Identification of a novel proinsulin-associated SNP and demonstration that proinsulin is unlikely to be a causal factor in subclinical vascular remodelling using Mendelian randomisation

Rona J. Strawbridge a, b, *, 2, Angela Silveira a, Marcel den Hoed c, Stefan Gustafsson d, Jian’an Luan e, Denis Rybin f, Josépe Dupuis g, h, Ruifang Li-Gao i, Maryam Kavousi i, Abbas Dehghan j, k, Kadri Halja l, Jari Lahti m, Jesper R. Gådin a, Alexandra Bäcklund a, Ulf de Faire n, o, Karl Gertow p, Phillipe Giral p, Anuj Goel q, r, Steve E. Humphries s, Sudhir Kurl t, Claudia Langenberg e, Lars L. Lannfelt u, Lars Lind c, v, Cecilia C.M. Lindgren q, Elmo Mannarino w, Dennis O. Mook-Kanamori i, x, Andrew P. Morris q, y, Renée de Mutsert i, Rainer Rauramaa z, aa, Peter Saliba-Gustafsson a, Bengt Sennblad a, ab, Andries J. Smit ac, ad, Ann-Christine Sylvén a, Elena Tremoli af, ag, Fabrizio Veglia a, Björn Zethelius ah, Hanna M. Björck a, Johan G. Eriksson al, aj, ak, Albert Hofman j, al, Oscar H. Franco j, Hugh Watkins q, r, J. Wouter Jukema am, an, Jose C. Florez ao, ap, aq, Nicholas J. Wareham a, James B. Meigs ap, aq, ar, Erik Ingelsson d, as, Damiano Baldassarre af, at, Anders Hamsten a, on behalf of the IMPROVE study group1

a Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
b Mental Health and Wellbeing, Institute of Mental Health and Wellbeing, University of Glasgow, Glasgow, UK
c Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
d Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
e MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK
f Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
g Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
h Department of Cardioiology, Karolinska University Hospital, Stockholm, Sweden
i Centre for Molecular Medicine, Karolinska Universitetssjukhuset Solna, Stockholm, 171 76, Sweden.
* Corresponding author. L8:03, Centre for Molecular Medicine, Karolinska Universitetssjukhuset Solna, Stockholm, 171 76, Sweden.

1 Additional members of the IMPROVE study group are listed in the Supplementary Data.

2 Present address: Department of Health and Wellbeing, University of Glasgow, 1 Lylebank Gardens, G12 8RZ.
**1. Introduction**

Hyperinsulinemia, most often a result of systemic insulin resistance, is associated with increased risk of cardiovascular disease (CVD) and severity of coronary artery disease (CAD) [1]. Under normo-glycemic conditions, proinsulin is fully processed to insulin and C-peptide before secretion, with very low concentrations of proinsulin being detectable in the blood. During hyperinsulinemia, increased demand for insulin production results in pancreatic β-cells failing to correctly process proinsulin to mature insulin [2]. Thus, circulating levels of proinsulin are increased (relative to insulin levels) and this might be a link between insulin resistance and other stages of CVD development, such as those reflected by cIMT.

However, it remains to be determined whether increased proinsulin levels merely reflect CVD risk-related processes, or whether proinsulin has a direct adverse functional effect on the vasculature.

Genetic predisposition to increased proinsulin levels [10] previously demonstrated no effect on CVD events, however we proposed that increased proinsulin levels could be important in earlier stages of CVD development, such as those reflected by cIMT.

The aims of this project were firstly to explore the existence of additional proinsulin-associated variants using a large panel of candidate loci for metabolic or inflammatory processes and secondly to use MR to assess causality of proinsulin on early-stage atherosclerosis, represented by cIMT. The IMPROVE cohort is uniquely positioned to address the second aim, having extensive genotyping, circulating proinsulin and insulin levels and detailed cIMT measurements in 2441 participants without type 2 diabetes (T2D). Demonstration that proinsulin has a causal role in cIMT would highlight insulin resistance as a target for CVD prevention. As insulin sensitisation by dietary, lifestyle and pharmaceutical interventions is possible, the role of proinsulin in CVD warrants further investigation.
2. Materials and methods

2.1. Cohort description: IMPROVE

The IMPROVE study recruitment and ultrasound protocol have been described in detail [11,12]. In short, IMPROVE consists of participants without symptoms or history of CVD at enrolment, who presented with at least three classical risk factors for CVD (including family history of CVD, diabetes or impaired fasting glucose, hyperalphalipoproteinemia, hypertension, hypertriglyceridemia, hypercholesterolemia, current smoking [11,12]). Between March 2004 and April 2005, 3711 participants were recruited from seven centres in five European countries (Finland, Sweden, the Netherlands, France and Italy). Extensive ultrasound examinations of cIMT were carried out at baseline and after 15 and 30 months, with linear regression of cIMT over time being used to calculate cIMT progression variables [11,12].

2.2. Genotyping in IMPROVE

IMPROVE has been genotyped using both the Illumina CardioMetabo 200k [13] and Immunochip 200k [14] bead array platforms. Single nucleotide polymorphisms (SNPs) were excluded for low call rate (<95%), low minor allele frequency (<0.01) and for deviation from Hardy-Weinberg equilibrium (p < 1*10^-5). Participants were excluded due to low call rate (<95%), aberrant sex assignment or cryptic relatedness (IBD >0.2). Multi-dimensional scaling (MDS) components 1–3 were calculated in PLINK [15] to enable adjustment for population structure. After quality control 2441 participants without T2D and 254,756 SNPs with minor allele frequency >1% were included in the analysis.

2.3. Identification of proinsulin-associated SNPs in IMPROVE

To identify proinsulin-associated variants [10], 254,756 SNPs were analysed for influence on proinsulin levels. This also enabled assessment of the reported [10] proinsulin-associated loci. The reported lead [10], the analysed proxy or the strongest SNP for each locus are presented in Supplementary Table 2.

Levels of proinsulin, insulin and glucose were natural log normalised prior to analysis. Linear regression, assuming an additive genetic model, was used to analyse the effect of SNPs on proinsulin levels. To enable comparison of SNP effects with the proinsulin genome-wide association study (GWAS) [10], the same three models explored in that paper were tested: insulin model (age, sex, population structure and insulin), glucose model (age, sex, population structure, insulin and glucose) and body-mass index (BMI) model (age, sex, population structure, insulin and BMI). Loci with p < 1*10^-5 (suggestive evidence of association) in any model were selected for replication. Of note, IMPROVE was not part of the prior proinsulin GWAS [10].

2.4. Replication of proinsulin-associated SNPs

Replication of proinsulin-associated loci was attempted in seven independent cohorts (Fenland (http://www.mrc-epid.cam.ac.uk/Research/Studies/Fenland), HBS [16–18], HBCS [19], PIVUS [20,21], ULSAM [22,23], PROCARDIS cases [24] and PROCARDIS controls [24], n = 18,773–20,003 participants). Replication cohorts are described in the Online Data Supplement and cohort characteristics are presented in Supplementary Table 3. The same analyses and models as the discovery analysis were applied. To improve power and examine consistency of effects, replication results were combined in an inverse variance meta-analysis, using GWAMA [25].

2.5. Expression quantitative trait locus analysis

To further characterize the proinsulin-associated locus on chromosome 15 (rs8029765), genotype-specific expression patterns were assessed. This was conducted in liver biopsies from participants undergoing elective surgery for aortic valve and/or ascending aortic disease in ASAP [26] (Supplementary Data) and biopsies from liver or heart and peripheral blood mononuclear cells from organ donation or autopsy specimens in the GTEx [27] database (Supplementary Data).

2.6. In silico analysis of effect on T2D

Proinsulin-increasing alleles of previously identified SNPs rather surprisingly demonstrated a mixture of positive, negative and null effects on T2D risk [10]. Thus to assess the impact of the novel proinsulin-associated locus on chromosome 15 (rs8029765), genotype-specific expression patterns were assessed. This was conducted in liver biopsies from participants undergoing elective surgery for aortic valve and/or ascending aortic disease in ASAP [26] (Supplementary Data) and biopsies from liver or heart and peripheral blood mononuclear cells from organ donation or autopsy specimens in the GTEx [27] database (Supplementary Data).

2.7. Effects of proinsulin on cIMT measures in IMPROVE

All phenotypes were assessed for normality and natural logarithmic transformation was applied where necessary, prior to further statistical analysis. The analysis of proinsulin levels for effect on cIMT measures were adjusted in basic (age, sex and population structure) and extended models (age, sex, population structure, BMI, systolic blood pressure (SBP), high density lipoprotein (HDL), triglycerides (TGs) and current smoking) [28]. As the mean cIMT measure encompasses the maximum cIMT measure, these variables were not considered independent and therefore related, multiple testing was not corrected for and p < 0.05 was considered significant. Secondary analyses explored the consistency of effects of proinsulin on other cIMT segments and on analysis of progression measures, where the baseline measure of the same segment was also included as a covariate. Because the secondary analyses were exploratory only, the size and direction of effects, rather than significance, is the focus of these analyses. Analyses were performed in STATA (STATAcorp LP, College Station, Texas, USA).
2.8. Mendelian randomization (MR) to assess causality of proinsulin on cIMT in IMPROVE

In order to be valid instruments for the MR analyses, the proinsulin-associated loci should demonstrate the expected associations with proinsulin levels. Accordingly, the effect size and direction were examined for all proinsulin-associated loci [10]. Previously published results were considered the gold standard (given the much larger sample size in the GWAS (n=20,000) compared with this study) and were thus referred to in further analyses. Only subjects with complete genotyping for the 11 proinsulin-associated SNPs (10 SNPs in 9 reported loci and the chromosome 15 SNP/locus identified here) were included in this analysis (N = 35 excluded).

As individual SNPs had only small effects on proinsulin levels, we calculated proinsulin-associated SNP scores by summing the number of alleles associated with higher proinsulin levels. SNP scores were calculated unweighted (sum of number of proinsulin-increasing alleles) and weighted (sum of the number of alleles multiplied by the effect size of the allele reported in the GWAS [10]).

We calculated proinsulin-associated SNP scores by summing the effect size of the allele reported in the GWAS for each SNP. The majority of proinsulin-associated SNPs have also been associated with other metabolic traits of relevance to CVD. This pleiotropy may introduce confounding or explain some of the proinsulin associations with cIMT. To determine whether this was the case, a further SNP score was constructed, including only the SNPs which had no reported metabolic effects (GWAS catalogue, https://www.ebi.ac.uk/gwas/, Supplementary Data). Analyses were adjusted using the basic model (age, sex and population structure). Instrument variable regression, with the proinsulin SNP scores as the instrument, was used to assess causality of proinsulin levels with future cIMT measures. The relevance of the chromosome 15 locus was assessed by replication in R using the rma function in the METAFOR package (https://cran.r-project.org/web/packages/metafor/index.html).

2.9. Replication of MR

Replication of the MR experiment was attempted in 12,113 participants from 5 cohorts (2 of which contributed to the SNP-proinsulin replication analysis) with IMTmean and IMTmax phenotypes. The replication cohorts are described in the Online Data Supplement and characteristics are presented in Supplementary Table 4. The cIMT measures were broadly comparable with IMPROVE, with multiple measurements of the carotid artery enabling assessment of both the maximum and mean values, both left and right carotid arteries being measured (except in ULSAM), and the 1st cm proximal to the bifurcation being omitted from common carotid artery measurements (except in PIVUS). Meta-analysis of SNP score associations with IMT measures was conducted in R using the rma function in the METAFOR package (https://cran.r-project.org/web/packages/metafor/index.html).

3. Results

3.1. Identification of proinsulin-associated SNPs

Analysis of the combined chip (adjusting for population structure, age, sex, insulin as well as BMI or glucose) identified a total of 16 SNPs in 3 loci reaching the suggestive level of significance for association with proinsulin (p < 1*10^-5) (Fig. 1 and Supplementary Tables 5 and 6). Consistent with the published proinsulin GWAS [10], more proinsulin-associated SNPs were identified using the glucose-adjusted model. Most (n = 13) of these SNPs were located in the ARAPI locus. This locus was identified previously [10], so this finding was taken to be a positive control and gave us confidence in the other loci identified. Conditional analysis for the two known signals (rs77756118 and rs11603334) in the ARAPI locus rendered these SNPs non-significant (Fig. 2). On chromosome 6, rs2323000 and rs6910151 represent the same signal. On chromosome 15 (rs8029765), the signal is located 29 Mb from the LARP6 and 20 Mb from the VPS13C loci. Conditioning on LARP6 (strongest SNP) and VPS13C (reported lead SNP) signals had negligible effects on chromosome 15 signal (Fig. 2 and Supplementary Table 7), which thus appeared to be a distinct signal. The reason that this SNP was not identified in the previous GWAS meta-analysis was most likely its relatively large effect size in IMPROVE (Beta = 0.09), compared with that in the GWAS discovery analysis (Beta = 0.01) [10]. A stronger effect size in the IMPROVE study compared with the GWAS is plausible: We believe proinsulin levels are important in CVD, and the IMPROVE cohort is enriched for CVD risk factors. In contrast, the GWAS included a much wider spectrum of CVD burden. Hence two proinsulin-associated SNPs, rs2323000 and rs8029765, were taken forward to replication.

3.2. Replication of proinsulin-associated SNPs

Replication results of the two proinsulin-associated SNPs from seven independent cohorts were meta-analysed (n = 18,773–20,004). The meta-analysis refuted the chromosome 6 locus association with proinsulin levels (Table 1). However, nominal significance was observed for the minor allele of rs8029765 (chromosome 15) being associated with higher proinsulin levels in most models (insulin model: Beta ± SE 0.013 ± 0.006, p = 0.04 (combined discovery and replication 0.020 ± 0.006, p = 0.001); BMI model: 0.014 ± 0.006, p = 0.03 (combined discovery and replication 0.020 ± 0.006, p = 0.002)). Additionally adjusting for glucose did not materially alter the effect size (0.012 ± 0.006, p = 0.0709 (combined discovery and replication 0.020 ± 0.006, p = 0.003)). The replication meta-analyses indicated that the effect sizes in women were not significantly different from those in men (p = 0.0659), with no evidence of heterogeneity between the strata (I² = 0, sex heterogeneity p = 0.80). These data provide evidence for a novel proinsulin-associated locus on chromosome 15, which did not influence risk of T2D (OR 1.00, confidence interval (CI) 0.95–1.06, p = 0.89, n = 51,411 [29]), rs8029765 (and the SNP in high LD, rs4450375, R² > 0.8, Supplementary Fig. 1) has not previously been reported to be associated with any of the metabolic or cardiovascular traits we studied.

3.3. Expression QTL for chromosome 15 locus

The relevance of the chromosome 15 locus was assessed by examining nearby genes for genotype-specific expression. In the GTEx heart (n = 190), liver (n = 97) or PBMC (n = 338) data, generated from relatively healthy individuals, no significant eQTLs were identified. However, in the ASAP cohort, which is significantly burdened with cardiovascular disease, levels of expression were detectable for eight genes in the chromosome 15 locus: MAN2A2, HDEC3, UNC45A, RCCD1, PRC1, VPS33B, SV2B, SLC3A1. In liver samples (n = 209), rs8029765 was significantly associated with expression of UNC45A (Supplementary Fig. 2, p = 5.6*10^-3) with the proinsulin-increasing allele demonstrating lower expression of UNC45A. The unc-45 myosin chaperone A encoded by the UNC45A gene is a widely expressed chaperone protein [30].
3.4. Effects of proinsulin on IMTmean and IMTmax

As anticipated, proinsulin levels demonstrated correlations with almost all established risk variables for CVD in the IMPROVE cohort (Supplementary Table 8). For further analysis of IMT measures, models were adjusted for relevant confounders [28]. No statistically significant associations were observed between proinsulin levels and the primary cIMT phenotypes IMTmean and IMTmax. Nominal associations (p < 0.05) were observed with the IMTmax progression, but not baseline measures, independently of established CVD risk factors (Table 2). Secondary analyses indicated that the common carotid (excluding the 1st cm proximal to the bifurcation) demonstrated (non-significant) associations between higher proinsulin levels and larger cIMT measures (Supplementary Table 9). All other segments demonstrated inverse associations.

3.5. Effect of proinsulin-increasing SNP scores on IMTmean and IMTmax

Reported proinsulin-associated SNPs demonstrated effects sizes generally comparable with those previously reported [10] (Supplementary Tables 2 and 10), thus were appropriate for use in the MR analysis. A total of 11 proinsulin-associated SNPs were included in the SNP score used for the MR analysis (10 previously reported SNPs in 9 loci and the chromosome 15 SNP, rs8029765). Instrument variable analysis demonstrated that the 11-SNP score was significantly and inversely associated with baseline IMTmean and IMTmax but not with progression measures (Table 3). However, MR-Eggers (which is more robust when pleiotropy is present [31]) refuted a significant association between the 11-SNP score and IMTmean (−0.020 ± 0.033 p = 0.5721) or IMTmax (−0.060 ± 0.054 p = 0.2738). As pleiotropic SNPs can bias estimates [31], a further score was constructed, using only non-pleiotropic SNPs (SNX7 rs1571500, LARP6 rs7163439 and UNC45A rs8029765). This SNP score was not associated with baseline or progression of cIMT measures (Table 3). Replication of the non-pleiotropic SNP score was attempted in 12,113 participants with baseline mean and maximum IMT measures comparable with those in IMPROVE. Meta-analysis of the effect of the 3-SNP score on baseline IMTmean and IMTmax demonstrated no association (Supplementary Fig. 4 and Supplementary Table 11). Analyses of secondary IMT phenotypes demonstrated that no single segment differed from the primary phenotypes analysed (Supplementary Table 12). Proinsulin levels according to SNP score are presented in Supplementary Table 13.

3.6. Mendelian randomization refutes causality of proinsulin on cIMT

Proinsulin levels vary over time with measurement of levels reflecting the exposure of the vasculature to proinsulin at a single point in time. In contrast, the proinsulin-increasing SNP score reflects a lifetime of exposure to proinsulin. For causality to be demonstrated, the direction of effect of the proinsulin-increasing SNP score on cIMT needs to be consistent with the effect of higher levels of proinsulin on cIMT. A further requirement is that the SNPs only act on cIMT through proinsulin levels. Here, the direction of effect was consistent between proinsulin levels and proinsulin-increasing SNP scores: higher levels of proinsulin were (non-significantly) associated with smaller cIMT measures. However, this finding does not imply causality, because when only the non-pleiotropic SNPs were included in the SNP score, no association with cIMT was observed. Taken together, proinsulin per se is unlikely to have a causal effect on cIMT.

4. Discussion

It has previously been reported that circulating levels of proinsulin are predictive of CVD [5–9]. To the best of our knowledge, this is the first study to examine whether proinsulin has a causal effect on early vascular processes, such as increased thickness of the carotid artery wall measured by cIMT. In this study we identified a proinsulin-associated locus and used this, and other previously reported proinsulin-associated SNPs [10], in a MR experiment,
Fig. 2. Regional plots of suggestive loci.
(A) Chromosome 11 locus; (B) chromosome 11 locus after conditioning on the previously reported ARAP1 loci (rs11603334 and rs77756118); (C) chromosome 15 locus; (D) chromosome 15 locus adjusted for the previously reported, nearby loci (LARP6 rs7163439 and VPS13C rs4502156); (E) chromosome 6 locus.
which demonstrated that proinsulin is unlikely to be a causal factor in determining cIMT.

The robust proinsulin-associated loci generally had sizes and directions of effect in IMPROVE that were comparable with those previously reported [10], which added support to the validity of the proinsulin-associated loci. However, these associations demonstrated strong effects on proinsulin levels but convincingly null effects on T2D. The MR analyses can be used to demonstrate causality if i) proinsulin levels are associated with cIMT; ii) SNPs demonstrate robust associations with proinsulin; and iii) SNPs influence cIMT only through proinsulin, and not via other mechanisms. Proinsulin levels have been shown to influence cIMT in a previous report from a cross-sectional study of healthy middle-aged men [4]. However, in IMPROVE the association was weak and non-significant. Robust associations between SNPs and proinsulin levels have been reported [10] and were consistent in IMPROVE. Whilst some association was observed between the 11-SNP score and cIMT, adjustment for pleiotropy (using MR-Eggers) or omission of the pleiotropic SNPs (3-SNP score) rendered the associations with cIMT non-significant. These results suggest that associations between proinsulin and cIMT are most likely due to confounding and pleiotropic effects of SNPs. Despite these findings, CVD prevention strategies which aim to reduce insulin resistance are still of value, as insulin resistance promotes other CVD risk factors including obesity and leads to T2D. In addition, proinsulin processing to insulin also releases c-peptide, which negatively correlates with HDL levels [34], therefore reducing hyperinsulinemia could enable increased HDL levels.

A limitation of this study is lack of power to detect an association, with much larger sample sizes (n > 350,000) being required to conclusively address this question. MR analyses are frequently weakened by small effect sizes of genetic variants and inability to collect large enough sample sizes to enable drawing of definitive conclusions. In the absence of sufficient sample sizes with cIMT measures, alternative methodologies would be required to confirm these findings. Metabolic phenotypes are particularly challenging because of the complexity of interlinking mechanisms.

A unique strength of this study is the detailed measurements of the cIMT. The secondary analyses in the present study demonstrated that a positive association between proinsulin and cIMT was limited to only one segment in the carotid tree, which is an interesting and novel observation. However the uniqueness of IMPROVE is also a limiting factor in that there are few available cohorts with comparable cIMT measures, which precludes replication of the cIMT segments considered in the secondary analysis. IMT measures are often rather variable, especially when considering progression over a relatively short time frame. We recognise this limitation, but note that the standard deviations for the studies included in this analysis are not excessive, that there is consistency across the studies and we have focused on the baseline measures and merely comment upon progression data. The weak association between proinsulin and cIMT is a weakness of this study, however there is prior independent evidence for the association between proinsulin and insulin resistance, via its association with myosin [32,33]. It is possible that genetically low UNC45A levels could influence proinsulin levels, either via aberrant trafficking of vesicles (leading to release of incompletely processed proinsulin) or as a result of a reduced number of β -cells in the pancreas (leaving the pancreas more vulnerable to damaging stress during insulin resistance and hyperinsulinemia). Five genetic variants in the UNC45A genomic region have been reported in association with metabolic traits (GWAS catalogue, dated 20161017). However, these are all ~500 kb from the chromosome 15 locus and are at most in low linkage disequilibrium with rs8029765 (r2 < 0.11, Supplementary Fig. 3). In addition, they show no association with proinsulin levels in this study (smallest p > 0.2411), suggesting that these signals are unrelated. The lack of effect of rs8029765 on T2D is consistent with other proinsulin-associated loci, for example PCSK1, which demonstrated strong effects on proinsulin levels but convincingly null effects on T2D [10].

The MR analyses can be used to demonstrate causality if i) proinsulin levels are associated with cIMT; ii) SNPs demonstrate robust associations with proinsulin; and iii) SNPs influence cIMT only through proinsulin, and not via other mechanisms. Proinsulin levels have been shown to influence cIMT in a previous report from a cross-sectional study of healthy middle-aged men [4]. However, in IMPROVE the association was weak and non-significant. Robust associations between SNPs and proinsulin levels have been reported [10] and were consistent in IMPROVE. Whilst some association was observed between the 11-SNP score and cIMT, adjustment for pleiotropy (using MR-Eggers) or omission of the pleiotropic SNPs (3-SNP score) rendered the associations with cIMT non-significant. These results suggest that associations between proinsulin and cIMT are most likely due to confounding and pleiotropic effects of SNPs. Despite these findings, CVD prevention strategies which aim to reduce insulin resistance are still of value, as insulin resistance promotes other CVD risk factors including obesity and leads to T2D. In addition, proinsulin processing to insulin also releases c-peptide, which negatively correlates with HDL levels [34], therefore reducing hyperinsulinemia could enable increased HDL levels.

A unique strength of this study is the detailed measurements of the cIMT. The secondary analyses in the present study demonstrated that a positive association between proinsulin and cIMT was limited to only one segment in the carotid tree, which is an interesting and novel observation. However the uniqueness of IMPROVE is also a limiting factor in that there are few available cohorts with comparable cIMT measures, which precludes replication of the cIMT segments considered in the secondary analysis. IMT measures are often rather variable, especially when considering progression over a relatively short time frame. We recognise this limitation, but note that the standard deviations for the studies included in this analysis are not excessive, that there is consistency across the studies and we have focused on the baseline measures and merely comment upon progression data. The weak association between proinsulin and cIMT is a weakness of this study, however there is prior independent evidence for the association between proinsulin and insulin resistance, via its association with myosin [32,33]. It is possible that genetically low UNC45A levels could influence proinsulin levels, either via aberrant trafficking of vesicles (leading to release of incompletely processed proinsulin) or as a result of a reduced number of β -cells in the pancreas (leaving the pancreas more vulnerable to damaging stress during insulin resistance and hyperinsulinemia). Five genetic variants in the UNC45A genomic region have been reported in association with metabolic traits (GWAS catalogue, dated 20161017). However, these are all ~500 kb from the chromosome 15 locus and are at most in low linkage disequilibrium with rs8029765 (r2 < 0.11, Supplementary Fig. 3). In addition, they show no association with proinsulin levels in this study (smallest p > 0.2411), suggesting that these signals are unrelated. The lack of effect of rs8029765 on T2D is consistent with other proinsulin-associated loci, for example PCSK1, which demonstrated strong effects on proinsulin levels but convincingly null effects on T2D [10].

The MR analyses can be used to demonstrate causality if i) proinsulin levels are associated with cIMT; ii) SNPs demonstrate robust associations with proinsulin; and iii) SNPs influence cIMT only through proinsulin, and not via other mechanisms. Proinsulin levels have been shown to influence cIMT in a previous report from a cross-sectional study of healthy middle-aged men [4]. However, in IMPROVE the association was weak and non-significant. Robust associations between SNPs and proinsulin levels have been reported [10] and were consistent in IMPROVE. Whilst some association was observed between the 11-SNP score and cIMT, adjustment for pleiotropy (using MR-Eggers) or omission of the pleiotropic SNPs (3-SNP score) rendered the associations with cIMT non-significant. These results suggest that associations between proinsulin and cIMT are most likely due to confounding and pleiotropic effects of SNPs. Despite these findings, CVD prevention strategies which aim to reduce insulin resistance are still of value, as insulin resistance promotes other CVD risk factors including obesity and leads to T2D. In addition, proinsulin processing to insulin also releases c-peptide, which negatively correlates with HDL levels [34], therefore reducing hyperinsulinemia could enable increased HDL levels.

A unique strength of this study is the detailed measurements of the cIMT. The secondary analyses in the present study demonstrated that a positive association between proinsulin and cIMT was limited to only one segment in the carotid tree, which is an interesting and novel observation. However the uniqueness of IMPROVE is also a limiting factor in that there are few available cohorts with comparable cIMT measures, which precludes replication of the cIMT segments considered in the secondary analysis. IMT measures are often rather variable, especially when considering progression over a relatively short time frame. We recognise this limitation, but note that the standard deviations for the studies included in this analysis are not excessive, that there is consistency across the studies and we have focused on the baseline measures and merely comment upon progression data. The weak association between proinsulin and cIMT is a weakness of this study, however there is prior independent evidence for the association between proinsulin and insulin resistance, via its association with myosin [32,33]. It is possible that genetically low UNC45A levels could influence proinsulin levels, either via aberrant trafficking of vesicles (leading to release of incompletely processed proinsulin) or as a result of a reduced number of β -cells in the pancreas (leaving the pancreas more vulnerable to damaging stress during insulin resistance and hyperinsulinemia). Five genetic variants in the UNC45A genomic region have been reported in association with metabolic traits (GWAS catalogue, dated 20161017). However, these are all ~500 kb from the chromosome 15 locus and are at most in low linkage disequilibrium with rs8029765 (r2 < 0.11, Supplementary Fig. 3). In addition, they show no association with proinsulin levels in this study (smallest p > 0.2411), suggesting that these signals are unrelated. The lack of effect of rs8029765 on T2D is consistent with other proinsulin-associated loci, for example PCSK1, which demonstrated strong effects on proinsulin levels but convincingly null effects on T2D [10].
proinsulin and cIMT [4]. Biomarker measurements reflect a snapshot in time, whereas genetic variants reflect the lifetime exposure to biomarker levels. Thus by using genetic variants robustly associated with proinsulin levels, MR analyses give a more stable lifetime risk assessment. For a genetic discovery analysis, the IMPROVE cohort is relatively small. However, the high CVD-risk profile of all subjects renders it a relevant cohort for association studies of CVD-associated biomarkers. In addition, all participants had at least 3 of 9 established CVD risk factors [12], meaning that the repertoire of potential confounders was very varied across the cohort and thereby limiting the potential for selection bias. Replication of the chromosome 15 proinsulin-associated SNP was conducted in a population of similar size to that utilised in the proinsulin GWAS [10], thus giving confidence to this finding.

In conclusion, we have identified a proinsulin-associated locus and used MR to demonstrate that proinsulin per se is unlikely to have a causal effect on cIMT, a proxy measurement for early atherosclerosis.

**Conflict of interest**

A.D has received consultancy and research support from Metagenics Inc. (outside the scope of submitted work). Metagenics Inc. had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript. E.I is an advisor and consultant for Precision Wellness, Inc., and advisor for Cellink for work unrelated to the present project. The other authors have nothing to disclose.

**Financial support**

IMPROVE was supported by the European Commission (Contract number: QLG1-CT-2002-00896), the Swedish Heart-Lung Foundation, the Swedish Research Council (projects 8691 and 09533), the Knut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Stockholm County Council (project 592229), the Strategic Cardiovascular and Diabetes Programmes of Karolinska Institutet and Stockholm County Council, the European Union Framework Programme 7 (FP7/2007–2013) for the Innovative Medicine Initiative under grant agreement n° IMI/115006 (the SUMMIT consortium), the Academy of Finland (Grant #110413), the British Heart Foundation (RG2008/08, RG2008/014) and the Italian Ministry of Health (Ricerca Corrente). The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers, General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. This research includes data and resources from the FHS of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute’s Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research was conducted using the Linux Clusters for Genetic Analysis (LinGA) computing resources at Boston University Medical Campus. Also supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) R01 DK078616 and NIDDK K24 DK080140, and a Career Development Award from the American Diabetes Association. The HBCS has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälso Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation. The NEO study thanks all participants, general practitioners and research nurses involved in this study. We thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. PIVUS acknowledges the support of the Knut och Alice Wallenberg Foundation (Wallenberg Academy Fellow), European Research Council (ERC-2013-StG; grant no. 335395), Swedish Diabetes Foundation, Swedish Heart-Lung Foundation (grant no. 20120197), Swedish Research Council (grant no. 2012-1397), and Wellcome Trust (grant no. WT098917, WT099352, and WT064890). The analyses were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b20113-10N, b20103-08. This facility is funded by the European Commission Framework 6 (FP6) program (LSHM-CT-2007-037273), the British Heart Foundation, AstaZe-neca, the Swedish Research Council (8691), the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation and the Torsten and Ragnar Söderberg Foundation. H.W acknowledges the support of the Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Research Excellence. A.G and H.W acknowledge European Union Seventh Framework Programme FP7/2007–2013 under grant agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & and A.G, the Wellcome Trust Institutional strategic support fund. RSI and RSI: The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; the Netherlands Organisation for Scientific Research; the Netherlands Organisation for Health Research and Development (ZonMW); the Research Institute for Diseases in the Elderly; The Netherlands Heart Foundation; the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. Support for genotyping was provided by the Netherlands Organisation for Scientific Research (NWO) (175.010.2005.011, 911.03.012), the Netherlands Genomics Initiative (NGI)/NWO project nr. 050-060-810 and Research Institute for Diseases in the Elderly (RIDE). ULSAM was supported by the Knut och Alice Wallenberg Foundation (Wallenberg Academy Fellow), European Research Council (ERC-2013-StG; grant no. 335395), Swedish Diabetes Foundation, Swedish Heart-Lung Foundation (grant no. 20120197), Swedish Research Council (grant no. 2012-1397), and Wellcome Trust (grant no. WT098917, WT099352, and WT064890). The analyses were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping of Swedish samples was performed by the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) which is part of the National Genomics Infrastructure at Science for Life Laboratory, Uppsala University Sweden. The SNP&SEQ Technology Platform is also supported by the Swedish Research Council for Infrastructures and the Knut and Alice Waldenborg Foundation. M.K is supported by The Netherlands Organisation for Scientific Research Grant VENI 91616909. B.S. is financially supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. M.D.H is supported by grants from the Swedish Heart-Lung Foundation (20140543), Swedish Research Council (2015-
03657), and NIH (R01DK106236). C.M.L. is a Wellcome Trust Research Career Development Fellow (086596/2/08/2). O.H.F works in ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd), Metagenics, Inc, and AXA. E.I. is supported by grants from the Knut and Alice Wallenberg Foundation, the European Research Council (ERC Starting Grant, 335395), the Swedish Heart-Lung Foundation (20120197), and the Swedish Research Council (2012-1397). J.B.M is supported by R01DK078616 and K24 DK080140. D.O.M-K is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023). J.W.J acknowledges that the research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-223004 PHASE. A.P.M is a Wellcome Trust Senior Fellow in Basic Biomedical Science (grant number WT098017). S.E.H. is funded by the National Institute for Health Research UCL Hospitals Biomedical Research Centre. None of the funding sources were involved in the design, collection, analysis or interpretation of the study.

Acknowledgments

We thank all participants and those involved in the recruitment of the cohorts included in this study.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.atherosclerosis.2017.09.031.

References