

Published in final edited form as:

*Diabetologia*. 2009 September ; 52(9): 1852–1857. doi:10.1007/s00125-009-1426-4.

## The *IRS-1* G972R polymorphism and Type 2 Diabetes: a paradigm for the difficult ascertainment of the contribution to disease susceptibility of “low frequency-low risk” variants

E. Morini<sup>1</sup>, S. Prudente<sup>1</sup>, E. Succurro<sup>2</sup>, M. Chandalia<sup>3</sup>, Y.-Y. Zhang<sup>4,5</sup>, S. Mammarella<sup>6,7</sup>, F. Pellegrini<sup>8,9</sup>, C. Powers<sup>4</sup>, V. Proto<sup>1</sup>, B. Dallapiccola<sup>1</sup>, A. Cama<sup>6,7</sup>, G. Sesti<sup>2</sup>, N. Abate<sup>3</sup>, A. Doria<sup>4,5</sup>, and V. Trischitta<sup>1,10,11</sup>

<sup>1</sup> IRCCS Casa Sollievo della Sofferenza – Mendel Institute, Rome, Italy

<sup>2</sup> Department of Experimental and Clinical Medicine, University Magna Gracia, Catanzaro, Italy

<sup>3</sup> Division of Endocrinology and Metabolism University of Texas Medical Branch, Galveston, TX, USA

<sup>4</sup> Research Division, Joslin Diabetes Center, Boston, MA, USA

<sup>5</sup> Department of Medicine, Harvard Medical School, Boston, MA, USA

<sup>6</sup> Department of Oncology and Neurosciences, University G. D’Annunzio, Chieti, Italy

<sup>7</sup> Center of Excellence on Aging, G. D’Annunzio University Foundation, Chieti, Italy

<sup>8</sup> Unit of Biostatistics, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

<sup>9</sup> Department of Clinical Pharmacology and Epidemiology, Consorzio Mario Negri Sud, S. Maria Imbaro, Italy

<sup>10</sup> Department of Medical Pathophysiology, “Sapienza” University, Rome, Italy

<sup>11</sup> Research Unit of Diabetes and Endocrine Diseases, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

### Abstract

**Aims and hypothesis**—Results on the association between the *IRS-1* G972R polymorphism and type 2 diabetes have been conflicting. To obtain further insights onto this topic, we performed a meta-analysis of all available case-control studies.

**Methods**—Meta-analysis of 32 studies (12,076 cases and 11,285 controls).

**Results**—The relatively infrequent R972 variant was not significantly associated with type 2 diabetes, (odds ratio [OR] =1.09, 95% confidence interval [CI] 0.96–1.23, p=0.184 under a dominant model). Some evidence of heterogeneity was observed across studies (p=0.1). In the 14 studies (9,713 individuals) in which the mean age at type 2 diabetes diagnosis was available, this variable explained 52% of heterogeneity (p=0.03). When these studies were subdivided into tertiles of mean age at diagnosis, the OR for diabetes was 1.48 (95% CI 1.17–1.87), 1.22 (95% CI 0.97–1.53), and 0.88

---

Correspondence should be addressed to: Vincenzo Trischitta, MD, CSS-Mendel Institute, Viale Regina Margherita 261, Postal code 00198, Rome-Italy, Tel: +39-06-44160531; Fax: +39-06-44160548, vincenzo.trischitta@uniroma1.it.

#### Duality of interest

The authors declare that there is no duality of interest associated with this manuscript.

(95% CI 0.68–1.13) in the youngest, intermediate and oldest tertile, respectively ( $p$  for trend of ORs=0.0022).

**Conclusion**—Our findings illustrate the difficulties of ascertaining the contribution of “low frequency-low risk” variants to type 2 diabetes susceptibility. In the specific context of the R972 variant, ~200,000 study subjects would be needed to have 80% power to identify a 9% increase in diabetes risk at genome-wide significance level. Under these circumstances, a strategy aimed at improving outcome definition and decreasing its heterogeneity may critically enhance our ability to detect genetic effects, thereby decreasing the required sample size. Our data suggest that focusing on early-onset diabetes, which is characterized by a stronger genetic background, may be part of such strategy.

## Keywords

Age at type 2 diabetes diagnosis; Genetics of type 2 diabetes; Insulin signalling; Insulin secretion; Meta-analysis

## Introduction

Despite the recent advancements resulting from genome-wide association studies (GWAS), most of the genetic factors contributing to type 2 diabetes remain undetermined [1]. Insulin receptor substrate-1 (IRS-1) is an important member of a protein family phosphorylated by the insulin receptor upon its binding with insulin [2]. Tissue-specific knockout mice have shown that IRS-1 is necessary for *in vivo* insulin action and secretion [2]. A relatively infrequent glycine to arginine substitution at position 972 of *IRS-1* (G972R or rs1801278, minor allele frequency [MAF] ranging 0.02–0.10 in the 4 different populations samples available from HapMap) has been extensively investigated as a determinant of type 2 diabetes susceptibility. *In vitro* studies have shown that the R972 allele results into a loss of IRS-1 function that impairs insulin signaling in several target tissues, including skeletal muscle, fat, and pancreatic  $\beta$ -cells [2–4]. *In vivo* studies have reported an association between *IRS-1* R972 variant and both insulin resistance [2,5] and reduced insulin secretion [2,6]. The deleterious role of the R972 variant on *in vivo* insulin action and glucose homeostasis has been recently confirmed by studies in transgenic mice [7]. In spite of such strong evidence for a functional role, the data concerning the association of this variant with type 2 diabetes have been, thus far, conflicting. An initial meta-analysis of 27 studies indicated that R972 carriers had a 25% increase in type 2 diabetes risk [8], but subsequent large case-control studies have failed to replicate this association (see Zeggini et al. [21]; Florez et al. [22]; van Dam et al. [23] in Table 1 of the Electronic supplementary material [ESM]). Unfortunately, neither the G972R variant nor good proxies in linkage disequilibrium with it (i.e.  $r^2 > 0.5$ ) were included in the publicly available GWAS meta-analysis DIAGRAM [9].

To obtain further insights about the role of R972 on type 2 diabetes, we performed an updated meta-analysis of all case-control studies available to date (see ESM Table 1). BMI and age at diabetes onset were analyzed as covariates in meta-regression.

## Methods

### Study design

All case-control studies reported in previous meta-analyses [8] and all papers found in the PubMed database as of January 2009 by using “Insulin receptor substrate-1”, “IRS-1”, “Gly972Arg”, “G972R”, “diabetes”, “variant”, “polymorphism” and “genotype” as keywords, were analyzed. In addition, we included five unpublished case-control studies in which all study subjects were self-reported Whites: four sets from the Genetics of type 2 diabetes in

Italy and the United States (GENIUS T2D) Consortium [10] and one set recruited in Chieti, Italy (ESM Table 1; Cama A. sample). Three of the published studies were excluded because they were subsets of these unpublished sets: Sigal et al. [9] of the GENIUS Boston sample, Mammarella et al. [8] and Esposito et al. [20] of the Cama A. sample (ESM Table 1).

### Study subjects in unpublished samples

Controls in all unpublished samples were non-diabetic subjects with fasting plasma glucose lower than 6.1 mmol/l and absence of drug treatment known to affect glucose metabolism. Cases were patients with type 2 diabetes defined according to the 2003 American Diabetes Association criteria.

### DNA extraction and genotyping

DNA from the unpublished sets was extracted from whole blood by standard methods. Genotyping details are described in Methods section of ESM.

### Statistical methods

Cases and controls of all studies were tested for Hardy-Weinberg equilibrium (HWE) by means of exact chi-square test. Between-study heterogeneity and the possible presence of publication bias were assessed by the Cochran's Q test and Macaskill's inverse pooled variance weighting method [11], respectively. Random-effects meta-analysis and meta-regression were used to estimate overall odds ratio (OR) and to explore heterogeneity [12]. Where appropriate, permutation resampling p values were calculated to address the risk of spurious significant results [13]. All the analyses were performed using SAS Statistical Package Release 9.1 (SAS Institute, Cary, NC, USA).

## Results

Of the 35 available studies, only those 32 that did not show significant deviations (exact  $p < 0.05$ ) from HWE in cases or controls were considered in the meta-analysis (ESM, Table 1). Given the small number of RR individuals (i.e. homozygous for the R972 variant) in 12 studies and their absence in the other 20 (a finding that could seriously bias the results of both additive and recessive models), we investigated only the dominant model, by comparing GR+RR (these latter when available) to GG individuals. Fig. 1 shows the individual results from the 32 case-control studies, along with those of the meta-analysis, which included 12,076 cases and 11,285 controls. As for any meta-analysis performed on published genetic data, we cannot exclude that some sample overlap have occurred; however, by carefully reading the description of samples analyzed in each study, this seems to be an unlikely event. No evidence of publication bias was observed ( $p = 0.27$ ). The OR for association between R972 and type 2 diabetes ranged from 0.55 to 4.75. In the meta-analysis, the R972 variant didn't show a significant association with type 2 diabetes, (OR 1.09, 95% confidence interval [95% CI] 0.96–1.23  $p = 0.184$ ). Some evidence of heterogeneity was observed across studies (Cochran's Q test  $p = 0.1$ ). In a meta-regression analysis, neither the mean BMI of cases nor that of controls (available in 23 studies corresponding to 20,114 individuals) significantly explained such heterogeneity ( $p = 0.58$  and  $p = 0.84$ , respectively). Similar data were obtained when analyses were carried after stratifying for BMI status (i.e.  $< \text{or} \geq 30 \text{ kg/m}^2$ ) ( $p = 0.77$ ). Also no effect of ethnicity (i.e. either White – 19,075 individuals from 20 studies –, or Asian – 2,699 individuals from 8 studies – or other – 1,587 individuals from 4 studies) was observed ( $p = 0.91$ ). Also, when only studies whose sample size was  $> 500$  individuals were analyzed, a similar OR to that obtained in the whole meta-analysis was observed (OR 1.08, 95% CI 0.93–1.24). By contrast, the mean age at type 2 diabetes diagnosis (available in 14 studies corresponding to 9,713 individuals) was significantly correlated with the magnitude of the genetic effect, explaining 52% of the heterogeneity ( $p = 0.03$ ) (Fig. 2a). When these studies were subdivided into tertiles of mean age

at diagnosis, the summary OR of type 2 diabetes was 1.48 (95% CI 1.17–1.87) for studies in the youngest tertile (39–44.9 yrs), 1.22 (95% CI 0.97–1.53) for studies in the intermediate tertile (45–50.9 yrs), and 0.88 (95% CI 0.68–1.13) for studies in the oldest tertile (51–58 yrs) (Fig. 2b). The standard p for the decreasing trend of ORs with increasing mean age at diagnosis was 0.0022, the permutation p was 0.014.

## Discussion

Our findings illustrate the difficulties of ascertaining contributions to type 2 diabetes susceptibility by “low frequency-low risk” variants. Despite the fact that this study included more than 23,000 individuals, the power to identify a 9% increase in type 2 diabetes risk associated with a variant having 0.06 frequency was only 58% at nominal significance levels ( $\alpha=0.05$ ) and virtually zero at genome-wide significance levels ( $\alpha=5\times 10^{-8}$ ). One can estimate that a total of ~40,000 and ~200,000 individuals would have been required to have 80% power at  $\alpha=0.05$  and  $\alpha=5\times 10^{-8}$ , respectively. Under these circumstances, improving the outcome definition and decreasing its heterogeneity may have critical effects on our ability to identify genetic effects.

In our meta-analysis, studies in which the mean age at type 2 diabetes diagnosis was less than 45 yrs showed an OR for type 2 diabetes of 1.48 – an effect size that a sample of “only” ~8,500 individuals would have 80% power to detect with genome-wide significance. Similar data, indicating a stronger effect on early abnormality of glucose homeostasis, were recently reported for *TCF7L2* [14] and for *TRIB3* [10]. Unfortunately, no data on the combined effect of several single nucleotide polymorphisms (SNPs) which are singly associated with early glucose abnormalities are so far available. Overall, focusing on forms of diabetes diagnosed relatively early in life, which are known to have a stronger genetic component [15,16], may be a useful strategy to facilitate the identification of SNPs associated with type 2 diabetes that are otherwise difficult to find, either because of their moderate effect or because of their low allele frequency, or because of both factors, as in the case of *IRS-1* G972R. The usefulness of this approach may also extend to truly rare variants (MAF<0.01), such as those that are believed to underlie the linkage peaks that are not explained by the common variants identified through GWAS. Indeed, in the linkage screen of the Diabetes UK Warren 2 sib pair collection, all seven linkage signals that were identified were stronger in families with an average age at diagnosis younger than 55 than in the families diagnosed at an older age [17].

In conclusion, studying early-onset forms is emerging as a critical tool to reach the “high-hanging” fruits of type 2 diabetes genetics, similar to what has been the approach with other complex disorders such as coronary artery disease [18]. Thus both, adequately powered new studies specifically targeted to early-onset cases and, further analyses of available GWAS data after stratification by age at onset, are needed.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was partly supported by the Italian Ministry of Health (Ricerca Corrente 2007 to S.P and 2009 to S.P and V.T) and by NIH (grants HL073168, DK055523, and DK036836 to A.D. and the Genetics Core of the Diabetes and Endocrinology Research Center at the Joslin Diabetes Center).

## Abbreviations

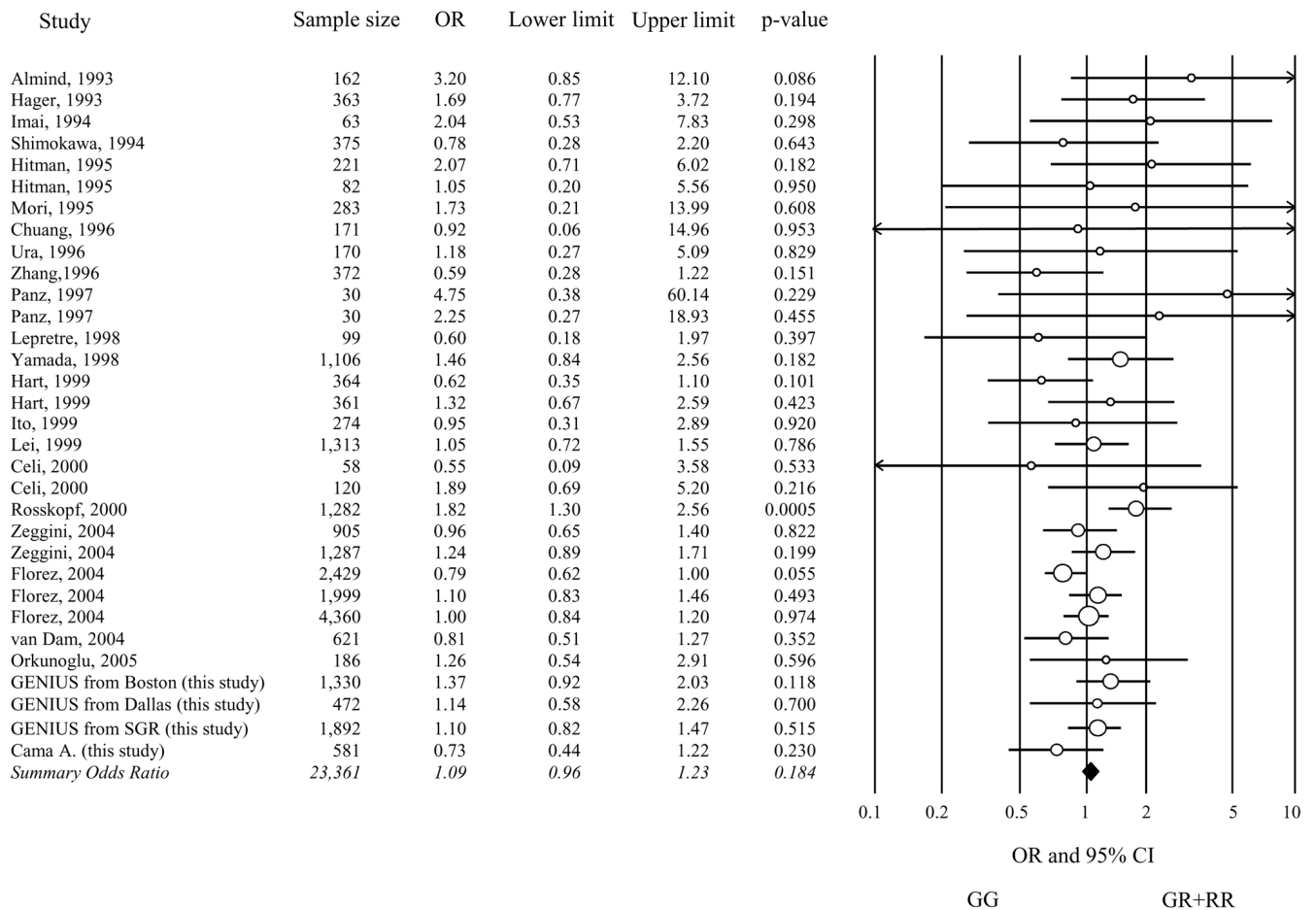
CI confidence interval

<b>ESM</b>	Electronic supplementary material
<b>GWAS</b>	genome-wide association studies
<b>HWE</b>	Hardy-Weinberg equilibrium
<b>IRS-1</b>	insulin receptor substrate-1
<b>MAF</b>	minor allele frequency
<b>OR</b>	odds ratio
<b>SNPs</b>	single nucleotide polymorphisms

## References

1. Doria A, Patti ME, Kahn CR. The emerging genetic architecture of type 2 diabetes. *Cell Metab* 2008;8:186–200. [PubMed: 18762020]
2. Sesti G, Federici M, Hribal ML, Lauro D, Sbraccia P, Lauro R. Defects of the insulin receptor substrate (IRS) system in human metabolic disorders. *FASEB J* 2001;15:2099–2111. [PubMed: 11641236]
3. Almind K, Inoue G, Pedersen O, Kahn CR. A common amino acid polymorphism in insulin receptor substrate-1 causes impaired insulin signaling. Evidence from transfection studies. *J Clin Invest* 1996;97:2569–2575. [PubMed: 8647950]
4. Marchetti P, Lupi R, Federici M, et al. Insulin Secretory Function Is Impaired in Isolated Human Islets Carrying the Gly972Arg IRS-1 Polymorphism. *Diabetes* 2002;51:1419–1424. [PubMed: 11978638]
5. Clausen JO, Hansen T, Bjorbaek C, et al. Insulin resistance: interactions between obesity and a common variant of insulin receptor substrate-1. *Lancet* 1995;346:397–402. [PubMed: 7623569]
6. Stumvoll M, Fritsche A, Volk A, et al. The Gly972Arg polymorphism in the insulin receptor substrate-1 gene contributes to the variation in insulin secretion in normal glucose-tolerant humans. *Diabetes* 2001;50:882–885. [PubMed: 11289056]
7. Hribal ML, Tornei F, Pujol A, et al. Transgenic mice overexpressing human G972R IRS-1 show impaired insulin action and insulin secretion. *J Cell Mol Med* 2008;12:2096–2106. [PubMed: 18208559]
8. Jellema A, Zeegers MP, Feskens EJ, Dagnelie PC, Mensink RP. Gly972Arg variant in the insulin receptor substrate-1 gene and association with Type 2 diabetes: a meta-analysis of 27 studies. *Diabetologia* 2003;46:990–995. [PubMed: 12819898]
9. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;40:638–645. [PubMed: 18372903]
10. Prudente S, Scarpelli D, Chandalia M, et al. The TRIB3 Q84R Polymorphism and Risk of Early-Onset Type 2 Diabetes. *J Clin Endocrinol Metab* 2009;94:190. [PubMed: 18984671]
11. Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med* 2001;20:641–654. [PubMed: 11223905]
12. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;21:589–624. [PubMed: 11836738]
13. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004;23:1663–1682. [PubMed: 15160401]
14. Korner A, Berndt J, Stumvoll M, Kiess W, Kovacs P. TCF7L2 gene polymorphisms confer an increased risk for early impairment of glucose metabolism and increased height in obese children. *J Clin Endocrinol Metab* 2007;92:1956–1960. [PubMed: 17311858]
15. Weijnen CF, Rich SS, Meigs JB, Krolewski AS, Warram JH. Risk of diabetes in siblings of index cases with Type 2 diabetes: implications for genetic studies. *Diabet Med* 2002;19:41–50. [PubMed: 11869302]
16. Mitchell BD, Kammerer CM, Reinhart LJ, Stern MP. NIDDM in Mexican-American families. Heterogeneity by age of onset. *Diabetes Care* 1994;17:567–573. [PubMed: 8082526]

17. Frayling TM, Wiltshire S, Hitman GA, et al. Young-onset type 2 diabetes families are the major contributors to genetic loci in the Diabetes UK Warren 2 genome scan and identify putative novel loci on chromosomes 8q21, 21q22, and 22q11. *Diabetes* 2003;52:1857–1863. [PubMed: 12829657]
18. Kathiresan S, Voight BF, et al. Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* 2009;41:334–341. [PubMed: 19198609]



**Fig. 1. Meta-analysis of 30 case-control studies**

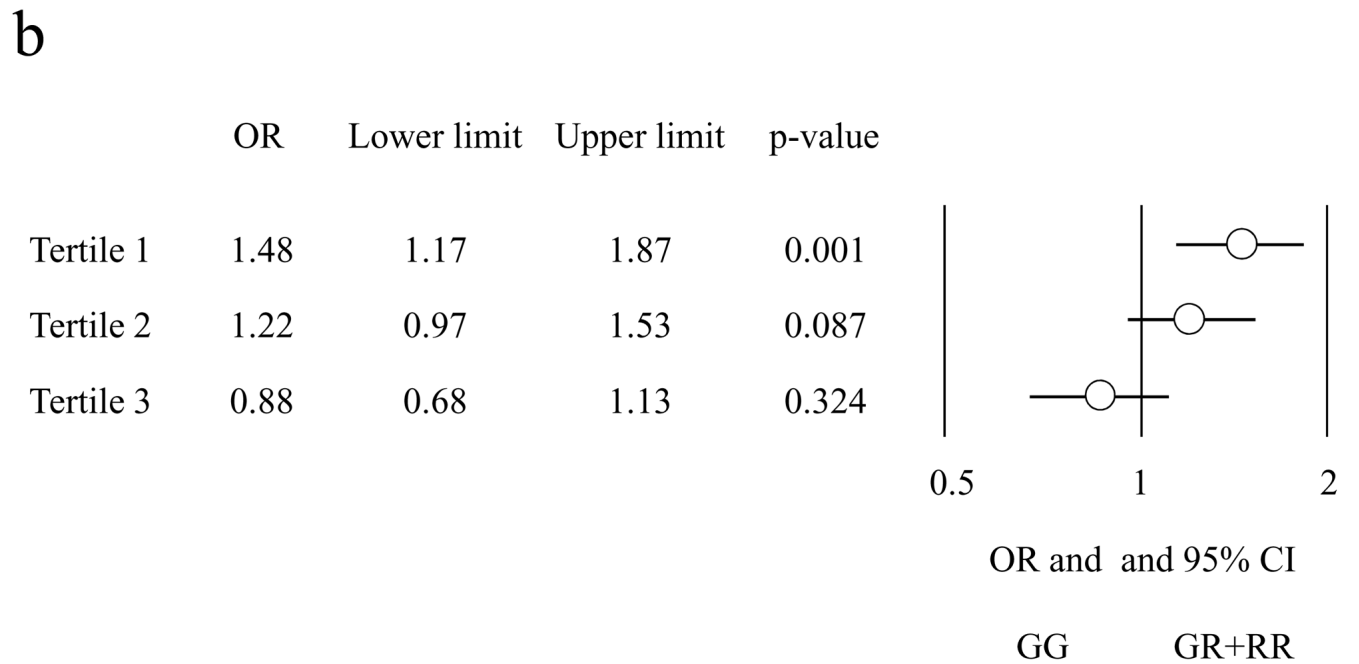
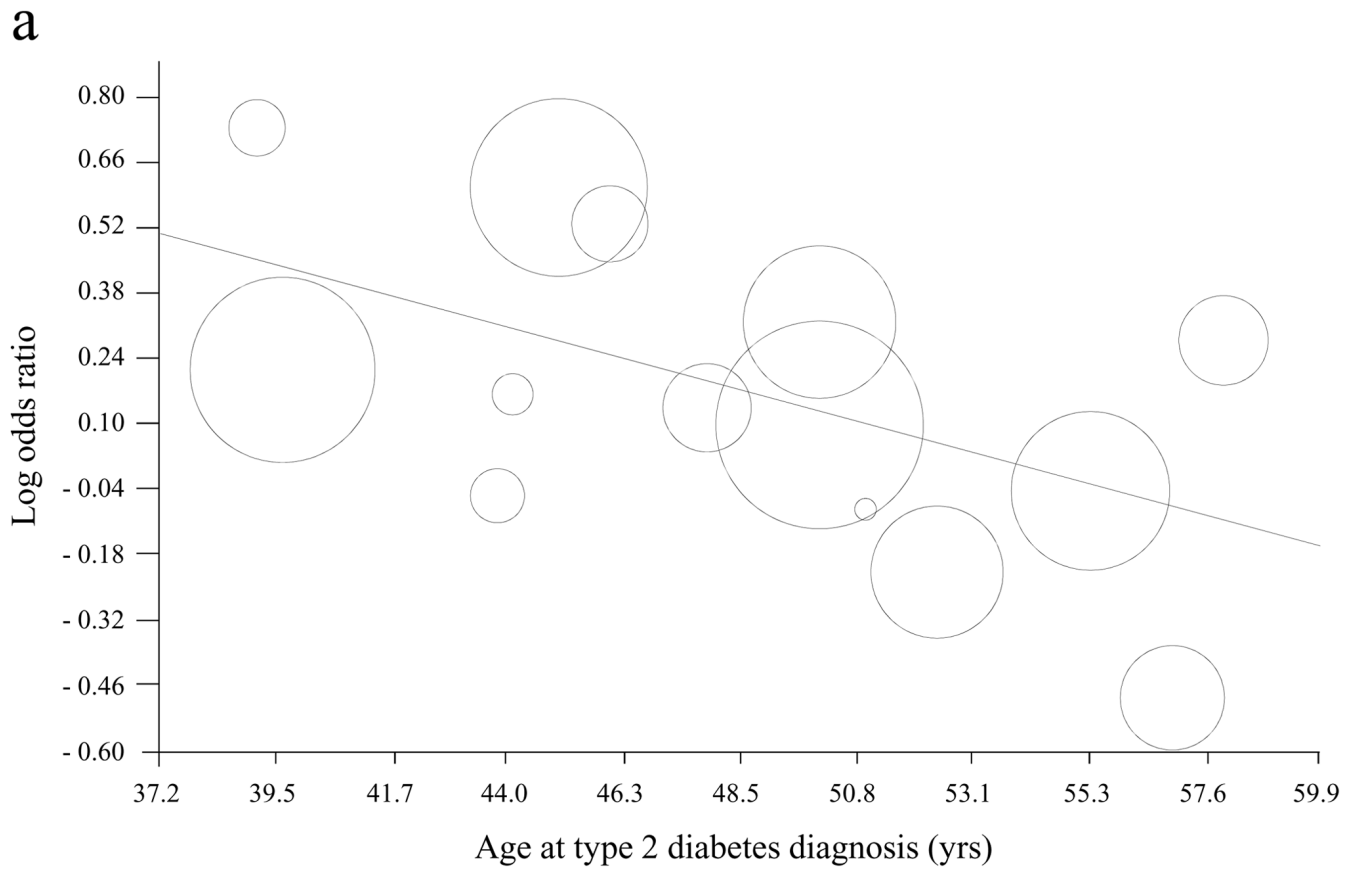
The cumulative effect of 32 published (ordered by publication date) and unpublished studies on the association between the *IRS-1* G972R polymorphism and type 2 diabetes was tested by random-effects model. A borderline significant heterogeneity was observed across studies (Cochran's Q test  $p=0.1$ ). Odds ratios and 95% confidence intervals for dominant genetic model are shown.

OR, odds ratios

CI, confidence intervals

SGR, San Giovanni Rotondo, Italy.





**Fig. 2. Relationship between odds ratio of type 2 diabetes and age at type 2 diabetes diagnosis in the 14 studies for which this information was available (n=9,713 individuals)**



**a)** Meta-regression of mean age at diagnosis of type 2 diabetes and Log odds ratio for type 2 diabetes of the R972 variant according to a dominant genetic model. There was a significant correlation ( $P=0.03$ ) explaining 52% of between-study heterogeneity.

**b)** Summary odds ratios (OR) of type 2 diabetes according to tertiles of age at type 2 diabetes diagnosis. The ranges of age at type 2 diabetes diagnosis were 39–44.9 yrs (5 studies,  $n=3,234$  individuals), 45–50.9 (5 studies,  $n=4,228$  individuals) and 51–58 (4 studies,  $n=2251$  individuals) in tertile 1, 2 and 3, respectively.

OR, odds ratios

CI, confidence intervals