Supporting Document

Hyperbolic relationship between insulin secretion and sensitivity

The relationship between two variables *x* and *y* is said to be hyperbolic if $x \times y = c$, where c is a constant. The *x* and *y* variables can be log-transformed ($\ln(x \times y) = \ln(x) + \ln(y) = \ln(c)$), and the hyperbolic relationship between these two variables can be re-expressed as linear model such that: $\ln(y) = \ln(c) - 1\ln(x)$, where the regression coefficient in the linear model of $\ln(y)$ as a function of $\ln(x)$ is $\beta = -1$.[1, 2] Note that β = -1 regardless of the units in which y and x are expressed and regardless of the base of logarithms used. The hypothesis that the relationship between insulin secretion and insulin sensitivity is hyperbolic is usually tested by determining if the slope of a linear model of *ln*(*insulin secretion*) as a function of *ln*(*insulin sensitivity*) is not significantly different from -1 (i.e., the 95% CI for the slope includes -1).[1-4]

Reasons for using SMA regression to model relationships between insulin secretion and sensitivity

We studied the relationship between insulin secretion and sensitivity by estimating *ln(insulin secretion)* as a function of *ln(insulin sensitivity)*. The models provided by the SMA regression were chosen after comparing their fit with models from linear and orthogonal regression using standard error of the residuals and confidence intervals for intercepts and slopes, Supplementary Tables S1 and S2. One reason for not considering linear regression models in this case is that this method only accounts for variability in *y*, while orthogonal and SMA regression models account for variability in both *x* and *y*.[5] Also, in our study, linear regression models had the highest standard error of the residuals in all cases, Supplementary Tables S1 and S2. Standard errors of the residuals were the same or slightly lower for orthogonal regression models when compared to those from SMA regression. However, SMA models were chosen because orthogonal regression is most useful when one is studying the relationship between two variables that are estimates of the same entity.[5] Another reason for choosing SMA models was that, their confidence intervals are estimated more efficiently than those for orthogonal regression models, and the confidence intervals for SMA slopes tend to be exact or close to exact in most practical instances.[6]

Generally, SMA and orthogonal regression models led to the same conclusions (same qualitative results). In the OGTT dataset, both SMA and orthogonal regression indicated that the slope for $\ln(CIR_{120})$ as a function of $\ln(ISI_0)$ was essentially -1 in the whole dataset (SMA: (95% CI -0·999 to -0·908); orthogonal: (95% CI -0·996 to -0·686)), significantly different from -1 in NGR subjects only (SMA: (95% CI -0·948 to -0·854); orthogonal: (95% CI -0·869 to -0·656)), and that it was not significantly different from -1 in the IGR subjects only (SMA: (95% CI -1·093 to -0·918); orthogonal: (95% CI -1·180 to -0·853)). In the IV/CLAMP dataset, both approaches suggested that the slope for ln(AIR) as a function of ln(M) was significantly different from -1 in the whole dataset (SMA: (95% CI -2·442 to -2·027); orthogonal: (95% CI -12·550 to -5·454)), NGR subjects only (SMA: (95% CI -2·333 to -1·898); orthogonal: (95% CI -6·745 to -3·701)) and IGR subjects only (SMA: (95% CI -4·146 to -2·803); orthogonal: (95% bootstrap CI -32·545 to -20·577)). In the OGTT dataset, both approaches indicated that the slope of line describing the relationship between $\ln(CIR_{120})$ and $\ln(ISI_0)$ in NGR subjects was significantly different from the slope describing this relationship in IGR subjects (SMA: LRS=4·26, p=0·04; orthogonal: LRS=6·69, p=0·01).

Conclusions from SMA and orthogonal regression models differed when testing whether BMI associated with the relationship between insulin secretion and sensitivity in the OGTTdataset. The test for common slopes in the SMA model indicated that the slopes for different categories of BMI (normal weight, overweight and obese) were significantly different (LRS=7·26, p=0·03), which was not the conclusion reached with the test for common slopes in the orthogonal regression model (suggested slopes were not significantly different; LRS=5·54, p=0·06). However, when plotting the slopes for both models the difference in slopes was more pronounced for the orthogonal regression model than for the SMA model (Supplementary Figure S1). This difference in results for the common slopes test is likely due to confidence intervals for orthogonal regression models being less efficient[6] (wider) than those of the SMA models.

Additionally, SMA regression models produced stable results for all datasets and their subsets (NGR, IGR, normal weight, overweight, obese). Meanwhile, orthogonal regression models provided unstable results (lower limit of CI higher than upper limit of CI) when fit using Warton's R package[7] for subjects in the IV/CLAMP dataset in the IGR, as well as overweight and obese subsets. We also fit orthogonal regression models for all subsets of the IV/CLAMP dataset using bootstrap estimation. For the subsets for which we obtained stable estimates using Warton's R package, we obtained similar (almost the same) intercept and slope estimates with bootstrap estimation. For the subsets for which unstable estimates were obtained with Warton's R package, we found that orthogonal regression models with bootstrap estimation had lower standard error of the residuals. However, for most subsets, the values of ln(AIR) obtained when fitting lines using orthogonal regression were way outside of range of the values from which the line was estimated in the extremes, which was not the case for lines obtained from linear and SMA regression models (Supplementary Figure S2).

A hyperbolic relationship between insulin secretion and sensitivity implies that the relationship between these two variables is symmetric and that we should obtain the same line if we model *ln(insulin secretion)* as a function of *ln(insulin sensitivity)* or *ln(insulin sensitivity)* as a function of *ln(insulin secretion)*, i.e. the slope of *ln(insulin sensitivity)* as a function of *ln(insulin secretion)* is also -1. Thus, we also modeled ln(ISI_0 or M) as a function of ln(CIR_{120} or AIR) using linear, orthogonal and SMA regression, Supplementary Tables S1 and S2. In both datasets and for all their subsets, the slopes of these models obtained using orthogonal and SMA regression were the reciprocals (=1/slope) of the slopes of those obtained when modeling ln(CIR_{120} or AIR) as a function of ln(ISI_0 or M). For example, for IGR subjects in the OGTT dataset, when modeling ln(CIR_{120}) as a function of ln(ISI_0) using SMA regression, the slope was –1·002 and when modeling ln(ISI_0) as a function of ln(CIR_{120}) the slope was -0·998= $\frac{1}{-1\cdot 002}$. This showed that orthogonal and SMA regression provided the same line if we modeled *ln(insulin secretion)* as a function of *ln(insulin sensitivity)* or *ln(insulin sensitivity)* as a function of *ln(insulin secretion)*. In all cases, results for testing the hypothesis of a hyperbolic relationship were the same when modeling ln(ISI_0 or M) as a function of ln(CIR_{120} or AIR) than when modeling ln(CIR_{120} or AIR) as a function of ln(ISI_0 or M). For example, for NGR subjects in the OGTT dataset, when modeling ln(CIR_{120}) as a function of ln(ISI_0) using SMA regression, the slope was not -1 (95% CI -0·948 to -0·854), which was also the case when modeling ln(ISI_0) as a function of ln(CIR_{120}) (95% CI -1·171 to -1·054). When we modeled ln(CIR_{120} or AIR) as a function of ln(ISI_0 or M) using linear regression, we obtained different lines than when modeling ln(ISI_0 or M) as a function of ln(CIR_{120} or AIR) (slopes were not reciprocal) for all subsets in both datasets. This highlighted another disadvantage of the linear regression models in our study, where we were interested in checking for a symmetric relationship.

Supplementary Table S1. Estimated relationship between $\ln(CIR_{120})$ and $\ln(ISI_0)$ in OGTT dataset using linear, orthogonal and SMA regression

		$\ln(CIR_{120})$ as function of $\ln(ISI_0)$ (Model of Interest)			$\ln(ISI_0)$ as function of $\ln(CIR_{120})$	
	Regression model	Parameters: Intercept (95% CI)	Parameters: Slope (95% CI)	S·E. Residuals	Slope (95% CI)	Slope is reciprocal to that of $\ln(CIR_{120})$ as function of $\ln(ISI_0)$
Glucose Tolerance Groups						
NGR	Linear regression	-4·206 (-4·290, -4·121)	-0·337 (-0·384, -0·290)	0·546	-0·416 (-0·475, -0·358)	No
	Orthogonal regression	-4·906 (-5·086, -4·726)	-0·757 (-0·869, -0·656)	0·487	-1·321 (-1·524, -1·150)	Yes
	SMA regression	-5·144 (-5·231, -5·057)	-0·900 (-0·948, -0·854)	0·490	-1·111 (-1·171, -1·054)	Yes
IGR	Linear regression	-5·381 (-5·576, -5·186)	-0·544 (-0·631, -0·456)	0·541	-0·542 (-0·629, -0·455)	No
	Orthogonal regression	-6·363 (-6·714, -6·012)	-1·003 (-1·180, -0·853)	0·435	-0·997 (-1·172, -0·847)	Yes
	SMA regression	-6·360 (-6·558, -6·162)	-1·002 (-1·093, -0·918)	0·435	-0·998 (-1·089, -0·915)	Yes
Body Mass Index Groups						
Normal weight	Linear regression	-4·274 (-4·456, -4·093)	-0·419 (-0·564, -0·273)	0·666	-0·270 (-0·363, -0·176)	No
	Orthogonal regression	-5·877 (-6·607, -5·147)	-1·856 (-2·763, -1·339)	0·500	-0·539 (-0·747, -0·362)	Yes
	SMA regression	-5·197 (-5·387, -5·006)	-1·245 (-1·399, -1·109)	0·516	-0·803 (-0·902, -0·715)	Yes
Overweight	Linear regression	-4·264 (-4·449, -4·078)	-0·262 (-0·375, -0·149)	0·618	-0·185 (-0·265, -1·05)	No
	Orthogonal regression	-7·066 (-8·455, -5·678)	-2·071 (-3·509, -1·394)	0·498	-0·483 (-0·717, -0·285)	Yes
	SMA regression	-5·701 (-5·893, -5·510)	-1·190 (-1·308, -1·082)	0·513	-0·840 (-0·924, -0·764)	Yes
Obese	Linear regression	-4·303 (-4·447, -4·158)	-0·277 (-0·344, -0·209)	0·611	-0·244 (-0·304, -0·185)	No
	Orthogonal regression	-6·347 (-6·987, -5·707)	-1·267 (-1·636, -0·995)	0·522	-0·789 (-1·005, -0·611)	Yes
	SMA regression	-5·928 (-6·076, -5·780)	-1·064 (-1·134, -0·999)	0·527	-0·940 (-1·001, -0·882)	Yes

[a] Comparisons done with calculations using unrounded numbers

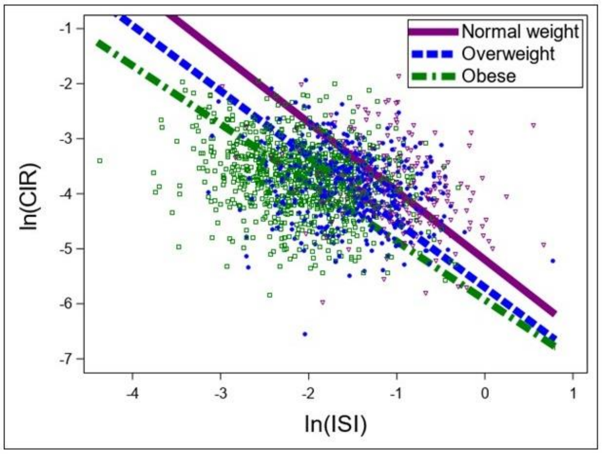
Supplementary Table S2. Estimated relationship between ln(AIR) and ln(M) in IV/CLAMP dataset using linear, orthogonal and SMA regression

		ln(AIR) as function of ln(M) (Model of Interest)			ln(M) as a function of ln(AIR)	
	Regression model	Parameters: Intercept (95% CI)	Parameters: Slope (95% CI)	S·E. Residuals	Slope (95% CI)	Slope is reciprocal to that of ln(AIR) as function of ln(M)[a]
Glucose Tolerance Groups						
NGR	Linear regression	6·374 (6·073, 6·674)	-0·747 (-0·965, -0·530)	0·585	-0·169 (-0·218, -0·120)	No
	Orthogonal regression with Warton's package	11·836 (9·945, 13·727)	-4·797 (-6·745, -3·701)	0·273	-0·208 (-0·270, -0·148)	Yes
	Orthogonal regression with bootstrap	11·858 (11·038, 12·993)	-4·813 (-5·684, -4·159)	0·274	-0·208 (-0·240, -0·176)	Yes
	SMA regression	8·204 (7·900, 8·508)	-2·104 (-2·333, -1·898)	0·347	-0·475 (-0·527, -0·429)	Yes
IGR	Linear regression	5·626 (4·857, 6·400)	-0·421 (-1·092, 0·250)	0·716	-0·036 (-0·094, 0·022)	No
	Orthogonal regression with Warton's package	33·607 (-11·736, 78·950)	-25·252 (Unstable CI)	0·210	-0·040 (-0·103, 0·024)	Yes
	Orthogonal regression with bootstrap	33·597 (28·398, 41·769)	-25·241 (-32·545, -20·577)	0·210	-0·040 (-0·049, -0·031)	Yes
	SMA regression	8·993 (8·214, 9·772)	-3·409 (-4·146, -2·803)	0·366	-0·293 (-0·357, -0·241)	Yes
Body Mass Index Groups						
Normal weight	Linear regression	5·991 (4·980, 7·002)	-0·641 (-1·228, -0·054)	0·625	-0·137 (-0·262, -0·012)	No
	Orthogonal regression with Warton's package	14·928 (5·740, 24·116)	-5·912 (-68·186, -3·013)	0·286	-0·169 (-0·332, -0·015)	Yes
	Orthogonal regression with bootstrap	14·918 (13·435, 17·116)	-5·904 (-7·232, -5·011)	0·286	-0·169 (-0·180, -0·158)	No (diff=0·01) [b]
	SMA regression	8·573 (7·554, 9·592)	-2·164 (-2·830, -1·655)	0·373	-0·462 (-0·604, -0·353)	Yes
Overweight	Linear regression	5·398 (4·782, 6·013)	-0·124 (-0·546, 0·298)	0·534	-0·030 (-0·131, 0·071)	No
	Orthogonal regression with Warton's package	41·892 (-82·628, 166·413)	-25·546 (Unstable CI)	0·261	-0·039 (-0·175, 0·095)	Yes
	Orthogonal regression with bootstrap	41·913 (32·446, 60·661)	-25·548 (-38·671, -18·934)	0·261	-0·039 (-0·050, -0·029)	No (diff=0·024) [b]
	SMA regression	8·151 (7·527, 8·775)	-2·042 (-2·508, -1·663)	0·362	-0·490 (-0·601, -0·399)	Yes
Obese	Linear regression	5·624 (5·158, 6·090)	-0·171 (-0·563, 0·222)	0·66	-0·016 (-0·052, 0·020)	No
	Orthogonal regression with Warton's package	72·909 (-82·271, 228·088)	-57·639 (Unstable CI)	0·201	-0·017 (-0·057, 0·023)	Yes
	Orthogonal regression with bootstrap	72·647 (55·087, 110·231)	-57·389 (-89·532, -42·396)	0·201	-0·017 (-0·022, -0·012)	No (diff=0·105) [b]
	SMA regression	9·277 (8·805, 9·749)	-3·291 (-3·707, -2·922)	0·355	-0·304 (-0·342, -0·269)	Yes

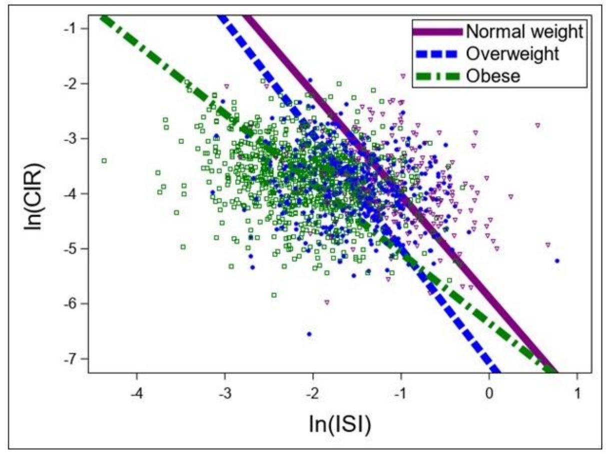
[a]Comparisons done with calculations using unrounded numbers. [b]Bootstrap distribution for ln(AIR) as function of ln(M) not symmetric, bootstrap distribution for ln(M) as function of ln(AIR) symmetric

Supplementary Figure S1. Visual test for common slopes between different BMI categories for SMA and orthogonal regression in OGTT dataset

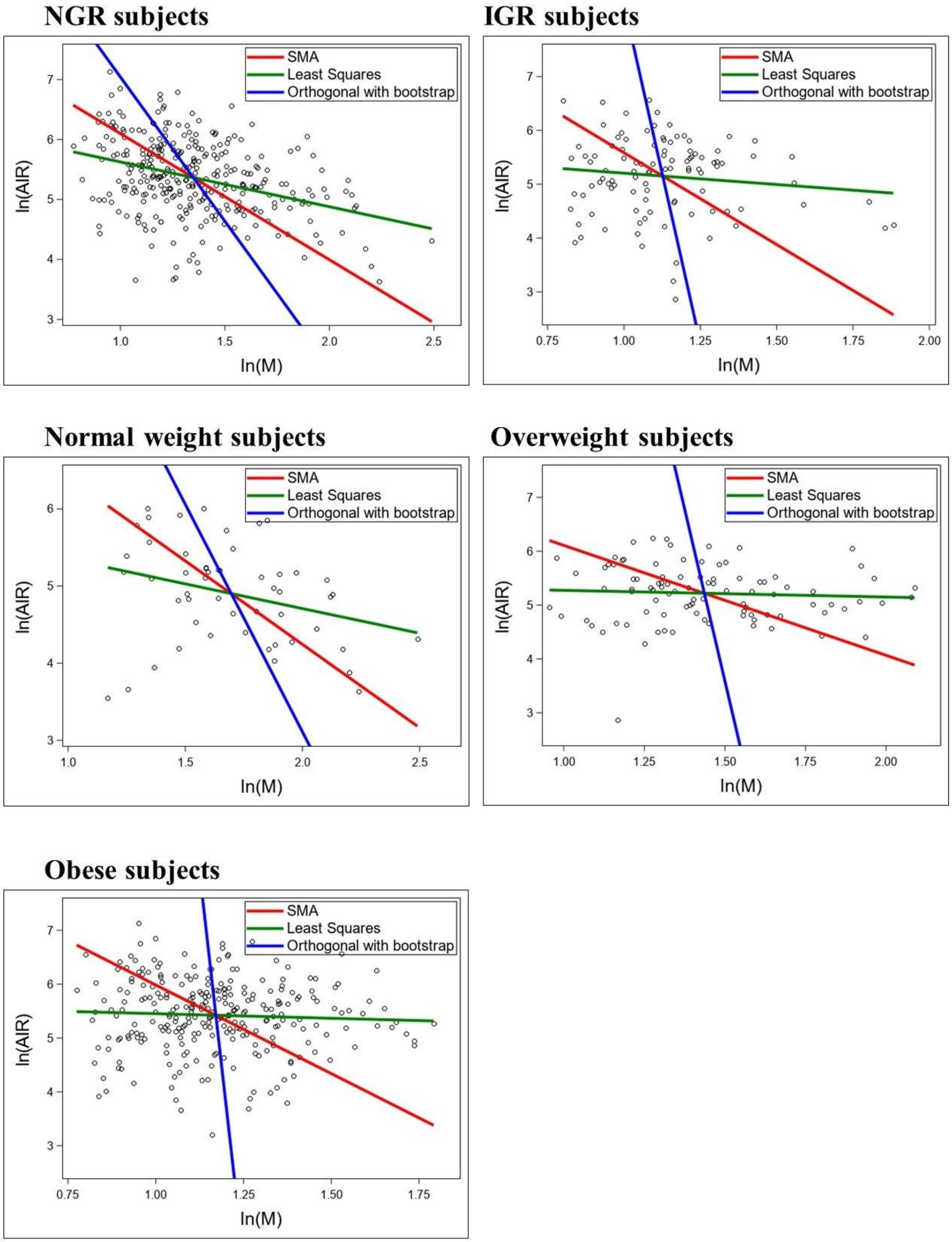
SMA regression



Orthogonal regression



Supplementary Figure S2. Lines estimated for NGR, IGR, normal weight, overweight, and obese subjects in IV/CLAMP dataset using linear regression, SMA regression, and orthogonal regression with bootstrap estimation



References

1. Kahn SE, Prigeon RL, Mcculloch DK, Boyko EJ, Bergman RN, Schwartz MW, et al. Quantification of the Relationship between Insulin Sensitivity and Beta-Cell Function in Human-Subjects - Evidence for a Hyperbolic Function. Diabetes. 1993;42(11):1663-72.
2. Kim SH, Silvers A, Viren J, Reaven GM. Relationship between insulin sensitivity and insulin secretion rate: not necessarily hyperbolic. Diabetic Med. 2016;33(7):961-7.
3. Utzschneider KM, Prigeon RL, Carr DB, Hull RL, Tong J, Shofer JB, et al. Impact of differences in fasting glucose and glucose tolerance on the hyperbolic relations between insulin sensitivity and insulin responses. Diabetes Care. 2006;29(2):356-62.
4. Utzschneider KM, Prigeon RL, Faulenbach MV, Tong J, Carr DB, Boyko EJ, et al. Oral disposition index predicts the development of future diabetes above and beyond fasting and 2-h glucose levels. Diabetes Care. 2009;32(2):335-41.
5. Patzer ABC, Bauer H, Chang C, Bolte J, Sulzle D. Revisiting the Scale-Invariant, Two-Dimensional Linear Regression Method. J Chem Educ. 2018;95(6):978-84.
6. Warton DI, Wright IJ, Falster DS, Westoby M. Bivariate line-fitting methods for allometry. Biol Rev Camb Philos Soc. 2006;81(2):259-91.
7. Warton DI, Duursma RA, Falster DS, Taskinen S. smatr 3-an R package for estimation and inference about allometric lines. Methods Ecol Evol. 2012;3(2):257-9.

Supplementary Table S3: Classifications of IFG, IGT, and NGR by WHO criteria in the OGTT dataset

	2-hr PG <140	140≤ 2-hr PG<200
FPG<110	NGR=1207	IGT=271
110≤FPG<126	IFG=29	IGT=59

Supplementary Table S4: Classifications of IFG, IGT, and NGR by WHO criteria in the IV/CLAMP dataset

	2-hr PG <140	140≤ 2-hr PG<200
FPG<110	NGR=318	IGT=96
110≤FPG<126	IFG=2	IGT=4

Supplementary Table S5: Correlations between ln(DI) with distance away from line, and between ln(BCDI) with distance along the line

		OGTT dataset		**IV/CLAMP dataset**	
Group		Distance away from line	Distance along the line	Distance away from line	Distance along the line
Whole	ln(DI)	0·999 (p<0·0001)	··	0·900 (p<0·0001)	··
	ln(BCDI)	··	1·000 (p<0·0001)	··	0·982 (p<0·0001)
NGR	ln(DI)	0·997 (p<0·0001)	··	0·889 (p<0·0001)	··
	ln(BCDI)	··	0·999 (p<0·0001)	··	0·986 (p<0·0001)
IGR	ln(DI)	1·000 (p<0·0001)	··	0·850 (p<0·0001)	··
	ln(BCDI)	··	1·000 (p<0·0001)	··	0·982 (p<0·0001)

Note: Correlations in NGR and IGR subjects only are calculated based on distances away from and along lines calculated in NGR and IGR subjects separately.

Supplementary Table S6: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the OGTT dataset for all subjects

Models		AUC	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	$\ln(CIR_{120})$	0·099 (0·010)	<0·0001
ln(DI)	$\ln(ISI_0)$	0·022 (0·010)	0·02
$\ln(CIR_{120})$ & $\ln(ISI_0)$	ln(DI)	0·008 (0·004)	0·08
ln(DI)	Distance away from the line	0·002 (0·001)	<0·0001
ln(DI)	Distance along the line	0·094 (0·014)	<0·0001
Distance away from the line and distance along the line	ln(DI)	0·008 (0·004)	0·08
$\ln(CIR_{120})$	**$\ln(ISI_0)$**	-0·077 (0·015)	<0·0001
$\ln(CIR_{120})$ & $\ln(ISI_0)$	$\ln(CIR_{120})$	0·106 (0·012)	<0·0001
$\ln(CIR_{120})$	**Distance away from the line**	-0·097 (0·010)	<0·0001
$\ln(CIR_{120})$	**Distance along the line**	-0·005 (0·016)	0·76
Distance away from the line and distance along the line	$\ln(CIR_{120})$	0·106 (0·012)	<0·0001
$\ln(CIR_{120})$ & $\ln(ISI_0)$	$\ln(ISI_0)$	0·030 (0·006)	<0·0001
$\ln(ISI_0)$	**Distance away from the line**	-0·020 (0·010)	0·05
$\ln(ISI_0)$	Distance along the line	0·072 (0·007)	<0·0001
Distance away from the line and distance along the line	$\ln(ISI_0)$	0·030 (0·006)	<0·0001
$\ln(CIR_{120})$ & $\ln(ISI_0)$	Distance away from the line	0·010 (0·005)	0·04
$\ln(CIR_{120})$ & $\ln(ISI_0)$	Distance along the line	0·102 (0·011)	<0·0001
$\ln(CIR_{120})$ & $\ln(ISI_0)$	Distances	0·000 (0·000)	>0·99
Distance away from the line	Distance along the line	0·092 (0·014)	<0·0001
Distance away from the line and distance along the line	Distance away from the line	0·010 (0·005)	0·04
Distance away from the line and distance along the line	Distance along the line	0·102 (0·012)	<0·0001

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 3.

Supplementary Table S7: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the OGTT dataset for NGR subjects only

Models		AUC	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	ln(CIR_{120})	0·083 (0·020)	<0·0001
ln(DI)	**ln(ISI_0)**	-0·014 (0·012)	0·23
ln(CIR_{120}) & ln(ISI_0)	ln(DI)	0·020 (0·009)	0·04
ln(DI)	Distance away from the line	0·006 (0·001)	<0·0001
ln(DI)	Distance along the line	0·021 (0·017)	0·21
Distance away from the line and distance along the line	ln(DI)	0·020 (0·009)	0·04
ln(CIR_{120})	**ln(ISI_0)**	-0·098 (0·015)	<0·0001
ln(CIR_{120}) & ln(ISI_0)	ln(CIR_{120})	0·103 (0·016)	<0·0001
ln(CIR_{120})	**Distance away from the line**	-0·078 (0·020)	<0·001
ln(CIR_{120})	**Distance along the line**	-0·063 (0·011)	<0·0001
Distance away from the line and distance along the line	ln(CIR_{120})	0·103 (0·016)	<0·0001
ln(CIR_{120}) & ln(ISI_0)	ln(ISI_0)	0·005 (0·003)	0·08
ln(ISI_0)	Distance away from the line	0·020 (0·013)	0·12
ln(ISI_0)	Distance along the line	0·035 (0·007)	<0·0001
Distance away from the line and distance along the line	ln(ISI_0)	0·005 (0·003)	0·08
ln(CIR_{120}) & ln(ISI_0)	Distance away from the line	0·025 (0·011)	0·02
ln(CIR_{120}) & ln(ISI_0)	Distance along the line	0·040 (0·009)	<0·0001
Distance away from the line	Distance along the line	0·015 (0·018)	0·39
Distance away from the line and distance along the line	Distance away from the line	0·025 (0·011)	0·02
Distance away from the line and distance along the line	Distance along the line	0·040 (0·009)	<0·0001

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 3.

Supplementary Table S8: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the OGTT dataset for IGR subjects only

Models		AUC	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	ln(CIR_{120})	0·080 (0·024)	0·001
ln(DI)	ln(ISI_0)	0·061 (0·023)	0·01
ln(CIR_{120}) & ln(ISI_0)	**ln(DI)**	-0·003 (0·007)	0·67
ln(DI)	Distance away from the line	0·000 (0·0002)	>0·99
ln(DI)	Distance along the line	0·117 (0·028)	<0·0001
Distance away from the line and distance along the line	**ln(DI)**	-0·003 (0·007)	0·67
ln(CIR_{120})	**ln(ISI_0)**	-0·019 (0·030)	0·54
ln(CIR_{120}) & ln(ISI_0)	ln(CIR_{120})	0·077 (0·026)	0·003
ln(CIR_{120})	**Distance away from the line**	-0·080 (0·024)	0·001
ln(CIR_{120})	Distance along the line	0·037 (0·028)	0·18
Distance away from the line and distance along the line	ln(CIR_{120})	0·077 (0·026)	0·003
ln(CIR_{120}) & ln(ISI_0)	ln(ISI_0)	0·058 (0·018)	0·001
ln(ISI_0)	**Distance away from the line**	-0·061 (0·023)	0·01
ln(ISI_0)	Distance along the line	0·055 (0·011)	<0·0001
Distance away from the line and distance along the line	ln(ISI_0)	0·058 (0·018)	0·001
ln(CIR_{120}) & ln(ISI_0)	**Distance away from the line**	-0·003 (0·006)	0·67
ln(CIR_{120}) & ln(ISI_0)	Distance along the line	0·114 (0·025)	<0·0001
Distance away from the line	Distance along the line	0·117 (0·028)	<0·0001
Distance away from the line and distance along the line	**Distance away from the line**	-0·003 (0·006)	0·67
Distance away from the line and distance along the line	Distance along the line	0·114 (0·025)	<0·0001

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 3.

Supplementary Table S9: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the IV/CLAMP dataset for all subjects

Models		AUC	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	ln(AIR)	0·054 (0·012)	<0·0001
ln(DI)	**ln(M)**	-0·036 (0·028)	0·21
ln(AIR) & ln(M)	ln(DI)	0·065 (0·019)	0·001
ln(DI)	**Distance away from the line**	-0·043 (0·011)	<0·0001
ln(DI)	Distance along the line	0·076 (0·019)	<0·0001
Distance away from the line and distance along the line	ln(DI)	0·065 (0·019)	0·001
ln(AIR)	**ln(M)**	-0·090 (0·032)	0·01
ln(AIR)&ln(M)	ln(AIR)	0·119 (0·026)	<0·0001
ln(AIR)	**Distance away from the line**	-0·098 (0·020)	<0·0001
ln(AIR)	Distance along the line	0·022 (0·009)	0·01
Distance away from the line and distance along the line	ln(AIR)	0·119 (0·026)	<0·0001
ln(AIR) & ln(M)	ln(M)	0·030 (0·014)	0·03
ln(M)	**Distance away from the line**	-0·008 (0·021)	0·71
ln(M)	Distance along the line	0·111 (0·032)	0·001
Distance away from the line and distance along the line	ln(M)	0·030 (0·014)	0·03
ln(AIR) & ln(M)	Distance away from the line	0·022 (0·010)	0·02
ln(AIR) & ln(M)	Distance along the line	0·141 (0·029)	<0·0001
Distance away from the line	Distance along the line	0·119 (0·025)	<0·0001
Distance away from the line and distance along the line	Distance away from the line	0·022 (0·010)	0·02
Distance away from the line and distance along the line	Distance along the line	0·141 (0·029)	<0·0001

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 4.

Supplementary Table S10: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the IV/CLAMP dataset for NGR subjects only

Models		AUC	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	ln(AIR)	0·056 (0·018)	0·002
ln(DI)	**ln(M)**	-0·056 (0·040)	0·16
ln(AIR) & ln(M)	ln(DI)	0·072 (0·029)	0·01
ln(DI)	**Distance away from the line**	-0·050 (0·016)	0·001
ln(DI)	Distance along the line	0·070 (0·029)	0·02
Distance away from the line and distance along the line	ln(DI)	0·072 (0·028)	0·01
ln(AIR)	**ln(M)**	-0·112 (0·046)	0·02
ln(AIR)&ln(M)	ln(AIR)	0·128 (0·039)	0·001
ln(AIR)	**Distance away from the line**	-0·106 (0·029)	0·0003
ln(AIR)	Distance along the line	0·015 (0·015)	0·33
Distance away from the line and distance along the line	ln(AIR)	0·128 (0·039)	0·001
ln(AIR)&ln(M)	ln(M)	0·016 (0·017)	0·34
ln(M)	Distance away from the line	0·006 (0·029)	0·84
ln(M)	Distance along the line	0·126 (0·046)	0·01
Distance away from the line and distance along the line	ln(M)	0·016 (0·017)	0·34
ln(AIR) & ln(M)	Distance away from the line	0·022 (0·015)	0·13
ln(AIR) & ln(M)	Distance along the line	0·142 (0·042)	0·001
Distance away from the line	Distance along the line	0·120 (0·036)	0·001
Distance away from the line and distance along the line	Distance away from the line	0·022 (0·015)	0·13
Distance away from the line and distance along the line	Distance along the line	0·142 (0·042)	0·001

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 4.

Supplementary Table S11: Differences in AUC between models with baseline covariates sex, age, and fraction of American Indian heritage compared in the IV/CLAMP dataset for IGR subjects only

Models		**AUC**	
Model 1	**Model 2**	**Difference (SE)**	**P-value**
ln(DI)	ln(AIR)	0·028 (0·014)	0·05
ln(DI)	ln(M)	0·034 (0·048)	0·48
ln(AIR) & ln(M)	ln(DI)	0·023 (0·024)	0·34
ln(DI)	**Distance away from the line**	-0·024 (0·023)	0·30
ln(DI)	Distance along the line	0·036 (0·018)	0·04
Distance away from the line and distance along the line	ln(DI)	0·023 (0·024)	0·34
ln(AIR)	ln(M)	0·007 (0·052)	0·90
ln(AIR)&ln(M)	ln(AIR)	0·050 (0·033)	0·12
ln(AIR)	**Distance away from the line**	-0·052 (0·032)	0·11
ln(AIR)	Distance along the line	0·009 (0·006)	0·18
Distance away from the line and distance along the line	ln(AIR)	0·050 (0·033)	0·12
ln(AIR)&ln(M)	ln(M)	0·057 (0·033)	0·09
ln(M)	**Distance away from the line**	-0·058 (0·034)	0·09
ln(M)	Distance along the line	0·002 (0·054)	0·97
Distance away from the line and distance along the line	ln(M)	0·057 (0·033)	0·09
ln(AIR)&ln(M)	**Distance away from the line**	-0·001 (0·003)	0·61
ln(AIR)&ln(M)	Distance along the line	0·059 (0·036)	0·10
Distance away from the line	Distance along the line	0·060 (0·035)	0·09
Distance away from the line and distance along the line	**Distance away from the line**	-0·0001 (0·003)	0·61
Distance away from the line and distance along the line	Distance along the line	0·059 (0·036)	0·10

Note: differences are calculated as Model1-Model2. Model with higher AUC is bolded. AUC for models included in Table 4.