmacological treatment in environments with cold temperature, is needed.

In conclusion, seasonal variations of insulin sensitivity were compensated by inverse variations of insulin secretion resulting in only small variations of plasma glucose in this population of Swedish elderly men. Insulin sensitivity measured with euglycaemic insulin clamp was directly associated with outdoor temperature independent of the winter/summer season effect. The seasonal or outdoor temperature effects could not be detected with the surrogate measure HOMA-IR. The results have implications on glycaemic control in diabetic patients, as they lack full ability to compensate with increased insulin secretion for worsened insulin sensitivity, i.e. resistance due to low temperature. Our results also have implications for design of clinical trials.

Conclusions

Paper I

- Selection of participants to a reliability study based on extreme values of the measurement of a predictor in a main study, compared with random sampling of participants, substantially improved precision of a corrected slope in a simple linear regression model, as displayed in Paper I. This precision gain was most pronounced when the true relation between response and predictor was strong and/or when the measurement error of the predictor was high.
- In an application on insulin sensitivity and fasting insulin the conclusion was that an individual with a high long-term average of fasting insulin was less insulin sensitive than the naive regression implied.

Paper II

- Extending the work in Paper I it was proved in Paper II that a maximum likelihood estimator, by utilizing the full information of the main study and the reliability study, increased the precision of the estimator of the corrected slope in a simple linear regression model, for both random sampling and extreme selection of participants to a reliability study.
- The maximum likelihood estimator had higher recision compared with the regression based estimator utilized in Paper I. The precision gain was strongest when there was a strong true linear relation between response and predictor and/or when there was poor reliability in measurement of the predictor.

Paper III

- Paper III was the first study that compared the relative importance of insulin sensitivity and of insulin secretion, corrected for measurement errors, on glucose concentrations, HbA1c and type 2 diabetes cross-sectional and longitudinally. It was established that, corrected for measurement errors, the partial effects of insulin sensitivity and insulin secretion were similar, expressed per one standard deviation decrease of each variable.

- The results of the study imply that interventions aimed at improvement of impaired insulin secretion can be as effective as interventions aimed at enhanced insulin sensitivity.

Paper IV

- Seasonal variations of insulin sensitivity with a decrease during the winter were compensated by inverse variations of insulin secretion resulting in only small variations of plasma glucose. Seasonal variations of insulin sensitivity were associated with the outdoor temperature. The seasonal effect could not be detected with the surrogate measure HOMA-IR.
- The study results have implications on glycaemic control in diabetic patients and for the design of clinical trials with insulin sensitivity as end point.

General discussion

Application of methods for correction of regression dilution bias

There is an extensive statistical literature on random measurement error of predictor variables in regression models (see, e.g. [1], [4] and [5]), the resulting bias in estimation of regression coefficients (regression dilution bias) and methods for bias correction based on validation study or reliability study data.

The rationale behind correction for regression dilution bias is that the usual level of a predictor has an impact on disease progression. The measured values of the predictor, being the usual levels with the addition of random fluctuations (which include both real but temporary deviations from the usual level and technical measurement errors) unrelated to disease or disease progression, will consequently yield an underestimation of the predictor's true impact ([4], [10], [14] and [85]).

Because most prospective studies use only one baseline measurement of each predictor systematic and substantial underestimations of the strengths of the real associations between predictors and outcome will occur for predictors with large intra-individual variability. This phenomenon leads to conservative estimates of a predictor's impact on disease risk in univariable models. As a consequence, in a situation with a positive association between a predictor with measurement error and disease risk, high predictor levels imply lower disease risk than in the case of no measurement error. In multivariable models, with ranking of the importance of two or more predictors with different degrees of intra-individual variability, it is essential to correct for the regression dilution bias in order to correctly rank the per se effects of the predictors.

There are only few examples of applications of methods for correction for regression dilution bias in epidemiological studies. The regression dilution bias effect is well known and methods for bias correction are available in standard statistical packages. In spite of this a literature survey by Jurek *et al.* [113] reveals that out of 57 published papers in three leading epidemiological journals 39 % do not mention regression dilution bias. None of the 57 papers corrects for this bias. The small number of practical applications of the methods for correction for regression dilution bias can be explained by limited training of biostatisticians and epidemiologists and of difficul-

ties in funding two-stage designs with a main study and a reliability or a validation study.

One of the first applications in medical research of methods for correction for regression dilution bias is MacMahon *et al.* [10]. The MacMahon paper makes correction of the effect of diastolic blood pressure (DBP) on the subsequent risk for stroke and coronary heart disease for the long-term variability in DBP. They use the idea of extreme selection implicitly when they estimate the reliability ratio ρ as $\hat{\rho} = (\overline{x}_{2U} - \overline{x}_{2L})/(\overline{x}_{1U} - \overline{x}_{1L})$ where U and L denote the upper and bottom quintile groups based on the first measurement X_1 (and $\overline{x}$ denotes a mean value). Their suggested estimator is unbiased regardless of choice of size of the upper and bottom groups but is less efficient than the regression based or the maximum likelihood estimator. In our experience the MacMahon paper and its followers Clarke *et al.* ([114] and [115]) are the only epidemiological examples of the use of data from participants with extreme first measurement values for correction of regression coefficients. MacMahon *et al.* use information in the tails of a distribution where all participants have replicate measurements. However, they do not explicitly recommend a design where only the participants with extreme first measurement values are selected for replicated measurements. We recommend a reliability study design such that only the 20 to 30 % of the participants with extreme predictor values in the main study are selected for a second measurement.

When is it correct to correct?

At first sight it seems appropriate to always correct for regression dilution bias when validation or reliability data are at hand. There are circumstances though, when correction is unappropriate or unnecessary.

When the aim of a study is to test the hypothesis of linear relation between two variables, but not to estimate the size of this relation, correction for regression dilution bias is not necessary. A test of the hypothesis $\beta_1 = 0$ in model 3 is at the same time a test for $\beta_{1c} = 0$ in model 1 as these parameters only differ by a factor. However, the power of the test decreases with increasing magnitude of measurement error.

If the reliability study is small its contribution to the length of the confidence interval for the corrected regression coefficient can be so large that the effort to collect replicates will not be worthwhile. In the introductory example with logarithmic transformed insulin sensitivity as response variable and logarithmic transformed fasting insulin as predictor variable the uncorrected estimated regression coefficient is −0.43 with a 95 % confidence interval from −0.47 to −0.39. If the fraction selected to the reliability study is 2 % of the participants in the main study the corrected estimate is −0.57 with a 95 % confidence interval from −0.78 to −0.36. The example illustrates that, under certain circumstances, the corrected interval information

will be that the upper limit is closer to zero compared with the uncorrected interval information.

The assumption of independence between residuals δ in model 1 and measurement errors u in model 2 is essential for the correction to be valid [76]. The assumption implies that the measurement error of the response and the predictor variables should be uncorrelated. A case in point is the relation between 24 hour sodium excretion and blood pressure, where the former variable has large intra-individual variation. In the INTERSALT study this relation is corrected for measurement error of 24 hour sodium excretion [116]. This correction is discussed by Smith and Phillips [117] who question if the assumption of unrelated measurement error is justified. The assumption implies that if on a certain day an individual has a urinary sodium concentration which is above his or her average then this does not imply that the blood pressure is likely to be above (or below) his or her average the same day.

In Paper III, the measurement error of the predictors M/I and EIR at age 71 years and the measurement error of the dependent variables were assumed to be independent [76] of each other. For the dependent variables measured at ages 77 and 82 years this assumption is plausible. For the dependent variables fasting plasma glucose and 2-h glucose measured at age 71 years the assumption can be questioned. However, similar results for fasting plasma glucose measured at all three ages were detected.

In conclusion, each putative correction for regression dilution bias should consider if the study objective is hypothesis testing or estimation, the conceived length of the confidence interval for the corrected regression coefficient, and ascertainment of the assumptions of the measurement error model.

Summary

The results in this thesis highlight the importance of considering measurement errors and its effects on risk assessments. In the studies, aspects of measurement errors in regression models, within the field of insulin secretion and insulin action, were investigated.

The novel extreme selection design was developed for a reliability study in order to efficiently estimate corrected regression coefficients in simple linear regression models with application to the relation between insulin sensitivity and fasting insulin (Paper I). This design was further improved by a maximum likelihood estimator (Paper II). Using the extreme selection design for a reliability study in combination with maximum likelihood estimation improves precision at a certain number of participants or alternatively decreases the required number of participants at a certain precision.

The bivariate regression models between the response variables fasting and 2-h glucose determined from an OGTT, and HBA1c, and the predictors insulin sensitivity and insulin secretion where the measurement error of the predictors have been taken into account, were estimated. We highlighted the importance of taking into account different magnitudes of measurement errors when evaluating the relative impacts of risk factors (Paper III).

An important source of biological variation over time is seasonality. Previous results on seasonal variation of insulin sensitivity are inconclusive ([57], [58], ([45], [59] and [60]). The study in Paper IV revealed seasonal variations of insulin sensitivity that were related to outdoor temperature and were compensated by inverse variations of insulin secretion. Paper IV was the first study that explored seasonal variations of insulin sensitivity measured with the gold-standard technique euglycaemic insulin clamp in a population-based sample.

Acknowledgements

Many people contributed to this work. I thank them all. I am especially indebted to:

Björn Zethelius, associate professor of medicine, my principal supervisor for all support, inspiring and constant enthusiasm, for sharing his profound knowledge of the epidemiology of T2DM, for always finding a solution to any problem and for appreciated friendship. Thanks also *Björn* for being an excellent tourist guide in Rome!

Hans Garmo, PhD in mathematical statistics, co-supervisor and stimulating conversation partner who never said no to long fruitful discussions about biostatistics in general and all my ideas.

Kurt Svärdsudd, professor of clinical epidemiology and family medicine, co-supervisor who, with his long experience of medical research, is a tower of strength to lean on for any PhD student.

Christian Berne, co-author who, with his deep knowledge of diabetes and his ability to present the core message in a scientific paper in an appetizing form, is a truly inspiring scientist and collaborator.

Johan Lindbäck, co-author who is a good colleague and a witty biostatistician. Thanks for introducing me to LATEX.

the late *Hans Lithell*, a truly great scientist, who employed me at the department of Geriatrics and introduced me to the fascinating world of medical research.

Bengt Vessby, an open-minded and proficient scientist who enlightened me about lipids and fatty acids. Thanks for friendly sharing your deep knowledge in the field of nutrition and cardiovascular diseases.

Rawya Mohsen, my colleague at the department of Geriatrics, who inspired me and supported me when I was new at the department. Rawya is the one who organized all the ULSAM data. What would we have done without you?

Lars Wallentin, for creating the scientifically inspiring environment UCR ('golden nugget' of Uppsala University), for employing me and for giving me the opportunity to write my thesis within my employment.

Bertil Lindahl, present director of UCR, for appreciated support during the end of the race.

All good and hard working people in the eminent Biometrics group at UCR: *Ulrika Andersson, Maria Axetun, Ulla Bengtsson, Anna Bornefalk Hermansson, Niclas Eriksson, Nermin Hadziosmanovic, Ulf Hellström, Karin*

Bibliography

[1] Fuller WA. *Measurement Error Models*. Wiley, New York, 1987.

[2] Rosner B and Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol*, 127(2):377–386, Feb 1988.

[3] Rosner B, Spiegelman D, and Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. *Am J Epidemiol*, 136(11):1400–1413, 1992.

[4] Carroll RJ, Ruppert D, Stefanski LA, and Crainiceanu CM. *Measurement Error in Nonlinear Models. A Modern Perspective*. Chapman & Hall, Boca Raton, Second edition, 2006.

[5] Gustafson P. *Measurement Error and Misclassification in Statistics and Epidemiology. Impacts and Bayesian Adjustments*. Chapman & Hall, Boca Raton, 2003.

[6] DeFronzo RA, Tobin JD, and Andres R. Glucose clamp technique: a method for quantifying insulin secretion and resistance. *Am J Physiol*, 237(3):E214–E223, Sep 1979.

[7] Sunehag AL, Treuth MS, Toffolo G, Butte NF, Cobelli C, Bier DM, and Haymond MW. Glucose production, gluconeogenesis, and insulin sensitivity in children and adolescents: An evaluation of their reproducibility. *Pediatr Res*, 50(1):115–123, Jul 2001.

[8] Mather KJ, Hunt AE, Steinberg HO, Paradisi G, Hook G, Katz A, Quon MJ, and Baron AD. Repeatability characteristics of simple indices of insulin resistance: implications for research applications. *J Clin Endocrinol Metab*, 86(11):5457–5464, Nov 2001.

[9] Andersson P, Lind L, Berne C, Berglund L, and Lithell HO. Insulin-mediated vasodilation and glucose uptake are independently related to fasting serum nonesterified fatty acids in elderly men. *J Intern Med*, 246(6):529–537, Dec 1999.

[10] MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A, and Stamler J. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet*, 335(8692):765–774, Mar 1990.

[11] Frost C and Thompson SG. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society: Series A*, 163:173–189, 2000.

[12] Mendoza JL and Mumford M. Corrections for attenuation and range restriction on the predictor. *Journal of Educational Statistics*, 12:282–293, 1987.

[13] Kendall M and Stuart A. *The Advanced Theory of Statistics, Volume 2.* Griffin, Fourth edition, 1979.

[14] Berglund L, Garmo H, Lindbäck J, and Zethelius B. Correction for regression dilution bias using replicates from subjects with extreme first measurements. *Stat Med*, 26(10):2246–2257, May 2007.

[15] Armstrong B. Measurement error in generalized linear models. *Communications in Statistics, Series B*, 14:529–544, 1985.

[16] Rosner B, Willett WC, and Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med*, 8(9):1051–69; discussion 1071–3, Sep 1989.

[17] Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, 69:331–342, 1982.

[18] Carroll RJ and Stefanski LA. Approximate quasilikelihood estimation in models with surrogate predictors. *Journal of the American Statistical Association*, 85:652–663, 1990.

[19] Gleser LJ. *Statistical Analysis of Measurement Error Models and Applications.*, chapter Improvements of the naive approach to estimation in nonlinear errors-in-variables regression models., pages 99–124. American Mathematics Society, Providence, 1990.

[20] Efron B and Tibshirani RJ. *An Introduction to the Bootstrap.* Chapman & Hall, 1993.

[21] Alberti KG and Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med*, 15(7):539–553, Jul 1998.

[22] Zimmet P, Alberti KG, and Shaw J. Global and societal implications of the diabetes epidemic. *Nature*, 414(6865):782–787, Dec 2001.

[23] Hossain P, Kawar B, and El Nahas M. Obesity and diabetes in the developing world – a growing challenge. *N Engl J Med*, 356(3):213–215, Jan 2007.

[24] Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, Hu ZX, Lin J, Xiao JZ, Cao HB, Liu PA, Jiang XG, Jiang YY, Wang JP, Zheng H, Zhang H, Bennett PH, and Howard BV. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. *Diabetes Care*, 20(4):537–544, Apr 1997.

[25] Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Keinänen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M, Salminen V, and Uusitupa M. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med*, 344(18):1343–1350, May 2001.

[26] Goodner CJ, Walike BC, Koerker DJ, Ensinck JW, Brown AC, Chideckel BW, Palmer J, and Kalnasy L. Insulin, glucagon, and glucose exhibit synchronous, sustained oscillations in fasting monkeys. *Science*, 195(4274):177–179, Jan 1977.

[27] Hellman B, Jansson L, Dansk H, and Grapengiesser E. Effects of external ATP on Ca(2+) signalling in endothelial cells isolated from mouse islets. *Endocrine*, 32(1):33–40, Aug 2007.

[28] Juhl C, Grøfte T, Butler PC, Veldhuis JD, Schmitz O, and Pørksen N. Effects of fasting on physiologically pulsatile insulin release in healthy humans. *Diabetes*, 51 Suppl 1:S255–S257, Feb 2002.

[29] Berman N, Genter P, Chou HF, Cortez C, Bowsher R, and Ipp B. Erratic oscillatory characteristics of plasma insulin concentrations in patients with insulinoma: mechanism for unpredictable hypoglycemia. *J Clin Endocrinol Metab*, 82(9):2899–2903, Sep 1997.

[30] Lang DA, Matthews DR, Peto J, and Turner RC. Cyclic oscillations of basal plasma glucose and insulin concentrations in human beings. *N Engl J Med*, 301(19):1023–1027, Nov 1979.

[31] Pørksen N, Munn S, Steers J, Vore S, Veldhuis J, and Butler P. Pulsatile insulin secretion accounts for 70 % of total insulin secretion during fasting. *Am J Physiol*, 269(3 Pt 1):E478–E488, Sep 1995.

[32] Pørksen N, Nyholm B, Veldhuis JD, Butler PC, and Schmitz O. In humans at least 75 % of insulin secretion arises from punctuated insulin secretory bursts. *Am J Physiol*, 273(5 Pt 1):E908–E914, Nov 1997.

[33] Widjaja A, Morris RJ, Levy JC, Frayn KN, Manley SB, and Turner RC. Within- and between-subject variation in commonly measured anthropometric and biochemical variables. *Clin Chem*, 45(4):561–566, Apr 1999.

[34] Poulsen F and Borres Jensen K. A luminescent oxygen channeling immunoassay for the determination of insulin in human plasma. *J Biomol Screen*, 12(2):240–247, Mar 2007.

[35] Keatinge WR, Coleshaw SR, and Holmes J. Changes in seasonal mortalities with improvement in home heating in England and Wales from 1964 to 1984. *Int J Biometeorol*, 33(2):71–76, Jun 1989.

[36] Auliciems A and Frost D. Temperature and cardiovascular deaths in Montreal. *Int J Biometeorol*, 33(3):151–156, Oct 1989.

[37] Wehr TA and Rosenthal NB. Seasonality and affective illness. *Am J Psychiatry*, 146(7):829–839, Jul 1989.

[38] Tanaka S, Konno A, Hashimoto A, Hayase A, Takagi Y, Kondo S, Nakamura Y, and Iimura O. The influence of cold temperatures on the progression of hypertension: an epidemiological study. *J Hypertens Suppl*, 7(1):S49–S51, Feb 1989.

[39] Alpérovitch A, Lacombe JM, Hanon O, Dartigues JF, Ritchie K, Ducimetière P, and Tzourio C. Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the Three-City study. *Arch Intern Med*, 169(1):75–80, Jan 2009.

[40] Gordon DJ, Trost DC, Hyde J, Whaley FS, Hannan PJ, Jacobs DR, and Ekelund LG. Seasonal cholesterol cycles: the Lipid Research Clinics Coronary Primary Prevention Trial placebo group. *Circulation*, 76(6):1224–1231, Dec 1987.

[41] Dawson-Hughes B, Dallal GB, Krall BA, Harris S, Sokoll LJ, and Falconer G. Effect of vitamin D supplementation on wintertime and overall bone loss in healthy postmenopausal women. *Ann Intern Med*, 115(7):505–512, Oct 1991.

[42] Malarkey WB, Hall JC, Pearl DK, Kiecolt-Glaser JK, and Glaser R. The influence of academic stress and season on 24-hour concentrations of growth hormone and prolactin. *J Clin Endocrinol Metab*, 73(5):1089–1092, Nov 1991.

[43] Ronkainen H, Pakarinen A, Kirkinen P, and Kauppila A. Physical exercise-induced changes and season-associated differences in the pituitary-ovarian function of runners and joggers. *J Clin Endocrinol Metab*, 60(3):416–422, Mar 1985.

[44] Konno N and Morikawa K. Seasonal variation of serum thyrotropin concentration and thyrotropin response to thyrotropin-releasing hormone in patients with primary hypothyroidism on constant replacement dosage of thyroxine. *J Clin Endocrinol Metab*, 54(6):1118–1124, Jun 1982.

[45] Nicolau GY, Lakatua D, Sackett-Lundeen L, and Haus E. Circadian and circannual rhythms of hormonal variables in elderly men and women. *Chronobiol Int*, 1(4):301–319, 1984.

[46] Simoni M, Velardo A, Montanini V, Faustini Fustini M, Seghedoni S, and Marrama P. Circannual rhythm of plasma thyrotropin in middle-aged and old euthyroid subjects. *Horm Res*, 33(5):184–189, 1990.

[47] Fregly MJ. Activity of the hypothalamic-pituitary-thyroid axis during exposure to cold. *Pharmacol Ther*, 41(1-2):85–142, 1989.

[48] Hamada N, Ohno M, Morii H, Jaeduk N, Yamakawa J, Inaba M, Ikeda S, and Wada M. Is it necessary to adjust the replacement dose of thyroid hormone to the season in patients with hypothyroidism? *Metabolism*, 33(3):215–218, Mar 1984.

[49] UK Prospective Diabetes Study. IV. Characteristics of newly presenting type 2 diabetic patients: male preponderance and obesity at different ages. Multicenter Study. *Diabet Med*, 5(2):154–159, Mar 1988.

[50] Padaiga Z, Tuomilehto J, Karvonen M, Dahlquist G, Podar T, Adojaan B, Urbonaite B, Zalinkevicius R, Brigis G, Virtala B, Kohtamäki K, Cepaitis Z, and Tuomilehto-Wolf B. Seasonal variation in the incidence of Type 1 diabetes mellitus during 1983 to 1992 in the countries around the Baltic Sea. *Diabet Med*, 16(9):736–743, Sep 1999.

[51] Asplund J. Seasonal variation of HbA1c in adult diabetic patients. *Diabetes Care*, 20(2):234, Feb 1997.

[52] Carney TA, Guy SP, and Helliwell CD. Seasonal variation in HbA1c in patients with Type 2 diabetes mellitus. *Diabet Med*, 17(7):554–555, Jul 2000.

[53] Sohmiya M, Kanazawa I, and Kato Y. Seasonal changes in body composition and blood HbA1c levels without weight change in male patients with type 2 diabetes treated with insulin. *Diabetes Care*, 27(5):1238–1239, May 2004.

[54] Dasgupta K, Chan C, Da Costa D, Pilote L, De Civita M, Ross N, Strachan I, Sigal R, and Joseph L. Walking behaviour and glycemic control in type 2 diabetes: seasonal and gender differences–study design and methods. *Cardiovasc Diabetol*, 6:1, 2007.

[55] Suarez L and Barrett-Connor E. Seasonal variation in fasting plasma glucose levels in man. *Diabetologia*, 22(4):250–253, Apr 1982.

[56] Jarrett RJ, Murrells TJ, Shipley MJ, and Hall T. Screening blood glucose values: effects of season and time of day. *Diabetologia*, 27(6):574–577, Dec 1984.

[57] Nicolau GY, Haus E, Lakatua DJ, Bogdan C, Popescu M, Petrescu E, Sackett-Lundeen L, Stelea P, and Stelea S. Circadian and circannual variations in plasma immunoreactive insulin (IRI) and C-peptide concentrations in elderly subjects. *Endocrinologie*, 21(4):243–255, 1983.

[58] Del Ponte A, Guagnano MT, and Sensi S. Time-related behaviour of endocrine secretion: circannual variations of FT3, cortisol, HGH and serum basal insulin in healthy subjects. *Chronobiol Int*, 1(4):297–300, 1984.

[59] Gravholt CH, Holck P, Nyholm B, Christiansen B, Erlandsen M, and Schmitz O. No seasonal variation of insulin sensitivity and glucose effectiveness in men. *Metabolism*, 49(1):32–38, Jan 2000.

[60] Alemzadeh R, Kichler J, Babar G, and Calhoun M. Hypovitaminosis D in obese children and adolescents: relationship with adiposity, insulin sensitivity, ethnicity, and season. *Metabolism*, 57(2):183–191, Feb 2008.

[61] Bunout D, Barrera G, de la Maza P, Gattas V, and Hirsch S. Seasonal variation in insulin sensitivity in healthy elderly people. *Nutrition*, 19(4):310–316, Apr 2003.

[62] Byberg L, Melhus H, Gedeborg R, Sundström J, Ahlbom A, Zethelius B, Berglund L, Wolk A, and Michaëlsson K. Total mortality after changes in leisure time physical activity in 50 year old men: 35 year follow-up of population based cohort. *BMJ*, 338:b688, 2009.

[63] Byberg L, Siegbahn A, Berglund L, McKeigue P, Reneland R, and Lithell H. Plasminogen activator inhibitor-1 activity is independently related to both insulin sensitivity and serum triglycerides in 70-year-old men. *Arterioscler Thromb Vasc Biol*, 18(2):258–264, Feb 1998.

[64] Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, and Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, 28(7):412–419, Jul 1985.

[65] Lithell HO, McKeigue PM, Berglund L, Mohsen R, Lithell UB, and Leon DA. Relation of size at birth to non-insulin dependent diabetes and insulin concentrations in men aged 50-60 years. *BMJ*, 312(7028):406–410, Feb 1996.

[66] Zethelius B, Hales CN, Lithell HO, and Berne C. Insulin resistance, impaired early insulin response, and insulin propeptides as predictors of the development of type 2 diabetes: a population-based, 7-year follow-up study in 70-year-old men. *Diabetes Care*, 27(6):1433–1438, Jun 2004.

[67] Smedman AE, Gustafsson IB, Berglund LG, and Vessby BO. Pentadecanoic acid in serum as a marker for intake of milk fat: relations between intake of milk fat and metabolic risk factors. *Am J Clin Nutr*, 69(1):22–29, Jan 1999.

[68] Hill WG. Design of experiments to estimate heritability by regression of offspring on selected parents. *Biometrics*, 26(3):566–571, Sep 1970.

[69] Curnow RN. The estimation of repeatability and heritability from records subject to culling. *Biometrics*, 17:553–566, 1961.

[70] Burrows PM. Variances of selection differentials in normal samples. *Biometrics*, 31(1)(1):125–133, Mar 1975.

[71] Chan LK and Mak TK. Maximum likelihood estimation of a linear structural relationship with replication. *Journal of the Royal Statistical Society, Series B*, 41:263–268, 1979.

[72] Cohen AC. Restriction and selection in multinormal distributions. *The Annals of Mathematical Statistics*, 28(3):731–741, 1957.

[73] Venzon DJ and Moolgavkar SH. A method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, 37:87–94, 1988.

[74] Moré JJ and Sorensen DC. Computing a trust-region step. *SIAM Journal on Scientific & Statistical Computing*, 4:553–572, 1983.

[75] Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1(8476):307–310, Feb 1986.

[76] Thoresen M and Laake P. On the simple linear regression model with correlated meaurement errors. *Journal of Statistical Planning and Inference*, 137(1):68–78, 2007.

[77] Zethelius B, Berglund L, Hänni A, and Berne C. The interaction between impaired acute insulin response and insulin resistance predict type 2 diabetes and impairment of fasting glucose: report from a 20-year follow-up in the Uppsala Longitudinal Study of Adult Men - ULSAM. *Ups J Med Sci*, 113(2):117–130, 2008.

[78] Spiegelman D, McDermott A, and Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr*, 65(4 Suppl):1179S–1186S, Apr 1997.

[79] Bring J. How to standardize regression coefficients. *The American Statistician*, 48:209–213, 1994.

[80] Bays H, McElhattan J, Bryzinski BS, and G. A. L. L. A. N. T. 6 Study Group. A double-blind, randomised trial of tesaglitazar versus pioglitazone in patients with type 2 diabetes mellitus. *Diab Vasc Dis Res*, 4(3):181–193, Sep 2007.

[81] Home PD, Jones NP, Pocock SJ, Beck-Nielsen H, Gomis R, Hanefeld M, Komajda M, Curtis P, and R. E. C. O. R. D. Study Group. Rosiglitazone RECORD study: glucose control outcomes at 18 months. *Diabet Med*, 24(6):626–634, Jun 2007.

[82] McClave JT and Benson PG. *Statistics for Business and Economics*. Prentice-Hall, Inc., Sixth edition, 1994.

[83] Blomqvist N. On the relationship between change and initial value. *Journal of the American Statistical Association*, 72:746–749, 1977.

[84] Edland SD. Blomqvist revisited: how and when to test the relationship between level and longitudinal rate of change. *Stat Med*, 19(11-12):1441–1452, 2000.

[85] Berglund L, Garmo H, Lindbäck J, Svärdsudd K, and Zethelius B. Maximum likelihood estimation of correction for dilution bias in simple linear regression using replicates from subjects with extreme first measurements. *Stat Med*, 27(22):4397–4407, 2008.

[86] Schafer DW and Purdy KG. Likelihood analysis for errors-in-variables regression with replicate measurements. *Biometrika*, 83(4):813–824, 1996.

[87] Song Y, Manson JE, Tinker L, Howard B, Kuller LH, Nathan L, Rifai N, and Liu S. Insulin sensitivity and insulin secretion determined by homeostasis model assessment and risk of diabetes in a multiethnic cohort of women: the Women's Health Initiative Observational Study. *Diabetes Care*, 30(7):1747–1752, Jul 2007.

[88] American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 32 Suppl 1:S62–S67, Jan 2009.

[89] Soop M, Nygren J, Brismar K, Thorell A, and Ljungqvist O. The hyperinsulinaemic-euglycaemic glucose clamp: reproducibility and metabolic effects of prolonged insulin infusion in healthy subjects. *Clin Sci (Lond)*, 98(4):367–374, Apr 2000.

[90] Bokemark L, Fródén A, Attvall S, Wikstrand J, and Fagerberg B. The euglycemic hyperinsulinemic clamp examination: variability and reproducibility. *Scand J Clin Lab Invest*, 60(1):27–36, Feb 2000.

[91] Utzschneider KM, Prigeon RL, Tong J, Gerchman F, Carr DB, Zraika S, Udayasankar J, Montgomery B, Mari A, and Kahn SE. Within-subject variability of measures of beta cell function derived from a 2 h OGTT: implications for research studies. *Diabetologia*, 50(12):2516–2525, Dec 2007.

[92] Cretti A, Lehtovirta M, Bonora E, Brunato B, Zenti MG, Tosi F, Caputo M, Caruso B, Groop LC, Muggeo M, and Bonadonna RC. Assessment of beta-cell function during the oral glucose tolerance test by a minimal model of insulin secretion. *Eur J Clin Invest*, 31(5):405–416, May 2001.

[93] Hedstrand H and Boberg J. Statistical analysis of the reproducibility of the intravenous glucose tolerance test and the serum insulin response to this test in the middle-aged men. *Scand J Clin Lab Invest*, 35(4):331–337, Jul 1975.

[94] Abbate SL, Fujimoto WY, Brunzell JD, and Kahn SE. Effect of heparin on insulin-glucose interactions measured by the minimal model technique: implications for reproducibility using this method. *Metabolism*, 42(3):353–357, Mar 1993.

[95] Hanley AJG, D'Agostino R Jr, Wagenknecht LB, Saad MF, Savage PJ, Bergman R, and Haffner SM. Increased proinsulin levels and decreased acute insulin response independently predict the incidence of type 2 diabetes in the insulin resistance atherosclerosis study. *Diabetes*, 51(4):1263–1270, Apr 2002.

[96] Zethelius B, Byberg L, Hales CN, Lithell H, and Berne C. Proinsulin and acute insulin response independently predict Type 2 diabetes mellitus in men–report from 27 years of follow-up study. *Diabetologia*, 46(1):20–26, Jan 2003.

[97] Kahn SE. The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of Type 2 diabetes. *Diabetologia*, 46(1):3–19, Jan 2003.

[98] Nagi DK, Knowler WC, Charles MA, Liu QZ, Hanson RL, McCance DR, Pettitt DJ, and Bennett PH. Early and late insulin response as predictors of NIDDM in Pima Indians with impaired glucose tolerance. *Diabetologia*, 38(2):187–192, Feb 1995.

[99] Haffner SM, Miettinen H, Gaskill SP, and Stern MP. Decreased insulin action and insulin secretion predict the development of impaired glucose tolerance. *Diabetologia*, 39(10):1201–1207, Oct 1996.

[100] Fahlén M, Odén A, Björntorp P, and Tibblin G. Seasonal influence on insulin secretion in man. *Clin Sci*, 41(5):453–458, Nov 1971.

[101] Harford RR, Reed HL, Morris MT, Sapien IB, Warden R, and D'Alesandro MM. Relationship between changes in serum thyrotropin and total and lipoprotein cholesterol with prolonged antarctic residence. *Metabolism*, 42(9):1159–1163, Sep 1993.

[102] Reed HL. Circannual changes in thyroid hormone physiology: the role of cold environmental temperatures. *Arctic Med Res*, 54 Suppl 2:9–15, 1995.

[103] Roos A, Bakker SJL, Links TP, Gans ROB, and Wolffenbuttel BHR. Thyroid function is associated with components of the metabolic syndrome in euthyroid subjects. *J Clin Endocrinol Metab*, 92(2):491–496, Feb 2007.

[104] Fernández-Real JM, López-Bermejo A, Castro A, Casamitjana R, and Ricart W. Thyroid function is intrinsically linked to insulin sensitivity and endothelium-dependent vasodilation in healthy euthyroid subjects. *J Clin Endocrinol Metab*, 91(9):3337–3343, Sep 2006.

[105] Rucker D, Allan JA, Fick GH, and Hanley DA. Vitamin D insufficiency in a population of healthy western Canadians. *CMAJ*, 166(12):1517–1524, Jun 2002.

[106] Liu E, Meigs JB, Pittas AG, McKeown NM, Economos CD, Booth SL, and Jacques PF. Plasma 25-hydroxyvitamin d is associated with markers of the insulin resistant phenotype in nondiabetic adults. *J Nutr*, 139(2):329–334, Feb 2009.

[107] Sarafidis PA, Lasaridis AN, Nilsson PM, Pikilidou MI, Stafilas PC, Kanaki A, Kazakos K, Yovos J, and Bakris GL. Validity and reproducibility of HOMA-IR, 1/HOMA-IR, QUICKI and McAuley's indices in patients with hypertension and type II diabetes. *J Hum Hypertens*, 21(9):709–716, Sep 2007.

[108] Borissova AM, Tankova TI, and Koev DJ. Insulin secretion, peripheral insulin sensitivity and insulin-receptor binding in subjects with different degrees of obesity. *Diabetes Metab*, 30(5):425–431, Nov 2004.

[109] Hoffman RP. Indices of insulin action calculated from fasting glucose and insulin reflect hepatic, not peripheral, insulin sensitivity in african-american and caucasian adolescents. *Pediatr Diabetes*, 9(3 Pt 2):57–61, Jun 2008.

[110] Gupte AA, Bomhoff GL, Swerdlow RH, and Geiger PC. Heat treatment improves glucose tolerance and prevents skeletal muscle insulin resistance in rats fed a high-fat diet. *Diabetes*, 58(3):567–578, Mar 2009.

[111] Kokura S, Adachi S, Manabe E, Mizushima K, Hattori T, Okuda T, Nakabe N, Handa O, Takagi T, Naito Y, Yoshida N, and Yoshikawa T. Whole body hyperthermia improves obesity-induced insulin resistance in diabetic mice. *Int J Hyperthermia*, 23(3):259–265, May 2007.

[112] Dormandy JA, Charbonnel B, Eckland DJA, Erdmann E, Massi-Benedetti M, Moules IK, Skene AM, Tan MH, Lefèbvre PJ, Murray GD, Standl E, Wilcox RG, Wilhelmsen L, Betteridge J, Birkeland K, Golay A, Heine RJ, Korányi L, Laakso M, Mokán M, Norkus A, Pirags V, Podar T, Scheen A, Scherbaum W, Schernthaner G, Schmitz O, Skrha J, Smith U, and Taton J. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. *Lancet*, 366(9493):1279–1289, Oct 2005.

[113] Jurek AM, Maldonado G, Greenland S, and Church TR. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *Eur J Epidemiol*, 21(12):871–876, 2006.

[114] Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, and Peto R. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol*, 150(4):341–353, Aug 1999.

[115] Clarke R, Lewington S, Youngman L, Sherliker P, Peto R, and Collins R. Underestimation of the importance of blood pressure and cholesterol for coronary heart disease mortality in old age. *Eur Heart J*, 23(4):286–293, Feb 2002.

[116] Elliott P, Stamler J, Nichols R, Dyer AR, Stamler R, Kesteloot H, and Marmot M. Intersalt revisited: further analyses of 24 hour sodium excretion and blood pressure within and across populations. Intersalt Cooperative Research Group. *BMJ*, 312(7041):1249–1253, May 1996.

[117] Smith GD and Phillips AN. Inflation in epidemiology: "the proof and measurement of association between two things" revisited. *BMJ*, 312(7047):1659–1661, Jun 1996.

Appendix. The regression calibration method

The regression calibration method uses validation or reliability data to obtain estimated true predictor values for all participants in the main study. These estimated true predictor values are used instead of the measurement error prone observed predictor values in ordinary regression estimation methods (e.g. ordinary least squares for linear models or maximum likelihood estimation for logistic models). This section describes the regression calibration method for linear and logistic regression models with three continuous predictors when data from a main study with n participants and data from a reliability study, where k participants made one replicate measurement each, are available.

Estimation of true predictor values

$\mathbf{X}_i$ is the vector of observed predictors for $i = 1, \ldots, n$ where n is the number of participants in the main study.

$$\mathbf{X}_i = \begin{pmatrix} X_{1,i} = \ln(\text{Insulin secretion})_i \\ X_{2,i} = (\text{Insulin sensitivity})_i^{0.5} \\ Z_i = (\ln(\text{Insulin secretion})(\text{Insulin sensitivity})^{0.5})_i \end{pmatrix}$$

A classical non-differential measurement error model $\mathbf{X}_i = \mathbf{x}_i + \mathbf{u}_i$, where $\mathbf{x}_i$ is the vector of true values and $\mathbf{u}_i$ is the vector of measurement errors, is assumed. The vector of true predictors and the vector of measurement errors are assumed to be normally distributed: $\mathbf{x}_i \sim N(\boldsymbol{\mu}, \Sigma_{\mathbf{xx}})$ and $\mathbf{u}_i \sim N(\mathbf{0}, \Sigma_{\mathbf{uu}})$, respectively.

The sample covariance matrix $\mathbf{S_{XX}}$ of the three predictor variables is calculated as $\mathbf{S_{XX}} = \frac{\sum_{i=1}^{n} \mathbf{X_i X_i^T}}{n-1}$.

The covariance matrix $\Sigma_{\mathbf{uu}}$ of the measurement errors of insulin secretion, insulin sensitivity and their product is estimated as the symmetric matrix

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{uu}} = \begin{bmatrix} \frac{\sum_{i=1}^{k}(X_{11,i}-X_{12,i})^2}{2k} & \frac{\sum_{i=1}^{k}(X_{11,i}-X_{12,i})(X_{21,i}-X_{22,i})}{2k} & \frac{\sum_{i=1}^{k}(X_{11,i}-X_{12,i})(Z_{1,i}-Z_{2,i})}{2k} \\ & \frac{\sum_{i=1}^{k}(X_{21,i}-X_{22,i})^2}{2k} & \frac{\sum_{i=1}^{k}(X_{21,i}-X_{22,i})(Z_{1,i}-Z_{2,i})}{2k} \\ & & \frac{\sum_{i=1}^{k}(Z_{1,i}-Z_{2,i})^2}{2k} \end{bmatrix}.$$

Let $\bar{\mathbf{X}}_i = \frac{\sum_{j=1}^{k_i}\mathbf{X}_{ij}}{k_i}$ where $k_i = 2$ for the k participants in the reliability study and $k_i = 1$ for the other $n - k$ participants and let $\hat{\boldsymbol{\mu}}_X = \frac{\sum_{i=1}^{n}\bar{\mathbf{X}}_i}{n}$.

Let $\hat{\boldsymbol{\Sigma}}_{\bar{X}\bar{X}} = \frac{\sum_{i=1}^{n}k_i\bar{\mathbf{X}}_i\bar{\mathbf{X}}_i^T}{v}$ where $v = \sum_{i=1}^{n}k_i - \sum_{i=1}^{n}k_i^2 / \sum_{i=1}^{n}k_i$.

Then $\Sigma_{\mathbf{xx}}$ is estimated as $\hat{\boldsymbol{\Sigma}}_{\mathbf{xx}} = \hat{\boldsymbol{\Sigma}}_{\bar{X}\bar{X}} - \frac{n-1}{v}\hat{\boldsymbol{\Sigma}}_{\mathbf{uu}}$.

A linear predictor of the true $\mathbf{x}_i$ vector is: $\hat{\mathbf{x}}_i = \hat{\boldsymbol{\mu}}_X + (\hat{\boldsymbol{\Sigma}}_{\mathbf{xx}} + \hat{\boldsymbol{\Sigma}}_{\mathbf{uu}}/k_i)^{-1}\hat{\boldsymbol{\Sigma}}_{\mathbf{xx}}(\bar{\mathbf{X}}_i - \hat{\boldsymbol{\mu}}_X)$.

The sample covariance matrix of the linear predictor is $\mathbf{S}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = \frac{\sum_{i=1}^{n}\hat{\mathbf{x}}_i\hat{\mathbf{x}}_i^T}{n-1}$.

Linear regression models

Y is a continuous response variable. The linear model relating Y to $\mathbf{x}$ is:

$$Y_i = \boldsymbol{\beta}^T\mathbf{x}_i + \delta_i$$

for $i = 1,\ldots,n$ where $\boldsymbol{\beta} = \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}Y} = \begin{pmatrix} \beta_{x_1} \\ \beta_{x_2} \\ \beta_z \end{pmatrix}$ and $\delta_i \sim N(0,\sigma_{\delta\delta})$ (where $\boldsymbol{\Sigma}_{\mathbf{x}Y}$ is the covariance matrix between the true $\mathbf{x}$ values and Y).

The corresponding model relating Y_i to $\mathbf{X}_i$ is:

$$Y_i = \boldsymbol{\beta}^{*T}\mathbf{X}_i + \epsilon_i$$

where $\epsilon_i \sim N(0,\sigma_{\epsilon\epsilon})$ and

$$\boldsymbol{\beta}^* = (\boldsymbol{\Sigma}_{\mathbf{xx}} + \boldsymbol{\Sigma}_{\mathbf{uu}})^{-1}\boldsymbol{\Sigma}_{\mathbf{x}Y} = \begin{pmatrix} \beta_{X_1}^* \\ \beta_{X_2}^* \\ \beta_Z^* \end{pmatrix}.$$

The measurement error corrected regression coefficients are estimated with the regression calibration method [4] as follows.

The sample covariance matrix between Y_i and $\mathbf{X}_i$ is $\mathbf{S}_{\mathbf{X}Y} = \frac{\sum_{i=1}^{n}\mathbf{X}_i Y_i^T}{n-1}$ and the sample covariance matrix between Y_i and $\hat{\mathbf{x}}_i$ is $\mathbf{S}_{\hat{\mathbf{x}}Y} = \frac{\sum_{i=1}^{n}\hat{\mathbf{x}}_i Y_i^T}{n-1}$.

The regression calibration estimator of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}}_{RC} = \mathbf{S}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1}\mathbf{S}_{\hat{\mathbf{x}}Y}$.

Logistic regression models

Y is a binary response variable (0/1). The model relating Y to $\mathbf{x}$ is:

$$P(Y_i = 1 | \mathbf{x}_i, \boldsymbol{\beta}) = \frac{\boldsymbol{\beta}^T \mathbf{x}_i}{1 + \boldsymbol{\beta}^T \mathbf{x}_i}.$$

The measurement error corrected regression coefficients are estimated with the regression calibration method by replacing $\mathbf{x}$ in model with $\hat{\mathbf{x}}_i$ and using the maximum likelihood method to estimate $\boldsymbol{\beta}$.

Standard errors and confidence intervals

Standard errors for the regression calibration corrected estimates will be underestimated by ordinary methods as these methods do not take into account the variance in the estimation of $\mathbf{x}$. Since the computation of explicit formulas for the standard errors is tedious [18], standard errors are typically obtained through bootstrapping [4].

Therefore, standard errors for the estimates, confidence intervals for the parameters and p values for tests of null hypotheses of zero effects are assessed with the bootstrap method [20]. B bootstrap samples are drawn from the n observations and all estimates are calculated in each bootstrap sample. 95% confidence intervals are calculated as the 2.5^{th} and the 97.5^{th} percentiles in the bootstrap distribution of the estimates. If the estimate > 0 the p-value is calculated as $2\#(\text{bootstrap estimates} < 0)/B$ and if the estimate < 0 the p-value is calculated as $2\#(\text{bootstrap estimates} > 0)/B$.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 440

Editor: The Dean of the Faculty of Medicine

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2009