Om in boxplots iets te kunnen vergelijken, wil je niet te veel data in een figuur hebben. Zoals hieronder. Je ziet dat bij sommige taken (beelden) van dit subject de KW’s grotere verschillen in outer oppervlakte hebben aangegeven. Maar voor de rest kun je hier denk ik niet veel zinnigs uit halen, want het doel is om op niveau van subject verschillen te kunnen ontdekken.

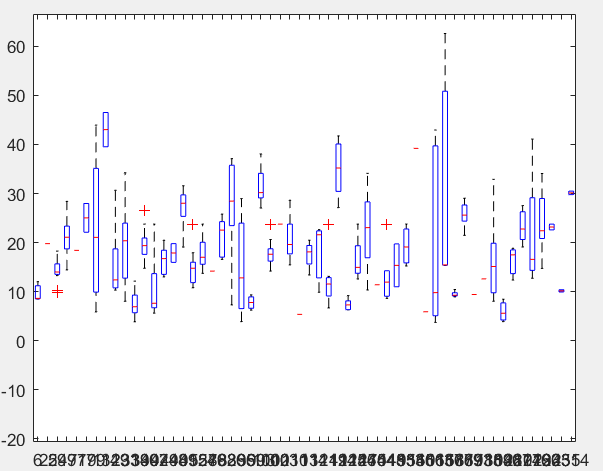


Figure 1 Een box geeft de outer area van meerdere annotaties weer. Van iedere taak van een subject is een box gemaakt.

Door voor met outer area en inner area het wall area percentage (WAP) te berekenen, krijg je per annotatie een getal dat iets over die annotatie verteld. Dit kan dan vergeleken worden met de WAP die de expert bij die bepaalde afbeelding heeft geannoteerd.

De gedachte was dat de gemiddelde WAP van alle tasks van een persoon ongeveer gelijk zou moeten zijn bij KW en expert.

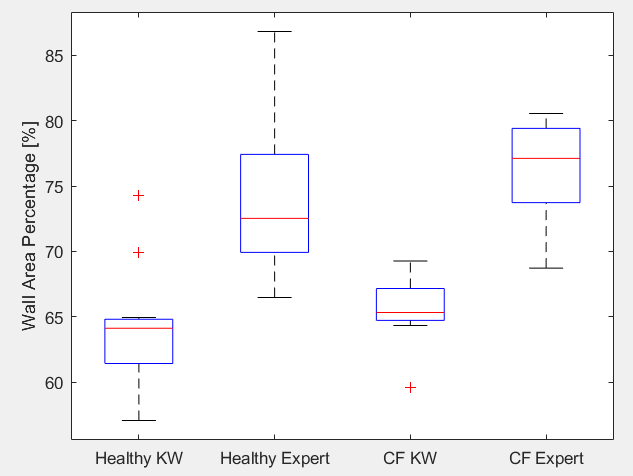
Ik heb een boxplot figuur gemaakt met de hoop dat KW en expert per groep bij elkaar in de buurt zouden liggen. Maar de WAP van de subject door de expert ligt hoger.

Figure 2 WAP of healthy subjects determined by KW, healthy subjects determined by expert and the same for subjects with CF

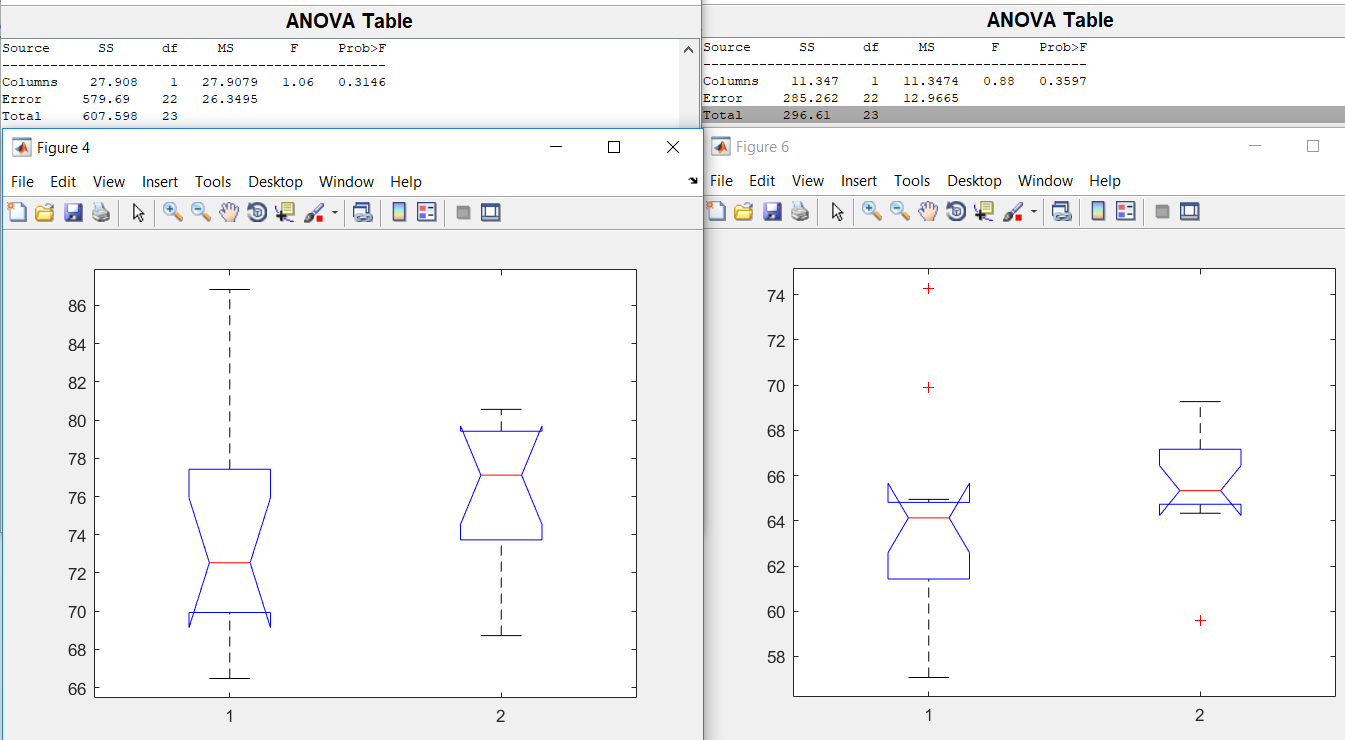
De Mann-Whitney U-test (ook wel wilcoxon rank sum test ) werd in een artikel gebruikt om te bepalen of de airway wall area significant verschilt tussen twee groepen.

De p-value bij zowel expert (p = 0.1749) als KW (p = 0.0606) geven aan dat dit er geen bewijs is voor een significant verschil tussen de WAP bij niet-CF en CF subjects.

Daarnaast heb ik nog een ANOVA test gedaan want bij matlab stond dat:

one-way ANOVA enables you to find out whether different groups of an independent variable have different effects on the response variable y.

Hier zijn de p values bij expert p=0.3146 (links) en KW p=0.3597 (rechts) ook boven 0.05. Dit bewijst dus ook dat de groepen hetzelfde zijn en er geen significant verschil te merken is tussen de WAP van beide groepen.



== Misschien nog een scatterplot maken van WAP KW tegen WAP Expert en waarbij ieder punt één task is ipv. Subject.

Omdat ik niet zo goed weet wat ik nog verder zou kunnen doen met WAP en de bijbehorende resultaten, ben ik gaan kijken naar MAPE, zoals voorgesteld was.

Mean absolute percentage error (MAPE) wordt normaal gebruikt om verschillende tijdpunten te vergelijken. Bij ons zijn de tijdpunten dan eigenlijk de annotaties door verschillende KW’s. In ‘mape.m’ bestand heb ik de formule ingevuld en voor ieder subject een MAPE berekend. Dit getal geldt voor het hele subject zodat op dat niveau vergeleken kan worden tussen CF en no-CF subjects.

De resultaten staan hieronder in twee tabellen.

Opvallend is dat een heel aantal hoger is dan 100, terwijl het een percentage is.

|  |  |
| --- | --- |
| No CF | MAPE |
| 1 | 97.1001060036229 |
| 3 | 61.5076497897558 |
| 6 | 76.9555364383915 |
| 8 | 141.690093883635 |
| 12 | 139.591842937706 |
| 16 | 110.520336151889 |
| 17 | 61.2538744115010 |
| 21 | 136.495613796560 |
| 23 | 77.7568513523065 |
| 24 | 79.6073338415467 |
| 25 | 102.667630031183 |
| 27 | 197.383284031532 |

|  |  |
| --- | --- |
| CF | MAPE |
| 2 | 137.462994352946 |
| 4 | 117.370826903098 |
| 5 | 76.8214678310694 |
| 7 | 111.510683554979 |
| 10 | 152.318488258162 |
| 11 | 86.7119241787485 |
| 13 | 100.850280996083 |
| 18 | 94.1251968945500 |
| 19 | 140.042387252625 |
| 26 | 75.6296432948038 |
| 28 | 171.734578180387 |
| 41 | 191.481833969997 |