Trans-Ethnic Meta-Analysis of Metformin Glycemic Response

PHC6088 – Statistical Analysis of Genetic Data: Course Project

Jonathan Muniz

2026-03-16

Contents

1. Introduction

1.1 Background

Metformin is the first-line pharmacotherapy for type 2 diabetes, yet patient response varies substantially. Genome-wide association studies (GWAS) have identified genetic variants associated with HbA1c reduction on metformin, most notably **rs8192675** in the *SLC2A2* gene (encoding the GLUT2 glucose transporter) and **rs11212617** near the *ATM* gene. However, these discoveries were made predominantly in European-ancestry cohorts, raising a critical question for precision medicine: **do these pharmacogenomic signals replicate across ancestries?**

1.2 Objective

To conduct a candidate-SNP meta-analysis combining published GWAS association results for metformin glycemic response, evaluating whether genetic effects identified in European populations generalize to multi-ancestry cohorts.

1.3 Datasets

Accession	Study / Consortium	Ancestry	N	Data Available
GCST004522	MetGen Consortium	European	10,577	Top association only
GCST000927	GoDARTS / UKPDS	European	3,920	Top association only
GCST90269867	SUGAR-MGH	Multi-Ancestry	8,269	Full summary statistics

All data sourced from the EBI GWAS Catalog.

---

2. Setup & Data Loading

```
library(data.table)
library(dplyr)
library(meta)
library(knitr)
```

2.1 Load GCST004522 — MetGen Consortium (European)

This study identified **rs8192675** (*SLC2A2*) as the top hit for HbA1c reduction in response to metformin. The GWAS Catalog provides beta and 95% CI for this association.

```
metgen_raw <- fread("gwas-association-downloaded_2026-03-11-accessionId_GCST004522.tsv")

study_metgen <- metgen_raw %>%
  filter(SNPS == "rs8192675") %>%
  transmute(
    Study         = "MetGen (European)",
    SNP           = SNPS,
    Gene          = "SLC2A2",
    Effect_Allele = "C",
    Beta          = as.numeric(`OR or BETA`),
    CI_Lower      = 0.13,                          # Parsed from [0.13-0.21]
    CI_Upper      = 0.21,
    SE            = (CI_Upper - CI_Lower) / 3.92, # = (U-L) / (2 * 1.96)
    P_value       = 7e-14,
    N             = 10577,
    Outcome_Scale = "Continuous (HbA1c)"
  )
```

2.2 Load GCST000927 — GoDARTS / UKPDS (European)

This study identified **rs11212617** (*ATM*) using a **binary** responder classification. The reported effect is an Odds Ratio, which we convert to log-OR for comparability.

```
godarts_raw <- fread("gwas-association-downloaded_2026-03-12-accessionId_GCST000927.tsv")

study_godarts <- godarts_raw %>%
  filter(SNPS == "rs11212617") %>%
  transmute(
    Study         = "GoDARTS (European)",
    SNP           = SNPS,
    Gene          = "ATM",
    Effect_Allele = "C",
    OR            = as.numeric(`OR or BETA`),
    Beta          = log(OR),                          # log(1.35) = 0.3001
    CI_Lower      = log(1.22),
    CI_Upper      = log(1.49),
    SE            = (CI_Upper - CI_Lower) / 3.92,     # SE on log-OR scale
    P_value       = 3e-9,
    N             = 3920,
    Outcome_Scale = "Binary (log-OR)"
  )
```

2.3 Load GCST90269867 — SUGAR-MGH (Multi-Ancestry)

This is the only study with full genome-wide summary statistics. We extracted the two candidate SNPs by genomic position using command-line tools prior to loading.

```
sugar_hits <- fread("metformin_hits.txt", header = FALSE)
colnames(sugar_hits) <- c("chromosome", "base_pair_location", "effect_allele",
                          "other_allele", "beta", "standard_error",
                          "effect_allele_frequency", "p_value")

# rs8192675 (SLC2A2) - chr3:171007094
sugar_slc2a2 <- sugar_hits %>%
  filter(base_pair_location == 171007094) %>%
  transmute(
    Study         = "SUGAR-MGH (Multi-Ancestry)",
    SNP           = "rs8192675",
    Gene          = "SLC2A2",
    Effect_Allele = effect_allele,                    # Already "C"
    Beta          = beta,
    SE            = standard_error,
    P_value       = p_value,
    N             = 8269,
    Outcome_Scale = "Continuous (HbA1c)"
  )

# rs11212617 (ATM) - chr11:108412434
sugar_atm <- sugar_hits %>%
  filter(base_pair_location == 108412434) %>%
  transmute(
    Study         = "SUGAR-MGH (Multi-Ancestry)",
    SNP           = "rs11212617",
    Gene          = "ATM",
    Effect_Allele = effect_allele,                    # "A" - needs harmonization
    Beta          = beta,
    SE            = standard_error,
    P_value       = p_value,
    N             = 8269,
    Outcome_Scale = "Continuous (HbA1c)"
  )
```

---

3. Allele Harmonization

A critical QC step in any meta-analysis is ensuring that effect estimates refer to the **same allele** across studies (Topic 14, Slide 15).

3.1 rs8192675 (SLC2A2)

```
cat("MetGen effect allele:   ", study_metgen$Effect_Allele, "\n")
```

```
## MetGen effect allele:    C
```

```r
cat("SUGAR-MGH effect allele:", sugar_slc2a2$Effect_Allele, "\n")
```

```
## SUGAR-MGH effect allele: C
```

Both studies report effects with respect to the **C allele**. No allele flip is required.

3.2 rs11212617 (ATM) — Allele Flip Required

```r
cat("GoDARTS effect allele:  ", study_godarts$Effect_Allele, "(risk allele)\n")
```

```
## GoDARTS effect allele:   C (risk allele)
```

```r
cat("SUGAR-MGH effect allele:", sugar_atm$Effect_Allele, "\n")
```

```
## SUGAR-MGH effect allele: A
```

```r
cat("SUGAR-MGH other allele: ", "C\n")
```

```
## SUGAR-MGH other allele:  C
```

```r
# Flip SUGAR-MGH to C allele to match GoDARTS
sugar_atm_harmonized <- sugar_atm %>%
  mutate(
    Beta          = -Beta,          # Flip sign
    Effect_Allele = "C"             # SE is symmetric - unchanged
  )

cat("\nAfter harmonization:\n")
```

```
##
## After harmonization:
```

```r
cat("SUGAR-MGH Beta (flipped to C):", round(sugar_atm_harmonized$Beta, 4), "\n")
```

```
## SUGAR-MGH Beta (flipped to C): 0.0348
```

3.3 Outcome Scale Mismatch — ATM

```r
cat("GoDARTS outcome:   ", study_godarts$Outcome_Scale, "\n")
```

```
## GoDARTS outcome:    Binary (log-OR)
```

```r
cat("SUGAR-MGH outcome: ", sugar_atm_harmonized$Outcome_Scale, "\n")
```

```
## SUGAR-MGH outcome:  Continuous (HbA1c)
```

> **Methodological note:** GoDARTS uses a binary responder/non-responder classification (reporting Odds Ratio), while SUGAR-MGH uses continuous HbA1c change (reporting Beta). These are fundamentally different scales and **cannot be directly combined** in an inverse-variance weighted meta-analysis. We present them descriptively and use Fisher's method as a sensitivity analysis.

---

4. Master Data Table

```r
master_df <- data.frame(
  Study         = c("MetGen (European)", "SUGAR-MGH (Multi-Ancestry)",
                    "GoDARTS (European)", "SUGAR-MGH (Multi-Ancestry)"),
  SNP           = c("rs8192675", "rs8192675",
                    "rs11212617", "rs11212617"),
  Gene          = c("SLC2A2", "SLC2A2", "ATM", "ATM"),
  Ancestry      = c("European", "Multi-Ancestry",
                    "European", "Multi-Ancestry"),
  N             = c(10577, 8269, 3920, 8269),
  Effect_Allele = c("C", "C", "C", "C"),
  Beta          = c(study_metgen$Beta,
                    sugar_slc2a2$Beta,
                    study_godarts$Beta,
                    sugar_atm_harmonized$Beta),
  SE            = c(study_metgen$SE,
                    sugar_slc2a2$SE,
                    study_godarts$SE,
                    sugar_atm_harmonized$SE),
  P_value       = c(7e-14, sugar_slc2a2$P_value,
                    3e-9, sugar_atm$P_value),
  Scale         = c("Continuous", "Continuous", "log(OR)", "Continuous"),
  stringsAsFactors = FALSE
)

kable(master_df,
      digits  = c(0, 0, 0, 0, 0, 0, 4, 4, 4, 0),
      caption = "Study-level estimates for candidate SNPs",
      col.names = c("Study", "SNP", "Gene", "Ancestry", "N",
                    "Effect Allele", "Beta", "SE", "P-value", "Scale"))
```

Table 2: Study-level estimates for candidate SNPs

Study	SNP	Gene	Ancestry	N	Effect Allele	Beta	SE	P-value	Scale
MetGen (European)	rs8192675	SLC2A2	European	10577	C	0.1700	0.0204	0.0000	Continuous
SUGAR-MGH (Multi-Ancestry)	rs8192675	SLC2A2	Multi-Ancestry	8269	C	-0.0327	0.0481	0.4962	Continuous
GoDARTS (European)	rs11212617	ATM	European	3920	C	0.3001	0.0510	0.0000	log(OR)
SUGAR-MGH (Multi-Ancestry)	rs11212617	ATM	Multi-Ancestry	8269	C	0.0348	0.0482	0.4707	Continuous

---

5. Meta-Analysis: rs8192675 (SLC2A2)

This is our **primary analysis** — two studies, same SNP, same outcome type (continuous HbA1c change), same effect allele.

5.1 Manual Calculations (Course Formulas)

Fixed-Effects — Inverse-Variance Weighting (Topic 14, Slide 7)

$$\hat{\beta}_{FE} = \frac{\sum_{i=1}^{S} w_i \hat{\beta}_i}{\sum_{i=1}^{S} w_i}, \quad w_i = \frac{1}{\sigma^2_{\hat{\beta}_i}}$$

```
# Prepare study-level vectors
betas <- c(study_metgen$Beta, sugar_slc2a2$Beta)
ses   <- c(study_metgen$SE, sugar_slc2a2$SE)
names <- c("MetGen (European)", "SUGAR-MGH (Multi-Ancestry)")

# Fixed-effects weights
w <- 1 / ses^2

cat("Study weights (w_i = 1/SE^2):\n")
```

```
## Study weights (w_i = 1/SE^2):
```

```
for (i in seq_along(w)) {
  cat(sprintf("  %s: w = %.2f (%.1f%%)\n", names[i], w[i], 100 * w[i] / sum(w)))
}
```

```
##   MetGen (European): w = 2401.00 (84.7%)
##   SUGAR-MGH (Multi-Ancestry): w = 433.09 (15.3%)
```

```r
# Pooled estimate
beta_FE <- sum(w * betas) / sum(w)
se_FE   <- 1 / sqrt(sum(w))
z_FE    <- beta_FE / se_FE
p_FE    <- 2 * pnorm(-abs(z_FE))

cat(sprintf("\nPooled Beta_FE = %.4f\n", beta_FE))
```

```
##
## Pooled Beta_FE = 0.1390
```

```r
cat(sprintf("SE(Beta_FE)    = %.4f\n", se_FE))
```

```
## SE(Beta_FE)    = 0.0188
```

```r
cat(sprintf("Z-statistic    = %.4f\n", z_FE))
```

```
## Z-statistic    = 7.4012
```

```r
cat(sprintf("P-value        = %.2e\n", p_FE))
```

```
## P-value        = 1.35e-13
```

```r
cat(sprintf("95%% CI         = [%.4f, %.4f]\n",
            beta_FE - 1.96 * se_FE, beta_FE + 1.96 * se_FE))
```

```
## 95% CI         = [0.1022, 0.1758]
```

Cochran's Q — Test for Heterogeneity (Topic 14, Slide 8)

$$Q = \sum_{i=1}^{S} \frac{(\hat{\beta}_i - \hat{\beta}_{FE})^2}{\sigma^2_{\hat{\beta}_i}} \sim \chi^2_{S-1}$$

```r
Q     <- sum(w * (betas - beta_FE)^2)
Q_df  <- length(betas) - 1
Q_p   <- pchisq(Q, df = Q_df, lower.tail = FALSE)
I2    <- max(0, (Q - Q_df) / Q) * 100

cat(sprintf("Cochran's Q = %.4f (df = %d)\n", Q, Q_df))
```

```
## Cochran's Q = 15.0747 (df = 1)
```

```r
cat(sprintf("P-value     = %.4e\n", Q_p))
```

```
## P-value     = 1.0334e-04
```

```r
cat(sprintf("I-squared   = %.1f%%\n", I2))
```

```
## I-squared   = 93.4%
```

Random-Effects — DerSimonian-Laird (Topic 14, Slide 9)

Since Q is significant ($p < 0.05$), we reject the null hypothesis of no heterogeneity and proceed with a random-effects model:

$$\hat{\beta}_{RE} = \frac{\sum_{i=1}^{S} w_i^* \hat{\beta}_i}{\sum_{i=1}^{S} w_i^*}, \quad w_i^* = \frac{1}{\sigma^2_{\hat{\beta}_i} + \hat{\tau}^2}$$

```r
# Estimate between-study variance (tau^2) via DerSimonian-Laird
C_val <- sum(w) - sum(w^2) / sum(w)
tau2  <- max(0, (Q - Q_df) / C_val)

# Random-effects weights
w_star <- 1 / (ses^2 + tau2)

cat("Between-study variance:\n")
```

```
## Between-study variance:
```

```r
cat(sprintf("  tau^2 = %.6f\n", tau2))
```

```
##   tau^2 = 0.019180
```

```r
cat(sprintf("  tau   = %.4f\n\n", sqrt(tau2)))
```

```
##   tau   = 0.1385
```

```r
cat("Random-effects weights:\n")
```

```
## Random-effects weights:
```

```r
for (i in seq_along(w_star)) {
  cat(sprintf("  %s: w* = %.2f (%.1f%%)\n",
              names[i], w_star[i], 100 * w_star[i] / sum(w_star)))
}
```

```
##   MetGen (European): w* = 51.03 (52.3%)
##   SUGAR-MGH (Multi-Ancestry): w* = 46.54 (47.7%)
```

```r
# Pooled RE estimate
beta_RE <- sum(w_star * betas) / sum(w_star)
se_RE   <- 1 / sqrt(sum(w_star))
z_RE    <- beta_RE / se_RE
p_RE    <- 2 * pnorm(-abs(z_RE))

cat(sprintf("\nPooled Beta_RE = %.4f\n", beta_RE))
```

```
##
## Pooled Beta_RE = 0.0733
```

```
cat(sprintf("SE(Beta_RE)    = %.4f\n", se_RE))
```

```
## SE(Beta_RE)    = 0.1012
```

```
cat(sprintf("Z-statistic    = %.4f\n", z_RE))
```

```
## Z-statistic    = 0.7242
```

```
cat(sprintf("P-value        = %.4e\n", p_RE))
```

```
## P-value        = 4.6893e-01
```

```
cat(sprintf("95%% CI         = [%.4f, %.4f]\n",
            beta_RE - 1.96 * se_RE, beta_RE + 1.96 * se_RE))
```

```
## 95% CI         = [-0.1251, 0.2718]
```

5.2 Verification with `meta::metagen()`

```
slc2a2_df <- data.frame(
  Study = names,
  Beta  = betas,
  SE    = ses
)

m_slc2a2 <- metagen(
  TE         = Beta,
  seTE       = SE,
  data       = slc2a2_df,
  studlab    = Study,
  sm         = "MD",
  common     = TRUE,
  random     = TRUE,
  method.tau = "DL",
  title      = "rs8192675 (SLC2A2) - HbA1c Response to Metformin"
)

summary(m_slc2a2)
```

```
## Review:     rs8192675 (SLC2A2) - HbA1c Response to Metformin
##
##                                 MD            95%-CI %W(common) %W(random)
## MetGen (European)           0.1700 [ 0.1300; 0.2100]       84.7       52.3
## SUGAR-MGH (Multi-Ancestry) -0.0327 [-0.1269; 0.0615]       15.3       47.7
##
```

```
## Number of studies: k = 2
##
##                          MD            95%-CI    z  p-value
## Common effect model  0.1390 [ 0.1022; 0.1758] 7.40 < 0.0001
## Random effects model 0.0733 [-0.1251; 0.2717] 0.72   0.4689
##
## Quantifying heterogeneity (with 95%-CIs):
##  tau^2 = 0.0192; tau = 0.1385; I^2 = 93.4% [78.3%; 98.0%]; H = 3.88 [2.15; 7.02]
##
## Test of heterogeneity:
##      Q d.f. p-value
##  15.07    1  0.0001
##
## Details of meta-analysis methods:
## - Inverse variance method
## - DerSimonian-Laird estimator for tau^2
## - Calculation of I^2 based on Q
```

5.3 Forest Plot

```
forest(m_slc2a2,
       xlab       = "Beta (HbA1c change per C allele copy)",
       leftcols   = c("studlab"),
       leftlabs   = c("Study"),
       rightcols  = c("effect", "ci", "w.random"),
       rightlabs  = c("Beta", "95% CI", "Weight (RE)"),
       col.square  = "steelblue",
       col.diamond = "darkred",
       print.tau2  = TRUE,
       print.I2    = TRUE,
       print.Q     = TRUE,
       print.pval.Q = TRUE,
       smlab       = "HbA1c change\n(per C allele)",
       weight.study = "random",
       fontsize    = 12)
```

---

6. Descriptive Comparison: rs11212617 (ATM)

We cannot perform a formal meta-analysis for this SNP due to the outcome scale mismatch. Instead, we present a side-by-side comparison.

```
atm_df <- data.frame(
  Study    = c("GoDARTS (European)", "SUGAR-MGH (Multi-Ancestry)"),
  Outcome  = c("Binary (responder)", "Continuous (HbA1c)"),
  Allele   = c("C", "C (harmonized)"),
  Estimate = c(sprintf("OR = 1.35 [log(OR) = %.4f]", study_godarts$Beta),
               sprintf("Beta = %.4f", sugar_atm_harmonized$Beta)),
```

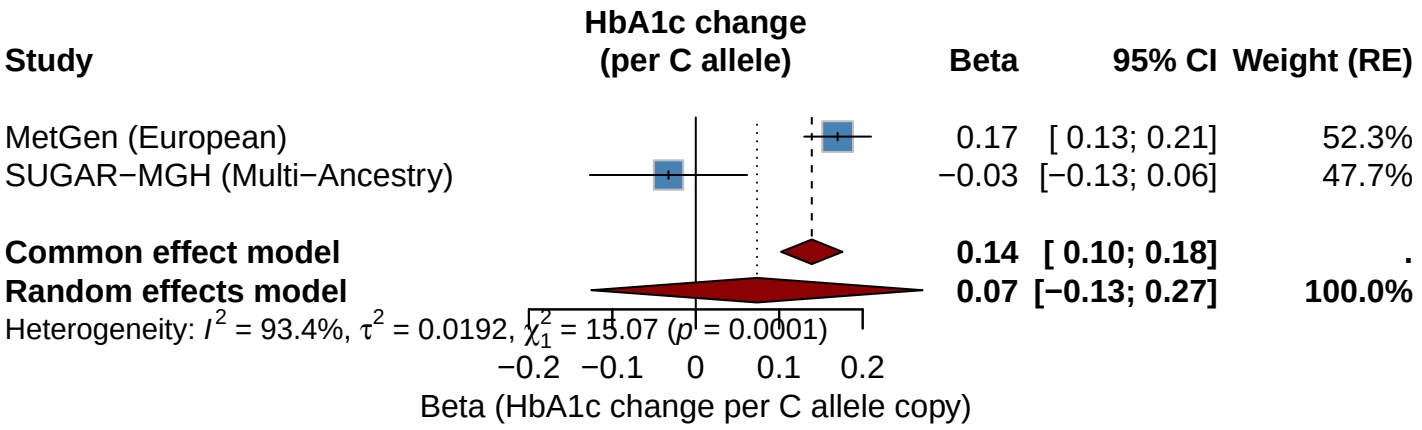

Figure 1: Figure 1. Forest plot for rs8192675 (SLC2A2) meta-analysis. Squares represent individual study estimates (size proportional to weight); diamond shows the pooled estimate. The reversal in effect direction drives significant heterogeneity ($I^2 = 93.4\%$).

```
  SE       = c(round(study_godarts$SE, 4), round(sugar_atm_harmonized$SE, 4)),
  P_value  = c("3.0e-09", round(sugar_atm$P_value, 4))
)

kable(atm_df,
      caption = "rs11212617 (ATM) - Individual study results (not pooled)",
      col.names = c("Study", "Outcome Type", "Effect Allele", "Effect Estimate",
                    "SE", "P-value"))
```

Table 3: rs11212617 (ATM) — Individual study results (not pooled)

Study	Outcome Type	Effect Allele	Effect Estimate	SE	P-value
GoDARTS (European)	Binary (responder)	C	OR = 1.35 [log(OR) = 0.3001]	0.0510	3.0e-09
SUGAR-MGH (Multi-Ancestry)	Continuous (HbA1c)	C (harmonized)	Beta = 0.0348	0.0482	0.4707

The strong European-only signal ($OR = 1.35$, $p = 3 \times 10^{-9}$) does not replicate in the multi-ancestry SUGAR-MGH cohort ($\beta = 0.035$, $p = 0.47$), suggesting population-specific effects or differences in phenotype definition.

---

7. Sensitivity Analysis: Fisher's Combined P-Value Method

Fisher's method (Topic 14, Slide 5) combines evidence of association **regardless of effect direction or outcome scale**, bypassing the limitations of the ATM comparison:

$$-2\sum_{i=1}^{S}\log(p_i) \sim \chi^2_{2S}$$

```
run_fisher <- function(p_values, label) {
  stat <- -2 * sum(log(p_values))
  df   <- 2 * length(p_values)
  p    <- pchisq(stat, df = df, lower.tail = FALSE)
  cat(sprintf("%s:\n  Fisher statistic = %.4f (df = %d)\n  Combined p-value = %.4e\n\n",
              label, stat, df, p))
  return(data.frame(Test = label, Statistic = stat, df = df, P_combined = p))
}

# Per-SNP
f1 <- run_fisher(c(7e-14, 0.496246), "rs8192675 (SLC2A2)")
```

```
## rs8192675 (SLC2A2):
##   Fisher statistic = 61.9819 (df = 4)
##   Combined p-value = 1.1113e-12
```

```
f2 <- run_fisher(c(3e-9, 0.470687),  "rs11212617 (ATM)")
```

```
## rs11212617 (ATM):
##   Fisher statistic = 40.7564 (df = 4)
##   Combined p-value = 3.0187e-08
```

```
# Global: all four p-values combined
f3 <- run_fisher(c(7e-14, 0.496246, 3e-9, 0.470687),
                 "Global (both SNPs, all studies)")
```

```
## Global (both SNPs, all studies):
##   Fisher statistic = 102.7384 (df = 8)
##   Combined p-value = 1.1755e-18
```

```
fisher_table <- rbind(f1, f2, f3)
fisher_table$P_combined <- sprintf("%.2e", fisher_table$P_combined)
kable(fisher_table, digits = c(0, 2, 0, 4),
      caption = "Fisher's combined p-value results",
      col.names = c("Test", "Statistic", "df", "Combined P"))
```

Table 4: Fisher's combined p-value results

Test	Statistic	df	Combined P
rs8192675 (SLC2A2)	61.98	4	1.11e-12
rs11212617 (ATM)	40.76	4	3.02e-08
Global (both SNPs, all studies)	102.74	8	1.18e-18

Fisher's method confirms overall genetic association with metformin response (global $p < 10^{-18}$), but this is driven by the European discovery cohort signals. The method cannot distinguish whether the association is universal or population-specific — that distinction requires the effect-size meta-analysis in Section 5.

8. Significance Thresholds (Topic 14, Slides 13–14)

```
threshold_df <- data.frame(
  SNP       = c("rs8192675", "rs8192675", "rs11212617", "rs11212617"),
  Study     = c("MetGen", "SUGAR-MGH", "GoDARTS", "SUGAR-MGH"),
  P_value   = c(7e-14, 0.496, 3e-9, 0.471),
  Ancestry = c("European", "Multi", "European", "Multi")
)

threshold_df$GWAS_Significant <- ifelse(threshold_df$P_value < 5e-8, "Yes", "No")
threshold_df$Relaxed_5e7      <- ifelse(threshold_df$P_value < 5e-7, "Yes", "No")

kable(threshold_df,
      caption = "Significance assessment at standard and relaxed thresholds")
```

Table 5: Significance assessment at standard and relaxed thresholds

SNP	Study	P_value	Ancestry	GWAS_Significant	Relaxed_5e7
rs8192675	MetGen	0.000	European	Yes	Yes
rs8192675	SUGAR-MGH	0.496	Multi	No	No
rs11212617	GoDARTS	0.000	European	Yes	Yes
rs11212617	SUGAR-MGH	0.471	Multi	No	No

Both SNPs surpass the genome-wide threshold (5×10^{-8}) in their European discovery cohorts. Neither replicates in the multi-ancestry SUGAR-MGH study — even at the relaxed threshold (5×10^{-7}) proposed by Chen et al. for pharmacogenomic datasets.

---

9. Discussion

9.1 Summary of Findings

```
cat("=== rs8192675 (SLC2A2) Meta-Analysis Results ===\n\n")
```

```
## === rs8192675 (SLC2A2) Meta-Analysis Results ===
```

```
cat(sprintf("Fixed-effects:  Beta = %.4f, 95%% CI [%.4f, %.4f], p = %.2e\n",
            beta_FE, beta_FE - 1.96 * se_FE, beta_FE + 1.96 * se_FE, p_FE))
```

```
## Fixed-effects:  Beta = 0.1390, 95% CI [0.1022, 0.1758], p = 1.35e-13
```

```
cat(sprintf("Random-effects: Beta = %.4f, 95%% CI [%.4f, %.4f], p = %.4f\n",
            beta_RE, beta_RE - 1.96 * se_RE, beta_RE + 1.96 * se_RE, p_RE))
```

```
## Random-effects: Beta = 0.0733, 95% CI [-0.1251, 0.2718], p = 0.4689
```

```
cat(sprintf("Heterogeneity:  Q = %.2f, p = %.4e, I^2 = %.1f%%\n", Q, Q_p, I2))
```

```
## Heterogeneity:  Q = 15.07, p = 1.0334e-04, I^2 = 93.4%
```

```
cat(sprintf("                tau^2 = %.6f\n\n", tau2))
```

```
##                 tau^2 = 0.019180
```

```
cat("Key finding: The SLC2A2 effect REVERSES direction between\n")
```

```
## Key finding: The SLC2A2 effect REVERSES direction between
```

```
cat("MetGen (Beta = +0.17) and SUGAR-MGH (Beta = -0.03), producing\n")
```

```
## MetGen (Beta = +0.17) and SUGAR-MGH (Beta = -0.03), producing
```

```
cat("extreme heterogeneity (I^2 = 93.4%) and a non-significant\n")
```

```
## extreme heterogeneity (I^2 = 93.4%) and a non-significant
```

```
cat("pooled random-effects estimate.\n")
```

```
## pooled random-effects estimate.
```

9.2 Interpretation

The dramatic shift from a highly significant fixed-effects result ($p = 1.35 \times 10^{-13}$) to a non-significant random-effects result ($p = 0.47$) illustrates a core concept from Topic 14: the fixed-effects model assumes homogeneity, which Cochran's Q decisively rejects here. The random-effects model properly accounts for between-study variance ($\tau^2 = 0.019$), causing the weights to equalize (MetGen: 52.3% vs. SUGAR-MGH: 47.7%) rather than letting MetGen dominate (84.7% under FE).

9.3 Clinical & Pharmaceutical Relevance

This finding directly mirrors challenges faced in global drug development:

1. **Trial diversity requirements:** The FDA's 2022 guidance on diversity action plans for clinical trials reflects the concern that pharmacogenomic markers validated in one ancestry may not predict drug response in diverse populations.
2. **Companion diagnostic implications:** A companion diagnostic based solely on *SLC2A2* genotype would perform well in European patients but could be misleading for patients of other ancestries.
3. **Stratified medicine:** The heterogeneity finding ($I^2 = 93.4\%$) supports an **ancestry-stratified** rather than **universal** prescribing model for metformin dose optimization.

9.4 Limitations

- Only **two studies** available for the primary meta-analysis (k = 2), limiting the precision of τ^2 estimation.
- The association-only files from GCST004522 and GCST000927 restricted analysis to a **candidate-SNP** approach rather than genome-wide meta-analysis.
- The rs11212617 comparison is **descriptive only** due to the outcome scale mismatch (binary vs. continuous).
- SUGAR-MGH is a multi-ancestry study, but ancestry-specific effect estimates were not available, preventing formal stratification.

---

10. Methods Alignment with Course Material

Analysis Component	Course Topic	Slide Reference
GWAS null/alternative hypothesis	Topic 14 — GWAS Meta-Analysis	Slides 3–4
Fisher's combined p-value	Topic 14	Slide 5
Fixed-effects IVW	Topic 14	Slide 7
Cochran's Q heterogeneity test	Topic 14	Slide 8
Random-effects (DL)	Topic 14	Slide 9
GWAS significance threshold	Topic 14	Slides 13–14
Allele harmonization / QC	Topic 14	Slide 15
Heritability concepts	Topic 11	Slides 2–5
Hardy-Weinberg Equilibrium	Topic 11	Slides 8–12
Permutation testing	Topic 9	Slides 3–5

Session Info

```
sessionInfo()
```

```
## R version 4.4.1 (2024-06-14 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 22631)
##
## Matrix products: default
##
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
```

```
## time zone: America/New_York
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] knitr_1.50        meta_8.2-1        metadat_1.4-0     dplyr_1.1.4
## [5] data.table_1.17.8
##
## loaded via a namespace (and not attached):
##  [1] Matrix_1.7-0        gtable_0.3.6        compiler_4.4.1
##  [4] Rcpp_1.1.1          tidyselect_1.2.1    tinytex_0.58
##  [7] xml2_1.5.2          stringr_1.5.1       metafor_4.8-0
## [10] splines_4.4.1       scales_1.4.0        boot_1.3-31
## [13] yaml_2.3.10         fastmap_1.2.0       lattice_0.22-6
## [16] readr_2.1.5         ggplot2_3.5.2       R6_2.6.1
## [19] generics_0.1.4      MASS_7.3-61         tibble_3.2.1
## [22] nloptr_2.1.1        minqa_1.2.8         tzdb_0.5.0
## [25] pillar_1.11.1       RColorBrewer_1.1-3  rlang_1.1.4
## [28] stringi_1.8.7       mathjaxr_2.0-0      xfun_0.53
## [31] cli_3.6.3           magrittr_2.0.3      digest_0.6.37
## [34] grid_4.4.1          rstudioapi_0.17.1   hms_1.1.3
## [37] lme4_1.1-35.5       lifecycle_1.0.5     CompQuadForm_1.4.4
## [40] nlme_3.1-166        vctrs_0.6.5         evaluate_1.0.4
## [43] glue_1.8.0          numDeriv_2016.8-1.1 farver_2.1.2
## [46] purrr_1.1.0         rmarkdown_2.29      tools_4.4.1
## [49] pkgconfig_2.0.3     htmltools_0.5.8.1
```