

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., *BMJQS* 2023

Supplemental File #2 for *Burden of Serious Harms from Diagnostic Error in the USA*
by Newman-Toker, et al. *BMJ Quality & Safety*, 2023 (doi:10.1136/bmjqs-2021-014130)

Enclosed below are additional Supplementary Materials (Supplement 2) to the peer-reviewed, scientific journal publication entitled *Burden of Serious Harms from Diagnostic Error in the USA* (2023).

This is the full statistical code (R v4.2.2, Vienna, Austria) for the Monte Carlo analysis.

```
# load packages
rm(list = ls())
library(sn)
library(rootSolve)
library(gdata)
library(MASS)
library(DEoptim)

set.seed(37)

logit <- function(x) {
  out <- log(x/(1-x))
  return(out)
}

expit <- function(x) {
  out <- exp(x) / (1 + exp(x))
  return(out)
}

#### primary analysis ####
dat <- read.csv(file = "big3_data_dx1.csv")

peh_general <- 374/1216 # general per-error-harm rate
```

Supplementary Materials (Supplemental File #2)

Page 1 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
peh_general_se <- sqrt(1/374 + 1/(1216-374)) # standard error on logit scale
peh_general_lower <- expit(logit(peh_general) + qnorm(0.025) * peh_general_se)
peh_general_upper <- expit(logit(peh_general) + qnorm(0.975) * peh_general_se)

perror_big3 <- 717 / (717 + 509) # proportion of big 3 in errors
pharm_big3 <- 157 / 207 # proportion of big 3 in harms

vascular_index <- which(dat$category == "vascular" & dat$disease != "OTHER Vascular")
vascular_other_index <- which(dat$disease == "OTHER Vascular")
infection_index <- which(dat$category == "infection" & dat$disease != "OTHER Infection")
infection_other_index <- which(dat$disease == "OTHER Infection")
cancer_index <- which(dat$category == "cancer" & dat$disease != "OTHER Cancer")
cancer_other_index <- which(dat$disease == "OTHER Cancer")

dat$nerror <- dat$incidence_point * dat$misrate_point
shp <- dat$proportion_highharm # severity harm proportion

peh_fun <- function(peh_general, nerror) {
  # function to calculate disease-specific per-error-harm rates
  peh_rate <- rep(NA, nrow(dat))
  peh_rate[c(vascular_index, infection_index, cancer_index)] <- peh_general * pharm_big3 /
perror_big3 *
  (sum(nerror) * shp[c(vascular_index, infection_index, cancer_index)]) /
  (sum(nerror[vascular_index] * shp[vascular_index]) * 1684 / 1344 +
  sum(nerror[infection_index] * shp[infection_index]) * 992 / 600 +
  sum(nerror[cancer_index] * shp[cancer_index]) * 2793 / 1529)
  peh_rate[vascular_other_index] <- sum(nerror[vascular_index] * peh_rate[vascular_index] *
334/1334) /
  nerror[vascular_other_index]
```

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
  peh_rate[infection_other_index] <- sum(nerror[infection_index] * peh_rate[infection_index] *
392/600) /
  nerror[infection_other_index]
  peh_rate[cancer_other_index] <- sum(nerror[cancer_index] * peh_rate[cancer_index] * 1264/1529)
/
  nerror[cancer_other_index]
  return(peh_rate)
}

# point estimate of per-error-serious harm rates
dat$peharmrate_point <- peh_fun(peh_general, dat$nerror)
dat$peharmrate_lb <- peh_fun(peh_general_lower, dat$nerror)
dat$peharmrate_ub <- peh_fun(peh_general_upper, dat$nerror)

# Monte Carlo simulation, to construct plausible intervals for harm rate and number of harm
snest <- function(u, lower, upper, lowerq = 0.025, upperq = 0.975) {
  solfun <- function(beta) {
    omega <- beta[1]
    alpha <- beta[2]
    out1 <- psn(lower, xi = u, omega = omega, alpha = alpha) - lowerq
    out2 <- psn(upper, xi = u, omega = omega, alpha = alpha) - upperq
    return((out1^2+out2^2))
  }
  start <- c(((upper-lower)/4)^2, 0) + mvrnorm(100, mu = rep(0, 2), Sigma = diag(1, 2))
  start[, 1] <- abs(start[, 1])
  tempfit <- t(apply(start, 1, function(x) {
    fit <- optim(fn = solfun, par = x,
      control = list(maxit = 500000))
    return(c(fit$par, fit$value))
  })
```

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
    )))  
  tempfit <- tempfit[complete.cases(tempfit), ]  
  return(tempfit[which(tempfit[, 3] == min(tempfit[, 3]))[1], 1:3])  
}
```

```
totl <- 316283434  
nmc <- 1*(10^7)  
numerical_error <- NULL  
dat$nerror_lb <- NA  
dat$nerror_ub <- NA  
dat$nharm_point <- dat$nerror * dat$peharmrate_point  
dat$nharm_lb <- NA  
dat$nharm_ub <- NA  
dat$harmrate_point <- dat$misrate_point * dat$peharmrate_point  
dat$harmrate_lb <- NA  
dat$harmrate_ub <- NA
```

Monte Carlo simulation based on skew-normal estimation

```
mc_result <- list()  
for (i in 1:nrow(dat)) {  
  mc_result[[i]] <- list()  
  if (dat$incidence_lb[i] == dat$incidence_ub[i]) {  
    mc_incidence <- rep(dat$incidence_point[i], nmc)  
  } else {  
    fit_incidence <- snest(logit(dat$incidence_point[i] / totl),  
                          lower = logit(dat$incidence_lb[i] / totl),  
                          upper = logit(dat$incidence_ub[i] / totl))  
    while(fit_incidence[3] > 10^(-10)) {
```

Supplementary Materials (Supplemental File #2)

Page 4 of 48

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
    fit_incidence <- snest(logit(dat$incidence_point[i] / totl),
      lower = logit(dat$incidence_lb[i] / totl),
      upper = logit(dat$incidence_ub[i] / totl))
  }
  mc_incidence <- expit(rsn(n = nmc, xi = logit(dat$incidence_point[i] / totl),
    omega = fit_incidence[1], alpha = fit_incidence[2])) * totl
}
temp <- (c(quantile(mc_incidence, c(0.025, 0.975)) - c(dat$incidence_lb[i], dat$incidence_ub[i]))) /
totl

if (i %in% c(vascular_index, infection_index, cancer_index)) {
  fit_misrate <- snest(logit(dat$misrate_point[i]),
    lower = logit(dat$misrate_lb[i]),
    upper = logit(dat$misrate_ub[i]))
  while(fit_misrate[3] > 10^(-10)) {
    fit_misrate <- snest(logit(dat$misrate_point[i]),
      lower = logit(dat$misrate_lb[i]),
      upper = logit(dat$misrate_ub[i]))
  }

  mc_misrate <- expit(rsn(n = nmc, xi = logit(dat$misrate_point[i]),
    omega = fit_misrate[1], alpha = fit_misrate[2]))
  temp <- c(temp, quantile(mc_misrate, c(0.025, 0.975)) - c(dat$misrate_lb[i], dat$misrate_ub[i]))

  mc_result[[i]]$fit_misrate <- fit_misrate
  mc_result[[i]]$mc_misrate <- mc_misrate
} else {
  temp <- c(temp, rep(NA, 2))
}
```

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
fit_peh <- snest(logit(dat$peharmrate_point[i]),
               lower = logit(dat$peharmrate_lb[i]),
               upper = logit(dat$peharmrate_ub[i]))
while(fit_peh[3] > 10^(-10)) {
  fit_peh <- snest(logit(dat$peharmrate_point[i]),
                 lower = logit(dat$peharmrate_lb[i]),
                 upper = logit(dat$peharmrate_ub[i]))
}
mc_peh <- expit(rsn(n = nmc, xi = logit(dat$peharmrate_point[i]),
                  omega = fit_peh[1], alpha = fit_peh[2]))
temp <- c(temp, quantile(mc_peh, c(0.025, 0.975)) - c(dat$peharmrate_lb[i], dat$peharmrate_ub[i]))

numerical_error <- rbind(numerical_error, temp)

mc_result[[i]]$fit_incidence <- fit_incidence
mc_result[[i]]$mc_incidence <- mc_incidence

mc_result[[i]]$fit_peh <- fit_peh
mc_result[[i]]$mc_peh <- mc_peh
mc_result[[i]]$numerical_error <- numerical_error
}

summary(numerical_error)

# calculate plausible intervals for "OTHER" vascular/infection/censor harm rate (misrate)
# because point estimate is calculated as weighted sum and plausible intervals should be calculated
accordingly
temp1 <- 0
```

*Supplementary Materials (Supplemental File #2)**Page 6 of 48*

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp2 <- 0
for (i in vascular_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
mc_result[[vascular_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[vascular_other_index] <- quantile(mc_result[[vascular_other_index]]$mc_misrate,
0.025)
dat$misrate_ub[vascular_other_index] <- quantile(mc_result[[vascular_other_index]]$mc_misrate,
0.975)

temp1 <- 0
temp2 <- 0
for (i in infection_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
mc_result[[infection_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[infection_other_index] <- quantile(mc_result[[infection_other_index]]$mc_misrate,
0.025)
dat$misrate_ub[infection_other_index] <- quantile(mc_result[[infection_other_index]]$mc_misrate,
0.975)

temp1 <- 0
temp2 <- 0
for (i in cancer_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
```

Supplementary Materials (Supplemental File #2)

Page 7 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
mc_result[[cancer_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[cancer_other_index] <- quantile(mc_result[[cancer_other_index]]$mc_misrate, 0.025)
dat$misrate_ub[cancer_other_index] <- quantile(mc_result[[cancer_other_index]]$mc_misrate, 0.975)

# calculate plausible intervals for number of errors, number of harms and harm rate
# based on Monte Carlo generated data
for (i in 1:nrow(dat)) {
  mc_incidence <- mc_result[[i]]$mc_incidence
  mc_misrate <- mc_result[[i]]$mc_misrate
  mc_peh <- mc_result[[i]]$mc_peh
  dat$nerror_lb[i] <- quantile(mc_incidence * mc_misrate, 0.025)
  dat$nerror_ub[i] <- quantile(mc_incidence * mc_misrate, 0.975)

  dat$nharm_lb[i] <- quantile(mc_incidence * mc_misrate * mc_peh, 0.025)
  dat$nharm_ub[i] <- quantile(mc_incidence * mc_misrate * mc_peh, 0.975)

  dat$harmrate_lb[i] <- quantile(mc_misrate * mc_peh, 0.025)
  dat$harmrate_ub[i] <- quantile(mc_misrate * mc_peh, 0.975)
}

# calculate subtotals and plausible intervals
# create structure for subtotals
temp <- c("Top 5 Vascular Subtotal",
         "Top 5 Infection Subtotal",
         "Top 5 Cancer Subtotal",
         "Top 15 Subtotal",
         "Total Vascular",
```

Supplementary Materials (Supplemental File #2)

Page 8 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
"Total Infection",
"Total Cancer",
"Big 3 Subtotal",
"All Other Dx Errors",
"Grand Total")
temp2 <- data.frame(matrix(NA, nrow = length(temp), ncol = ncol(dat)))
colnames(temp2) <- colnames(dat)
temp2[, 2] <- temp
temp2[, 1] <- rep("Subtotals", length(temp))
dat <- rbind(dat, temp2)

dat[, -c(1:2)] <- apply(dat[, -c(1:2)], c(1, 2), as.numeric)

# calculate subtotal incidence
dat$incidence_point[19] <- sum(dat$incidence_point[1:5])
dat$incidence_point[20] <- sum(dat$incidence_point[7:11])
dat$incidence_point[21] <- sum(dat$incidence_point[13:17])
dat$incidence_point[22] <- sum(dat$incidence_point[19:21])

dat$incidence_point[23] <- sum(dat$incidence_point[1:6])
dat$incidence_point[24] <- sum(dat$incidence_point[7:12])
dat$incidence_point[25] <- sum(dat$incidence_point[13:18])
dat$incidence_point[26] <- sum(dat$incidence_point[1:18])

temp <- lapply(1:18, function(i) return(mc_result[[i]]$mc_incidence))

temp1 <- rep(0, nmc)
for (i in 1:5) {
```

Supplementary Materials (Supplemental File #2)

Page 9 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[19] <- quantile(temp1, 0.025)
dat$incidence_ub[19] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 7:11) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[20] <- quantile(temp1, 0.025)
dat$incidence_ub[20] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 13:17) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[21] <- quantile(temp1, 0.025)
dat$incidence_ub[21] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in c(1:5, 7:11, 13:17)) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[22] <- quantile(temp1, 0.025)
dat$incidence_ub[22] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
```

Supplementary Materials (Supplemental File #2)

Page 10 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
for (i in 1:6) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[23] <- quantile(temp1, 0.025)
dat$incidence_ub[23] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 7:12) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[24] <- quantile(temp1, 0.025)
dat$incidence_ub[24] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 13:18) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[25] <- quantile(temp1, 0.025)
dat$incidence_ub[25] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 1:18) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[26] <- quantile(temp1, 0.025)
dat$incidence_ub[26] <- quantile(temp1, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 11 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat[27:28, 3:5] <- NA

# calculate subtotal number of errors
dat$ncerror[19] <- sum(dat$ncerror[1:5])
dat$ncerror[20] <- sum(dat$ncerror[7:11])
dat$ncerror[21] <- sum(dat$ncerror[13:17])
dat$ncerror[22] <- sum(dat$ncerror[19:21])

dat$ncerror[23] <- sum(dat$ncerror[1:6])
dat$ncerror[24] <- sum(dat$ncerror[7:12])
dat$ncerror[25] <- sum(dat$ncerror[13:18])
dat$ncerror[26] <- sum(dat$ncerror[1:18])

dat$ncerror[27] <- dat$ncerror[26] / pncerror_big3 * (1 - pncerror_big3)
dat$ncerror[28] <- dat$ncerror[26] / pncerror_big3

temp1 <- rep(0, nmc)
for (i in 1:5) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
}
dat$ncerror_lb[19] <- quantile(temp1, 0.025)
dat$ncerror_ub[19] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 7:11) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

Supplementary Materials (Supplemental File #2)

Page 12 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
}  
dat$nerror_lb[20] <- quantile(temp1, 0.025)  
dat$nerror_ub[20] <- quantile(temp1, 0.975)  
  
temp1 <- rep(0, nmc)  
for (i in 13:17) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[21] <- quantile(temp1, 0.025)  
dat$nerror_ub[21] <- quantile(temp1, 0.975)  
  
temp1 <- rep(0, nmc)  
for (i in c(1:5, 7:11, 13:17)) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[22] <- quantile(temp1, 0.025)  
dat$nerror_ub[22] <- quantile(temp1, 0.975)  
  
temp1 <- rep(0, nmc)  
for (i in 1:6) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[23] <- quantile(temp1, 0.025)  
dat$nerror_ub[23] <- quantile(temp1, 0.975)  
  
temp1 <- rep(0, nmc)
```

Supplementary Materials (Supplemental File #2)

Page 13 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
for (i in 7:12) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
}
dat$nerror_lb[24] <- quantile(temp1, 0.025)
dat$nerror_ub[24] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 13:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
}
dat$nerror_lb[25] <- quantile(temp1, 0.025)
dat$nerror_ub[25] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 1:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
}
dat$nerror_lb[26] <- quantile(temp1, 0.025)
dat$nerror_ub[26] <- quantile(temp1, 0.975)

dat$nerror_lb[27] <- quantile(temp1 / perror_big3 * (1 - perror_big3), 0.025)
dat$nerror_ub[27] <- quantile(temp1 / perror_big3 * (1 - perror_big3), 0.975)

dat$nerror_lb[28] <- quantile(temp1 / perror_big3, 0.025)
dat$nerror_ub[28] <- quantile(temp1 / perror_big3, 0.975)
```

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
# calculate subtotal misrate
dat$misrate_point[19:28] <- dat$nerror[19:28] / dat$incidence_point[19:28]

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 1:5) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[19] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[19] <- quantile(temp1 / temp2, 0.975)

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 7:11) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[20] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[20] <- quantile(temp1 / temp2, 0.975)

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 13:17) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
```

Supplementary Materials (Supplemental File #2)

Page 15 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$misrate_lb[21] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[21] <- quantile(temp1 / temp2, 0.975)

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in c(1:5, 7:11, 13:17)) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[22] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[22] <- quantile(temp1 / temp2, 0.975)

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 1:6) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[23] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[23] <- quantile(temp1 / temp2, 0.975)

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 7:12) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
```

Supplementary Materials (Supplemental File #2)

Page 16 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
}  
dat$misrate_lb[24] <- quantile(temp1 / temp2, 0.025)  
dat$misrate_ub[24] <- quantile(temp1 / temp2, 0.975)  
  
temp1 <- rep(0, nmc)  
temp2 <- rep(0, nmc)  
for (i in 13:18) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
  temp2 <- temp2 + mc_result[[i]]$mc_incidence  
}  
dat$misrate_lb[25] <- quantile(temp1 / temp2, 0.025)  
dat$misrate_ub[25] <- quantile(temp1 / temp2, 0.975)  
  
temp1 <- rep(0, nmc)  
temp2 <- rep(0, nmc)  
for (i in 1:18) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
  temp2 <- temp2 + mc_result[[i]]$mc_incidence  
}  
dat$misrate_lb[26] <- quantile(temp1 / temp2, 0.025)  
dat$misrate_ub[26] <- quantile(temp1 / temp2, 0.975)  
  
# calculate subtotal nharm, peh, and harm rate  
dat$nharm_point[19] <- sum(dat$nharm_point[1:5])  
dat$nharm_point[20] <- sum(dat$nharm_point[7:11])
```

Supplementary Materials (Supplemental File #2)

Page 17 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$nharm_point[21] <- sum(dat$nharm_point[13:17])
dat$nharm_point[22] <- sum(dat$nharm_point[19:21])

dat$nharm_point[23] <- sum(dat$nharm_point[1:6])
dat$nharm_point[24] <- sum(dat$nharm_point[7:12])
dat$nharm_point[25] <- sum(dat$nharm_point[13:18])
dat$nharm_point[26] <- sum(dat$nharm_point[1:18])

dat$nharm_point[27] <- dat$nharm_point[26] / pharm_big3 * (1 - pharm_big3)
dat$nharm_point[28] <- dat$nharm_point[26] / pharm_big3

dat$harmrate_point[19:28] <- dat$nharm_point[19:28] / dat$incidence_point[19:28]
dat$peharmrate_point[19:28] <- dat$nharm_point[19:28] / dat$neror[19:28]

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 1:5) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}

dat$nharm_lb[19] <- quantile(temp3, 0.025)
dat$nharm_ub[19] <- quantile(temp3, 0.975)

dat$harmrate_lb[19] <- quantile(temp3 / temp1, 0.025)
```

Supplementary Materials (Supplemental File #2)

Page 18 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$harmrate_ub[19] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[19] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[19] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 7:11) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[20] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[20] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[20] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[20] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[20] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[20] <- quantile(temp3 / temp2, 0.975)
```

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 13:17) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}

dat$harm_lb[21] <- quantile(temp3, 0.025)
dat$harm_ub[21] <- quantile(temp3, 0.975)

dat$harmrate_lb[21] <- quantile(temp3 / temp1, 0.025)
dat$harmrate_ub[21] <- quantile(temp3 / temp1, 0.975)

dat$peharmrate_lb[21] <- quantile(temp3 / temp2, 0.025)
dat$peharmrate_ub[21] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in c(1:5, 7:11, 13:17)) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

Supplementary Materials (Supplemental File #2)

Page 20 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
}
```

```
dat$nharm_lb[22] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[22] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[22] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[22] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[22] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[22] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 1:6) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[23] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[23] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[23] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[23] <- quantile(temp3 / temp1, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 21 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$peharmrate_lb[23] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[23] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 7:12) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[24] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[24] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[24] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[24] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[24] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[24] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

Supplementary Materials (Supplemental File #2)

Page 22 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 13:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}
```

```
dat$nharm_lb[25] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[25] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[25] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[25] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[25] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[25] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 1:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}
```

Supplementary Materials (Supplemental File #2)

Page 23 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$nharm_lb[26] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[26] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[26] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[26] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[26] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[26] <- quantile(temp3 / temp2, 0.975)
```

```
dat$nharm_lb[27] <- quantile(temp3 / pharm_big3 * (1 - pharm_big3), 0.025)
```

```
dat$nharm_ub[27] <- quantile(temp3 / pharm_big3 * (1 - pharm_big3), 0.975)
```

```
dat$nharm_lb[28] <- quantile(temp3 / pharm_big3, 0.025)
```

```
dat$nharm_ub[28] <- quantile(temp3 / pharm_big3, 0.975)
```

```
dat$peharmrate_lb[27] <- quantile((temp3 / pharm_big3 * (1 - pharm_big3)) /  
                                (temp2 / perror_big3 * (1 - perror_big3)), 0.025)
```

```
dat$peharmrate_ub[27] <- quantile((temp3 / pharm_big3 * (1 - pharm_big3)) /  
                                (temp2 / perror_big3 * (1 - perror_big3)), 0.975)
```

```
dat$peharmrate_lb[28] <- quantile((temp3 / pharm_big3) / (temp2 / perror_big3), 0.025)
```

```
dat$peharmrate_ub[28] <- quantile((temp3 / pharm_big3) / (temp2 / perror_big3), 0.975)
```

```
dat$weight <- dat$peharmrate_point / peh_general
```

```
dat <- dat[, c(1:2, 9:10, 23, 3:5,
```

```
             6:8, 12:14, 20:22,
```

```
             11, 15, 16, 17:19)]
```

Supplementary Materials (Supplemental File #2)

Page 24 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
all(dat$misrate_lb < dat$misrate_point, na.rm = T)
all(dat$misrate_ub > dat$misrate_point, na.rm = T)
all(dat$harmrate_lb < dat$harmrate_point, na.rm = T)
all(dat$harmrate_ub > dat$harmrate_point, na.rm = T)
all(dat$nharm_lb < dat$nharm_point, na.rm = T)
all(dat$nharm_ub > dat$nharm_point, na.rm = T)
all(dat$error_lb < dat$error, na.rm = T)
all(dat$error_ub > dat$error, na.rm = T)

write.csv(dat, file = "big3_results_dx1_updated.csv")

#### secondary analysis ####
dat <- read.csv(file = "big3_data_combo.csv")
set.seed(37)
peh_general <- 374/1216 # general per-error-harm rate
peh_general_se <- sqrt(1/374 + 1/(1216-374)) # standard error on logit scale
peh_general_lower <- expit(logit(peh_general) + qnorm(0.025) * peh_general_se)
peh_general_upper <- expit(logit(peh_general) + qnorm(0.975) * peh_general_se)

perror_big3 <- 717 / (717 + 509) # big 3 error rate
pharm_big3 <- 157 / 207 # harm rate

vascular_index <- which(dat$category == "vascular" & dat$disease != "OTHER Vascular")
vascular_other_index <- which(dat$disease == "OTHER Vascular")
infection_index <- which(dat$category == "infection" & dat$disease != "OTHER Infection")
```

Supplementary Materials (Supplemental File #2)

Page 25 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
infection_other_index <- which(dat$disease == "OTHER Infection")
cancer_index <- which(dat$category == "cancer" & dat$disease != "OTHER Cancer")
cancer_other_index <- which(dat$disease == "OTHER Cancer")

dat$nerror <- dat$incidence_point * dat$misrate_point
shp <- dat$proportion_highharm # severity harm proportion

peh_fun <- function(peh_general, nerror) {
  # function to calculate disease-specific per-error-harm rates
  peh_rate <- rep(NA, nrow(dat))
  peh_rate[c(vascular_index, infection_index, cancer_index)] <- peh_general * pharm_big3 /
  perror_big3 *
  (sum(nerror) * shp[c(vascular_index, infection_index, cancer_index)]) /
  (sum(nerror[vascular_index] * shp[vascular_index]) * 1684 / 1344 +
  sum(nerror[infection_index] * shp[infection_index]) * 992 / 600 +
  sum(nerror[cancer_index] * shp[cancer_index]) * 2793 / 1529)
  peh_rate[vascular_other_index] <- sum(nerror[vascular_index] * peh_rate[vascular_index] *
  334/1334) /
  nerror[vascular_other_index]
  peh_rate[infection_other_index] <- sum(nerror[infection_index] * peh_rate[infection_index] *
  392/600) /
  nerror[infection_other_index]
  peh_rate[cancer_other_index] <- sum(nerror[cancer_index] * peh_rate[cancer_index] * 1264/1529)
  /
  nerror[cancer_other_index]
  return(peh_rate)
}

# point estimate of per-error-serious harm rates
dat$peharmrate_point <- peh_fun(peh_general, dat$nerror)
```

Supplementary Materials (Supplemental File #2)

Page 26 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$peharmrate_lb <- peh_fun(peh_general_lower, dat$nerror)
dat$peharmrate_ub <- peh_fun(peh_general_upper, dat$nerror)

# Monte Carlo simulation, to construct plausible intervals for harm rate and number of harm
snest <- function(u, lower, upper, lowerq = 0.025, upperq = 0.975) {
  solfun <- function(beta) {
    omega <- beta[1]
    alpha <- beta[2]
    out1 <- psn(lower, xi = u, omega = omega, alpha = alpha) - lowerq
    out2 <- psn(upper, xi = u, omega = omega, alpha = alpha) - upperq
    return((out1^2+out2^2))
  }
  start <- c(((upper-lower)/4)^2, 0) + mvrnorm(100, mu = rep(0, 2), Sigma = diag(1, 2))
  start[, 1] <- abs(start[, 1])
  tempfit <- t(apply(start, 1, function(x) {
    fit <- optim(fn = solfun, par = x,
      control = list(maxit = 500000))
    return(c(fit$par, fit$value))
  })))
  tempfit <- tempfit[complete.cases(tempfit), ]
  return(tempfit[which(tempfit[, 3] == min(tempfit[, 3]))[1], 1:3])
}

totl <- 316283434
nmc <- 1*(10^7)
numerical_error <- NULL
dat$nerror_lb <- NA
dat$nerror_ub <- NA
dat$nharm_point <- dat$nerror * dat$peharmrate_point
```

Supplementary Materials (Supplemental File #2)

Page 27 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$nharm_lb <- NA
dat$nharm_ub <- NA
dat$harmrate_point <- dat$misrate_point * dat$peharmrate_point
dat$harmrate_lb <- NA
dat$harmrate_ub <- NA

# Monte Carlo simulation based on skew-normal estimation
mc_result <- list()
for (i in 1:nrow(dat)) {
  mc_result[[i]] <- list()
  if (dat$incidence_lb[i] == dat$incidence_ub[i]) {
    mc_incidence <- rep(dat$incidence_point[i], nmc)
  } else {
    fit_incidence <- snest(logit(dat$incidence_point[i] / totl),
                          lower = logit(dat$incidence_lb[i] / totl),
                          upper = logit(dat$incidence_ub[i] / totl))
    while(fit_incidence[3] > 10^(-10)) {
      fit_incidence <- snest(logit(dat$incidence_point[i] / totl),
                            lower = logit(dat$incidence_lb[i] / totl),
                            upper = logit(dat$incidence_ub[i] / totl))
    }
    mc_incidence <- expit(rsn(n = nmc, xi = logit(dat$incidence_point[i] / totl),
                             omega = fit_incidence[1], alpha = fit_incidence[2])) * totl
  }
  temp <- (c(quantile(mc_incidence, c(0.025, 0.975)) - c(dat$incidence_lb[i], dat$incidence_ub[i]))) /
  totl

  if (i %in% c(vascular_index, infection_index, cancer_index)) {
    fit_misrate <- snest(logit(dat$misrate_point[i]),
```

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
        lower = logit(dat$misrate_lb[i]),
        upper = logit(dat$misrate_ub[i])
while(fit_misrate[3] > 10^(-10)) {
  fit_misrate <- snest(logit(dat$misrate_point[i]),
    lower = logit(dat$misrate_lb[i]),
    upper = logit(dat$misrate_ub[i]))
}

mc_misrate <- expit(rsn(n = nmc, xi = logit(dat$misrate_point[i]),
  omega = fit_misrate[1], alpha = fit_misrate[2]))
temp <- c(temp, quantile(mc_misrate, c(0.025, 0.975)) - c(dat$misrate_lb[i], dat$misrate_ub[i]))

mc_result[[i]]$fit_misrate <- fit_misrate
mc_result[[i]]$mc_misrate <- mc_misrate
} else {
  temp <- c(temp, rep(NA, 2))
}

fit_peh <- snest(logit(dat$peharmrate_point[i]),
  lower = logit(dat$peharmrate_lb[i]),
  upper = logit(dat$peharmrate_ub[i]))
while(fit_peh[3] > 10^(-10)) {
  fit_peh <- snest(logit(dat$peharmrate_point[i]),
    lower = logit(dat$peharmrate_lb[i]),
    upper = logit(dat$peharmrate_ub[i]))
}

mc_peh <- expit(rsn(n = nmc, xi = logit(dat$peharmrate_point[i]),
  omega = fit_peh[1], alpha = fit_peh[2]))
temp <- c(temp, quantile(mc_peh, c(0.025, 0.975)) - c(dat$peharmrate_lb[i], dat$peharmrate_ub[i]))
```

*Supplementary Materials (Supplemental File #2)**Page 29 of 48*

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
numerical_error <- rbind(numerical_error, temp)

mc_result[[i]]$fit_incidence <- fit_incidence
mc_result[[i]]$mc_incidence <- mc_incidence

mc_result[[i]]$fit_peh <- fit_peh
mc_result[[i]]$mc_peh <- mc_peh
mc_result[[i]]$numerical_error <- numerical_error
}

summary(numerical_error) # 10e-5

# calculate plausible intervals for "OTHER" vascular/infection/censor harm rate (misrate)
# because point estimate is calculated as weighted sum and plausible intervals should be calculated
accordingly
temp1 <- 0
temp2 <- 0
for (i in vascular_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
mc_result[[vascular_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[vascular_other_index] <- quantile(mc_result[[vascular_other_index]]$mc_misrate,
0.025)
dat$misrate_ub[vascular_other_index] <- quantile(mc_result[[vascular_other_index]]$mc_misrate,
0.975)

temp1 <- 0
temp2 <- 0
```

*Supplementary Materials (Supplemental File #2)**Page 30 of 48*

*Burden of Serious Harms from Diagnostic Error in the USA**Newman-Toker et al., BMJQS 2023*

```
for (i in infection_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
mc_result[[infection_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[infection_other_index] <- quantile(mc_result[[infection_other_index]]$mc_misrate,
0.025)
dat$misrate_ub[infection_other_index] <- quantile(mc_result[[infection_other_index]]$mc_misrate,
0.975)

temp1 <- 0
temp2 <- 0
for (i in cancer_index) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
mc_result[[cancer_other_index]]$mc_misrate <- temp1 / temp2
dat$misrate_lb[cancer_other_index] <- quantile(mc_result[[cancer_other_index]]$mc_misrate, 0.025)
dat$misrate_ub[cancer_other_index] <- quantile(mc_result[[cancer_other_index]]$mc_misrate, 0.975)

# calculate plausible intervals for number of errors, number of harms and harm rate
# based on Monte Carlo generated data
for (i in 1:nrow(dat)) {
  mc_incidence <- mc_result[[i]]$mc_incidence
  mc_misrate <- mc_result[[i]]$mc_misrate
  mc_peh <- mc_result[[i]]$mc_peh
  dat$nerror_lb[i] <- quantile(mc_incidence * mc_misrate, 0.025)
```

*Supplementary Materials (Supplemental File #2)**Page 31 of 48*

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$error_ub[i] <- quantile(mc_incidence * mc_misrate, 0.975)

dat$nharm_lb[i] <- quantile(mc_incidence * mc_misrate * mc_peh, 0.025)
dat$nharm_ub[i] <- quantile(mc_incidence * mc_misrate * mc_peh, 0.975)

dat$harmrate_lb[i] <- quantile(mc_misrate * mc_peh, 0.025)
dat$harmrate_ub[i] <- quantile(mc_misrate * mc_peh, 0.975)
}

# calculate subtotals and plausible intervals
# create structure for subtotals
temp <- c("Top 5 Vascular Subtotal",
         "Top 5 Infection Subtotal",
         "Top 5 Cancer Subtotal",
         "Top 15 Subtotal",
         "Total Vascular",
         "Total Infection",
         "Total Cancer",
         "Big 3 Subtotal",
         "All Other Dx Errors",
         "Grand Total")
temp2 <- data.frame(matrix(NA, nrow = length(temp), ncol = ncol(dat)))
colnames(temp2) <- colnames(dat)
temp2[, 2] <- temp
temp2[, 1] <- rep("Subtotals", length(temp))
dat <- rbind(dat, temp2)

dat[, -c(1:2)] <- apply(dat[, -c(1:2)], c(1, 2), as.numeric)
```

Supplementary Materials (Supplemental File #2)

Page 32 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
# calculate subtotal incidence
dat$incidence_point[19] <- sum(dat$incidence_point[1:5])
dat$incidence_point[20] <- sum(dat$incidence_point[7:11])
dat$incidence_point[21] <- sum(dat$incidence_point[13:17])
dat$incidence_point[22] <- sum(dat$incidence_point[19:21])

dat$incidence_point[23] <- sum(dat$incidence_point[1:6])
dat$incidence_point[24] <- sum(dat$incidence_point[7:12])
dat$incidence_point[25] <- sum(dat$incidence_point[13:18])
dat$incidence_point[26] <- sum(dat$incidence_point[1:18])

temp <- lapply(1:18, function(i) return(mc_result[[i]]$mc_incidence))

temp1 <- rep(0, nmc)
for (i in 1:5) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[19] <- quantile(temp1, 0.025)
dat$incidence_ub[19] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
for (i in 7:11) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[20] <- quantile(temp1, 0.025)
dat$incidence_ub[20] <- quantile(temp1, 0.975)

temp1 <- rep(0, nmc)
```

Supplementary Materials (Supplemental File #2)

Page 33 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
for (i in 13:17) {  
  temp1 <- temp1 + temp[[i]]  
}  
dat$incidence_lb[21] <- quantile(temp1, 0.025)  
dat$incidence_ub[21] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in c(1:5, 7:11, 13:17)) {  
  temp1 <- temp1 + temp[[i]]  
}  
dat$incidence_lb[22] <- quantile(temp1, 0.025)  
dat$incidence_ub[22] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in 1:6) {  
  temp1 <- temp1 + temp[[i]]  
}  
dat$incidence_lb[23] <- quantile(temp1, 0.025)  
dat$incidence_ub[23] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in 7:12) {  
  temp1 <- temp1 + temp[[i]]  
}  
dat$incidence_lb[24] <- quantile(temp1, 0.025)  
dat$incidence_ub[24] <- quantile(temp1, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 34 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
for (i in 13:18) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[25] <- quantile(temp1, 0.025)
dat$incidence_ub[25] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)
for (i in 1:18) {
  temp1 <- temp1 + temp[[i]]
}
dat$incidence_lb[26] <- quantile(temp1, 0.025)
dat$incidence_ub[26] <- quantile(temp1, 0.975)
```

```
dat[27:28, 3:5] <- NA
```

```
# calculate subtotal number of errors
dat$nerror[19] <- sum(dat$nerror[1:5])
dat$nerror[20] <- sum(dat$nerror[7:11])
dat$nerror[21] <- sum(dat$nerror[13:17])
dat$nerror[22] <- sum(dat$nerror[19:21])
```

```
dat$nerror[23] <- sum(dat$nerror[1:6])
dat$nerror[24] <- sum(dat$nerror[7:12])
dat$nerror[25] <- sum(dat$nerror[13:18])
dat$nerror[26] <- sum(dat$nerror[1:18])
```

Supplementary Materials (Supplemental File #2)

Page 35 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$nerror[27] <- dat$nerror[26] / perror_big3 * (1 - perror_big3)
```

```
dat$nerror[28] <- dat$nerror[26] / perror_big3
```

```
temp1 <- rep(0, nmc)
```

```
for (i in 1:5) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
}
```

```
dat$nerror_lb[19] <- quantile(temp1, 0.025)
```

```
dat$nerror_ub[19] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
for (i in 7:11) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
}
```

```
dat$nerror_lb[20] <- quantile(temp1, 0.025)
```

```
dat$nerror_ub[20] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
for (i in 13:17) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
}
```

```
dat$nerror_lb[21] <- quantile(temp1, 0.025)
```

```
dat$nerror_ub[21] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)
```

Supplementary Materials (Supplemental File #2)

Page 36 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
for (i in c(1:5, 7:11, 13:17)) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[22] <- quantile(temp1, 0.025)  
dat$nerror_ub[22] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in 1:6) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[23] <- quantile(temp1, 0.025)  
dat$nerror_ub[23] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in 7:12) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[24] <- quantile(temp1, 0.025)  
dat$nerror_ub[24] <- quantile(temp1, 0.975)
```

```
temp1 <- rep(0, nmc)  
for (i in 13:18) {  
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate  
}  
dat$nerror_lb[25] <- quantile(temp1, 0.025)  
dat$nerror_ub[25] <- quantile(temp1, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 37 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
for (i in 1:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
}
dat$nerror_lb[26] <- quantile(temp1, 0.025)
dat$nerror_ub[26] <- quantile(temp1, 0.975)

dat$nerror_lb[27] <- quantile(temp1 / perror_big3 * (1 - perror_big3), 0.025)
dat$nerror_ub[27] <- quantile(temp1 / perror_big3 * (1 - perror_big3), 0.975)

dat$nerror_lb[28] <- quantile(temp1 / perror_big3, 0.025)
dat$nerror_ub[28] <- quantile(temp1 / perror_big3, 0.975)

# calculate subtotal misrate
dat$misrate_point[19:28] <- dat$nerror[19:28] / dat$incidence_point[19:28]

temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 1:5) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[19] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[19] <- quantile(temp1 / temp2, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 38 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 7:11) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[20] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[20] <- quantile(temp1 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 13:17) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[21] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[21] <- quantile(temp1 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in c(1:5, 7:11, 13:17)) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[22] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[22] <- quantile(temp1 / temp2, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 39 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 1:6) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[23] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[23] <- quantile(temp1 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 7:12) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[24] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[24] <- quantile(temp1 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 13:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[25] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[25] <- quantile(temp1 / temp2, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 40 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
for (i in 1:18) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp2 <- temp2 + mc_result[[i]]$mc_incidence
}
dat$misrate_lb[26] <- quantile(temp1 / temp2, 0.025)
dat$misrate_ub[26] <- quantile(temp1 / temp2, 0.975)

# calculate subtotal nharm, peh, and harm rate
dat$nharm_point[19] <- sum(dat$nharm_point[1:5])
dat$nharm_point[20] <- sum(dat$nharm_point[7:11])
dat$nharm_point[21] <- sum(dat$nharm_point[13:17])
dat$nharm_point[22] <- sum(dat$nharm_point[19:21])

dat$nharm_point[23] <- sum(dat$nharm_point[1:6])
dat$nharm_point[24] <- sum(dat$nharm_point[7:12])
dat$nharm_point[25] <- sum(dat$nharm_point[13:18])
dat$nharm_point[26] <- sum(dat$nharm_point[1:18])

dat$nharm_point[27] <- dat$nharm_point[26] / pharm_big3 * (1 - pharm_big3)
dat$nharm_point[28] <- dat$nharm_point[26] / pharm_big3

dat$harmrate_point[19:28] <- dat$nharm_point[19:28] / dat$incidence_point[19:28]
```

Supplementary Materials (Supplemental File #2)

Page 41 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$peharmrate_point[19:28] <- dat$nharm_point[19:28] / dat$nerror[19:28]
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 1:5) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[19] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[19] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[19] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[19] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[19] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[19] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 7:11) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

Supplementary Materials (Supplemental File #2)

Page 42 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}
```

```
dat$nharm_lb[20] <- quantile(temp3, 0.025)
dat$nharm_ub[20] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[20] <- quantile(temp3 / temp1, 0.025)
dat$harmrate_ub[20] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[20] <- quantile(temp3 / temp2, 0.025)
dat$peharmrate_ub[20] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 13:17) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}
```

```
dat$nharm_lb[21] <- quantile(temp3, 0.025)
dat$nharm_ub[21] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[21] <- quantile(temp3 / temp1, 0.025)
```

Supplementary Materials (Supplemental File #2)

Page 43 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$harmrate_ub[21] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[21] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[21] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in c(1:5, 7:11, 13:17)) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[22] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[22] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[22] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[22] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[22] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[22] <- quantile(temp3 / temp2, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 44 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 1:6) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
}

dat$harm_lb[23] <- quantile(temp3, 0.025)
dat$harm_ub[23] <- quantile(temp3, 0.975)

dat$harmrate_lb[23] <- quantile(temp3 / temp1, 0.025)
dat$harmrate_ub[23] <- quantile(temp3 / temp1, 0.975)

dat$peharmrate_lb[23] <- quantile(temp3 / temp2, 0.025)
dat$peharmrate_ub[23] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
temp2 <- rep(0, nmc)
temp3 <- rep(0, nmc)
for (i in 7:12) {
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

Supplementary Materials (Supplemental File #2)

Page 45 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

}

```
dat$nharm_lb[24] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[24] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[24] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[24] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[24] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[24] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 13:18) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[25] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[25] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[25] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[25] <- quantile(temp3 / temp1, 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 46 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$peharmrate_lb[25] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[25] <- quantile(temp3 / temp2, 0.975)
```

```
temp1 <- rep(0, nmc)
```

```
temp2 <- rep(0, nmc)
```

```
temp3 <- rep(0, nmc)
```

```
for (i in 1:18) {
```

```
  temp1 <- temp1 + mc_result[[i]]$mc_incidence
```

```
  temp2 <- temp2 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate
```

```
  temp3 <- temp3 + mc_result[[i]]$mc_incidence * mc_result[[i]]$mc_misrate * mc_result[[i]]$mc_peh
```

```
}
```

```
dat$nharm_lb[26] <- quantile(temp3, 0.025)
```

```
dat$nharm_ub[26] <- quantile(temp3, 0.975)
```

```
dat$harmrate_lb[26] <- quantile(temp3 / temp1, 0.025)
```

```
dat$harmrate_ub[26] <- quantile(temp3 / temp1, 0.975)
```

```
dat$peharmrate_lb[26] <- quantile(temp3 / temp2, 0.025)
```

```
dat$peharmrate_ub[26] <- quantile(temp3 / temp2, 0.975)
```

```
dat$nharm_lb[27] <- quantile(temp3 / pharm_big3 * (1 - pharm_big3), 0.025)
```

```
dat$nharm_ub[27] <- quantile(temp3 / pharm_big3 * (1 - pharm_big3), 0.975)
```

Supplementary Materials (Supplemental File #2)

Page 47 of 48

Burden of Serious Harms from Diagnostic Error in the USA

Newman-Toker et al., BMJQS 2023

```
dat$nharm_lb[28] <- quantile(temp3 / pharm_big3, 0.025)
dat$nharm_ub[28] <- quantile(temp3 / pharm_big3, 0.975)

dat$peharmrate_lb[27] <- quantile((temp3 / pharm_big3 * (1 - pharm_big3)) /
                                (temp2 / perror_big3 * (1 - perror_big3)), 0.025)
dat$peharmrate_ub[27] <- quantile((temp3 / pharm_big3 * (1 - pharm_big3)) /
                                (temp2 / perror_big3 * (1 - perror_big3)), 0.975)

dat$peharmrate_lb[28] <- quantile((temp3 / pharm_big3) / (temp2 / perror_big3), 0.025)
dat$peharmrate_ub[28] <- quantile((temp3 / pharm_big3) / (temp2 / perror_big3), 0.975)

dat$weight <- dat$peharmrate_point / peh_general
dat <- dat[, c(1:2, 9:10, 23, 3:5,
              6:8, 12:14, 20:22,
              11, 15, 16, 17:19)]

all(dat$misrate_lb < dat$misrate_point, na.rm = T)
all(dat$misrate_ub > dat$misrate_point, na.rm = T)
all(dat$harmrate_lb < dat$harmrate_point, na.rm = T)
all(dat$harmrate_ub > dat$harmrate_point, na.rm = T)
all(dat$nharm_lb < dat$nharm_point, na.rm = T)
all(dat$nharm_ub > dat$nharm_point, na.rm = T)
all(dat$nerror_lb < dat$nerror, na.rm = T)
all(dat$nerror_ub > dat$nerror, na.rm = T)

write.csv(dat, file = "big3_results_combo_updated.csv")
```