

# Project 4 – Trabeculectomy Outcomes: IOP and VA

## 0. Overview

This report describes trabeculectomy outcomes in a small cohort of glaucoma patients. We summarise intraocular pressure (IOP) and visual acuity (VA) at key timepoints, define surgical success based on IOP thresholds and percentage reduction from baseline, and describe medication burden at last follow up.

## 1. Load cleaned data

```
data <- readr::read_csv("data/trab_clean.csv")
glimpse(data)

## # Rows: 33
## # Columns: 18
## $ Patient_ID      <chr> "PX001", "PX002", "PX003", "PX004", "PX005", "PX0~  
## $ Sex            <chr> "F", "F", "M", "f", "F", "M", "M", "M", "F", ~  
## $ Age            <dbl> 35, 80, 45, 75, 62, 48, 56, 56, 54, 58, 62, 6~  
## $ Eye             <chr> "RE", "RE", "RE", "LE", "RE", "RE", "RE", "LE", "~  
## $ surgery_clean   <chr> "RICCE + Trab", "RICCE + Trab", "Trab + SICS", "T~  
## $ surgery_date    <date> 2022-08-02, 2022-09-13, 2022-11-29, 2022-12-13, ~  
## $ first_visit_date <date> NA, NA, 2022-11-28, 2016-09-28, 2021-09-30, NA, ~  
## $ last_visit_date  <date> NA, NA, NA, 2024-07-02, 2025-07-30, NA, NA, 2023~  
## $ 'Pre-Op medication' <chr> NA, NA, "IV Mannitol, Diamox", "misopt", "misopt, ~  
## $ 'Current medication' <chr> NA, NA, NA, "hypromellose", "brimonidine and dorz~  
## $ va_preop_logmar   <dbl> NA, NA, 2.00, 1.78, 1.00, 2.00, NA, NA, 0.60, 0.3~  
## $ va_1dpo_logmar    <dbl> NA, NA, 2.00, 0.18, 0.78, 1.00, NA, NA, 0.30, 1.7~  
## $ va_6mo_logmar     <dbl> NA, NA, NA, 1.78, 0.78, NA, 1.08, 0.18, 0.30, 0.4~  
## $ va_1yr_logmar     <dbl> NA, NA, NA, 1.78, 1.18, NA, NA, NA, 0.78, 0.1~  
## $ iop_preop          <dbl> NA, NA, 30.0, 11.0, 32.0, NA, NA, NA, 12.0, 36.0, ~  
## $ iop_1dpo           <dbl> NA, NA, NA, 2.0, 12.0, NA, NA, NA, 6.0, 17.0, 10.~  
## $ iop_6mo            <dbl> NA, NA, NA, 6.0, 55.0, NA, 10.0, 11.0, 13.7, 10.0~  
## $ iop_1yr            <dbl> NA, NA, NA, 9, 33, NA, NA, NA, 36, 16, NA, NA~
```

## 2. Derive outcome variables

```
data <- data %>%
  mutate(
    iop_drop_6mo = iop_preop - iop_6mo,
    iop_drop_1yr = iop_preop - iop_1yr,
    iop_pct_drop_6mo = (iop_preop - iop_6mo) / iop_preop,
    iop_pct_drop_1yr = (iop_preop - iop_1yr) / iop_preop,
    success_6mo_30pct = if_else(!is.na(iop_preop) & !is.na(iop_6mo) & iop_6mo <= 21 & iop_pct_drop_6mo >
```

```
success_1yr_30pct = if_else(!is.na(iop_preop) & !is.na(iop_1yr) & iop_1yr <= 21 & iop_pct_drop_1yr <= 30, 1, 0)
va_change_6mo = va_6mo_logmar - va_preop_logmar,
va_change_1yr = va_1yr_logmar - va_preop_logmar,
va_status_6mo = case_when(
  is.na(va_change_6mo) ~ NA_character_,
  va_change_6mo <= -0.10 ~ "Improved",
  va_change_6mo >= 0.10 ~ "Worsened",
  TRUE ~ "Stable"
),
va_status_1yr = case_when(
  is.na(va_change_1yr) ~ NA_character_,
  va_change_1yr <= -0.10 ~ "Improved",
  va_change_1yr >= 0.10 ~ "Worsened",
  TRUE ~ "Stable"
)
)
```

### 3. Plots

### 3.1 IOP over time – summary and plot

```

iop_long <- data %>%
  select(Patient_ID, Eye, Age, Sex, iop_preop, iop_1dpo, iop_6mo, iop_1yr) %>%
  pivot_longer(cols = starts_with("iop_"), names_to = "timepoint", values_to = "iop") %>%
  mutate(timepoint = factor(timepoint,
                            levels = c("iop_preop", "iop_1dpo", "iop_6mo", "iop_1yr"),
                            labels = c("Pre-op", "1 DPO", "6/12", "1 year")))

iop_summary <- iop_long %>%
  group_by(timepoint) %>%
  summarise(n = sum(!is.na(iop)),
            mean_iop = mean(iop, na.rm = TRUE),
            med_iop = median(iop, na.rm = TRUE),
            sd_iop = sd(iop, na.rm = TRUE)) %>%
  mutate(se_iop = sd_iop / sqrt(n),
        lower = mean_iop - 1.96 * se_iop,
        upper = mean_iop + 1.96 * se_iop)

iop_summary

## # A tibble: 4 x 8
##   timepoint      n  mean_iop  med_iop  sd_iop  se_iop  lower  upper
##   <fct>    <int>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
## 1 Pre-op       28     24.6    22.5    10.6    2.00    20.7    28.5
## 2 1 DPO        22     10.8     9.0    6.65    1.42    7.97    13.5
## 3 6/12         22     13.8    11.5    10.2    2.18    9.55    18.1
## 4 1 year       9      18.8    15.0    9.39    3.13   12.6    24.9

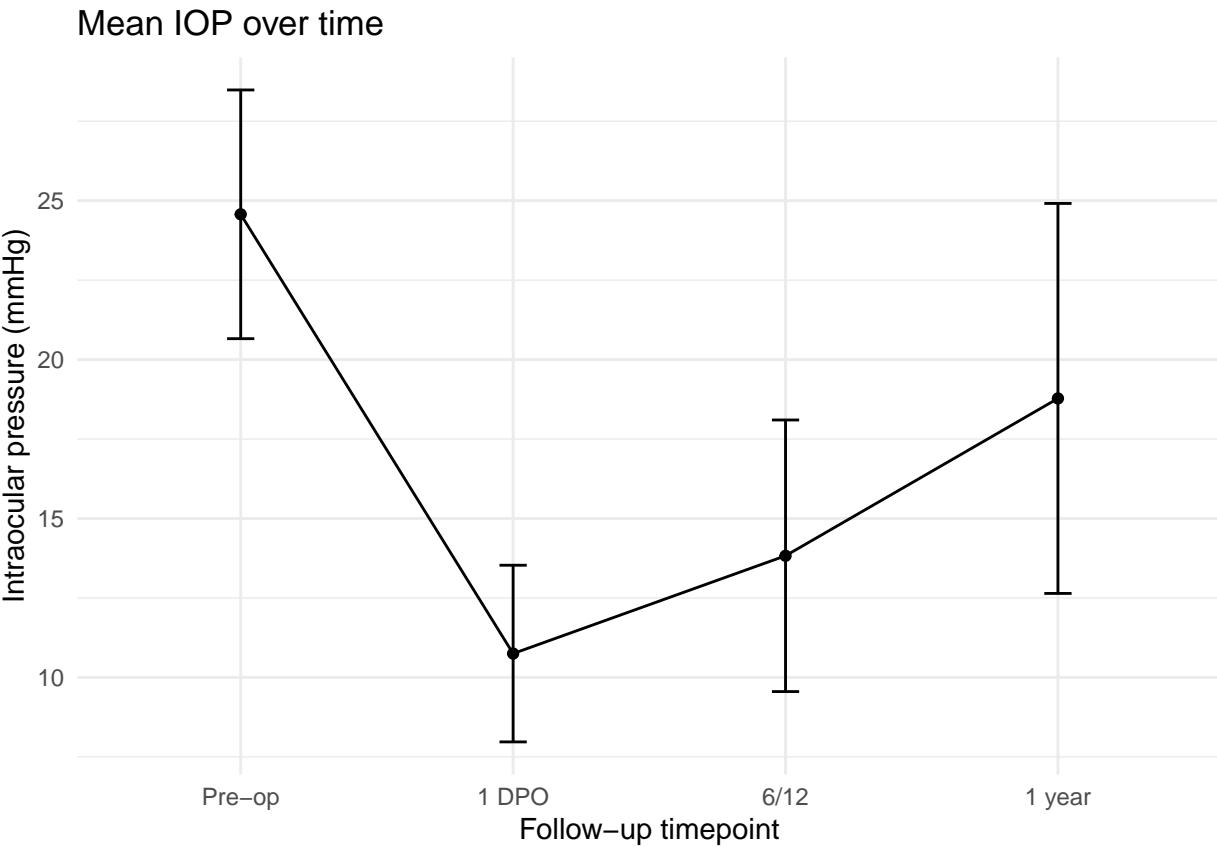
ggplot(iop_summary, aes(x = timepoint, y = mean_iop, group = 1)) +
  geom_line() +
  geom_point() +

```

```

geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.1) +
  labs(title = "Mean IOP over time",
       x = "Follow-up timepoint",
       y = "Intraocular pressure (mmHg)") +
  theme_minimal()

```



### 3.2. VA Over Time (logMAR)

```

#Plot 2: Boxplot of VA logMAR at Each Timepoint
va_long <- data %>%
  select(Patient_ID, va_preop_logmar, va_1dpo_logmar, va_6mo_logmar, va_1yr_logmar) %>%
  pivot_longer(
    cols = starts_with("va_"),
    names_to = "timepoint",
    values_to = "logmar"
  ) %>%
  mutate(timepoint = factor(timepoint,
    levels = c("va_preop_logmar", "va_1dpo_logmar", "va_6mo_logmar", "va_1yr_logmar"),
    labels = c("Pre-op", "1DPO", "6MO", "1YR"))
))

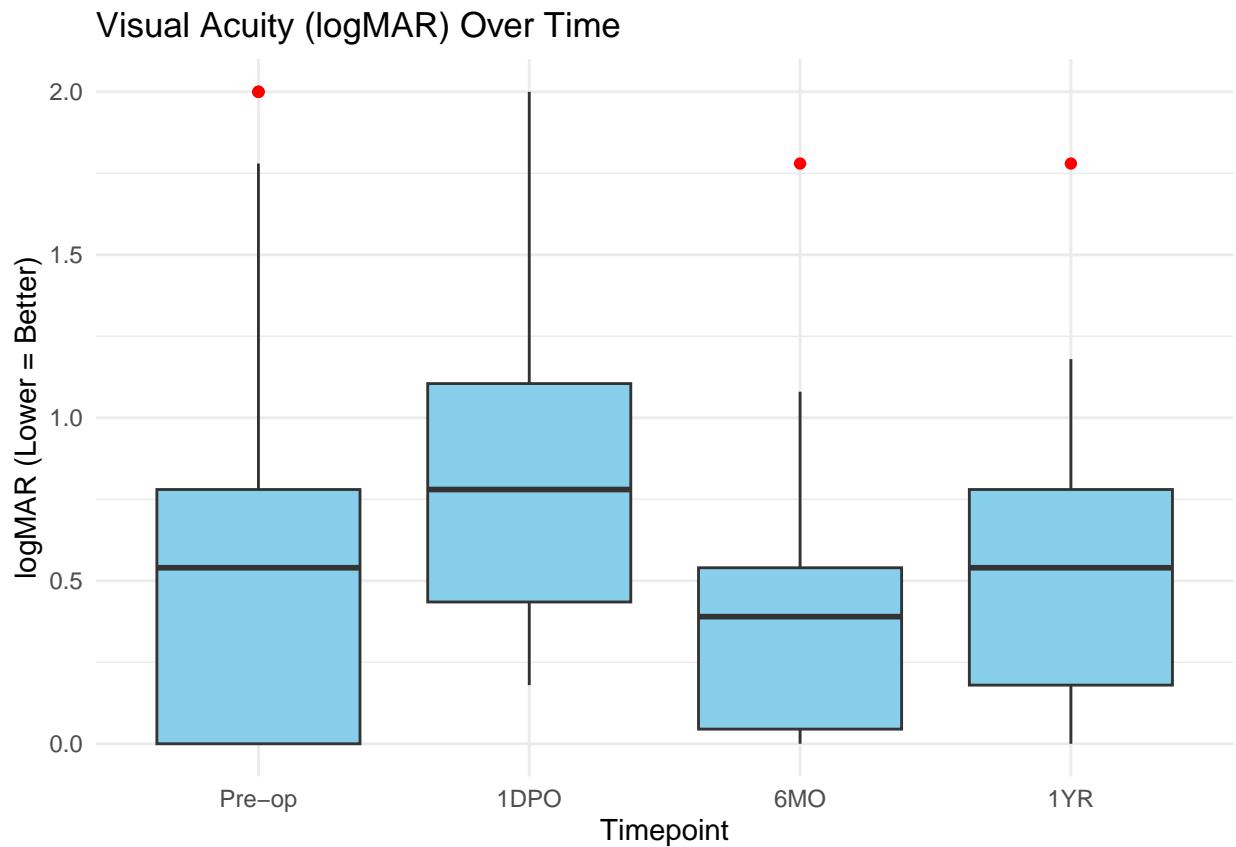
ggplot(va_long, aes(x = timepoint, y = logmar)) +
  geom_boxplot(fill = "skyblue", outlier.color = "red") +

```

```

  labs(
    title = "Visual Acuity (logMAR) Over Time",
    x = "Timepoint",
    y = "logMAR (Lower = Better)"
  ) +
  theme_minimal()

```



#### 4. IOP success rates

```

success_tab <- data %>%
  summarise(
    n_6mo = sum(!is.na(iop_6mo)),
    n_1yr = sum(!is.na(iop_1yr)),
    prop_6mo_30pct = mean(success_6mo_30pct, na.rm = TRUE),
    prop_1yr_30pct = mean(success_1yr_30pct, na.rm = TRUE)
  ) %>%
  mutate(
    prop_6mo_30pct = percent(prop_6mo_30pct, accuracy = 0.1),
    prop_1yr_30pct = percent(prop_1yr_30pct, accuracy = 0.1)
  )

success_tab

```

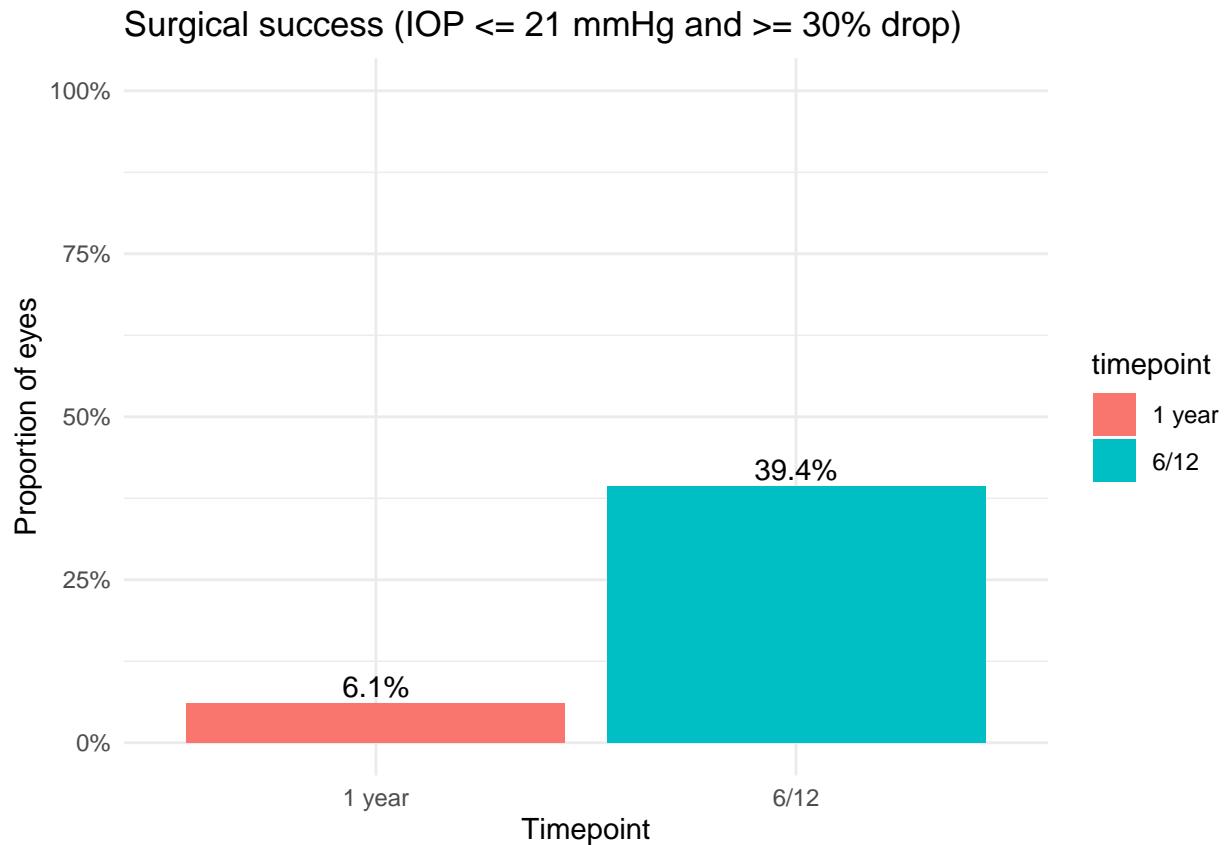
```

## # A tibble: 1 x 4
##   n_6mo n_1yr prop_6mo_30pct prop_1yr_30pct
##   <int> <int> <chr>          <chr>
## 1     22      9  39.4%         6.1%

success_long <- data %>%
  summarise(`6/12` = mean(success_6mo_30pct, na.rm = TRUE),
            `1 year` = mean(success_1yr_30pct, na.rm = TRUE)) %>%
  pivot_longer(everything(), names_to = "timepoint", values_to = "prop") %>%
  mutate(label = percent(prop), accuracy = 0.1)

ggplot(success_long, aes(x = timepoint, y = prop, fill = timepoint)) +
  geom_col() +
  geom_text(aes(label = label), vjust = -0.3) +
  scale_y_continuous(labels = percent_format(accuracy = 1), limits = c(0, 1)) +
  labs(title = "Surgical success (IOP <= 21 mmHg and >= 30% drop)",
       x = "Timepoint", y = "Proportion of eyes") +
  theme_minimal()

```



## 5. VA stability

```

va_status_6mo_tab <- data %>%
  count(va_status_6mo) %>%

```

```

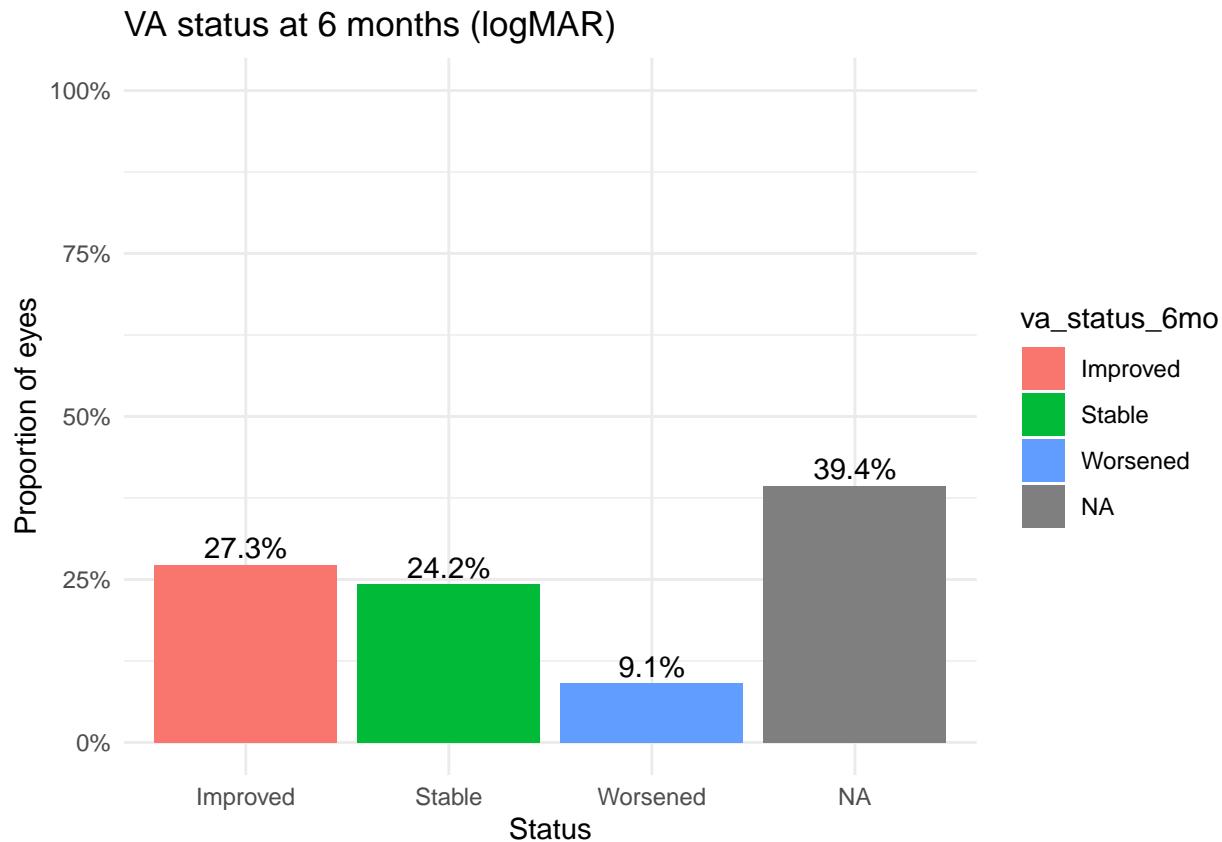
    mutate(prop = n / sum(n), prop_pct = percent(prop, accuracy = 0.1))

va_status_6mo_tab

## # A tibble: 4 x 4
##   va_status_6mo     n   prop prop_pct
##   <chr>       <int>  <dbl> <chr>
## 1 Improved        9 0.273  27.3%
## 2 Stable          8 0.242  24.2%
## 3 Worsened        3 0.0909 9.1%
## 4 <NA>           13 0.394  39.4%

ggplot(va_status_6mo_tab, aes(x = va_status_6mo, y = prop, fill = va_status_6mo)) +
  geom_col() +
  geom_text(aes(label = prop_pct), vjust = -0.3) +
  scale_y_continuous(labels = percent_format(accuracy = 1), limits = c(0, 1)) +
  labs(title = "VA status at 6 months (logMAR)",
       x = "Status", y = "Proportion of eyes") +
  theme_minimal()

```



## 6. Medication burden and success type

```

# Pre-op: n/a = unknown, not "no meds"
med_preop_any <- function(x) {
  x_clean <- tolower(trimws(as.character(x)))

```

```

# Standard missing / unknown
x_clean[x_clean %in% c("na", "n/a", "", "-", "--")] <- NA

dplyr::case_when(
  is.na(x_clean) ~ NA_integer_,

  # Explicit "no meds" text if it ever appears
  str_detect(x_clean, "no med") ~ 0L,
  str_detect(x_clean, "no medication") ~ 0L,
  str_detect(x_clean, "none") ~ 0L,

  # Everything else: assume on at least one glaucoma med
  TRUE ~ 1L
)
}

# Current: "nil" means truly no meds
med_current_any <- function(x) {
  x_clean <- tolower(trimws(as.character(x)))

  dplyr::case_when(
    is.na(x_clean) ~ NA_integer_,

    # Here "nil" means no medication
    x_clean == "nil" ~ 0L,
    str_detect(x_clean, "no med") ~ 0L,
    str_detect(x_clean, "no medication") ~ 0L,
    str_detect(x_clean, "none") ~ 0L,

    # Everything else: assume on at least one glaucoma med
    TRUE ~ 1L
  )
}

data <- data %>%
  mutate(
    meds_preop_any = med_preop_any(`Pre-Op medication`),
    meds_current_any = med_current_any(`Current medication`),

    med_free_last = dplyr::if_else(
      !is.na(meds_current_any) & meds_current_any == 0L,
      1L, 0L
    )
  )

```

## 6.1 Complete vs qualified success

```

# 1-year success type
data <- data %>%
  mutate(
    # Among eyes that meet IOP target (success_1yr_30pct),

```

```

# split into complete vs qualified at last visit
complete_success_last = dplyr::if_else(
  success_1yr_30pct == 1L & med_free_last == 1L,
  1L, 0L
),
qualified_success_last = dplyr::if_else(
  success_1yr_30pct == 1L & med_free_last == 0L,
  1L, 0L
)
)

med_summary <- data %>%
  summarise(
    n_total      = n(),
    n_preop_any   = sum(!is.na(meds_preop_any)),
    n_last_any    = sum(!is.na(meds_current_any)),
    n_med_free_last = sum(med_free_last == 1, na.rm = TRUE),
    n_success_1yr   = sum(success_1yr_30pct == 1, na.rm = TRUE),
    n_complete_last = sum(complete_success_last == 1, na.rm = TRUE),
    n_qualified_last = sum(qualified_success_last == 1, na.rm = TRUE)
  ) %>%
  mutate(
    prop_med_free_last_overall =
      scales::percent(n_med_free_last / n_total, accuracy = 0.1),
    prop_qualified_overall =
      scales::percent(n_qualified_last / n_total, accuracy = 0.1),
    prop_qualified_among_success =
      ifelse(n_success_1yr > 0,
        scales::percent(n_qualified_last / n_success_1yr, accuracy = 0.1),
        NA_character_)
  )
)

med_summary

## # A tibble: 1 x 10
##   n_total n_preop_any n_last_any n_med_free_last n_success_1yr n_complete_last
##       <int>       <int>       <int>           <int>       <int>           <int>
## 1       33         26         30           21         2             1
## # i 4 more variables: n_qualified_last <int>, prop_med_free_last_overall <chr>,
## #   prop_qualified_overall <chr>, prop_qualified_among_success <chr>

success_type_last <- data %>%
  summarise(
    complete  = mean(complete_success_last, na.rm = TRUE),
    qualified = mean(qualified_success_last, na.rm = TRUE)
  ) %>%
  tidyr::pivot_longer(dplyr::everything(), names_to = "type", values_to = "prop") %>%
  mutate(label = scales::percent(prop, accuracy = 0.1))

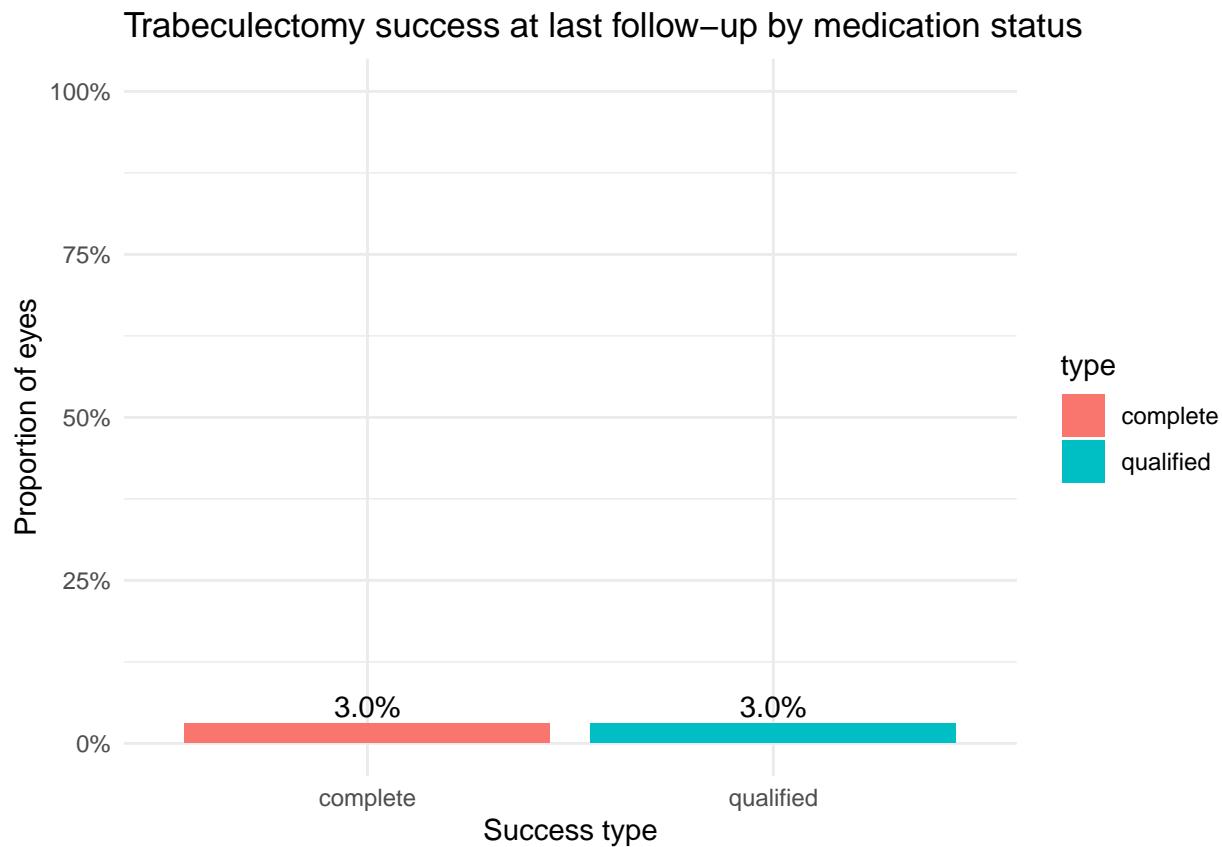
ggplot(success_type_last, aes(x = type, y = prop, fill = type)) +
  geom_col() +
  geom_text(aes(label = label), vjust = -0.3) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1), limits = c(0, 1)) +

```

```

  labs(
    title = "Trabeculectomy success at last follow-up by medication status",
    x = "Success type",
    y = "Proportion of eyes"
  ) +
  theme_minimal()

```



## 7. Additional visualisations (IOP and VA)

### 7.1 VA trajectory (mean logMAR over time)

```

# Long format for VA (logMAR)
va_long <- data %>%
  select(Patient_ID, Eye,
         va_preop_logmar, va_1dpo_logmar,
         va_6mo_logmar, va_1yr_logmar) %>%
  pivot_longer(
    cols = starts_with("va_"),
    names_to = "timepoint",
    values_to = "va_logmar"
  ) %>%
  mutate(
    timepoint = factor(

```

```

    timepoint,
    levels = c("va_preop_logmar",
              "va_1dpo_logmar",
              "va_6mo_logmar",
              "va_1yr_logmar"),
    labels = c("Pre-op", "1 DPO", "6/12", "1 year")
  )
)

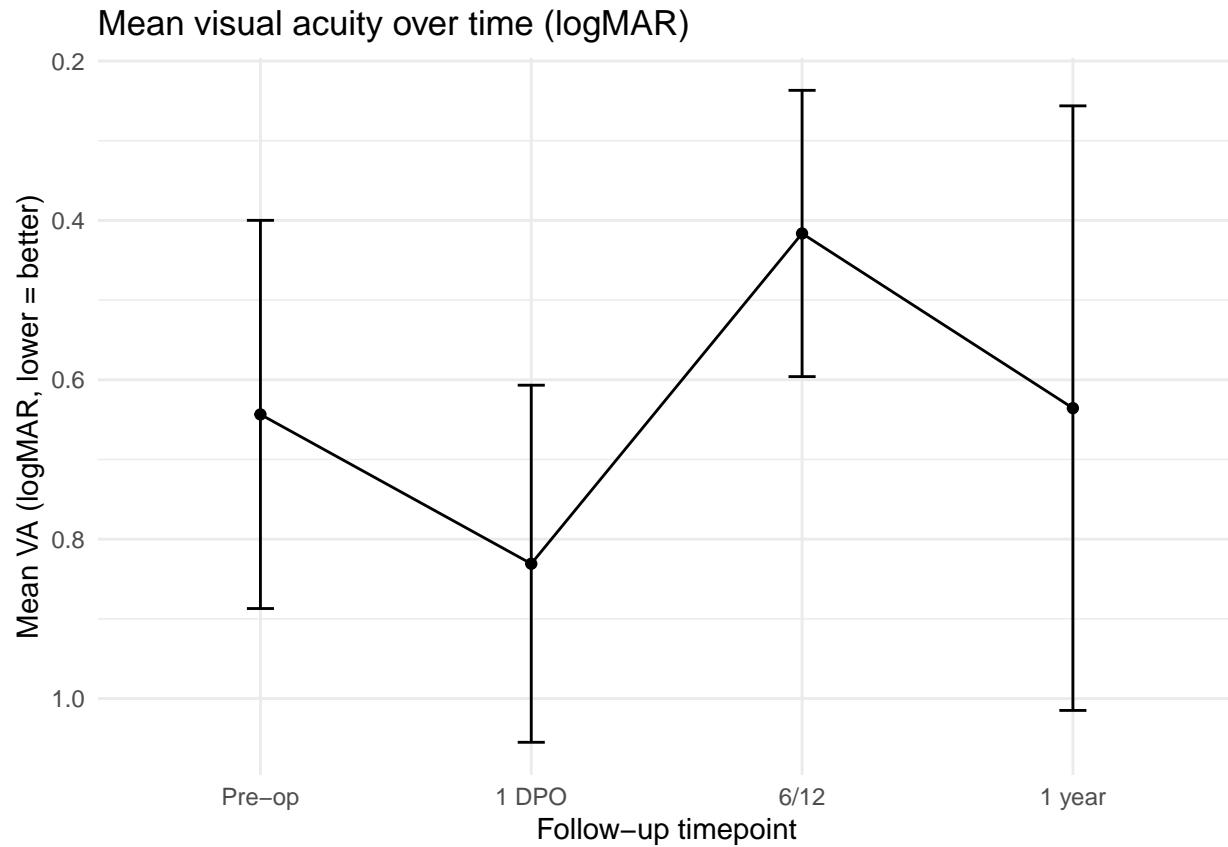
va_summary <- va_long %>%
  group_by(timepoint) %>%
  summarise(
    n      = sum(!is.na(va_logmar)),
    mean_va = mean(va_logmar, na.rm = TRUE),
    med_va  = median(va_logmar, na.rm = TRUE),
    sd_va   = sd(va_logmar, na.rm = TRUE)
  ) %>%
  mutate(
    se_va = sd_va / sqrt(n),
    lower = mean_va - 1.96 * se_va,
    upper = mean_va + 1.96 * se_va
  )

va_summary

## # A tibble: 4 x 8
##   timepoint     n mean_va med_va sd_va  se_va lower upper
##   <fct>     <int>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 Pre-op       29    0.643    0.54  0.669  0.124  0.400  0.887
## 2 1 DPO        24    0.831    0.78  0.560  0.114  0.607  1.05
## 3 6/12         22    0.416    0.39  0.430  0.0916 0.237  0.596
## 4 1 year       9     0.636    0.54  0.581  0.194  0.256  1.01

# VA trajectory plot (logMAR, reversed axis so better VA is "up")
ggplot(va_summary, aes(x = timepoint, y = mean_va, group = 1)) +
  geom_line() +
  geom_point() +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.1) +
  scale_y_reverse() +
  labs(
    title = "Mean visual acuity over time (logMAR)",
    x = "Follow-up timepoint",
    y = "Mean VA (logMAR, lower = better)"
  ) +
  theme_minimal()

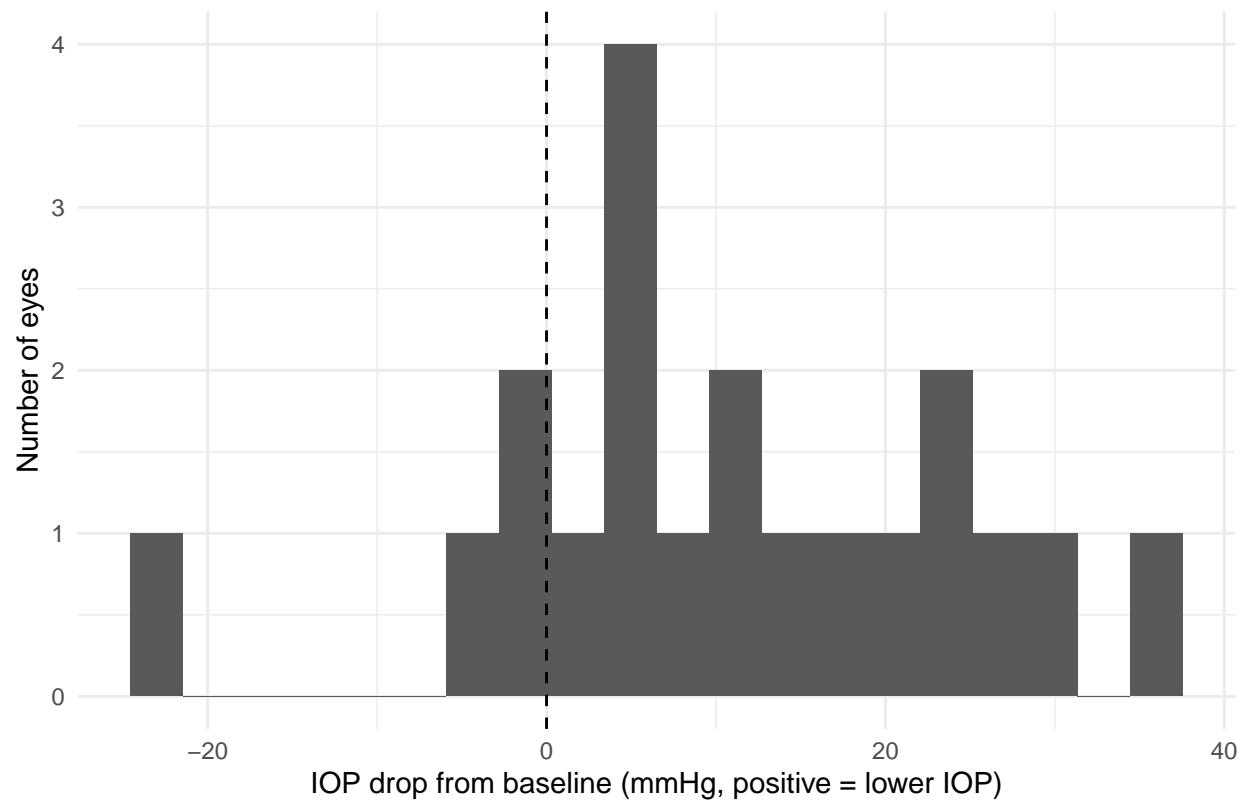
```



## 7.2 Distribution of IOP drop and VA change

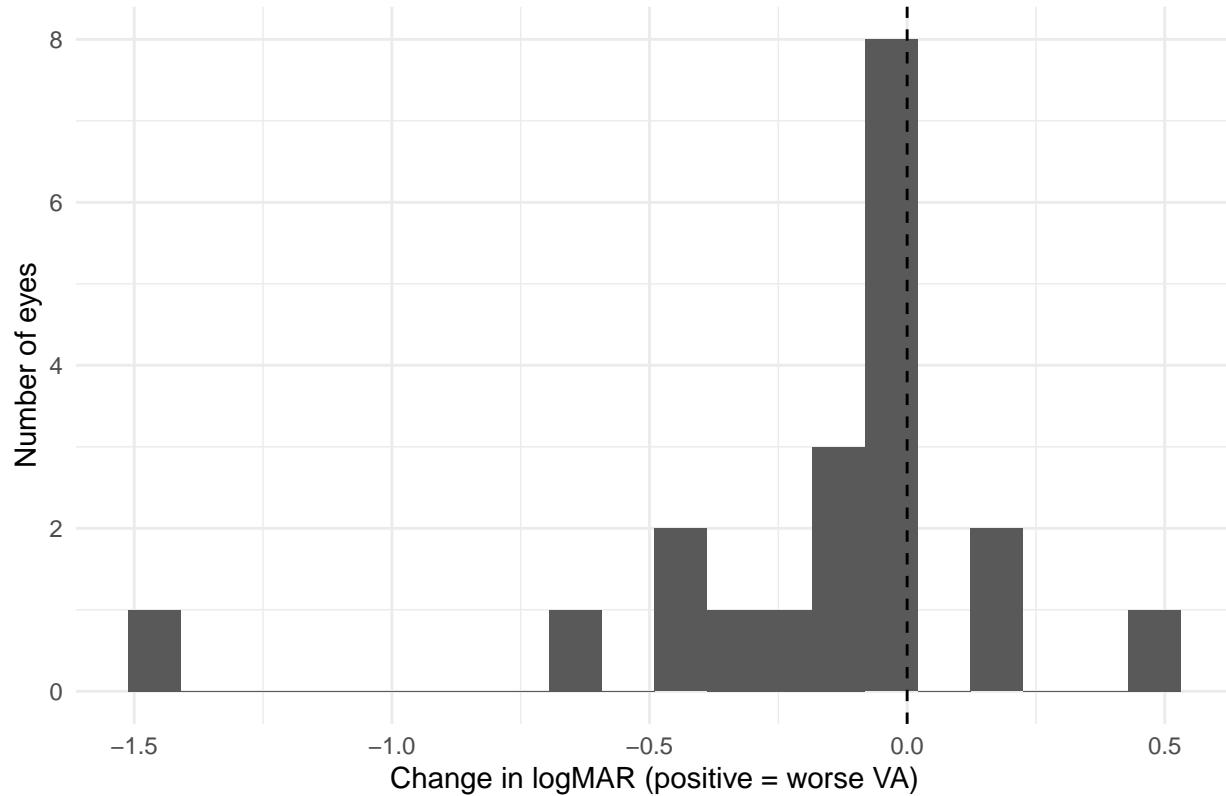
```
# Histogram of 6-month IOP drop from baseline
ggplot(data, aes(x = iop_drop_6mo)) +
  geom_histogram(bins = 20) +
  geom_vline(xintercept = 0, linetype = "dashed") +
  labs(
    title = "Distribution of IOP change at 6 months",
    x = "IOP drop from baseline (mmHg, positive = lower IOP)",
    y = "Number of eyes"
  ) +
  theme_minimal()
```

### Distribution of IOP change at 6 months



```
# Histogram of VA change at 6 months
ggplot(data, aes(x = va_change_6mo)) +
  geom_histogram(bins = 20) +
  geom_vline(xintercept = 0, linetype = "dashed") +
  labs(
    title = "Distribution of change in VA at 6 months",
    x = "Change in logMAR (positive = worse VA)",
    y = "Number of eyes"
  ) +
  theme_minimal()
```

Distribution of change in VA at 6 months



## 8. Inferential analysis

### 8.1 IOP change over time

```
# Preop vs 6 months
iop_paired_6mo <- data %>%
  select(Patient_ID, iop_preop, iop_6mo) %>%
  filter(!is.na(iop_preop), !is.na(iop_6mo))

nrow(iop_paired_6mo)

## [1] 20

# Sanity check
nrow(iop_paired_6mo)

## [1] 20

summary(iop_paired_6mo$iop_preop - iop_paired_6mo$iop_6mo)

##    Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## -23.000   4.175   9.000  10.440  20.275  36.000
```

```

# Paired t test: 6 months
iop_t_6mo <- t.test(
  iop_paired_6mo$iop_6mo,
  iop_paired_6mo$iop_preop,
  paired = TRUE
)
iop_t_6mo

##
## Paired t-test
##
## data: iop_paired_6mo$iop_6mo and iop_paired_6mo$iop_preop
## t = -3.4655, df = 19, p-value = 0.00259
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
## -16.745279 -4.134721
## sample estimates:
## mean difference
## -10.44

# If distribution looks ugly or n is small, also do Wilcoxon
iop_wilcox_6mo <- wilcox.test(iop_paired_6mo$iop_6mo,
                                 iop_paired_6mo$iop_preop,
                                 paired = TRUE,
                                 exact = FALSE)
iop_wilcox_6mo

##
## Wilcoxon signed rank test with continuity correction
##
## data: iop_paired_6mo$iop_6mo and iop_paired_6mo$iop_preop
## V = 24, p-value = 0.00264
## alternative hypothesis: true location shift is not equal to 0

```

## 8.2 Visual acuity change (logMAR)

```

va_paired_6mo <- data %>%
  select(Patient_ID, va_preop_logmar, va_6mo_logmar) %>%
  filter(!is.na(va_preop_logmar), !is.na(va_6mo_logmar))

nrow(va_paired_6mo)

## [1] 20

summary(va_paired_6mo$va_6mo_logmar - va_paired_6mo$va_preop_logmar)

##    Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## -1.460 -0.240  0.000 -0.156  0.000  0.480

```

```

# Paired t test
va_t_6mo <- t.test(
  va_paired_6mo$va_6mo_logmar,
  va_paired_6mo$va_preop_logmar,
  paired = TRUE
)
va_t_6mo

##
##  Paired t-test
##
## data:  va_paired_6mo$va_6mo_logmar and va_paired_6mo$va_preop_logmar
## t = -1.7674, df = 19, p-value = 0.09322
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
## -0.34074108 0.02874108
## sample estimates:
## mean difference
##                 -0.156

# Wilcoxon as a check
va_wilcox_6mo <- wilcox.test(
  va_paired_6mo$va_6mo_logmar,
  va_paired_6mo$va_preop_logmar,
  paired = TRUE,
  exact = FALSE
)
va_wilcox_6mo

##
##  Wilcoxon signed rank test with continuity correction
##
## data:  va_paired_6mo$va_6mo_logmar and va_paired_6mo$va_preop_logmar
## V = 16.5, p-value = 0.08366
## alternative hypothesis: true location shift is not equal to 0

```