

# CTPA Interobserver Agreement Analysis

Weixin Xiao

2025-10-05

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    4.0.0      v tibble     3.2.1
## v lubridate  1.9.4      v tidyr      1.3.1
## v purrr      1.0.4
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
ctpa_data <- read.csv("CTPAData.csv", stringsAsFactors = FALSE)
```

```
# Create the base contingency table
contingency_table <- table(ctpa_data$Resident, ctpa_data$Staff)
```

```
# Calculate totals and percentages
total_cases <- sum(contingency_table)
row_totals <- rowSums(contingency_table)
col_totals <- colSums(contingency_table)
```

```
# Create Contingency Table
```

```
simple_table <- matrix(
  c(
    paste0(contingency_table["N", "N"], " (", round(contingency_table["N", "N"]/total_cases*100, 1), "%"),
    paste0(contingency_table["N", "P"], " (", round(contingency_table["N", "P"]/total_cases*100, 1), "%"),
    paste0(row_totals["N"], " (", round(row_totals["N"]/total_cases*100, 1), "%)"),
    paste0(contingency_table["P", "N"], " (", round(contingency_table["P", "N"]/total_cases*100, 1), "%"),
    paste0(contingency_table["P", "P"], " (", round(contingency_table["P", "P"]/total_cases*100, 1), "%")
  )
```

```

paste0(row_totals["P"], " (", round(row_totals["P"]/total_cases*100, 1), "%)"),
paste0(col_totals["N"], " (", round(col_totals["N"]/total_cases*100, 1), "%)"),
paste0(col_totals["P"], " (", round(col_totals["P"]/total_cases*100, 1), "%)"),
paste0(total_cases, " (100%)")
),
nrow = 3,
ncol = 3,
byrow = TRUE
)

colnames(simple_table) <- c("Staff: Negative", "Staff: Positive", "Row Total")
rownames(simple_table) <- c("Resident: Negative", "Resident: Positive", "Column Total")

kable(simple_table, align = 'c') %>%
  kable_styling(bootstrap_options = "striped", full_width = FALSE) %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(3, bold = TRUE) %>%
  column_spec(4, bold = TRUE)

```

	Staff: Negative	Staff: Positive	Row Total
Resident: Negative	463 (77.7%)	9 (1.5%)	<b>472 (79.2%)</b>
Resident: Positive	12 (2%)	112 (18.8%)	<b>124 (20.8%)</b>
<b>Column Total</b>	<b>475 (79.7%)</b>	<b>121 (20.3%)</b>	<b>596 (100%)</b>

```

# Calculate Cohen's Kappa
observed_agreement <- sum(diag(contingency_table)) / total_cases
expected_agreement <- sum((row_totals/total_cases) * (col_totals/total_cases))
kappa_value <- (observed_agreement - expected_agreement) / (1 - expected_agreement)

# Calculate 95% CI
n <- total_cases
p_o <- observed_agreement
p_e <- expected_agreement
var_kappa <- (p_o * (1 - p_o)) / (n * (1 - p_e)^2)
se_kappa <- sqrt(var_kappa)
ci_lower <- kappa_value - 1.96 * se_kappa
ci_upper <- kappa_value + 1.96 * se_kappa

kappa_summary <- data.frame(
  Statistic = c("Cohen's Kappa", "95% CI Lower", "95% CI Upper"),
  Value = c(
    round(kappa_value, 3),
    round(ci_lower, 3),
    round(ci_upper, 3)
  )
)

kable(kappa_summary, align = c('l', 'c')) %>%
  kable_styling(full_width = FALSE)

```

Statistic	Value
Cohen's Kappa	0.892
95% CI Lower	0.847
95% CI Upper	0.937

```

cat("## Interpretation\n\n")

## ## Interpretation

cat("**Data Summary:**\n")

## **Data Summary:**

cat("- Total CTPA studies:", total_cases, "\n")

## - Total CTPA studies: 596

cat("- Overall agreement:", round(observed_agreement * 100, 1), "%\n")

## - Overall agreement: 96.5 %

cat("- Discordant cases:", sum(contingency_table["P", "N"] + contingency_table["N", "P"]), "\n\n")

## - Discordant cases: 21

cat("**Key Findings:**\n")

## **Key Findings:**

cat("- Cohen's Kappa =", round(kappa_value, 3), "\n")

## - Cohen's Kappa = 0.892

cat("- 95% CI: [", round(ci_lower, 3), ",", round(ci_upper, 3), "]\n")

## - 95% CI: [ 0.847 , 0.937 ]

library(ggplot2)
library(scales)

##
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':
##
##   discard

```

```
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
# Prepare data with percentages
contingency_df <- as.data.frame(contingency_table)
names(contingency_df) <- c("Resident", "Staff", "Count")
contingency_df$Percentage <- contingency_df$Count / total_cases * 100
contingency_df$Label <- paste0(contingency_df$Count, "\n(", round(contingency_df$Percentage, 1), "%)")

# Generate Agreement Heatmap
ggplot(contingency_df, aes(x = Staff, y = Resident, fill = Percentage)) +
  geom_tile(color = "white", linewidth = 1.5, width = 0.9, height = 0.9) +
  geom_text(aes(label = Label), color = "black", size = 5, fontface = "bold") +
  scale_fill_gradientn(
    colors = c("#f7fbff", "#4292c6", "#08306b"),
    values = rescale(c(0, 10, 80)),
    name = "Percentage (%)",
    limits = c(0, 80)
  ) +
  scale_x_discrete(labels = c("Negative", "Positive")) +
  scale_y_discrete(labels = c("Negative", "Positive")) +
  labs(
    title = "CTPA Interpretation Agreement Matrix",
    subtitle = "Resident vs Staff Radiologist",
    x = "Staff Radiologist Interpretation",
    y = "Resident Radiologist Interpretation",
    caption = paste("Total cases:", total_cases)
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
    plot.subtitle = element_text(hjust = 0.5, size = 12),
    axis.title = element_text(face = "bold"),
    legend.position = "right",
    panel.grid = element_blank(),
    plot.caption = element_text(face = "italic")
  ) +
  coord_fixed(ratio = 1)
```

CTPA Interpretation Agreement Matrix

Resident vs Staff Radiologist

