efa\_to\_run.RMD

# Filter data, remove ids with too many missings, also remove outliers

set.seed(2124)  
res = filter\_multi\_df\_outliers\_missing(  
 df=df,  
 value\_cols=q\_cols)

##   
## iter imp variable  
## 1 1  
## 1 2  
## 1 3  
## 1 4  
## 1 5  
## 2 1  
## 2 2  
## 2 3  
## 2 4  
## 2 5  
## 3 1  
## 3 2  
## 3 3  
## 3 4  
## 3 5  
## 4 1  
## 4 2  
## 4 3  
## 4 4  
## 4 5  
## 5 1  
## 5 2  
## 5 3  
## 5 4  
## 5 5

df = res[["df"]]  
res[["missing\_df"]]

## [1] id q7 q8 q9 q10 q11 q13 q15 q16 q17 q18 q19 q20 q22 q23 q24 q25 q26 q27  
## [20] q28 q29 q30 q31 q32 q33 q35 q36 q37 q39 q40  
## <0 rows> (or 0-length row.names)

res[["outlier\_df"]][["id"]]

## [1] "patient12" "patient15" "patient17" "patient18" "patient21"   
## [6] "patient51" "patient59" "patient84" "patient93" "patient94"   
## [11] "patient95" "patient105" "patient127" "patient129" "patient151"  
## [16] "patient155" "patient162" "patient167" "patient171" "patient196"

# Check EFA assumptions

res = check\_efa\_assumptions(  
 df=df,  
 value\_cols=q\_cols,  
 figs\_path=NULL)  
res[["barlet\_test"]][["p.value"]]

## [1] 0

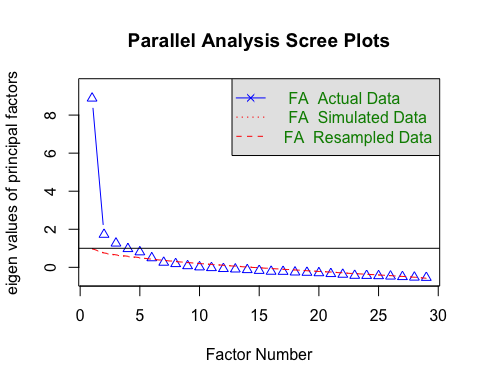
res[["kaiser\_sampling\_adequacy"]]

## Kaiser-Meyer-Olkin factor adequacy  
## Call: psych::KMO(r = cor\_mat)  
## Overall MSA = 0.88  
## MSA for each item =   
## q7 q8 q9 q10 q11 q13 q15 q16 q17 q18 q19 q20 q22 q23 q24 q25   
## 0.88 0.91 0.86 0.84 0.87 0.88 0.91 0.90 0.91 0.79 0.78 0.86 0.91 0.84 0.93 0.81   
## q26 q27 q28 q29 q30 q31 q32 q33 q35 q36 q37 q39 q40   
## 0.88 0.90 0.82 0.86 0.90 0.90 0.93 0.90 0.91 0.91 0.92 0.85 0.84

# res[["plt\_func"]]()

# Get number of factors using parallel analysis

res = get\_factor\_num(  
 df=df,  
 value\_cols=q\_cols,  
 figs\_path=NULL)



## Parallel analysis suggests that the number of factors = 6 and the number of components = NA

factor\_num = res[["kaiser\_factor\_num\_new"]]  
factor\_num

## [1] 5

# Fit the EFA with those number of factors and remove items

res = fit\_fa\_model(  
 df=df,  
 factor\_num=factor\_num,  
 value\_cols=q\_cols)  
  
df\_filtered = res[["df\_filtered"]]  
res[["tuker\_lewis\_index"]]

## [1] 0.8713995

res[["rmse"]]

## [1] 0.04108489

res[["rejected\_items"]]

## [1] "q15"

res[["cfi"]]

## [1] 0.9161316

accepted\_items = res[["accepted\_items"]]

# Fit the EFA using filtered data, this time no item is rejected  
res = fit\_fa\_model(  
 df=df\_filtered,  
 factor\_num=factor\_num,  
 value\_cols=accepted\_items,  
 rotate="oblimin",  
 fm="ml")  
res[["tuker\_lewis\_index"]]

## [1] 0.8679193

res[["rmse"]]

## [1] 0.04045727

res[["rejected\_items"]]

## character(0)

res[["cfi"]]

## [1] 0.9153118

res[["factor\_item\_list"]]

## [[1]]  
## [1] "q35" "q36" "q37" "q39" "q40"  
##   
## [[2]]  
## [1] "q7" "q8" "q9" "q10" "q11" "q13"  
##   
## [[3]]  
## [1] "q16" "q17" "q30" "q31" "q32" "q33"  
##   
## [[4]]  
## [1] "q18" "q19" "q20" "q26" "q27" "q28" "q29"  
##   
## [[5]]  
## [1] "q22" "q23" "q24" "q25"