Notes on calculating combined GAM estimates within the Rapid Assessment Method (RAM)

15 February 2020

Contents

1 Background

2 Possible approach

PROBIT gives a probability so we look to combining two probabilities:

$$P(GAM_{\text{MUAC}} \cup GAM_{\text{WHZ}}) = P(GAM_{\text{MUAC}}) + P(GAM_{\text{WHZ}})$$

However, the problem is that we do not have `independent` probabilities. We overestimate because the intersection gets counted twice. Therefore we need:

$$P(GAM_{\text{MUAC}} \cup GAM_{\text{WHZ}}) = P(GAM_{\text{MUAC}}) + P(GAM_{\text{WHZ}}) - P(GAM_{\text{MUAC}} \cap GAM_{\text{WHZ}})$$

We have the first two terms but not the third. We can estimate the third term from a 2 by 2 table:

	WHZ < -2	**WHZ** ≥ -2
MUAC < 125	a	b
MUAC ≥ 125	c	d

and

$$P(GAM_{\text{MUAC}} \cap GAM_{\text{WHZ}}) = \frac{a}{a + b + c + d}$$

We have a small sample size so the estimate will lack precision but I think that being "clever" and using something like:

$$P(GAM_{\text{MUAC}} \cap GAM_{\text{WHZ}}) = P(GAM_{\text{MUAC}}) \times P(GAM_{\text{WHZ}})$$

will not work as it assumes independence.

We can try to move forward with this hybrid method.

We try this in R using a dataset from Uganda.

```
## Read dataset
x <- read.table(file = "data/ugan01.csv", header = TRUE, sep = ",")
```

We then create case definitions.

```
## Case definitions
x$gamWHZ <- ifelse(x$whz < -2, 1, 2)                ## GAM by WHZ
x$gamMUAC <- ifelse(x$muac < 125, 1, 2)             ## GAM by MUAC
x$cGAM <- ifelse(x$whz < -2 | x$muac < 125, 1, 2)  ## GAM by WHZ and MUAC
```

Calculating for prevalence in the classical approach we get:

```
## Classic prevalence for GAM by MUAC
round(prop.table(table(x$gamMUAC))[1] * 100, 2)
```

```
##    1
## 13.8
```

```
## Classic prevalence for GAM by WHZ
round(prop.table(table(x$gamWHZ))[1] * 100, 2)
```

```
##    1
## 9.05
```

```
## Classic prevalence for GAM by WHZ and MUAC
round(prop.table(table(x$cGAM))[1] * 100, 2)
```

```
##    1
## 15.5
```

We can test whether GAM cases by MUAC and GAM cases by WHZ are independent.

```
## Test if the two case definitions are independent
chisq.test(table(x$gamMUAC, x$gamWHZ))
```

The chi-square test has a p-value of $8.8108361 \times 10^{-74}$ indicating that the two case definitions are not independent.

We then proceed with our proposed hybrid approach using simple PROBIT prevalence estimation.

```
## Simple PROBIT prevalence for GAM by MUAC
pMUAC <- pnorm(125, mean(x$muac), sd(x$muac))
```

```
## [1] 13.65
```

```
## Simple PROBIT prevalence for GAM by WHZ
pWHZ <- pnorm(-2, mean(x$whz), sd(x$whz))
```

```
## [1] 7.61
```

```
## Estimate the UNION probability
pUNION <- table(x$gamMUAC, x$gamWHZ)[1,1] / sum(table(x$gamMUAC, x$gamWHZ))
```

```
## [1] 7.35
```

```
## Estimate of GAM by MUAC and WHZ by PROBIT
round((pMUAC + pWHZ - pUNION) * 100, 2)
```

```
## [1] 13.91
```