

# NLA and Microcystin

*Jeff W. Hollister*

*3/1/2016*

Review of some numbers from the 2007 NLA data on microcystin.

```
nla_url <- "http://www2.epa.gov/sites/production/files/2014-10/nla2007_recreational_conditionestimates_1"
nla_dat <- read.csv(nla_url)
```

The raw dataset from the `nla_url` has 1252 rows and 48.

Some filtering of this data is required. For our analyses we prefer to use the probabilistic draw in case we want to use the weighted values. Do this we need to only use `PROB_Lakes` and the first visit.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
nla_dat <- nla_dat %>%
  filter(SITE_TYPE == "PROB_Lake") %>%
  filter(VISIT_NO == 1) %>%
  select(SITE_ID, CHLA, MCYST_TL_UGL)
```

Filtering this out results in 1033 rows and 3.

From this we can see how many lakes had microcystin present.

```
sum(nla_dat$MCYST_TL_UGL > 0.05)
```

```
## [1] 344
```

```
sum(nla_dat$MCYST_TL_UGL > 0.05)/nrow(nla_dat)
```

```
## [1] 0.3330106
```