

brain CTP differences (local plots)

Chloe Yap

30 December 2022

Contents

Set-up	1
Read-in	2
Linear models for individual cell-type differences	2
Sub-types	2
ggstatsplot for significant hits	3
Plot extracting diagnosis effect	4
All in one go (multivariate)	4
Extract diagnosis effect	5
Repeat for Jill's analysis	8
Extract sex effect	9
Extract age effect	11
Extract age ² effect (this model doesn't also include age)	13
Extract mega-analysis results only	14
CTL only	16

Set-up

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5    v purrr  0.3.4
## v tibble  3.1.6    v dplyr  1.0.9
## v tidyr   1.2.0    v stringr 1.4.0
## v readr   2.1.2    v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

library(ggstatsplot)

## You can cite this package as:
## Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach.
## Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167

library(ggrepel)
library(wesanderson)
library(DT)
```

Read-in

```
ctp.clr.1e3 <- read.delim("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/ctp_clr1e-3_1e3.csv")
asd_exclude <- read.delim("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/asd_exclude.csv")
scz_exclude <- read.delim("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/scz_exclude.csv")
colnames(asd_exclude) <- c("FID", "IID")
```

Linear models for individual cell-type differences

Sub-types

clr, no adjustment for oligodendrocytes

```
asd_exc_nocov <- lm(Exc ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_inh_nocov <- lm(Inh ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_astro_nocov <- lm(NonN_Astro_FGF3R ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_endo_nocov <- lm(NonN_Endo ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_micro_nocov <- lm(NonN_Micro ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_oligo_nocov <- lm(NonN_Oligo_MBP ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_opc_nocov <- lm(NonN_OPC ~ dx, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))

asd_exc_cov <- lm(Exc ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_inh_cov <- lm(Inh ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_astro_cov <- lm(NonN_Astro_FGF3R ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_endo_cov <- lm(NonN_Endo ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_micro_cov <- lm(NonN_Micro ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_oligo_cov <- lm(NonN_Oligo_MBP ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
asd_opc_cov <- lm(NonN_OPC ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ASDbrain") %>% filter(!IID %in% asd_exclude$IID))
```

ASDbrain

```

scz_exc_nocov <- lm(Exc ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_inh_nocov <- lm(Inh ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_astro_nocov <- lm(NonN_Astro_FGF3R ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_endo_nocov <- lm(NonN_Endo ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_micro_nocov <- lm(NonN_Micro ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_oligo_nocov <- lm(NonN_Oligo_MBP ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_opc_nocov <- lm(NonN_OPC ~ dx, data = ctp.clr.1e3 %>% filter(study == "LIBD"))

scz_exc_cov <- lm(Exc ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_inh_cov <- lm(Inh ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_astro_cov <- lm(NonN_Astro_FGF3R ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_endo_cov <- lm(NonN_Endo ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_micro_cov <- lm(NonN_Micro ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_oligo_cov <- lm(NonN_Oligo_MBP ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))
scz_opc_cov <- lm(NonN_OPC ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "LIBD"))

```

LIBD

```

azd_exc_nocov <- lm(Exc ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_inh_nocov <- lm(Inh ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_astro_nocov <- lm(NonN_Astro_FGF3R ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_endo_nocov <- lm(NonN_Endo ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_micro_nocov <- lm(NonN_Micro ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_oligo_nocov <- lm(NonN_Oligo_MBP ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_opc_nocov <- lm(NonN_OPC ~ dx, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))

azd_exc_cov <- lm(Exc ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_inh_cov <- lm(Inh ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_astro_cov <- lm(NonN_Astro_FGF3R ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_endo_cov <- lm(NonN_Endo ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_micro_cov <- lm(NonN_Micro ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_oligo_cov <- lm(NonN_Oligo_MBP ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))
azd_opc_cov <- lm(NonN_OPC ~ dx + age + sex + batch, data = ctp.clr.1e3 %>% filter(study == "ROSMAP"))

```

ROSMAP

ggstatsplot for significant hits

- ASD: none
- SCZ: Inh, Oligo
- AZD: Astro

```

lm_ctp_sig <- ggstatsplot::combine_plots(
  plotgrid.args = list(nrow = 1, ncol = 3),
  plotlist = list(
    # SCZ: Inh

```

```

ggstatsplot::ggcoefstats(scz_inh_cov,
  title = "Inhibitory neurons (clr) ~ SCZ",
  ylab = NULL, stats.label.args = c(nudge_y = 0.3), exclude.intercept = TRUE) + theme(text = element_text(size = 12))
# SCZ: Oligo
ggstatsplot::ggcoefstats(scz_oligo_cov,
  title = "Oligodendrocytes (clr) ~ SCZ",
  ylab = NULL, stats.label.args = c(nudge_y = 0.3), exclude.intercept = TRUE) + theme(text = element_text(size = 12))
# AZD: Astro
ggstatsplot::ggcoefstats(azd_astro_cov,
  title = "Astrocytes (clr) ~ AZD",
  ylab = NULL, stats.label.args = c(nudge_y = 0.3), exclude.intercept = TRUE) + theme(text = element_text(size = 12))

lm_ctp_sig
ggsave(paste("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_sig.png"), lm_ctp_sig)
ggsave(paste("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_sig.svg"), lm_ctp_sig)

```

Plot extracting diagnosis effect

All in one go (multivariate)

```

# Take CTL as reference
ctp.clr.1e3$dx <- relevel(as.factor(ctp.clr.1e3$dx), ref = "CTL")

# No covariates
tidy_df_nocov_asd <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_nocov_scz <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_nocov_azd <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_nocov_all <- rbind(tidy_df_nocov_asd, tidy_df_nocov_scz, tidy_df_nocov_azd)

# With covariates: including age2
tidy_df_cov_asd <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_scz <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_azd <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_all_sex <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_all_age2 <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))

# With covariates: excluding age2
tidy_df_cov_asd_noage2 <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_scz_noage2 <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))
tidy_df_cov_azd_noage2 <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo, Oxy) ~ 1, data = ctp.clr.1e3))

```

```

tidy_df_cov_all_age <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo) ~ 1, data = tidy_df_cov_all))
tidy_df_cov_all <- rbind(tidy_df_cov_asd, tidy_df_cov_scz, tidy_df_cov_azd)
tidy_df_cov_all_noage2 <- rbind(tidy_df_cov_asd_noage2, tidy_df_cov_scz_noage2, tidy_df_cov_azd_noage2)
tidy_df_cov_all_mega <- rbind(tidy_df_cov_all_sex, tidy_df_cov_all_age, tidy_df_cov_all_age2)

```

Coefficients: covariates

```
datatable(tidy_df_cov_all)
```

```
datatable(tidy_df_cov_all_noage2)
```

```
datatable(tidy_df_cov_all_mega)
```

Coefficients: no covariates

```
datatable(tidy_df_nocov_all)
```

Extract diagnosis effect

Adjusted for covariates

```

tidy_df_cov_all_dx <- tidy_df_cov_all %>% filter(grepl("dx", term))

# Add label
tidy_df_cov_all_dx$lab <- ifelse(tidy_df_cov_all_dx$p.value < 0.06, paste(
  "b=", ifelse(abs(tidy_df_cov_all_dx$estimate) < 0.001, formatC(tidy_df_cov_all_dx$estimate, format = "f", digits = 3),
    "p=", ifelse(tidy_df_cov_all_dx$p.value < 0.001, formatC(tidy_df_cov_all_dx$p.value, format = "e", digits = 3),
      "df=", tidy_df_cov_all_dx$df.error, sep = ""), NA)

# Change cell-type names
tidy_df_cov_all_dx$celltype <- tidy_df_cov_all_dx$response
#tidy_df_cov_all_dx$celltype <- gsub("_FGF3R", "", tidy_df_cov_all_dx$celltype)
#tidy_df_cov_all_dx$celltype <- gsub("_TYROBP", "", tidy_df_cov_all_dx$celltype)
#tidy_df_cov_all_dx$celltype <- gsub("_MBP", "", tidy_df_cov_all_dx$celltype)

# Change diagnosis name
tidy_df_cov_all_dx$dx <- gsub("dx", "", tidy_df_cov_all_dx$term)
tidy_df_cov_all_dx$variable <- "diagnosis"

# Specific order
tidy_df_cov_all_dx$MatchDx <- match(tidy_df_cov_all_dx$dx, c("CTL", "ASD", "SCZ", "AZD"))
tidy_df_cov_all_dx$MatchCell <- match(tidy_df_cov_all_dx$celltype, unique(tidy_df_cov_all_dx$celltype))
tidy_df_cov_all_dx$MatchTerm <- match(tidy_df_cov_all_dx$term, unique(tidy_df_cov_all_dx$term))
tidy_df_cov_all_dx$MatchStudy <- match(tidy_df_cov_all_dx$study, c("ASDbrain (ASD)", "LIBD (SCZ)", "ROSI (ASD)"))

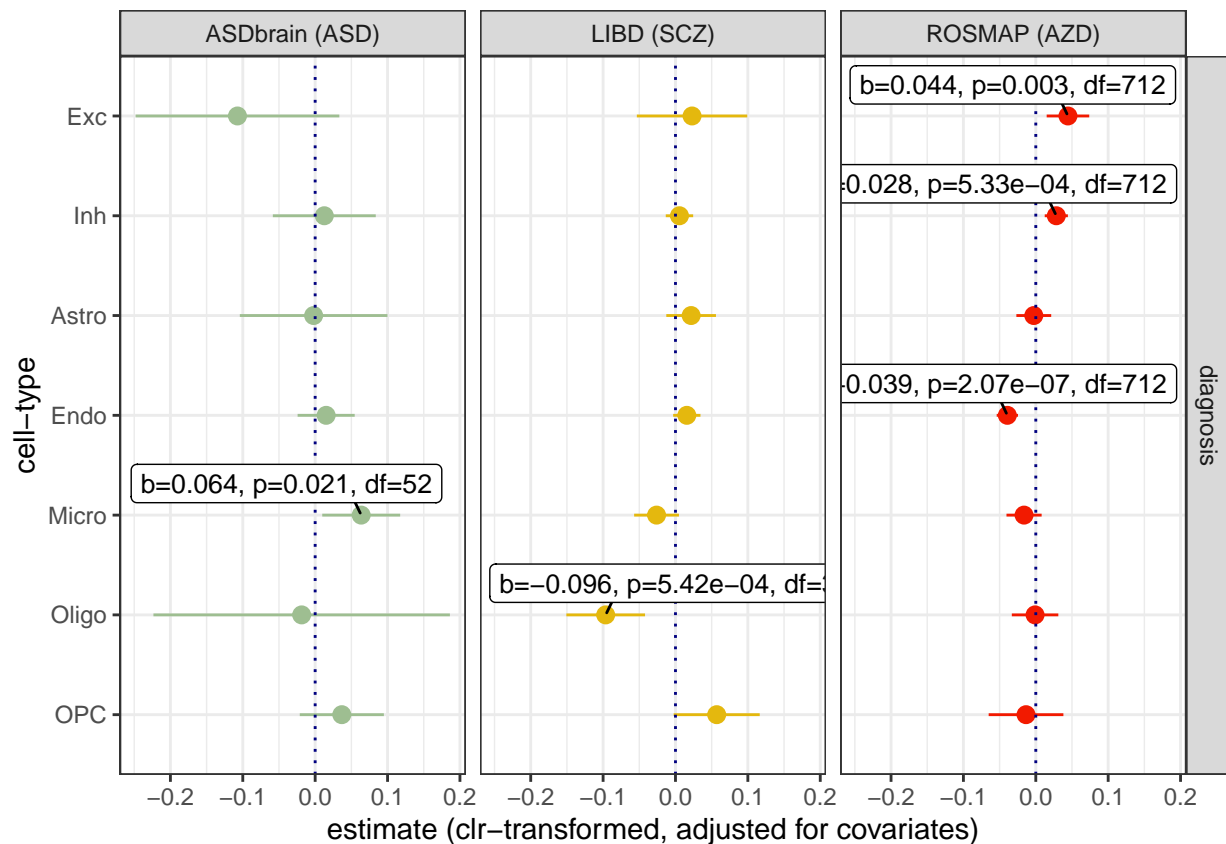
```

```

# Plot
tidy_df_cov_all_dx %>%
  mutate(dx = fct_reorder(dx, MatchDx)) %>%
  mutate(term = fct_reorder(term, MatchTerm)) %>%
  mutate(study = fct_reorder(study, MatchStudy)) %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = term, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
    # geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
    scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4)], name = "term")
    theme_bw() + theme(legend.position = "none") +
    xlab("cell-type") +
    ylab("estimate (clr-transformed, adjusted for covariates)") +
    coord_flip() +
    facet_grid(variable ~ study)

```

Warning: Removed 16 rows containing missing values (geom_label_repel).



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_dx_cov_all.png",
```

Warning: Removed 16 rows containing missing values (geom_label_repel).

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_dx_cov_all.svg",
```

```
## Warning: Removed 16 rows containing missing values (geom_label_repel).
```

Not adjusted for covariates

```
tidy_df_nocov_all_dx <- tidy_df_nocov_all %>% filter(grepl("dx", term))

# Add label
tidy_df_nocov_all_dx$lab <- ifelse(tidy_df_nocov_all_dx$p.value < 0.05, paste(
  "b=", ifelse(abs(tidy_df_nocov_all_dx$estimate) < 0.001, formatC(tidy_df_nocov_all_dx$estimate, format = "e",
  "p=", ifelse(tidy_df_nocov_all_dx$p.value < 0.001, formatC(tidy_df_nocov_all_dx$p.value, format = "e",
  "df=", tidy_df_nocov_all_dx$df.error, sep = ""), NA)

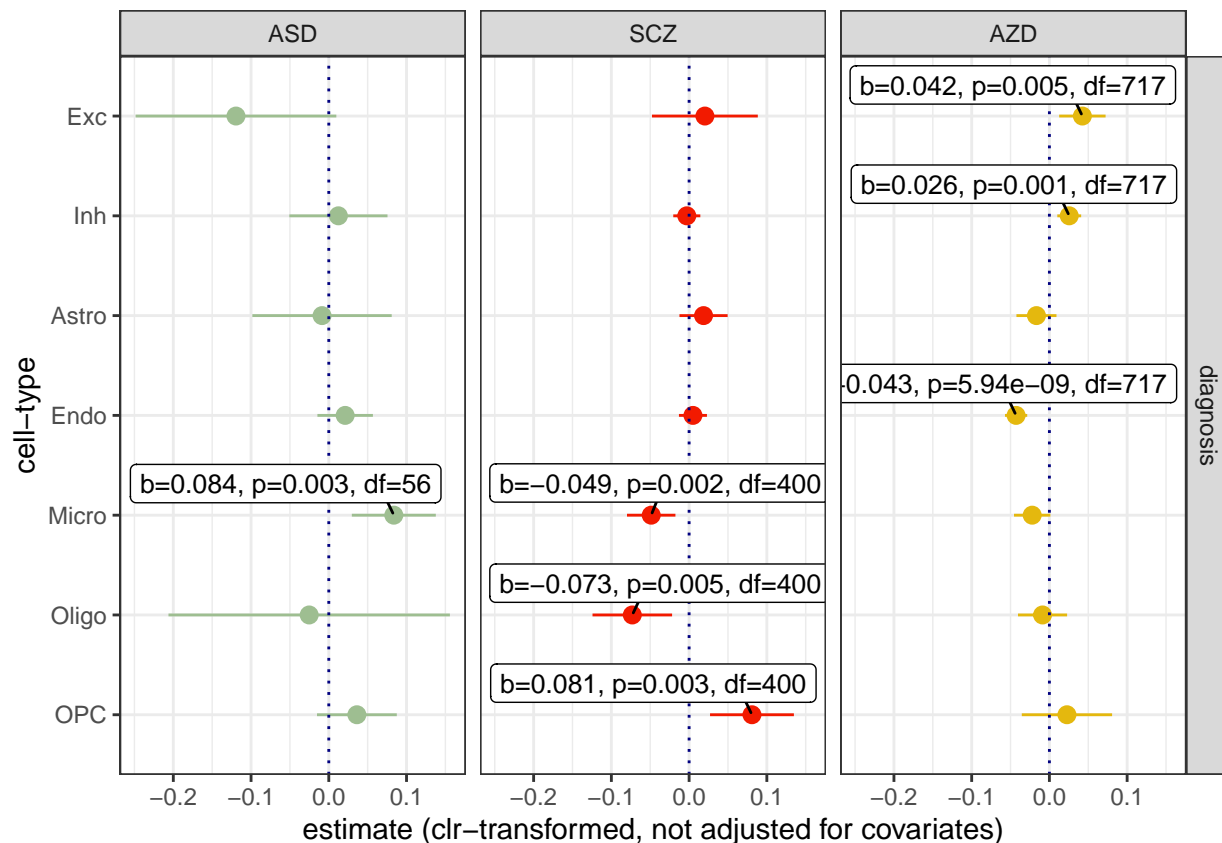
# Change cell-type names
tidy_df_nocov_all_dx$celltype <- tidy_df_nocov_all_dx$response

# Change diagnosis name
tidy_df_nocov_all_dx$dx <- gsub("dx", "", tidy_df_nocov_all_dx$term)
tidy_df_nocov_all_dx$variable <- "diagnosis"

# Specific order
tidy_df_nocov_all_dx$MatchDx <- match(tidy_df_nocov_all_dx$dx, c("CTL", "ASD", "SCZ", "AZD"))
tidy_df_nocov_all_dx$MatchCell <- match(tidy_df_nocov_all_dx$celltype, unique(tidy_df_nocov_all_dx$celltype))

# Plot
tidy_df_nocov_all_dx %>%
  mutate(dx = fct_reorder(dx, MatchDx)) %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = term, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
    # geom_pointrange(aes(ymin = mxy, ymax = miny), position = position_jitter()) +
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
    scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4)], name = "term") +
    theme_bw() + theme(legend.position = "none") +
    xlab("cell-type") +
    ylab("estimate (clr-transformed, not adjusted for covariates)") +
    coord_flip() +
    facet_grid(variable ~ dx)
```

```
## Warning: Removed 14 rows containing missing values (geom_label_repel).
```



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_dx_nocov_all.png")
```

```
## Warning: Removed 14 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_dx_nocov_all.svg")
```

```
## Warning: Removed 14 rows containing missing values (geom_label_repel).
```

Repeat for Jill's analysis

```
load("~/Documents/Research/brain_CTP/Data/methylation/ASD_methylation_brain/analysis/ctp_pfc_tc.clr.lm.l")

# Flip direction of effect around and filter for region
tidy_df_cov_asd_pfc <- tidy_df_cov_asd_pfc %>% mutate(term = ifelse(term == "ASDCTL", "groupASD", term))
tidy_df_cov_asd_tc <- tidy_df_cov_asd_tc %>% mutate(term = ifelse(term == "ASDCTL", "groupASD", term))

# Filter for region
tidy_df_cov_asd_pfc_dx <- tidy_df_cov_asd_pfc %>% filter(term == "groupASD")
tidy_df_cov_asd_tc_dx <- tidy_df_cov_asd_tc %>% filter(term == "groupASD")

# clean
tidy_df_cov_asd_pfc_dx$celltype <- tidy_df_cov_asd_pfc_dx$response
tidy_df_cov_asd_tc_dx$celltype <- tidy_df_cov_asd_tc_dx$response
```



```

tidy_df_cov_asd_pfc_dx$Region <- "BA9"
tidy_df_cov_asd_tc_dx$Region <- "BA41-42-22"
tidy_df_cov_asd_pfc_tc_dx <- rbind(tidy_df_cov_asd_pfc_dx, tidy_df_cov_asd_tc_dx)

# add label
tidy_df_cov_asd_pfc_dx$lab <- ifelse(tidy_df_cov_asd_pfc_dx$p.value < 0.05, paste(
  "b=", ifelse(tidy_df_cov_asd_pfc_dx$estimate < 0.001, formatC(tidy_df_cov_asd_pfc_dx$estimate, format =
    "p=", ifelse(tidy_df_cov_asd_pfc_dx$p.value < 0.001, formatC(tidy_df_cov_asd_pfc_dx$p.value, format =
      "df=", tidy_df_cov_asd_pfc_dx$df.error, sep = "")), NA)

# add match
tidy_df_cov_asd_pfc_tc_dx$MatchCell <- match(tidy_df_cov_asd_pfc_tc_dx$celltype, unique(tidy_df_cov_asd_pfc_tc_dx$celltype))
tidy_df_cov_asd_pfc_tc_dx$MatchRegion <- match(tidy_df_cov_asd_pfc_tc_dx$Region, unique(tidy_df_cov_asd_pfc_tc_dx$Region))

tidy_df_cov_asd_pfc_tc_dx %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  mutate(Region = fct_reorder(Region, MatchRegion)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = Region, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
    scale_colour_manual(values = wes_palette(2, name = "Zissou1", type = "continuous"), name = "") +
    theme_bw() + theme(legend.position = "none") +
    xlab("cell-type") +
    ylab("estimate (clr-transformed, adjusted for covariates)") +
    coord_flip() +
    facet_wrap(~ Region)

ggsave("~/Documents/Research/brain_CTP/Write-up/Progressive/Jill_Pancortical/ASD_brain_PFC_TC_lm_ctp_clr.png")

# Write output table
out_jill_pfc <- data.frame(celltype = tidy_df_cov_asd_pfc_dx$celltype, Region = tidy_df_cov_asd_pfc_dx$Region, estimate = tidy_df_cov_asd_pfc_dx$estimate, conf.low = tidy_df_cov_asd_pfc_dx$conf.low, conf.high = tidy_df_cov_asd_pfc_dx$conf.high)
out_jill_tc <- data.frame(celltype = tidy_df_cov_asd_tc_dx$celltype, Region = tidy_df_cov_asd_tc_dx$Region, estimate = tidy_df_cov_asd_tc_dx$estimate, conf.low = tidy_df_cov_asd_tc_dx$conf.low, conf.high = tidy_df_cov_asd_tc_dx$conf.high)

write.table(out_jill_pfc, "~/Documents/Research/brain_CTP/Write-up/Progressive/Jill_Pancortical/ASD_brain_PFC_TC_lm_ctp_clr_pfc.csv", as.is = TRUE)
write.table(out_jill_tc, "~/Documents/Research/brain_CTP/Write-up/Progressive/Jill_Pancortical/ASD_brain_PFC_TC_lm_ctp_clr_tc.csv", as.is = TRUE)
write.table(tidy_df_cov_asd_pfc, "~/Documents/Research/brain_CTP/Write-up/Progressive/Jill_Pancortical/ASD_brain_PFC_lm_ctp_clr_pfc.csv", as.is = TRUE)
write.table(tidy_df_cov_asd_tc, "~/Documents/Research/brain_CTP/Write-up/Progressive/Jill_Pancortical/ASD_brain_PFC_lm_ctp_clr_tc.csv", as.is = TRUE)

```

Extract sex effect

Adjusted for covariates

```

tidy_df_cov_all_sex <- rbind(
  tidy_df_cov_all %>% filter(grepl("sex", term)),
  tidy_df_cov_all_mega %>% filter(study == "sex_mega-analysis") %>% filter(grepl("sex", term)) %>% mutate(study = "sex_mega-analysis")

# Add label
tidy_df_cov_all_sex$lab <- ifelse(tidy_df_cov_all_sex$p.value < 0.05, paste(
  "b=", ifelse(abs(tidy_df_cov_all_sex$estimate) < 0.001, formatC(tidy_df_cov_all_sex$estimate, format =
    "p=", ifelse(tidy_df_cov_all_sex$p.value < 0.001, formatC(tidy_df_cov_all_sex$p.value, format =
      "df=", tidy_df_cov_all_sex$df.error, sep = "")), NA)

```

```

"p=", ifelse(tidy_df_cov_all_sex$p.value < 0.001, formatC(tidy_df_cov_all_sex$p.value, format = "e",
"df=", tidy_df_cov_all_sex$df.error, sep = ""), NA)

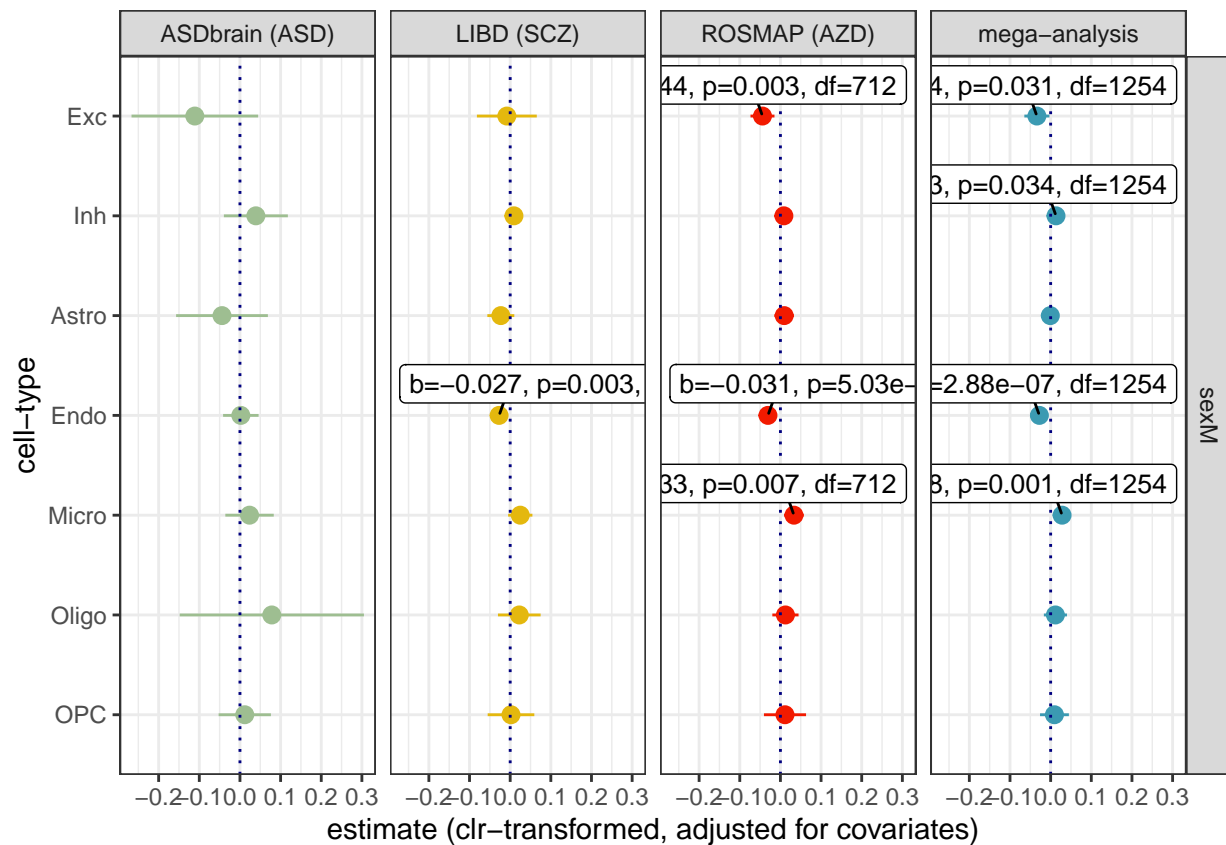
# Change cell-type names
tidy_df_cov_all_sex$celltype <- tidy_df_cov_all_sex$response

# Specific order
tidy_df_cov_all_sex$MatchCell <- match(tidy_df_cov_all_sex$celltype, unique(tidy_df_cov_all_sex$celltype))
tidy_df_cov_all_sex$MatchStudy <- match(tidy_df_cov_all_sex$study, unique(tidy_df_cov_all_sex$study))

# Plot
tidy_df_cov_all_sex %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  mutate(study = fct_reorder(study, MatchStudy)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
  #   geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
  geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
  geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
  scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4,1)], name = "Study") +
  theme_bw() + theme(legend.position = "none") +
  xlab("cell-type") +
  ylab("estimate (clr-transformed, adjusted for covariates)") +
  coord_flip() +
  facet_grid(term ~ study)

```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_cov_all.png")
```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_cov_all.svg")
```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```

Extract age effect

Adjusted for covariates

```
tidy_df_cov_all_age <- rbind(
  tidy_df_cov_all_noage2 %>% filter(grepl("age", term)),
  tidy_df_cov_all_mega %>% filter(study == "age_mega-analysis") %>% filter(grepl("age", term)))

# Add label
tidy_df_cov_all_age$lab <- ifelse(tidy_df_cov_all_age$p.value < 0.05, paste(
  "b=", ifelse(abs(tidy_df_cov_all_age$estimate) < 0.01, formatC(tidy_df_cov_all_age$estimate, format =
  "p=", ifelse(tidy_df_cov_all_age$p.value < 0.01, formatC(tidy_df_cov_all_age$p.value, format = "e", d
  "df=", tidy_df_cov_all_age$df.error, sep = "")), NA)

# Change cell-type names
```

```

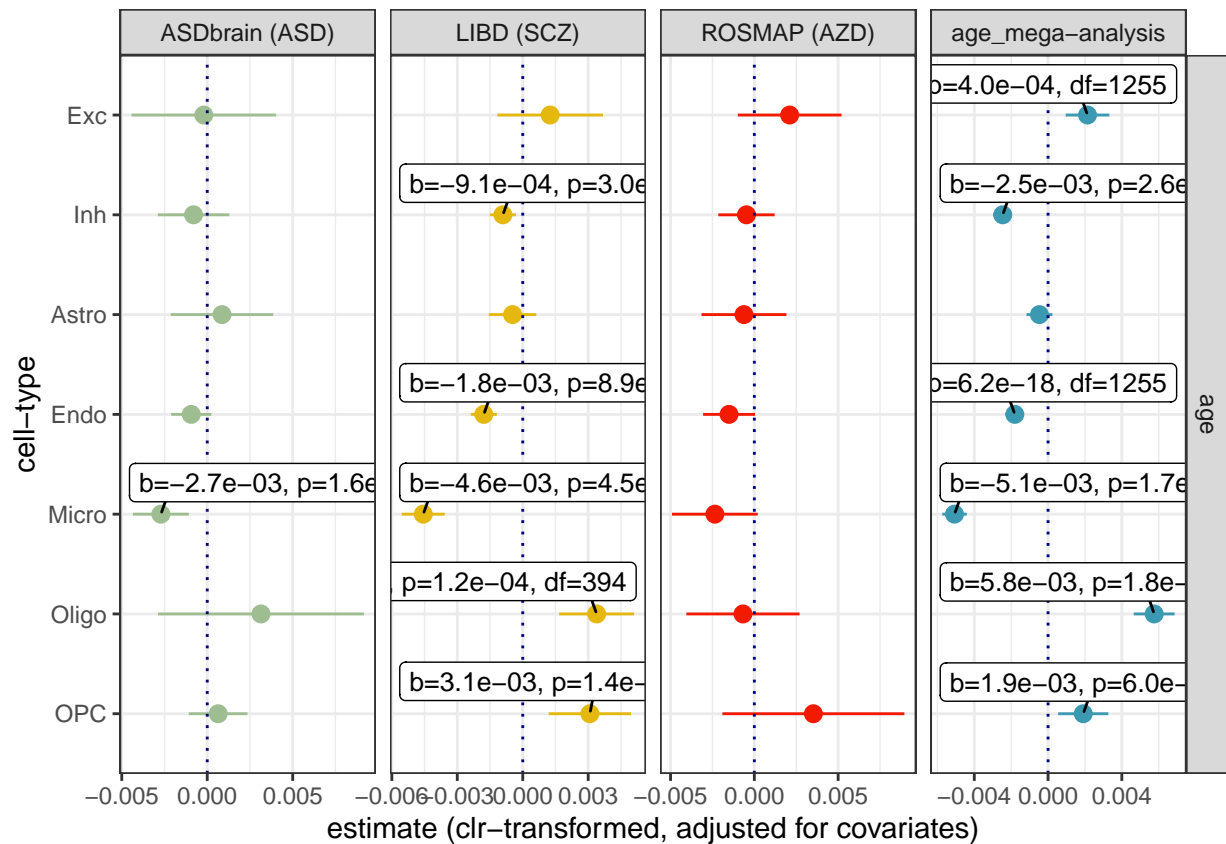
tidy_df_cov_all_age$celltype <- tidy_df_cov_all_age$response

# Specific order
tidy_df_cov_all_age$MatchCell <- match(tidy_df_cov_all_age$celltype, unique(tidy_df_cov_all_age$celltype))
tidy_df_cov_all_age$MatchStudy <- match(tidy_df_cov_all_age$study, unique(tidy_df_cov_all_age$study))

# Plot
tidy_df_cov_all_age %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  mutate(study = fct_reorder(study, MatchStudy)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
    # geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
    scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4,1)], name = "study") +
    theme_bw() + theme(legend.position = "none") +
    xlab("cell-type") +
    ylab("estimate (clr-transformed, adjusted for covariates)") +
    coord_flip() +
    facet_grid(term ~ study, scales = "free_x")

```

Warning: Removed 16 rows containing missing values (geom_label_repel).



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age_cov_all.png")
```

```
## Warning: Removed 16 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age_cov_all.svg")
```

```
## Warning: Removed 16 rows containing missing values (geom_label_repel).
```

Extract age² effect (this model doesn't also include age)

Adjusted for covariates

```
tidy_df_cov_all_age2 <- rbind(
  tidy_df_cov_all %>% filter(grepl("age2", term)),
  tidy_df_cov_all_mega %>% filter(study == "age2_mega-analysis") %>% filter(grepl("age2", term)))

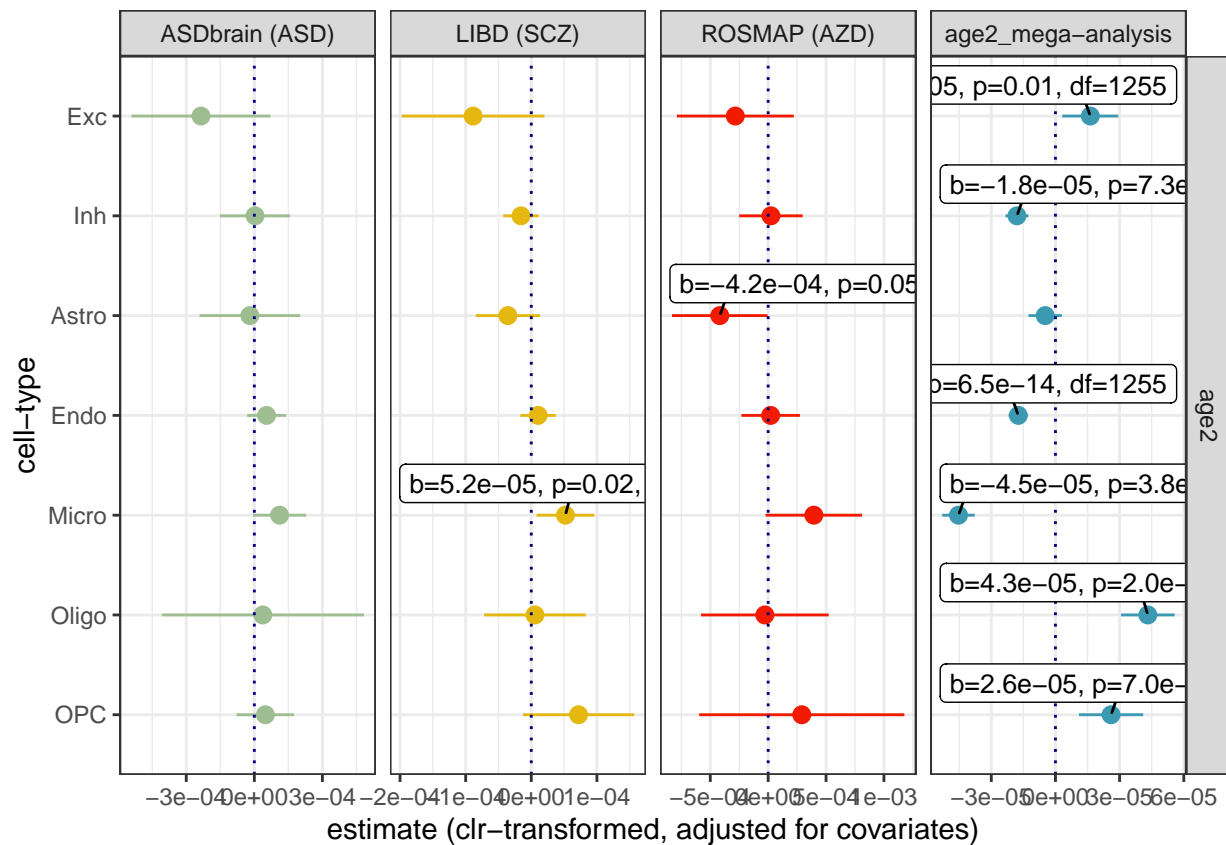
# Add label
tidy_df_cov_all_age2$lab <- ifelse(tidy_df_cov_all_age2$p.value < 0.05, paste(
  "b=", ifelse(tidy_df_cov_all_age2$estimate < 0.01, formatC(tidy_df_cov_all_age2$estimate, format = "e"),
  "p=", ifelse(tidy_df_cov_all_age2$p.value < 0.01, formatC(tidy_df_cov_all_age2$p.value, format = "e"),
  "df=", tidy_df_cov_all_age2$df.error, sep = ""), NA)

# Change cell-type names
tidy_df_cov_all_age2$celltype <- gsub("NonN_", "", tidy_df_cov_all_age2$response)
tidy_df_cov_all_age2$celltype <- gsub("_FGF3R", "", tidy_df_cov_all_age2$celltype)
tidy_df_cov_all_age2$celltype <- gsub("_TYROBP", "", tidy_df_cov_all_age2$celltype)
tidy_df_cov_all_age2$celltype <- gsub("_MBP", "", tidy_df_cov_all_age2$celltype)

# Specific order
tidy_df_cov_all_age2$MatchCell <- match(tidy_df_cov_all_age2$celltype, unique(tidy_df_cov_all_age2$celltype))
tidy_df_cov_all_age2$MatchStudy <- match(tidy_df_cov_all_age2$study, unique(tidy_df_cov_all_age2$study))

# Plot
tidy_df_cov_all_age2 %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  mutate(study = fct_reorder(study, MatchStudy)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +
  geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
  # geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
  geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
  geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
  scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4,1)], name = "Study") +
  theme_bw() + theme(legend.position = "none") +
  xlab("cell-type") +
  ylab("estimate (clr-transformed, adjusted for covariates)") +
  coord_flip() +
  facet_grid(term ~ study, scales = "free_x")
```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age2_cov_all.png")
```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age2_cov_all.svg")
```

```
## Warning: Removed 20 rows containing missing values (geom_label_repel).
```

Extract mega-analysis results only

```
tidy_df_cov_mega_sex_age <- rbind(
  tidy_df_cov_all_mega %>% filter(study == "age_mega-analysis") %>% filter(grepl("age", term)),
  tidy_df_cov_all_mega %>% filter(study == "sex_mega-analysis") %>% filter(grepl("sex", term))) %>% mutate(
    # Add label
    tidy_df_cov_mega_sex_age$lab <- ifelse(tidy_df_cov_mega_sex_age$p.value < 0.05, paste(
      "b=", ifelse(tidy_df_cov_mega_sex_age$estimate < 0.01, formatC(tidy_df_cov_mega_sex_age$estimate, format="e", digits=2),
        "p=", ifelse(tidy_df_cov_mega_sex_age$p.value < 0.01, formatC(tidy_df_cov_mega_sex_age$p.value, format="e", digits=2),
          "df=", tidy_df_cov_mega_sex_age$df.error, sep = "")), NA)

    # Change cell-type names
    tidy_df_cov_mega_sex_age$celltype <- tidy_df_cov_mega_sex_age$response
```

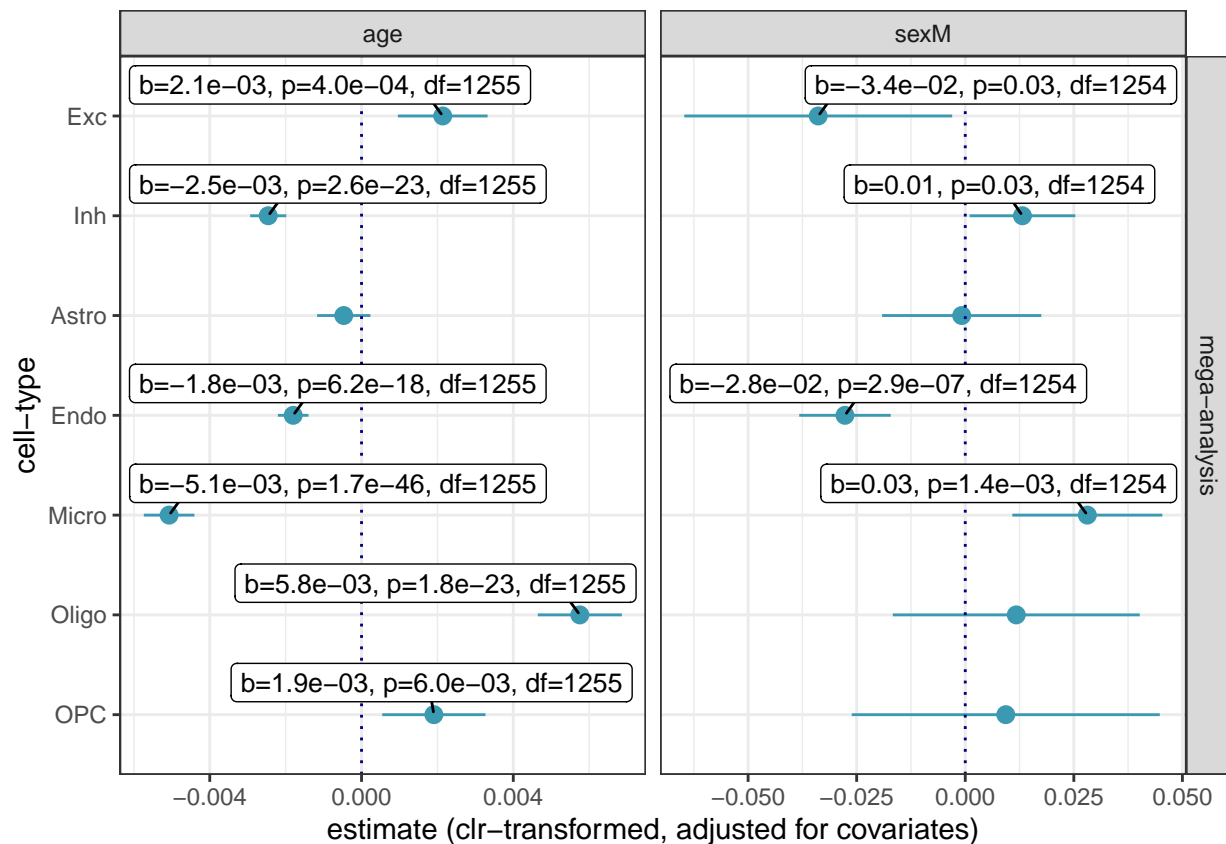
```

# Specific order
tidy_df_cov_mega_sex_age$MatchCell <- match(tidy_df_cov_mega_sex_age$celltype, unique(tidy_df_cov_mega_sex_age$celltype))

# Plot
tidy_df_cov_mega_sex_age %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +
  geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
  # geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
  geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
  geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
  scale_colour_manual(values = wes_palette(6, name = "Zissou1", type = "continuous")[c(1,5,6)], name = "study") +
  theme_bw() + theme(legend.position = "none") +
  xlab("cell-type") +
  ylab("estimate (clr-transformed, adjusted for covariates)") +
  coord_flip() +
  facet_grid(study ~ term, scales = "free_x")

```

```
## Warning: Removed 4 rows containing missing values (geom_label_repel).
```



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_age_cov_mega_sex_age")
```

```
## Warning: Removed 4 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_age_cov_mega
```

```
## Warning: Removed 4 rows containing missing values (geom_label_repel).
```

CTL only

- N.B. ASDbrain all on batch1 now (after excluding some people and taking only CTL)

```
# Take CTL as reference
ctp_clr.1e3$dx <- relevel(as.factor(ctp_clr.1e3$dx), ref = "CTL")

# With covariates
# - ASD: N.B. no batch variable included as all CTL in one batch
tidy_df_cov_asd_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_scz_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_azd_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_all_sex_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_all_age2_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_sex_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,

# With covariates: excluding age2
tidy_df_cov_asd_noage2_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_scz_noage2_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_azd_noage2_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,
tidy_df_cov_all_age_CTL <- parameters::model_parameters(model = lm(cbind(Exc, Inh, Astro, Endo, Micro, Oligo,

tidy_df_cov_all_CTL <- rbind(tidy_df_cov_asd_CTL, tidy_df_cov_scz_CTL, tidy_df_cov_azd_CTL)
tidy_df_cov_all_noage2_CTL <- rbind(tidy_df_cov_asd_noage2_CTL, tidy_df_cov_scz_noage2_CTL, tidy_df_cov_azd_noage2_CTL)
tidy_df_cov_all_mega_CTL <- rbind(tidy_df_cov_all_sex_CTL, tidy_df_cov_all_age_CTL, tidy_df_cov_all_age2_CTL)
```

Age

```
tidy_df_cov_all_age_CTL <- rbind(
  tidy_df_cov_all_noage2_CTL %>% filter(grepl("age", term)),
  tidy_df_cov_all_mega_CTL %>% filter(study == "age_mega-analysis CTL") %>% filter(grepl("age", term))

# Add label
tidy_df_cov_all_age_CTL$lab <- ifelse(tidy_df_cov_all_age_CTL$p.value < 0.05, paste(
  "b=", ifelse(abs(tidy_df_cov_all_age_CTL$estimate) < 0.01, formatC(tidy_df_cov_all_age_CTL$estimate,
  "p=", ifelse(tidy_df_cov_all_age_CTL$p.value < 0.01, formatC(tidy_df_cov_all_age_CTL$p.value, format = "e",
  "df=", tidy_df_cov_all_age_CTL$df.error, sep = ""), NA)
```



```

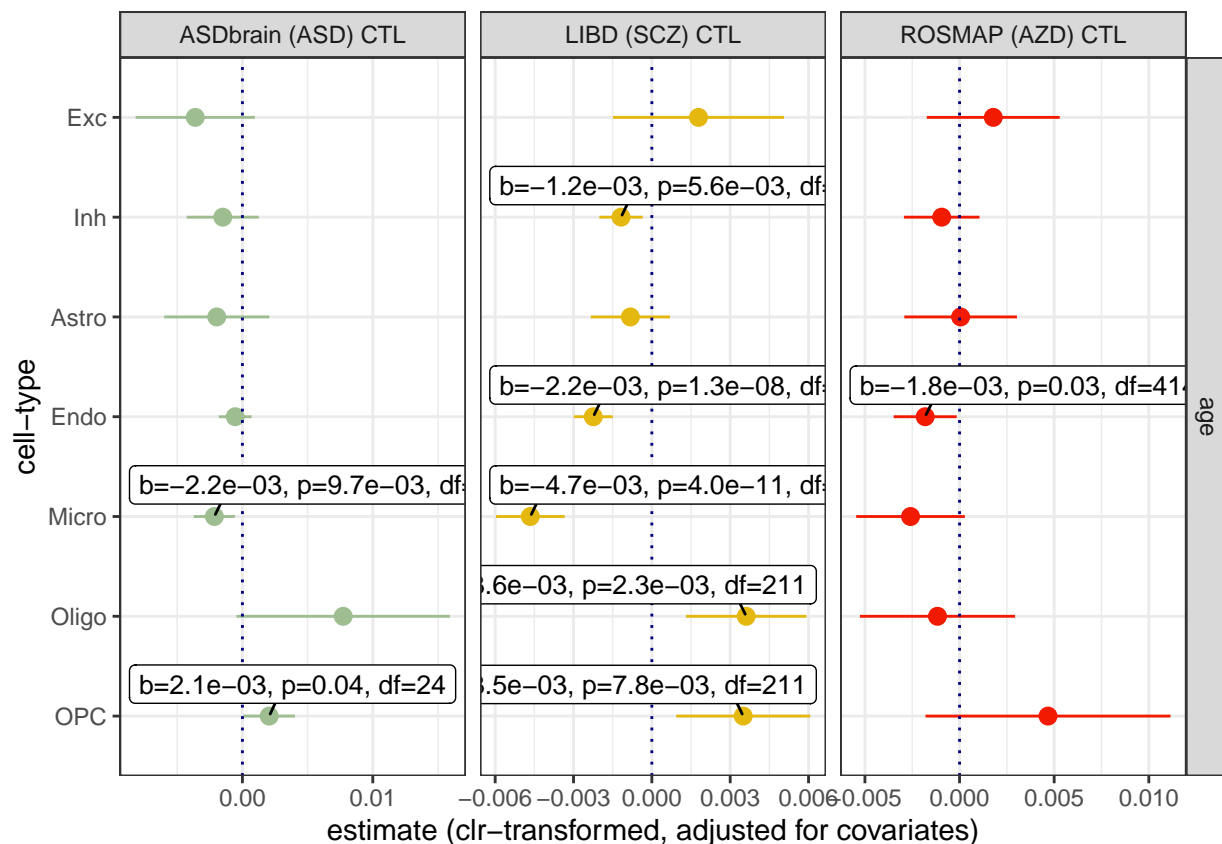
# Change cell-type names
tidy_df_cov_all_age_CTL$celltype <- tidy_df_cov_all_age_CTL$response

# Specific order
tidy_df_cov_all_age_CTL$MatchCell <- match(tidy_df_cov_all_age_CTL$celltype, unique(tidy_df_cov_all_age_CTL$celltype))
tidy_df_cov_all_age_CTL$MatchStudy <- match(tidy_df_cov_all_age_CTL$study, unique(tidy_df_cov_all_age_CTL$study))

# Plot
tidy_df_cov_all_age_CTL %>%
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%
  mutate(study = fct_reorder(study, MatchStudy)) %>%
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +
    # geom_pointrange(aes(ymin = maxy, ymax = miny), position = position_jitter()) +
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +
    scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4,1)], name = "study") +
    theme_bw() + theme(legend.position = "none") +
    xlab("cell-type") +
    ylab("estimate (clr-transformed, adjusted for covariates)") +
    coord_flip() +
    facet_grid(term ~ study, scales = "free_x")

```

Warning: Removed 13 rows containing missing values (geom_label_repel).



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age_cov_all_CTL.
```

```
## Warning: Removed 13 rows containing missing values (geom_label_repel).
```

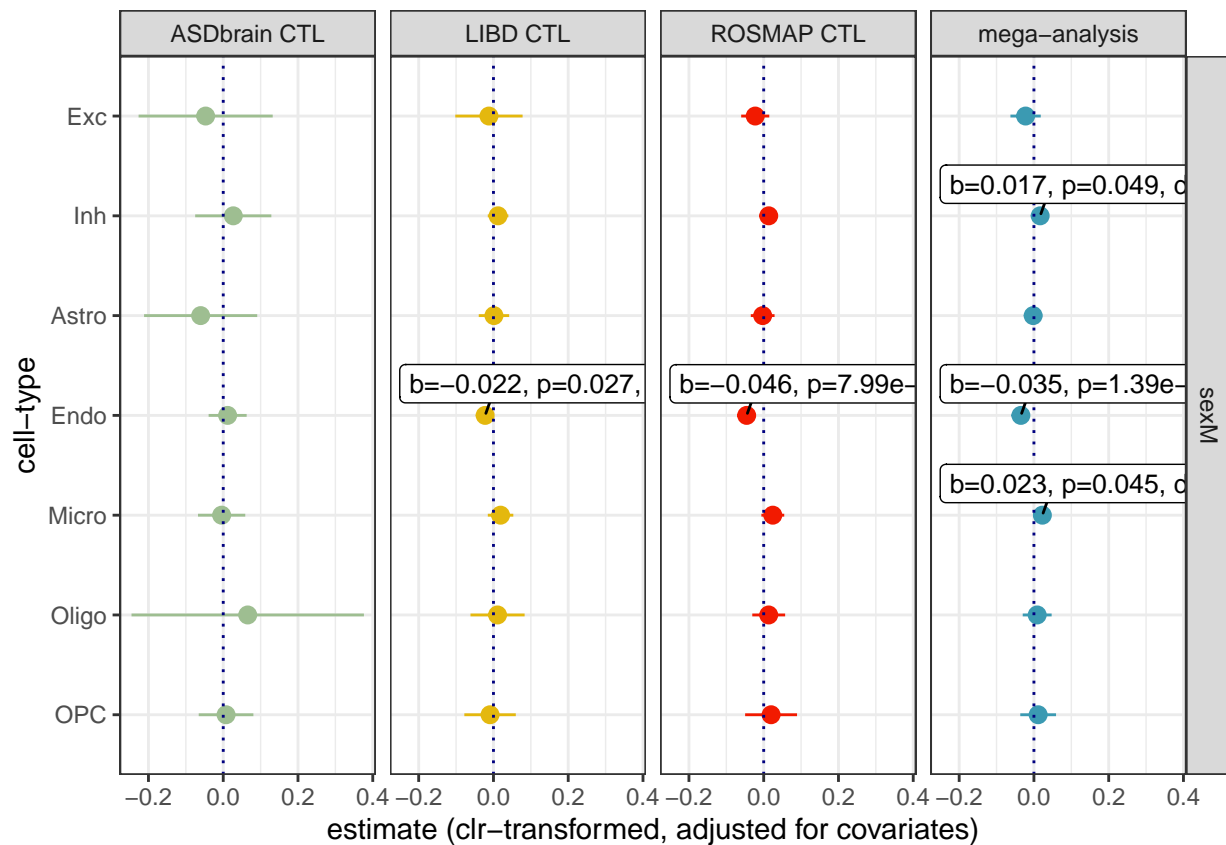
```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_age_cov__all_CTL.
```

```
## Warning: Removed 13 rows containing missing values (geom_label_repel).
```

Sex

```
tidy_df_cov_all_sex_CTL <- rbind(  
  tidy_df_cov_all_CTL %>% filter(grepl("sex", term)),  
  tidy_df_cov_all_mega_CTL %>% filter(study == "sex_mega-analysis CTL") %>% filter(grepl("sex", term)) )  
  
# Add label  
tidy_df_cov_all_sex_CTL$lab <- ifelse(tidy_df_cov_all_sex_CTL$p.value < 0.05, paste(  
  "b=", ifelse(abs(tidy_df_cov_all_sex_CTL$estimate) < 0.001, formatC(tidy_df_cov_all_sex_CTL$estimate,  
  "p=", ifelse(tidy_df_cov_all_sex_CTL$p.value < 0.001, formatC(tidy_df_cov_all_sex_CTL$p.value, format  
  "df=", tidy_df_cov_all_sex_CTL$df.error, sep = ""), NA)  
  
# Change cell-type names  
tidy_df_cov_all_sex_CTL$celltype <- tidy_df_cov_all_sex_CTL$response  
  
# Specific order  
tidy_df_cov_all_sex_CTL$MatchCell <- match(tidy_df_cov_all_sex_CTL$celltype, unique(tidy_df_cov_all_sex_CTL$celltype))  
tidy_df_cov_all_sex_CTL$MatchStudy <- match(tidy_df_cov_all_sex_CTL$study, unique(tidy_df_cov_all_sex_CTL$study))  
  
# Plot  
tidy_df_cov_all_sex_CTL %>%  
  mutate(celltype = fct_reorder(celltype, MatchCell)) %>%  
  mutate(study = fct_reorder(study, MatchStudy)) %>%  
  ggplot(aes(x = fct_rev(celltype), y = estimate, colour = study, label = lab)) +  
    geom_pointrange(aes(ymin = conf.low, ymax = conf.high)) +  
  #   geom_pointrange(aes(ymin = mxy, ymax = miny), position = position_jitter()) +  
    geom_hline(aes(yintercept = 0), linetype = "dotted", colour = "navy") +  
    geom_label_repel(size = 3.5, nudge_x = 0.15, force = 8, colour = "black", min.segment.length = 0) +  
    scale_colour_manual(values = wes_palette(4, name = "Zissou1", type = "continuous")[c(2,3,4,1)], name = "Study") +  
    theme_bw() + theme(legend.position = "none") +  
    xlab("cell-type") +  
    ylab("estimate (clr-transformed, adjusted for covariates)") +  
    coord_flip() +  
    facet_grid(term ~ study)
```

```
## Warning: Removed 23 rows containing missing values (geom_label_repel).
```



```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_cov_all_CTL..")
```

```
## Warning: Removed 23 rows containing missing values (geom_label_repel).
```

```
ggsave("~/Documents/Research/brain_CTP/Data/integration/ctp_differences/lm_ctp_clr1e-3_sex_cov_all_CTL..")
```

```
## Warning: Removed 23 rows containing missing values (geom_label_repel).
```