AOSIS

# Assessing the test–retest repeatability of insulin resistance measures: Homeostasis model assessment 2 and oral glucose insulin sensitivity

CrossMark
click for updates

**Authors:**
Catherine A.P. Crofts[1]
Mark C. Wheldon[1]
Caryn Zinn[1]
Xiaomiao Lan-Pidhainy[2]
Thomas M.S. Wolever[3]
Grant Schofield[1]

**Affiliation:**
[1]Faculty of Health and Environmental Sciences, Auckland University of Technology, New Zealand

[2]Division of Nutritional Sciences, Cornell University, United States

[3]Department of Nutritional Sciences, University of Toronto, Canada

**Corresponding author:**
Catherine Crofts,
ccrofts@aut.ac.nz

**Background:** Insulin resistance is commonly assessed using the homeostasis model assessment (HOMA) variants. HOMA is potentially insensitive to change because of its high coefficient of variation. The repeatability coefficient is an alternative means of assessing test repeatability. To be confident of clinical change, rather than biological variation, a subsequent test needs to differ from the former by more than the repeatability coefficient using the equation.

Test 1 = Test 2 ± repeatability coefficient.

The repeatability coefficients for measures of insulin resistance are unknown.

**Aim:** To compare the repeatability coefficient of HOMA2 variants (Beta-cell function [%B], insulin sensitivity [%S], insulin resistance [IR]) to a dynamic measure of insulin resistance, and the oral glucose insulin sensitivity (OGIS) test.

**Setting:** The raw data from a previously used data set were reanalysed.

**Methods:** Glycaemic and insulinaemic tests were performed on 32 men and women both with ($n = 10$) and without type 2 diabetes ($n = 22$). From these data, eight fasting tests and three 50-g oral glucose tolerance tests were used to calculate HOMA2 and OGIS. The methods of Bland and Altman assessed repeatability.

**Results:** Repeatability coefficients for all participants for the HOMA2 %B, %S and IR variants were 72.91, 189.75 and 0.9, which equates to 89%, 135% and 89% of their respective grand means. By contrast, OGIS had a repeatability coefficient of 87.13, which equates to 21% of the grand mean.

**Conclusion:** Because of the high repeatability coefficient relative to the grand mean, use of HOMA2 measures for assessing insulin resistance in small population studies should be reconsidered.

## Introduction

It is projected that by 2030, nearly 7% of the global population will have type 2 diabetes with many of them living in developing countries.[1] Although insulin resistance is commonly undiagnosed, it is accepted as a key factor in the aetiology of type 2 diabetes and other non-communicable diseases.[2,3,4] This means that insulin resistance needs further investigation as a means for diagnosing and managing at-risk populations. However, quantifying insulin resistance is challenging. Various methods are used to assess insulin resistance. Generally, these are validated against the resource-intensive hyperinsulinaemic–euglycaemic clamp[5] and include methods based on a single fasting blood glucose and insulin sample[6] or those using a dynamic assessment of insulin and glucose response following an oral glucose tolerance test. However, assessing insulin resistance in clinical practice has been discouraged because of a lack of clinical benefit.[7]

An example of a method based on a fasting insulin and glucose sample is the homeostasis model assessment (HOMA), or the second-generation HOMA2. HOMA is the collective term for the assessment of three different aspects of insulin resistance: Beta-cell function (HOMA %B), insulin sensitivity (HOMA %S) and insulin resistance (HOMA IR). The original HOMA methods are calculated using the following equations (Figure 1).

$$HOMA\text{-}IR = \frac{Glucose\ x\ Insulin}{22.5},\ HOMA\text{-}\beta = \frac{20x\ Insulin}{Glucose-3.5}\%,\ HOMA\text{-}S = \frac{1}{HOMA\text{-}\beta}$$

Source: Wallace et al.[8]

**FIGURE 1:** Homeostasis model assessment calculations where glucose is mmol/L and insulin is mU/L.

By contrast, HOMA2 is a computer model that accounts for variations in hepatic and peripheral insulin resistance. The formulae are not available, but the calculator can be downloaded from the Diabetes Trials Unit, University of Oxford.[9]

HOMA methods are commonly used to assess insulin resistance with studies ranging from large-scale epidemiological assessments to smaller interventional studies. HOMA methods are practical instruments as they are cost-effective and easy to use. However, previous studies have shown HOMA to have coefficients of variation (CVs) ranging from 10% to 50%.[10,11,12] This suggests that HOMA methods may be insufficiently sensitive to accurately detect, or monitor clinical progression in, insulin resistance, especially for individuals or small populations.

Emerging research suggests that fasting insulin levels may not predict post-prandial insulin levels,[13] and these latter levels are more important for diagnosing risk.[14] The oral glucose insulin sensitivity (OGIS) test is an alternative method for assessing insulin resistance.[15] It is based on a model derived from a multiple-sampled oral glucose tolerance test with insulin assays. This allows the post-prandial insulin and glucose levels to be included in the model. As with the HOMA2 model, the formulae are not available but an online calculator is available.[16] Furthermore, there is also limited repeatability data on these dynamic tests of insulin resistance with one study reporting a CV for duplicate OGIS tests as 7.1%.[15]

Instruments assessing change need an acceptable level of repeatability. Repeatability is how much variation can be expected among repeat measurements on the same subject under identical conditions. This enables subsequent test results (e.g. blood tests) to either indicate clinical change or biological variation (noise). Repeatability is often assessed using CV (the ratio of standard deviation compared to the mean) and expressed by percentage, using the equation (see Equation 1).

$$c_v = \frac{\sigma}{\mu} \qquad \text{[Eqn 1]}$$

By contrast, the repeatability coefficient defines the range within which 95% of the differences between two measurements in the same subject by the same measurement method are likely to fall, assuming there is no change in clinical condition between the tests.[17,18] It can be expressed by equation 2

*Test 1 ≈ Test 2 ± repeatability coefficient*     [Eqn 2]

Therefore, to be confident of clinical change, a subsequent test needs to have changed by an amount that is greater than the repeatability coefficient (either larger or smaller).

Otherwise, the result will still lie within the range of typical biological variation. All else being equal, tests with a small repeatability coefficient are more sensitive to clinical change. Conversely, tests with a large repeatability coefficient generally require greater clinical change to occur before they can confidently be taken as an indication of such change. This makes tests with a large repeatability coefficient less desirable for clinical use as they are less sensitive to clinical change. Although CV is more commonly used to assess repeatability, use of the repeatability coefficient may be more practical for clinicians as it may be easier to interpret and incorporate into clinical practice.

The aim of this study was to assess the repeatability coefficients for the HOMA2 variants and OGIS by evaluating repeated oral glucose tolerance tests with insulin assays in a small group of subjects with or without type 2 diabetes.

# Research methods and design
## Study design

The raw data from a glycaemic and insulinaemic index study were reanalysed.[19] From these data, measures, including HOMA2 and OGIS, were calculated and repeatability assessed. The original protocol conformed to standard glycaemic index testing protocols to minimise inter-individual variation. The models HOMA2 and OGIS were chosen as representative of a fasting and dynamic model available without any specialised technologies.

## Study population and sampling strategy

Full methods and methodology are available.[19] To summarise, 22 participants without and 10 participants with type 2 diabetes were recruited. All participants were aged between 18 and 70 years; had a body mass index (BMI) < 35 kg/m²; had no recent history of hospitalisation; or had any history of gastrointestinal, hepatic or renal disease. All participants with type 2 diabetes used medication: eight used metformin only, one used a combination of metformin and pioglitazone and one used a combination of metformin and a sulphonylurea. These patients took their usual medication on study days after the fasting blood sample but before commencing the test meal.

Originally, the 22 healthy participants were divided into 'control' and 'hyperinsulinaemic' groups based on a fasting insulin of 40 pmol/L.[19] As there is no consensus in the literature for defining hyperinsulinaemia from fasting insulin,[20,21] these groups were combined into a single group termed 'No Diabetes'.

## Data collection

Blood samples were collected according to standard protocols. Each participant had fasting blood samples drawn on eight separate mornings. On three of those mornings, they then consumed 50 g anhydrous glucose in 250 mL water and on the other five mornings they consumed 50 g available carbohydrate

from sucrose, instant mashed potato, white bread, polished rice or pearled barley. Venous blood samples were then drawn at 15, 30, 45, 60, 90 and 120 min for participants without type 2 diabetes and at 30, 60, 90, 120 and 180 min for participants with type 2 diabetes. Timing commenced after starting to eat. This study analysed all results from the eight fasting tests and from the three glucose meals.

## Data analysis

### Blood analysis

Venous blood samples were collected in BD vacutainer SST tubes. Serum glucose was measured by the glucose oxidase method (Synchron LX Systems) with inter-assay CV of 1.9%. Insulin was measured using one-step immunoenzymatic assay (Beckman Access Ultrasensitive Insulin Assay) with inter-assay CV of 2.5% – 4.3%. Insulin has no cross-reactivity with proinsulin.

### Calculations and statistical analysis

As individual height and weight data were not available for each person, the standards of 1.7 m for height and 70 kg for weight were used for each person. The fasting glucose and insulin values from each test were used to calculate each of the HOMA2 variants (HOMA2 %S, HOMA2 %B and HOMA2 IR) via the available online calculator.[22]

The glucose and insulin values from each of the three oral glucose tests were used to calculate OGIS via the available spreadsheet.[16]

For each test, within-subject means were plotted against within-subject standard deviations to determine if there was a mean–variance relationship. Ordinary least squares regression was used to assess the strength of such relationships. If the slope coefficient was significant at the 0.05 significance level, the process was repeated for the mean and standard deviation of the natural log of the variable.

If a significant mean–variance relationship was determined, participants were divided into subgroups according to test results. The intent was to reduce the mean–variance relationship and therefore bias in the repeatability coefficient at each end of the range while maintaining a clinically meaningful result.

Repeatability was quantified by estimating repeatability coefficients according to the methods of Bland and Altman.[17] Repeatability coefficients were derived from the square root of the residual mean square errors $\left(\sqrt{s_w^2}\right)$ from one-way analyses of variance with subjects as factors fitted to the raw or logged responses for each outcome variable. The 95% repeatability coefficient is $1.96\sqrt{2}s_w$.[11,17]

The following calculations defined the ranges within which two repeat measurements could be expected to fall under the assumption of no clinical change between repeat tests (equation 3 and 4):

*Test 1* ≈ *Test 2* ± *repeatability coefficient* for non–log-transformed data                    [Eqn 3]

or as

*Test 1* ≈ *Test 2* ×/÷ exp(*repeatability coefficient*) for log-transformed data.                    [Eqn 4]

## Ethical consideration

Ethical permission for data collection was previously granted by Research Ethics Boards at the University of Toronto and St Michael's Hospital. All participants gave written informed consent.

# Results

Figure 2 displays the raw data for Control (left) and Diabetes (right) for fasting glucose, fasting insulin, HOMA2 %B, HOMA2 IR, HOMA2 %S and OGIS. Across all participants, fasting glucose has a narrow spread, especially for the Control participants. By contrast, the spread for all HOMA2 variants is greater for both the Control and Diabetes groups. There is also variation among participants; some have a tight cluster of results, others have a wider spread; there are also notable outliers. These vary between individuals and measures. Overall, the spread for OGIS does not appear to be as tight as that for fasting glucose, but tighter than the HOMA2 variants. This may reflect the difference in spread between fasting glucose and fasting insulin.

## Mean–variance relationships

Mean–variance relationships were positive and significant for all tests with the exception of OGIS for both groups, and fasting insulin and HOMA2 IR for the Diabetes group (Table 1). With the exception of fasting glucose in the 'No Diabetes' group, this relationship remained unchanged for all variables following natural logarithmic transformation (Table 1). This implied heterogeneity among members of the 'No Diabetes' group. To account for this, we created a 'Hyperinsulinaemic' and 'Control' group. Although the definition of hyperinsulinaemia is commonly based on fasting insulin levels, we believed this was insufficient given a lack of consensus on fasting insulin and the evidence suggesting post-prandial insulin is crucial.[13,14] Therefore, we developed our own criteria based on examination of the graphed insulin response versus time for all 'No Diabetes' participants. This led to the division of the 'No Diabetes' group into 'Control' (*n* = 19) and 'Hyperinsulinaemic' (*n* = 3). The latter consisted of three participants who each had the combination of a fasting insulin > 72 pmol/L and a 2-h insulin > 3× the fasting value. The final analysis sets were 'All participants', 'Diabetes', 'No Diabetes' and 'Control', with the latter being a subset of 'No Diabetes'.

Re-examination of the mean–variance relationships identified a non-significant relationship for fasting glucose (*p* = 0.17) and HOMA2 %S (*p* = 0.061) for the Control participants (Table 1).

With the exception of log fasting glucose for the No Diabetes set, the only measures that demonstrated a non-significant, log-transformed mean–variance relationship also had a non-significant mean–variance relationship for the raw data (Table 1). This implied there was little to be gained by using the log-transformed results. Instead, we used raw data in all subsequent analyses and expressed repeatability coefficients as both the raw figure and as a percentage of the grand means for each measure.

## Repeatability coefficients

The repeatability coefficients for HOMA2 variants, fasting insulin, fasting glucose and OGIS by participant sets are presented in Table 2. The repeatability coefficient for fasting insulin was 7.9 µU/mL for all participant sets with the exception of the Control subset, where it was 5.9 µU/mL. Across the analysis subsets, the repeatability coefficient was approximately 90% of the Grand mean of the participants' fasting insulin. For
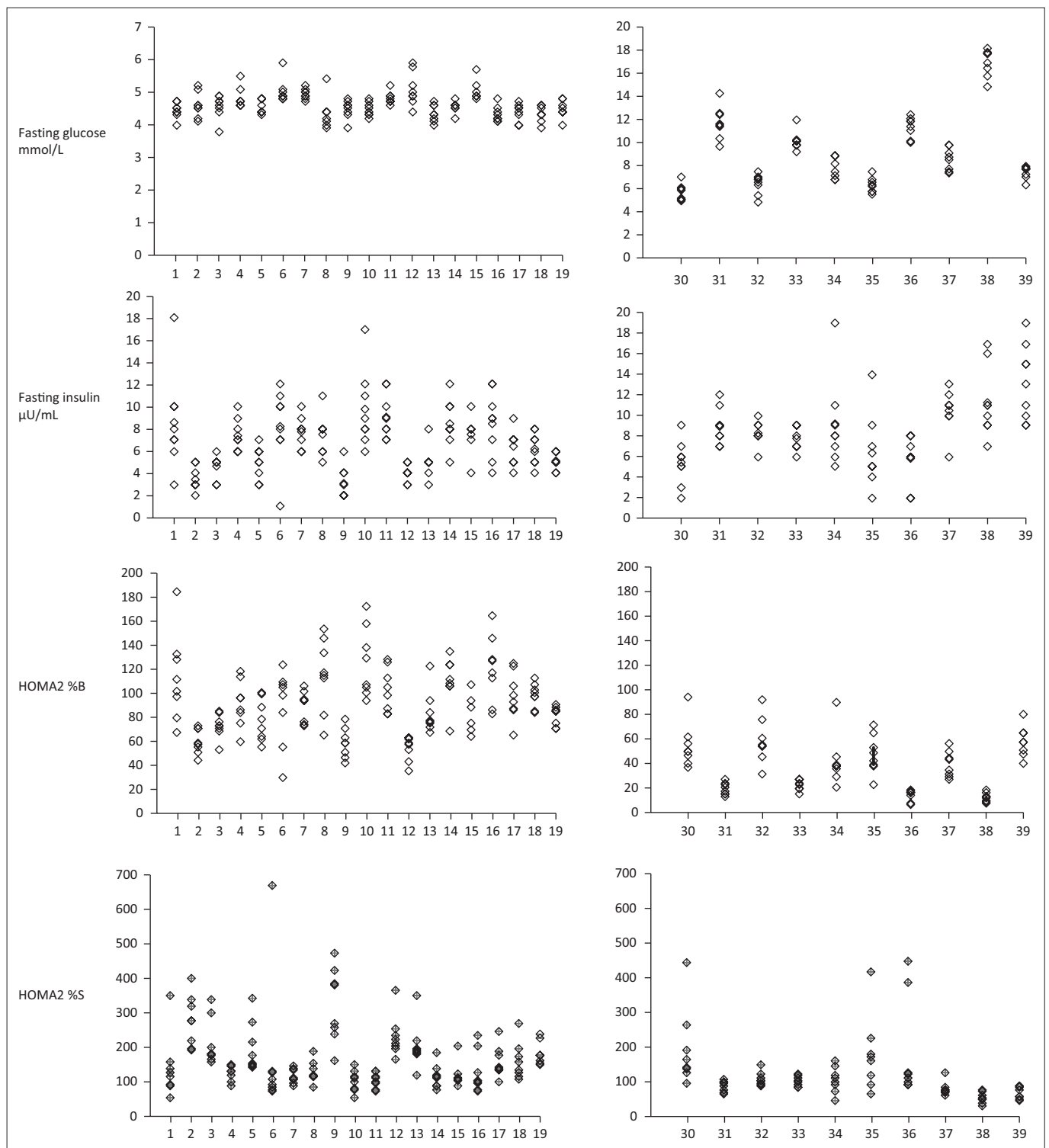


**FIGURE 2:** Raw data for control (left) and diabetes (right) groups for fasting glucose, fasting insulin, HOMA2 %B, HOMA2 %S, HOMA2 IR and OGIS.

Figure 2 continues on the next page →
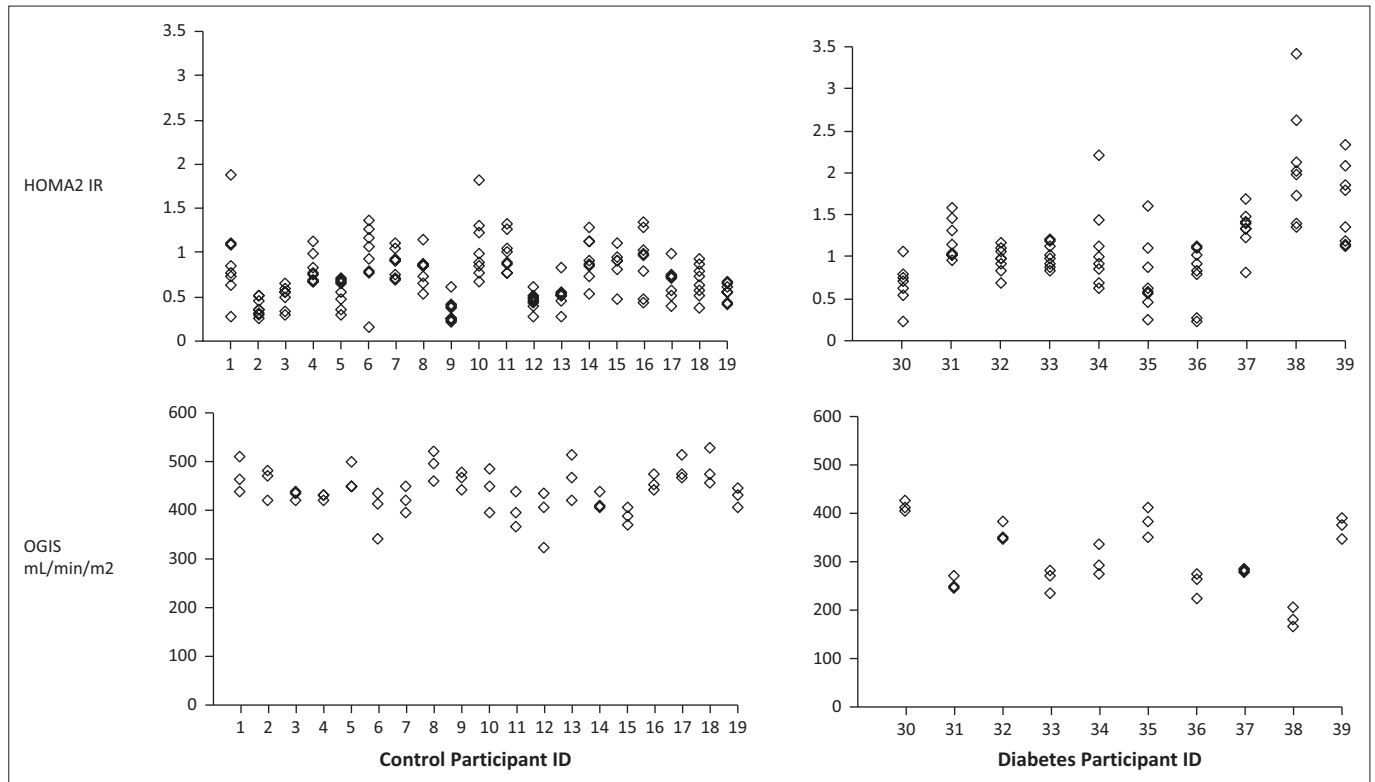
**FIGURE 2 (Continues...):** Raw data for control (left) and diabetes (right) groups for fasting glucose, fasting insulin, HOMA2 %B, HOMA2 %S, HOMA2 IR and OGIS.

**TABLE 1:** Regression coefficient and *p*-value for mean–variance relationships.

| Variable | All participants | | Diabetes | | No Diabetes | | Control | |
|---|---|---|---|---|---|---|---|---|
| | Reg Coef | *p* | Reg Coef | *p* | Reg Coef | *p* | Reg Coef | *p* |
| Fasting glucose (mmol/L) | 0.097 | < 0.001 | 0.049 | 0.035 | 0.165 | 0.039 | 0.119 | 0.17 |
| log fasting glucose | 1.474 | < 0.001 | 0.54 | 0.034 | 2.219 | 0.069 | – | – |
| Fasting insulin (μU/mL) | 0.156 | < 0.001 | −0.23 | 0.635 | 0.221 | < 0.001 | 0.388 | < 0.001 |
| log fasting insulin | 0.942 | < 0.001 | – | – | 0.997 | < 0.001 | 1.182 | < 0.001 |
| HOMA2 %B | 0.306 | < 0.001 | 0.318 | 0.001 | 0.409 | < 0.001 | 0.296 | < 0.001 |
| log HOMA2 %B | 0.864 | < 0.001 | 1.06 | < 0.001 | 1.366 | < 0.001 | 1.328 | < 0.001 |
| HOMA2 %S | 0.406 | < 0.001 | 0.825 | < 0.001 | 0.333 | 0.007 | 0.303 | 0.061 |
| log HOMA2 %S | 1.227 | < 0.001 | 1.40 | 0.004 | 1.220 | < 0.001 | – | – |
| HOMA2 IR | 0.194 | < 0.001 | 0.236 | 0.067 | 0.180 | < 0.001 | 0.363 | <0.001 |
| log HOMA2 IR | 0.897 | < 0.001 | – | – | 0.921 | < 0.001 | 1.105 | < 0.001 |
| OGIS (mL/min/m²) | 0.450 | 0.099 | 0.005 | 0.904 | −0.003 | 0.962 | −0.350 | 0.731 |

%B, beta cell function; HOMA2, homeostasis model assessment 2; IR, insulin resistance; OGIS, oral glucose insulin sensitivity test; Reg Coef, regression coefficient; %S, insulin sensitivity.

**TABLE 2:** Repeatability coefficients for simple measures of insulin resistance (all data).

| Variable | All participants (*n* = 32) | | | | Diabetes (*n* = 10) | | | | No Diabetes (*n* = 22) | | | | Control (*n* = 18) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $s_w$ | ±Rep Coef | $\hat{\mu}$ | Change % | $s_w$ | ±Rep Coef | $\hat{\mu}$ | Change % | $s_w$ | ±Rep Coef | $\hat{\mu}$ | Change % | $s_w$ | ±Rep Coef | $\hat{\mu}$ | Change % |
| Fasting glucose (mmol/L) | 0.59 | 1.62 | 6.08 | 26.64 | 0.93 | 2.59 | 9.20 | 28.15 | 0.32 | 0.88 | 4.65 | 18.92 | 0.30† | 0.84 | 4.06 | 20.69 |
| Fasting insulin (μU/mL) | 2.85 | 7.91 | 8.61 | 91.87 | 2.86† | 7.90 | 8.59 | 91.97 | 2.85 | 7.90 | 8.62 | 91.65 | 2.14 | 5.93 | 6.52 | 90.96 |
| HOMA2 %B | 26.31 | 72.91 | 82.31 | 88.57 | 13.08 | 36.26 | 36.41 | 99.45 | 30.53 | 84.61 | 103.18 | 82.00 | 20.67 | 57.31 | 90.62 | 63.24 |
| HOMA2 %S | 68.46 | 189.75 | 140.05 | 135.49 | 70.02 | 194.09 | 111.19 | 174.56 | 67.72 | 187.72 | 153.17 | 122.56 | 72.68† | 201.45 | 169.33 | 118.97 |
| HOMA2 IR | 0.32 | 0.90 | 1.01 | 89.11 | 0.39† | 1.07 | 1.15 | 93.04 | 0.29 | 0.81 | 0.95 | 85.26 | 0.23 | 0.64 | 0.72 | 88.89 |
| OGIS (mL/min/m²) | 31.43† | 87.13 | 413.10 | 21.1 | 21.88† | 60.67 | 303.69 | 19.98 | 34.91† | 96.77 | 462.84 | 20.91 | 34.90† | 96.74 | 475.79 | 20.33 |

%B, beta cell function; Change % = per cent change of the repeatability coefficient relative to the Grand mean; HOMA2, homeostasis model assessment; IR, insulin resistance; OGIS, oral glucose insulin sensitivity test; Rep Coef, repeatability coefficient; %S, insulin sensitivity; $s_w$, residual mean square error; $\hat{\mu}$ = Grand mean.
†Denotes a non-significant mean–variance relationship from Table 1.
OGIS was conducted with three repeated tests. All other variables had eight repeated tests.

HOMA2 variants, the repeatability coefficient ranged between 60% and 170% of their respective Grand means. Within the HOMA2 variants, only HOMA2 IR had a relatively consistent percentage change throughout the participant sets (~90%).

OGIS was the only index to have a non-significant mean–variance relationship across all four participant sets (Table 1). The magnitude of the repeatability coefficient for the Diabetes set (60.7 mL/min/m²) was markedly different to both the

Control (96.7 mL/min/m²) and No Diabetes sets (96.8 mL/min/m²). However, relative to their respective Grand means, this represents a relatively consistent percentage change of approximately 20% in all analysis sets.

## Discussion

We examined the reproducibility of fasting and dynamic measures of insulin resistance by calculating the repeatability coefficient for the HOMA2 variants and OGIS in people with and without type 2 diabetes. HOMA2 and OGIS were chosen as representative of techniques assessing insulin resistance that did not require specialised methods or monitoring. The computer-modelled HOMA2 was also chosen above the original HOMA calculations as we believed that HOMA2 would become the more widely used model.

Although the absolute value of the repeatability coefficients varied by participant subset, measures based on fasting insulin, including all HOMA2 variants, had a large repeatability coefficient, meaning that they would require a large change relative to the population mean in order to indicate clinical change with confidence. This suggests that the main driver of the variation is fasting insulin given its wider spread as compared to fasting glucose (Figure 2). It is unclear why OGIS had a narrower spread. The results may be confounded by a fewer number of repeated measures for OGIS but also may reflect normal homeostasis maintaining a narrower blood glucose level compared to blood insulin levels. (OGIS is based on three glucose blood levels and two different insulin levels.)

We chose to express figures as percentages of the grand mean for two main reasons. Firstly, to determine whether there was consistency in the magnitude of change required when examining the different participant subsets. It was also believed that calculating percentages may be easier for clinicians than using different values based on glycaemic status.

However, when the repeatability coefficient was converted to a percentage of the grand mean of that subset, only certain measures showed this consistency. For example, in order to be confident clinical change has occurred, a subsequent test of either fasting insulin or HOMA2 IR needs to differ from a former test by approximately 90% irrespective of the participant subset. Depending on the individual's glycaemic status, this equates to a change in fasting insulin of between 6 µU/mL and 8 µU/mL or HOMA2 IR values of 0.64–1.07. Similarly, OGIS only required a 20% (60.67 mL/min/m²–96.74 mL/min/m²) difference. By contrast, HOMA2 %B required a change of 99.45% (36.26) for the Diabetes set, compared to only 63.24% (57.31) for the Control set.

These findings suggest that OGIS should be preferred over either fasting insulin or any variant of HOMA2, when assessing individuals or small populations for insulin resistance for the reason that OGIS appears to be more sensitive to clinical change.

There are limited data on the test–retest repeatability of measures of insulin resistance, especially including the use of the repeatability coefficient. The original HOMA IR has been reported as needing to change by +90% or -47% in patients with type 2 diabetes to ensure that the second sample is clinically significant when compared to a previous sample.[23] A direct comparison between studies is inappropriate given the two different HOMA models used; however, both studies show that large changes are needed in HOMA IR in people with type 2 diabetes to ensure that there is clinical change. The degree of change that is required suggests that HOMA IR is an impractical clinical measure in people with diabetes.

Using the repeatability coefficient rather than CV meant that it was harder to compare our results to the existing literature. However, CV can be derived from the repeatability coefficient using the following calculation in Equation 5:

$$CV \approx \frac{s_w}{\hat{\mu}} \approx \frac{Rep\,Coef}{2.77 \times \hat{\mu}} \qquad \text{[Eqn 5]}$$

Using this conversion, our results align with current CV reports. Gordon and colleagues reported the CV for OGIS to be 7.8% (range, 4.2% – 14.2%) for eight people with four repeated tests.[24] This is comparable to our CV of approximately 7%. Widjaja et al. reported a within-subject CV for fasting insulin of 26% for daily measures taken over 12 days.[12] This is also comparable to our findings of 32%. Higher CV results for HOMA were also noted by Mather and colleagues, who reported a CV of 58% for HOMA IR in subjects with a BMI > 27 kg/m² compared to 24% in subjects with a BMI of < 27 kg/m².[11] Although not directly comparable given the two different models, we recorded a lower CV of 32% for HOMA2 IR.

Another challenge with comparing repeatability studies is the assessment methods for quantifying insulin concentrations. Insulin concentrations will vary depending on analytical method, including the use of plasma or serum.[25,26] Although some discrepancies may have resolved with improving technologies, they should be considered if older studies are referenced. Participant factors such as age, sex, body fat distribution and health status may also affect insulin sensitivity.[27] It remains unknown whether these different groups would have similar repeatability coefficients. Repeatability studies may also focus on biological variation by excluding analytical variation from the overall variation.[12] While excluding analytical variation may be appropriate for the research paradigm, it is impractical in clinical practice. The number of repeated tests also varied with some studies using duplicated measures, while others used three or more measures. These factors may explain why some studies show good repeatability for HOMA variants, while others show a much wider variation. The differences in these factors may also impede direct comparisons or generalisations among studies.

This study highlighted that for many variables, the repeatability coefficient differed according to the subset of the study population. For this reason, we converted the absolute

figures of the repeatability coefficient to a percentage of the Grand mean to determine if there were consistencies throughout the subgroupings (Table 2). It was believed that should a consistency be found, then percentages may be more practical as (1) fewer figures would need to be remembered and (2) the patient would not have to be subclassified.

A notable finding of our study was the maintenance of a positive and significant mean-variance relationship for almost all the study variables, including fasting glucose; OGIS was the only variable that consistently lacked a positive and significant mean–variance relationship. These positive mean–variance relationships mean that the repeatability coefficient may be over- or underestimated at the extremes of the ranges of observed test results. Given that measures based on fasting insulin required 60% – 175% difference in results to ensure clinical change, the influence of the bias may not matter. What was clear is that OGIS did not have a positive and significant mean–variance relationship for any subgrouping tested, and although the repeatability coefficient altered depending on the subgroup, it remained a consistent 20% of the population grand mean.

The large number of repeated tests of fasting measures ($n = 8$) was a particular strength for our study. We were also able to assess these measures in people both with type 2 diabetes ($n = 10$) and with normal glucose tolerance ($n = 22$).

There were a number of limitations to our study. While we were able to assess HOMA through eight repeated tests, we only had data from three repeated tests by which to assess OGIS. However, previous research has only assessed the repeatability of OGIS via duplicated tests.[15] This suggests that further research should assess the repeatability of OGIS using a larger sample of repeated tests. OGIS has only been validated against the hyperinsulinaemic–euglycaemic clamp test for the glucose 75-g, 3-h test.[15] This study used 50 g glucose, and while the participants with diabetes had a 3-h test, those without diabetes only had a 2-h test. The original participant height and weight data were no longer available. Therefore, we applied a standard height and weight for each participant. Although this would not have affected the within-subject variability, it would have reduced the between-subject variability. People with diabetes took their regular medication (metformin, pioglitazone and an unspecified sulphonylurea) as part of the study. While this reflects their normal post-prandial response, the medication plausibly decreased the within-subject variation. Further research should investigate within- and between-subject variation in people with type 2 diabetes without the influence of medication.

The results from our study highlight that HOMA2 variants are insufficient to detect small, but clinically significant changes within an individual. This suggests that HOMA2, and potentially other measures based on fasting insulin, may be impractical for use within clinical practice or small-scale research projects. The implications of high degree of variation for larger scale research projects are not yet known. HOMA

has been used to either classify participants or assess the effects of an intervention. These results suggest that using measures based on fasting insulin for baseline stratification may not be effective as participants may have different results on different testing occasions. If these variables are to be used as a primary outcome, then power calculations should ensure that the study has a sufficient sample size in order to accurately detect change. Many studies do not use HOMA as a primary outcome, and this would be reflected in sample size. In a placebo-controlled intervention study, in order to detect a 15% change in HOMA2 IR in people with normal glucose tolerance, a target sample size of 55 people in each arm is needed to provide 80% power at the 0.05 level of significance using a two-sample $t$-test. This assumes our detected standard deviation of 0.28 applies to both arms. In people with type 2 diabetes, the increased standard deviation of 0.38 then requires a target population of 100 people in each arm. Many studies do not have these participant numbers; therefore, we could not be confident that any documented changes in HOMA2 variables resulting from different interventions would be legitimate outcomes.

We aimed this study towards clinical practice and translational research. HOMA and HOMA2 variants are widely used in many areas of research including the assessment of interventions aimed at improving insulin sensitivity. Although measuring insulin resistance is discouraged in medical practice as is does not enhance disease risk calculations,[7] use of HOMA as a medical diagnostic tool is becoming more common. We also chose to use the repeatability coefficient, rather than the more commonly used CV as it was considered to be more easily interpreted by clinicians. The repeatability coefficient allows an easier interpretation of whether the differences between two measures are biological variation ('noise') or clinical change. Ease of use by the clinician was also the reason we chose to include the analytical variation as part of the within-subject variation.

We accept that measures based on fasting insulin are much cheaper and less demanding than those based on the results derived from an oral glucose tolerance test. This may partially explain the popularity of HOMA. We further recognise that only recommending tests based on an oral glucose tolerance test would likely result in fewer assessments of insulin resistance. But should we settle for convenience over accuracy?

There may be a number of physiological explanations for our observations. Insulin is secreted from the pancreas via a pulsatile pattern leading to oscillatory blood concentrations. These oscillations have a slow ultradian periodicity (~140 min), which is modulated by a small-amplitude, high-frequency oscillation (periodicity 3–10 min).[28,29] Collectively, these oscillations have the potential to confound clinical sampling, as it will be unknown where the insulin level lies in these periodic cycles. It is believed that the periodicity of the high-frequency oscillations may vary depending on the individual's glycaemic status,[30] but there is insufficient

evidence to be able to predict an individual's oscillation periodicity or their intra-individual variation. Other factors known to confound insulin resistance such as cortisol,[31] glucagon[30] or sunlight or vitamin D exposure[32] were unable to be assessed in this study. Our results also lend support to the theory that given the high intra-individual variation, fasting insulin tests should be based on the mean of three samples taken at 5-min intervals.[8] However, this method of multiple samples rarely occurs in practice and was not employed in this study.

## Conclusion

We believe that early detection of insulin resistance is important given it precedes or is highly associated with a range of metabolic diseases. Although HOMA measures are a convenient method for assessing insulin resistance, their high variability impedes accuracy in diagnosis and monitoring clinical change. A subsequent HOMA2 test needs to change by approximately 90% to be confident that clinical change has occurred. Dynamic methods such as OGIS may have a significantly higher degree of repeatability, although they need further study. Considering the global incidence of insulin resistance and its impact on non-communicable disease, our findings support the inclusion of OGIS in clinical settings.

## Acknowledgements

### Competing interest

Grant Schofield is the Chief Education Health and Nutrition Advisor for the New Zealand Ministry of Education and serves as a board member for Health Promotion Agency of New Zealand, a government agency.

Thomas Wolever and his wife are employees and part-owners of GI Labs, Inc., a contract research organisation, and of Glycaemic Index Testing, Inc., a corporation that provides statistical services related to calculation of the glycaemic index.

Mark C. Wheldon, Xiaomiao Lan-Pidhainy and Caryn Zinn have no conflicts of interest to declare.

### Author contributions

C.A.P.C. was the project lead and writer and was responsible for concept and design, data analysis and interpretation. M.C.W. was the reviewer, was responsible for concept and design, and provided statistical expertise. C.Z. was the reviewer and performed data analysis and interpretation. X.L.-P. was responsible for concept and design and data acquisition. T.M.S.W. was responsible for concept and design and data acquisition. G.S. was the reviewer and performed data analysis and interpretation.

## References

1. Chen L, Magliano DJ, Zimmet PZ. The worldwide epidemiology of type 2 diabetes mellitus – present and future perspectives. Nat Rev Endocrinol. 2012;8(4):228–236. https://doi.org/10.1038/nrendo.2011.183

2. Zethelius B, Cederholm J. Comparison between indexes of insulin resistance for risk prediction of cardiovascular diseases or development of diabetes. Diabetes Res Clin Pract. 2015;110(2):183–192. https://doi.org/10.1016/j.diabres.2015.09.003

3. Moghetti P. Insulin resistance and polycystic ovary syndrome. Curr Pharm Des. 2016;22(36):5526–5534. https://doi.org/10.2174/1381612822666160720155855

4. Bugianesi E, Moscatiello S, Ciaravella MF, Marchesini G. Insulin resistance in nonalcoholic fatty liver disease. Curr Pharm Des. 2010;16(17):1941–1951. https://doi.org/10.2174/138161210791208875

5. DeFronzo RA, Tobin JD, Andres R. Glucose clamp technique: A method for quantifying insulin secretion and resistance. Am J Physiol Endocrinol Metab. 1979;237(3):E214–E223.

6. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: Comparison with the euglycemic insulin clamp. Diabetes Care. 1999;22(9):1462–1470. https://doi.org/10.2337/diacare.22.9.1462

7. Samaras K, McElduff A, Twigg SM, et al. Insulin levels in insulin resistance: Phantom of the metabolic opera? Med J Aust. 2006;185(3):159.

8. Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. Diabetes Care. 2004;27(6):1487–1495. https://doi.org/10.2337/diacare.27.6.1487

9. Diabetes Trials Unit. HOMA calulator [homepage on the Internet]. 2015 [cited 2017 May 12]. Available from: http://www.dtu.ox.ac.uk/homacalculator/

10. Lotz TF, Chase JG, McAuley KA, et al. Monte Carlo analysis of a new model-based method for insulin sensitivity testing. Comput Methods Programs Biomed. 2008;89(3):215–225. https://doi.org/10.1016/j.cmpb.2007.03.007

11. Mather KJ, Hunt AE, Steinberg HO, et al. Repeatability characteristics of simple indices of insulin resistance: Implications for research applications. J Clin Endocrinol Metab. 2001;86(11):5457–5464. https://doi.org/10.1210/jcem.86.11.7880

12. Widjaja A, Morris RJ, Levy JC, et al. Within- and between-subject variation in commonly measured anthropometric and biochemical variables. Clin Chem. 1999;45(4):561–566.

13. Crofts C, Schofield G, Zinn C, Wheldon M, Kraft J. Identifying hyperinsulinaemia in the absence of impaired glucose tolerance: An examination of the Kraft database. Diabetes Res Clin Pract. 2016;118:50–57. https://doi.org/10.1016/j.diabres.2016.06.007

14. Hayashi T, Boyko EJ, Sato KK, et al. Patterns of insulin concentration during the OGTT predict the risk of type 2 diabetes in Japanese Americans. Diabetes Care. 2013;36(5):1229–1235. https://doi.org/10.2337/dc12-0246

15. Mari A, Pacini G, Murphy E, Ludvik B, Nolan J. A model-based method for assessing insulin sensitivity from the oral glucose tolerance test. Diabetes Care. 2001;24(3):539–548. https://doi.org/10.2337/diacare.24.3.539

16. Mari A. OGIS: Insulin sensitivity from the oral glucose test [homepage on the Internet]. n.d. [cited 2017 May 12]Available from: http://webmet.pd.cnr.it/ogis/index.php

17. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8(2):135–160. https://doi.org/10.1177/096228029900800204

18. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1–15. https://doi.org/10.2165/00007256-200030010-00001

19. Lan-Pidhainy X, Wolever T. Are the glycemic and insulinemic index values of carbohydrate foods similar in healthy control, hyperinsulinemic and type 2 diabetic patients? Eur J Clin Nutr. 2011;65(6):727–734. https://doi.org/10.1038/ejcn.2011.28

20. McAuley KA, Williams SM, Mann JI, et al. Diagnosing insulin resistance in the general population. Diabetes Care. 2001;24(3):460–464. https://doi.org/10.2337/diacare.24.3.460

21. Kraft JR. Detection of diabetes mellitus in situ (occult diabetes). Lab Med. 1975;6(2):10–22. https://doi.org/10.1093/labmed/6.2.10

22. Diabetes Trials Unit. HOMA calculator [homepage on the Internet]. 2004 [cited 2013 June 07]. Available from: http://www.dtu.ox.ac.uk/homacalculator/index.php

23. Jayagopal V, Kilpatrick ES, Jennings PE, Hepburn DA, Atkin SL. Biological variation of homeostasis model assessment-derived insulin resistance in type 2 diabetes. Diabetes Care. 2002;25(11):2022–2025. https://doi.org/10.2337/diacare.25.11.2022

24. Gordon B, Fraser S, Bird S, Benson A. Reproducibility of multiple repeated oral glucose tolerance tests. Diabetes Res Clin Pract. 2011;94(3):e78–e82. https://doi.org/10.1016/j.diabres.2011.08.025

25. Henderson JR. Serum-insulin or plasma-insulin? Lancet. 1970;296(7672):545–547. https://doi.org/10.1016/S0140-6736(70)91348-6

26. Manley SE, Stratton IM, Clark PM, Luzio SD. Comparison of 11 human insulin assays: Implications for clinical investigation and research. Clin Chem. 2007;53(5):922–932. https://doi.org/10.1373/clinchem.2006.077784

27. Karakelides H, Irving BA, Short KR, O'Brien P, Nair KS. Age, obesity, and sex effects on insulin sensitivity and skeletal muscle mitochondrial function. Diabetes. 2010;59(1):89–97. https://doi.org/10.2337/db09-0591

28. Hellman B. Pulsatility of insulin release – A clinically important phenomenon. Upsala J Med Sci. 2009;114(4):193–205. https://doi.org/10.3109/03009730903366075

29. Schmitz O, Rungby J, Edge L, Juhl CB. On high-frequency insulin oscillations. Ageing Res Rev. 2008;7(4):301–305. https://doi.org/10.1016/j.arr.2008.04.002

30. Rohrer S, Menge BA, Grüber L, et al. Impaired crosstalk between pulsatile insulin and glucagon secretion in prediabetic individuals. J Clin Endocrinol Metab. 2012;97(5):E791–E795. https://doi.org/10.1210/jc.2011-3439

31. Anagnostis P, Athyros VG, Tziomalos K, Karagiannis A, Mikhailidis DP. The pathogenetic role of cortisol in the metabolic syndrome: A hypothesis. J Clin Endocrinol Metab. 2009;94(8):2692–2701. https://doi.org/10.1210/jc.2009-0370

32. Belenchia AM, Tosh AK, Hillman LS, Peterson CA. Correcting vitamin D insufficiency improves insulin sensitivity in obese adolescents: A randomized controlled trial. Am J Clin Nutr. 2013;97(4):774–781. https://doi.org/10.3945/ajcn.112.050013