

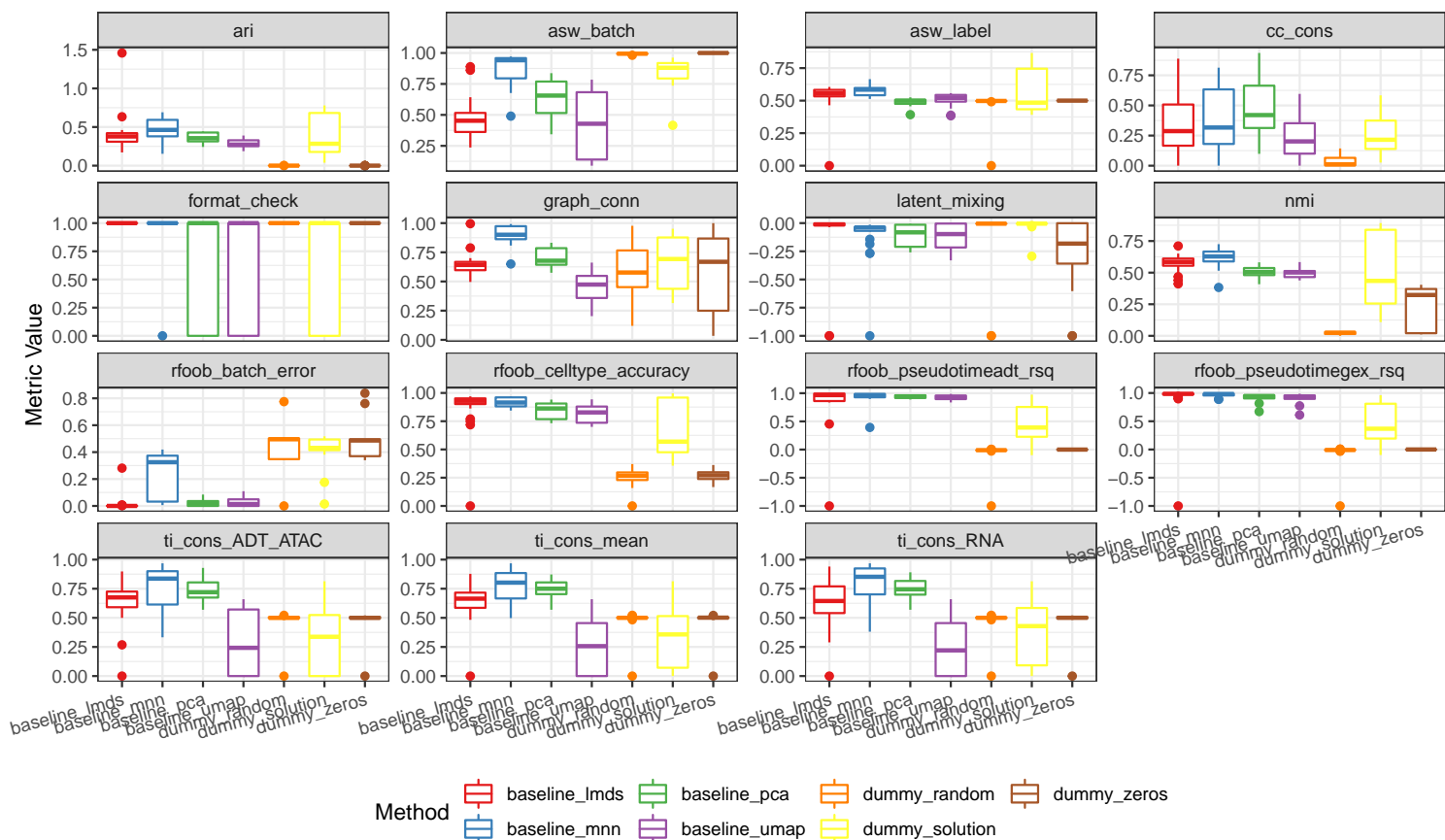
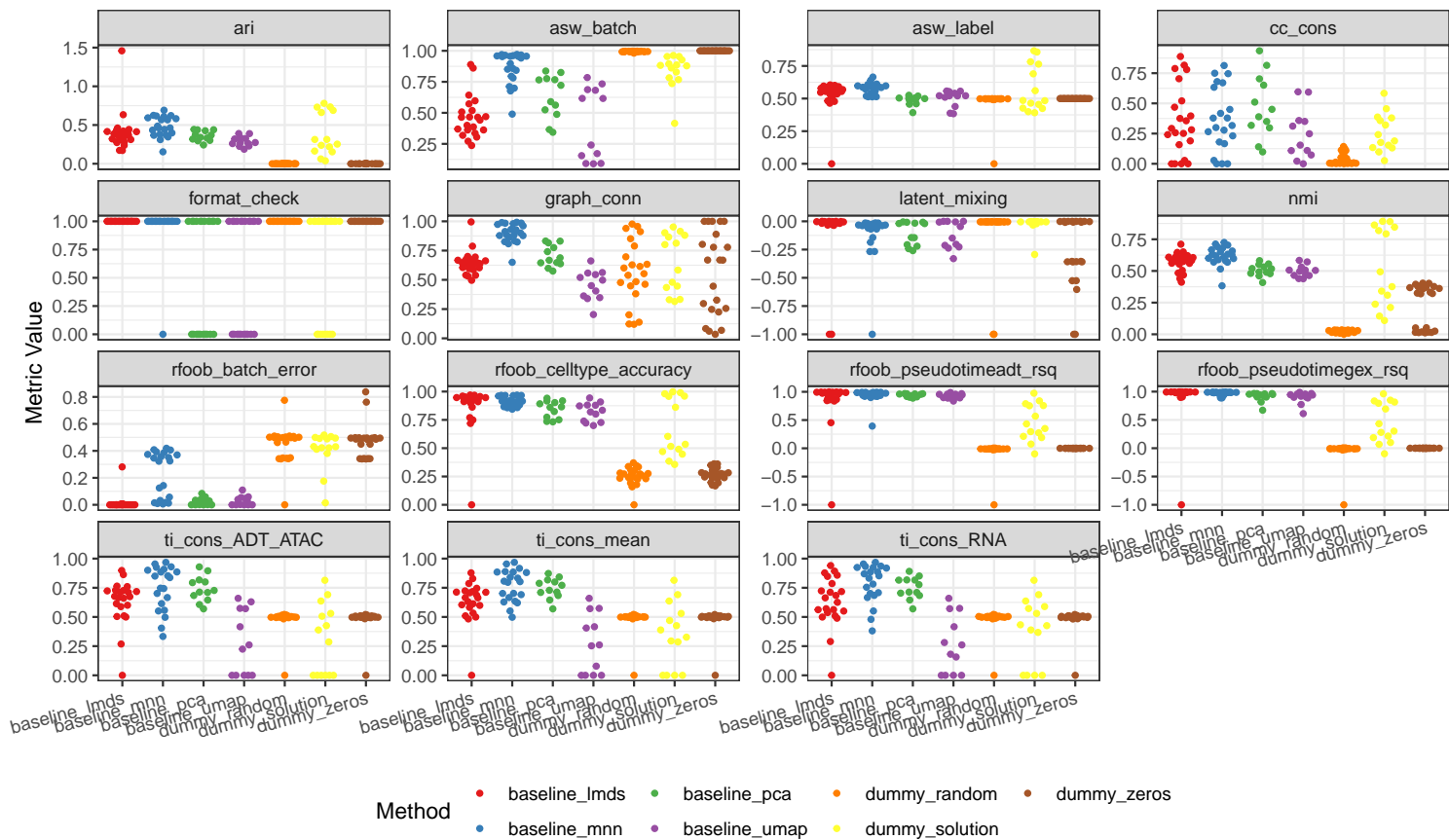
## Joint Embedding - Pilot Analysis

## Visualise results

Colour by method.

## Warning: Removed 22 rows containing missing values (position\_quasirandom).

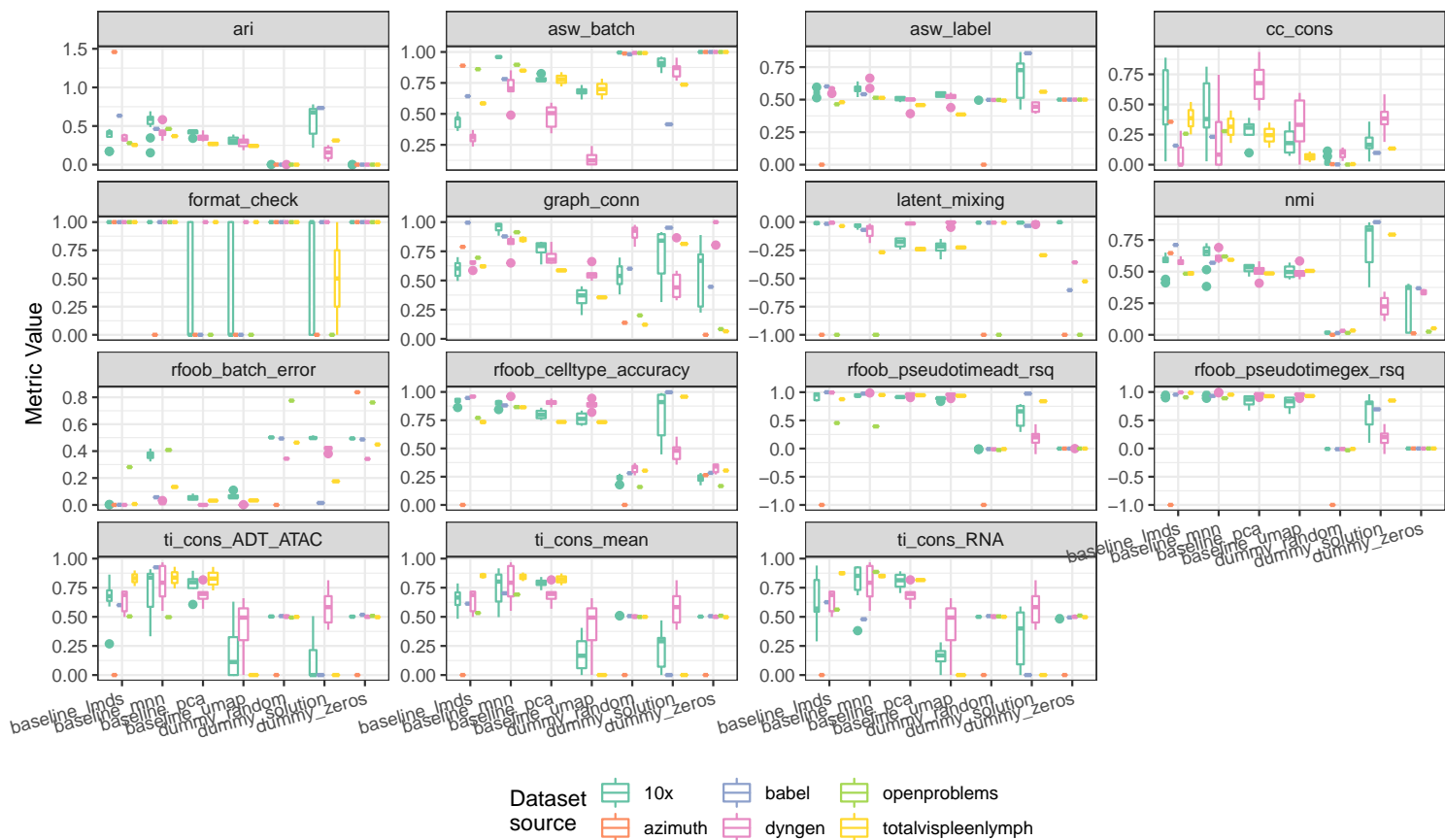
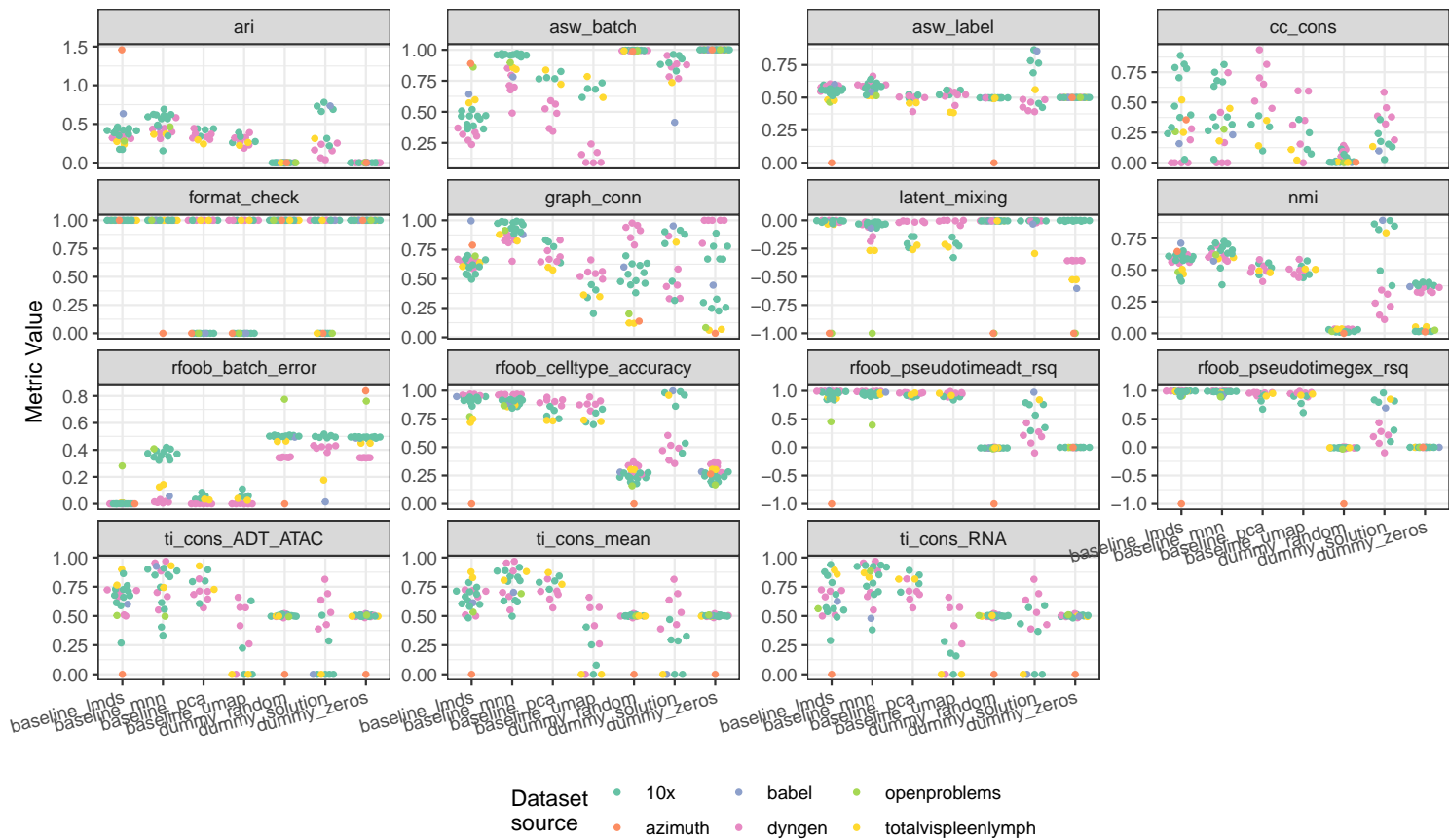
## Warning: Removed 22 rows containing non-finite values (stat\_boxplot).



Colour by dataset source.

## Warning: Removed 22 rows containing missing values (position\_quasirandom).

## Warning: Removed 22 rows containing non-finite values (stat\_boxplot).



Comparing most interesting metrics.

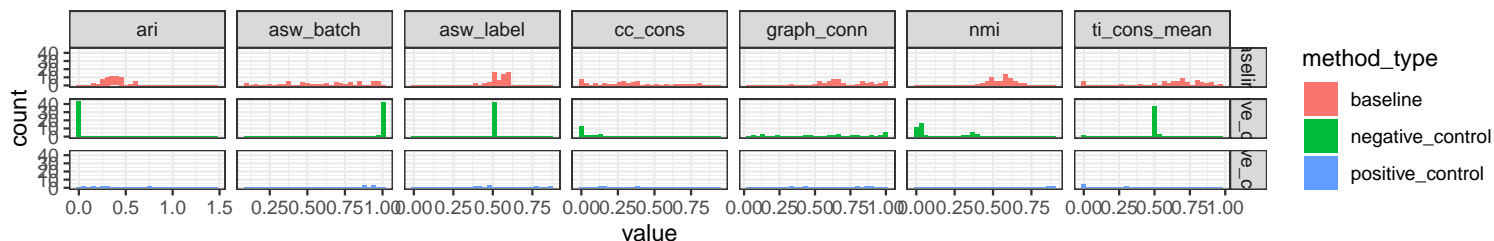
```
# unique(dfg$metric_id)
patchwork::wrap_plots(
  ggplot(dfg %>% filter(metric_id %in% c("ari", "nmi", "asw_label", "ti_cons_mean", "cc_cons", "asw_batch", "graph_conn", "latent_mixing", "rfoob_batch_error", "rfoob_celltype_accuracy", "rfoob_pseudotimeadt_rsq", "rfoob_pseudotimegex_rsq")) +
    geom_histogram(aes(value, fill = method_type)) +
    facet_grid(method_type~metric_id, scales = "free_x") +
    labs(title = "Metrics included in paper") +
    theme_bw(),
  ggplot(dfg %>% filter(!metric_id %in% c("format_check", "ari", "nmi", "asw_label", "ti_cons_mean", "cc_cons", "asw_batch", "graph_conn", "latent_mixing", "rfoob_batch_error", "rfoob_celltype_accuracy", "rfoob_pseudotimeadt_rsq", "rfoob_pseudotimegex_rsq")) +
    geom_histogram(aes(value, fill = method_type)) +
    facet_grid(method_type~metric_id, scales = "free_x") +
    labs(title = "Metrics not included in paper") +
    theme_bw(),
  ncol = 1
)
```

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

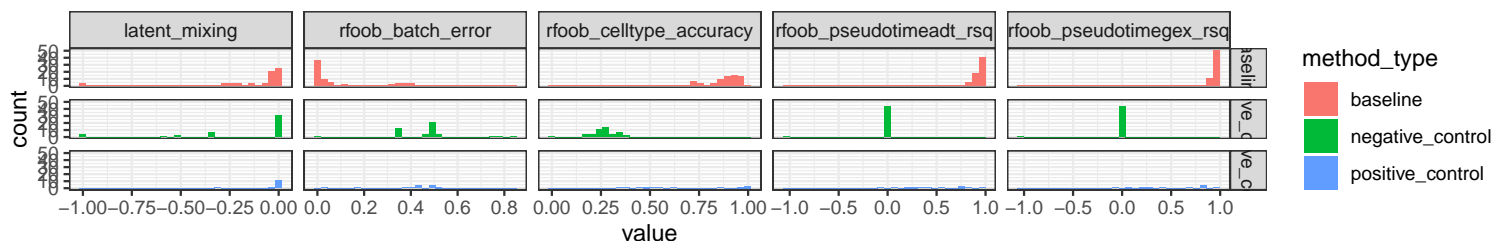
## Warning: Removed 22 rows containing non-finite values (stat\_bin).

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

Metrics included in paper



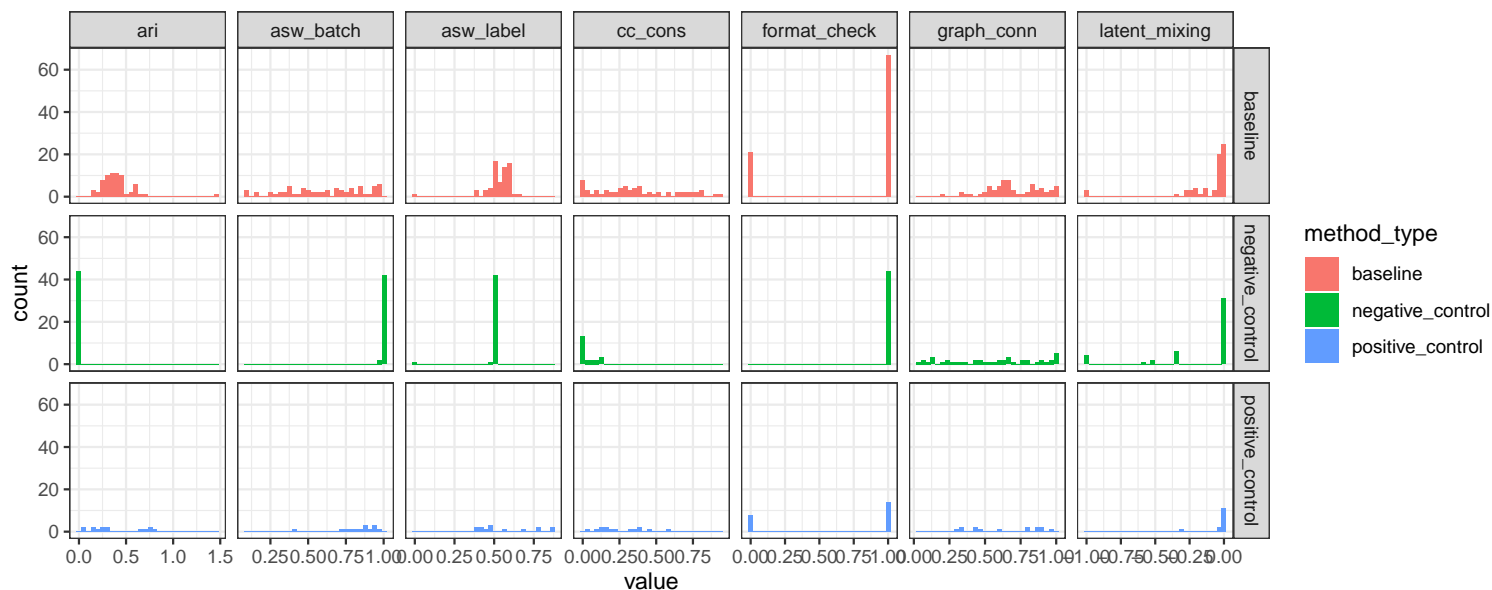
Metrics not included in paper



```
ggplot(dfg %>% filter(metric_id <= "latent_mixing")) +
  geom_histogram(aes(value, fill = method_type)) +
  facet_grid(method_type~metric_id, scales = "free_x") +
  theme_bw()
```

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

## Warning: Removed 22 rows containing non-finite values (stat\_bin).



```
ggplot(dfg %>% filter(metric_id > "latent_mixing")) +
  geom_histogram(aes(value, fill = method_type)) +
  facet_grid(method_type~metric_id, scales = "free_x") +
  theme_bw()
```

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

