

Creating a support system for people with diabetes: how to deal with diabetes distress from social sources

Underlying Analyses

Mitchell Kesteloo & Merijn Bruijnes

09 augustus, 2021

```
knitr::opts_chunk$set(echo = TRUE)

library(mediation) # Mediation package

## Warning: package 'mediation' was built under R version 3.6.3
## Loading required package: MASS
## Loading required package: Matrix
## Loading required package: mvtnorm
## Loading required package: sandwich
## mediation: Causal Mediation Analysis
## Version: 4.5.0

library(rockchalk) # Graphing simple slopes; moderation

## Warning: package 'rockchalk' was built under R version 3.6.3
##
## Attaching package: 'rockchalk'
## The following object is masked from 'package:MASS':
##
##      mvrnorm

library(multilevel) # Sobel Test

## Warning: package 'multilevel' was built under R version 3.6.3
## Loading required package: nlme

library(gvlma) # Testing Model Assumptions
library(stargazer) # Handy regression tables

##
## Please cite as:
## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
## R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

library(ggplot2) # to check distribution visually
library(dplyr) # to check difference between groups visually
```

```

##
## Attaching package: 'dplyr'
## The following object is masked from 'package:nlme':
##
## collapse
## The following object is masked from 'package:rockchalk':
##
## summarize
## The following object is masked from 'package:MASS':
##
## select
## The following objects are masked from 'package:stats':
##
## filter, lag
## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union
library(ggpubr) # boxplots are easier with this lib

## Warning: package 'ggpubr' was built under R version 3.6.3
library(lsr) # calculate effect size
library(pwr) # power analysis

## Warning: package 'pwr' was built under R version 3.6.3
library(psych) #alpha function

##
## Attaching package: 'psych'
## The following objects are masked from 'package:ggplot2':
##
## %+%, alpha
## The following object is masked from 'package:mediation':
##
## mediate
library(car) # Levene's test

## Loading required package: carData
## Warning: package 'carData' was built under R version 3.6.3
##
## Attaching package: 'car'
## The following object is masked from 'package:psych':
##
## logit
## The following object is masked from 'package:dplyr':
##
## recode

```

```
library(pander)
panderOptions("table.alignment.default", "left")

set.seed(20210809) # set the seed so that (monte carlo) simulations will yeild same results for all, se
```

Contents

1	Introduction	3
2	Data files and data preparation	4
2.1	File data_chatbot.csv	4
3	Participants	7
3.1	Quality of data	7
3.2	Participants profile	8
4	Results	8
4.1	Reliability testing of questionnaires	8
4.1.1	Reliability diabetes distress scale	9
4.1.2	Reliability CSQ-8	10
4.1.3	Reliability feeling of being heard	10
4.1.4	Reliability involvement	11
4.2	Testing system usability using SUS	12
4.3	Testing hypothesis 1: People using the conversational agent will have a larger reduction in diabetes distress than the control group	12
4.3.1	Testing assumptions for H1	12
4.3.2	Result for hypothesis 1	14
4.4	Testing hypothesis 2: The effect of the social help program type on the diabetes distress difference is mediated by the attitude towards the social help program	16
4.4.1	Testing assumptions for H2	16
4.4.2	Results H2	24
4.5	Testing hypothesis 3: People who have higher initial diabetes distress will have a larger re-duction in diabetes distress than people with low initial diabetes distress	25
4.5.1	Testing assumptions for H3	25
4.5.2	Results H3	29
4.6	Testing hypothesis 4: People who use the conversational agent have a higher feeling of being heard than the control group	30
4.6.1	Testing assumptions for H4	30
4.6.2	Result for hypothesis 4	31
4.7	Exploratory research	31
5	R version information	40

1 Introduction

This document presents the statistical analysis of the participants' diabetes-related distress difference based on the type of intervention. Furthermore, a moderated mediation analysis is done as reported in the thesis report:

Creating a support system for people with diabetes: how to deal with diabetes distress from social sources.

Authored by Mitchell Kesteloo, edited by Merijn Bruijnes.

2 Data files and data preparation

To make sure the resulting data set does not contain any sensitive information, we combined the csv files generated by Prolific, session 1, session 2 and session 3 into one csv file. For each session, we have two csv files: the control group stayed in one Qualtrics survey, but the treatment group was redirected to a new one since they were first redirected to the frontend on our TU Delft server. The participants in the treatment group were redirected to another Qualtrics survey after finishing the conversation. The resulting data set and its variables are shown in Section 2.1 below.

2.1 File data_chatbot.csv

Data containing information about the participants (age, gender), their answers to the DD-scale before exposure to the social help program type and their answers to the DD-scale, CSQ-8, FBH, SUS and involvement questions after the last session. It also includes any comments and tips given from each session.

Table 1: Fields and labels from csv file data_chatbot.csv

variable	label
id	identification participant
age	age of the participant
sex	sex of participant
program_type	program type appointed to participant
pre_diabetes_distress	the diabetes distress score before exposure to the social help program
diabetes_distress_diff	the difference between the diabetes distress scores from before exposure to the program and after exposure
post_diabetes_distress	the diabetes distress score after exposure to the social help program
attitude_to_program	the attitude of the participant towards the social help program
feeling_of_being_heard	participant's feeling of being heard by the social help program
pre_DD.scale_1	participants' answer to pre exposure diabetes distress question: Feeling that your friends and family are not supportive enough of your diabetes management efforts?
pre_DD.scale_2	participants' answer to pre exposure diabetes distress question: Feeling that friends or family don't appreciate how difficult living with diabetes can be
pre_DD.scale_3	participants' answer to pre exposure diabetes distress question: Feeling that friends or family don't give me the emotional support that I would like
pre_DD.scale_4	participants' answer to pre exposure diabetes distress question: Feeling that my family and friends make a bigger deal out of diabetes than they should
pre_DD.scale_5	participants' answer to pre exposure diabetes distress question: Feeling that my friends and family worry more about hypoglycemia than I want them to

variable	label
pre_DD.scale_6	participants' answer to pre exposure diabetes distress question: Feeling that my friends or family treat me as if I were more fragile or sicker than I really am
pre_DD.scale_7	participants' answer to pre exposure diabetes distress question: Feeling that my friends or family act like "diabetes police" (bother me too much)
pre_DD.scale_8	participants' answer to pre exposure diabetes distress question: Feeling that people treat me differently when they find out I have diabetes
pre_DD.scale_9	participants' answer to pre exposure diabetes distress question: Feeling like I have to hide my diabetes from other people
pre_DD.scale_10	participants' answer to pre exposure diabetes distress question: Feeling that people will think less of me if they knew I had diabetes
pre_DD.scale_11	participants' answer to pre exposure diabetes distress question: Feeling concerned that diabetes may make me less attractive to employers
post_DD.scale_1	participants' answer to post exposure diabetes distress question: Feeling that your friends and family are not supportive enough of your diabetes management efforts?
post_DD.scale_2	participants' answer to post exposure diabetes distress question: Feeling that friends or family don't appreciate how difficult living with diabetes can be
post_DD.scale_3	participants' answer to post exposure diabetes distress question: Feeling that friends or family don't give me the emotional support that I would like
post_DD.scale_4	participants' answer to post exposure diabetes distress question: Feeling that my family and friends make a bigger deal out of diabetes than they should
post_DD.scale_5	participants' answer to post exposure diabetes distress question: Feeling that my friends and family worry more about hypoglycemia than I want them to
post_DD.scale_6	participants' answer to post exposure diabetes distress question: Feeling that my friends or family treat me as if I were more fragile or sicker than I really am
post_DD.scale_7	participants' answer to post exposure diabetes distress question: Feeling that my friends or family act like "diabetes police" (bother me too much)
post_DD.scale_8	participants' answer to post exposure diabetes distress question: Feeling that people treat me differently when they find out I have diabetes

variable	label
post_DD.scale_9	participants' answer to post exposure diabetes distress question: Feeling like I have to hide my diabetes from other people
post_DD.scale_10	participants' answer to post exposure diabetes distress question: Feeling that people will think less of me if they knew I had diabetes
post_DD.scale_11	participants' answer to post exposure diabetes distress question: Feeling concerned that diabetes may make me less attractive to employers
fbh_1	participants' answer to the question: The social help program really addressed your needs
fbh_2	participants' answer to the question: The social help program took you seriously
fbh_3	participants' answer to the question: You're satisfied with the (emotional) support you received from the social help program
fbh_4	participants' answer to the question: The social help program replied appropriately to you
fbh_5	participants' answer to the question: The social help program paid full attention to what you were trying to tell him/her
fbh_6	participants' answer to the question: The social help program listened to your preferences
fbh_7	participants' answer to the question: The social help program only thought about what is best for you
csq.8_1	participants' answer to the question: How would you rate the quality of the services you have received?
csq.8_2	participants' answer to the question: Did you get the kind of service you wanted?
csq.8_3	participants' answer to the question: To what extent has our program met your needs?
csq.8_4	participants' answer to the question: If a friend were in need of similar help, would you recommend our program to him or her?
csq.8_5	participants' answer to the question: How satisfied are you with the amount of help you have received?
csq.8_6	participants' answer to the question: Has the service you received helped you to deal more effectively with problems/difficulties?
csq.8_7	participants' answer to the question: In an overall general sense, how satisfied are you with the services you have received?
csq.8_8	participants' answer to the question: If you were to seek help again, would you come back to our program?
involvement_1	participants' answer to the question: I find the topic of socially related diabetes distress interesting

variable	label
involvement_2	participants' answer to the question: I find the topic of socially related diabetes distress involving
involvement_3	participants' answer to the question: I find the topic of socially related diabetes distress personally relevant
sus_1	treatment group participants' answer to the question: I think that I would like to use this system frequently
sus_2	treatment group participants' answer to the question: I found the system unnecessarily complex
sus_3	treatment group participants' answer to the question: I thought the system was easy to use
sus_4	treatment group participants' answer to the question: I think that I would need the support of a technical person to be able to use this system
sus_5	treatment group participants' answer to the question: I found the various functions in this system were well integrated
sus_6	treatment group participants' answer to the question: I thought there was too much inconsistency in this system
sus_7	treatment group participants' answer to the question: I would imagine that most people would learn to use this system very quickly
sus_8	treatment group participants' answer to the question: I found the system very cumbersome to use
sus_9	treatment group participants' answer to the question: I felt very confident using the system
sus_10	treatment group participants' answer to the question: I needed to learn a lot of things before I could get going with this system
comments_sess1	participants' comments on the first session
comments_sess2	participants' comments on the second session
comments_sess3	participants' comments on the third session
tips_sess3	participants' tips on the complete study

3 Participants

3.1 Quality of data

Participants were recruited via the Prolific platform. The platform is similar to Amazon Mechanical Turk, but Prolific is purely focused on recruiting participants for scientific research. Every participant received a small monetary payment for their participation (8.4 pounds per hour).

As explained in the thesis report, there are always people on platforms like these who are interested in making money as quickly as possible. To counter this, most studies include some form of attention checks. In this study, there were multiple attention checks: 1 in the first session, 1 in the second session and 2 in the last session. If someone failed any of the attention checks, their submission was rejected and their data was

removed from the data set.

We investigated whether there are any significant differences between the two groups by using Kruskal-Wallis tests. There is one significant difference between groups: there are more men in the control group than in the agent group. We checked whether this difference had any influence on the results. A linear model was fitted where the diabetes distress difference is explained by the gender of the participant. The gender-coefficient's p-value is 0.863, implicating that this variable does not influence the results.

3.2 Participants profile

Table 2: Participants profile

Participants	Control group	Treatment group	Total
Number, n	77	79	156
-Male, n(%)	48 (62.3 %)	34 (43 %)	82 (52.6 %)
Age			
-Mean (SD)	48 (62.3 %)	34 (43 %)	NA (NA)
-Range	48 (62.3 %)	34 (43 %)	(28,38] - [18,28]
Pre diabetes distress			
-Mean (SD)	2.4 (1)	2.7 (1.1)	2.6 (1.1)
-Range	1 - 5.4	1 - 5.1	1 - 5.4

Table 3: Differences between groups on age, gender and pre-measurement diabetes distress (continued below)

statistic	parameter	p.value	method
4.331771	5	0.5027023	Kruskal-Wallis rank sum test
5.78768	1	0.01613888	Kruskal-Wallis rank sum test
43.92594	42	0.3899089	Kruskal-Wallis rank sum test

data.name

program_type by age
program_type by sex
program_type by
pre_diabetes_distress

```
##          Df Sum Sq Mean Sq F value Pr(>F)
## sex      1    0.02  0.0202    0.03  0.863
## Residuals 154 104.03  0.6755
```

4 Results

4.1 Reliability testing of questionnaires

All questionnaires are checked on their internal reliability. As seen below, all have a Cronbach alpha higher than 0.7 showing decent internal reliability.

4.1.1 Reliability diabetes distress scale

```
##
## Reliability analysis
## Call: alpha(x = subset(data, select = c("pre_DD.scale_1", "pre_DD.scale_2",
##     "pre_DD.scale_3", "pre_DD.scale_4", "pre_DD.scale_5", "pre_DD.scale_6",
##     "pre_DD.scale_7", "pre_DD.scale_8", "pre_DD.scale_9", "pre_DD.scale_10",
##     "pre_DD.scale_11")))
##
##      raw_alpha std.alpha G6(smc) average_r S/N   ase mean  sd median_r
##      0.89      0.89      0.92      0.42 8.1 0.013  2.6 1.1      0.4
##
## lower alpha upper      95% confidence boundaries
## 0.86 0.89 0.92
##
## Reliability if an item is dropped:
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
## pre_DD.scale_1      0.88      0.88      0.91      0.42 7.4      0.014 0.023 0.40
## pre_DD.scale_2      0.88      0.88      0.90      0.42 7.1      0.015 0.024 0.40
## pre_DD.scale_3      0.88      0.88      0.91      0.42 7.1      0.015 0.025 0.40
## pre_DD.scale_4      0.89      0.89      0.91      0.44 7.8      0.014 0.022 0.41
## pre_DD.scale_5      0.87      0.87      0.90      0.41 6.9      0.015 0.024 0.40
## pre_DD.scale_6      0.89      0.89      0.91      0.44 7.8      0.014 0.024 0.41
## pre_DD.scale_7      0.89      0.89      0.91      0.44 7.9      0.014 0.020 0.41
## pre_DD.scale_8      0.89      0.89      0.91      0.44 7.8      0.014 0.024 0.41
## pre_DD.scale_9      0.88      0.88      0.91      0.42 7.1      0.015 0.025 0.40
## pre_DD.scale_10     0.88      0.88      0.90      0.41 7.0      0.015 0.025 0.40
## pre_DD.scale_11     0.88      0.88      0.91      0.42 7.1      0.015 0.026 0.40
##
## Item statistics
##      n raw.r std.r r.cor r.drop mean  sd
## pre_DD.scale_1 156 0.67 0.68 0.66 0.60 2.4 1.3
## pre_DD.scale_2 156 0.73 0.74 0.73 0.66 2.5 1.5
## pre_DD.scale_3 156 0.75 0.74 0.72 0.68 2.7 1.6
## pre_DD.scale_4 156 0.61 0.60 0.57 0.51 2.5 1.5
## pre_DD.scale_5 156 0.79 0.78 0.77 0.73 2.8 1.6
## pre_DD.scale_6 156 0.60 0.60 0.56 0.51 2.3 1.5
## pre_DD.scale_7 156 0.58 0.59 0.55 0.49 2.4 1.5
## pre_DD.scale_8 156 0.62 0.61 0.57 0.53 2.7 1.5
## pre_DD.scale_9 156 0.75 0.74 0.72 0.67 3.0 1.8
## pre_DD.scale_10 156 0.76 0.75 0.73 0.69 2.7 1.6
## pre_DD.scale_11 156 0.72 0.73 0.70 0.65 2.6 1.5
##
## Non missing response frequency for each item
##      1 2 3 4 5 6 miss
## pre_DD.scale_1 0.29 0.34 0.19 0.09 0.06 0.03 0
## pre_DD.scale_2 0.33 0.29 0.13 0.12 0.08 0.04 0
## pre_DD.scale_3 0.35 0.17 0.16 0.15 0.13 0.04 0
## pre_DD.scale_4 0.41 0.12 0.24 0.10 0.07 0.06 0
## pre_DD.scale_5 0.28 0.24 0.19 0.12 0.12 0.06 0
## pre_DD.scale_6 0.45 0.23 0.11 0.08 0.08 0.05 0
## pre_DD.scale_7 0.38 0.22 0.15 0.13 0.08 0.03 0
## pre_DD.scale_8 0.28 0.22 0.20 0.15 0.09 0.06 0
## pre_DD.scale_9 0.35 0.10 0.17 0.13 0.12 0.13 0
```

```
## pre_DD.scale_10 0.28 0.26 0.17 0.10 0.13 0.06    0
## pre_DD.scale_11 0.31 0.22 0.17 0.17 0.08 0.05    0
```

4.1.2 Reliability CSQ-8

```
##
## Reliability analysis
## Call: alpha(x = subset(data, select = c("csq.8_1", "csq.8_2", "csq.8_3",
##     "csq.8_4", "csq.8_5", "csq.8_6", "csq.8_7", "csq.8_8")))
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.92      0.92    0.92      0.59 11 0.0098 3.1 0.55      0.58
##
## lower alpha upper      95% confidence boundaries
## 0.9 0.92 0.94
##
## Reliability if an item is dropped:
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## csq.8_1      0.91      0.91    0.91      0.60 10.4  0.011 0.0033 0.58
## csq.8_2      0.91      0.91    0.91      0.59 10.1  0.011 0.0036 0.58
## csq.8_3      0.91      0.91    0.91      0.59 10.2  0.011 0.0035 0.58
## csq.8_4      0.91      0.91    0.90      0.59 10.2  0.011 0.0030 0.59
## csq.8_5      0.91      0.91    0.90      0.59 10.0  0.011 0.0034 0.58
## csq.8_6      0.91      0.91    0.90      0.59 10.1  0.011 0.0027 0.58
## csq.8_7      0.90      0.90    0.90      0.57  9.4  0.012 0.0029 0.57
## csq.8_8      0.90      0.90    0.90      0.58  9.5  0.012 0.0029 0.57
##
## Item statistics
##      n raw.r std.r r.cor r.drop mean   sd
## csq.8_1 156 0.76 0.77 0.72 0.69 3.1 0.66
## csq.8_2 156 0.78 0.79 0.75 0.72 3.1 0.60
## csq.8_3 156 0.79 0.78 0.74 0.71 2.9 0.78
## csq.8_4 156 0.79 0.79 0.75 0.72 3.3 0.73
## csq.8_5 156 0.80 0.80 0.76 0.73 3.2 0.69
## csq.8_6 156 0.79 0.79 0.75 0.72 3.1 0.65
## csq.8_7 156 0.84 0.85 0.83 0.79 3.2 0.66
## csq.8_8 156 0.85 0.84 0.82 0.78 3.2 0.79
##
## Non missing response frequency for each item
##      1    2    3    4 miss
## csq.8_1 0.00 0.16 0.54 0.29  0
## csq.8_2 0.01 0.08 0.67 0.24  0
## csq.8_3 0.04 0.21 0.51 0.24  0
## csq.8_4 0.03 0.06 0.47 0.44  0
## csq.8_5 0.00 0.15 0.49 0.35  0
## csq.8_6 0.01 0.15 0.59 0.25  0
## csq.8_7 0.00 0.14 0.54 0.32  0
## csq.8_8 0.04 0.12 0.46 0.38  0
```

4.1.3 Reliability feeling of being heard

```
##
## Reliability analysis
## Call: alpha(x = subset(data, select = c("fbh_1", "fbh_2", "fbh_3",
##     "fbh_4", "fbh_5", "fbh_6", "fbh_7")))
##
```

```
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##     0.79      0.8      0.8      0.36   4 0.026  4.7 0.93     0.34
##
##   lower alpha upper      95% confidence boundaries
## 0.74 0.79 0.84
##
## Reliability if an item is dropped:
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## fbh_1      0.78      0.79      0.77      0.38 3.7   0.027 0.0097  0.37
## fbh_2      0.76      0.76      0.75      0.35 3.2   0.030 0.0114  0.33
## fbh_3      0.76      0.77      0.76      0.35 3.3   0.031 0.0124  0.33
## fbh_4      0.79      0.79      0.79      0.39 3.9   0.026 0.0096  0.38
## fbh_5      0.75      0.75      0.73      0.33 3.0   0.032 0.0080  0.33
## fbh_6      0.76      0.77      0.76      0.35 3.3   0.030 0.0146  0.33
## fbh_7      0.77      0.78      0.78      0.38 3.6   0.028 0.0157  0.38
##
## Item statistics
##      n raw.r std.r r.cor r.drop mean  sd
## fbh_1 156 0.63 0.61 0.53 0.45 4.5 1.5
## fbh_2 156 0.70 0.71 0.67 0.58 5.0 1.2
## fbh_3 156 0.71 0.71 0.65 0.57 4.7 1.4
## fbh_4 156 0.59 0.58 0.47 0.41 4.8 1.5
## fbh_5 156 0.75 0.76 0.74 0.64 4.8 1.3
## fbh_6 156 0.71 0.70 0.64 0.57 4.7 1.4
## fbh_7 156 0.62 0.63 0.52 0.48 4.7 1.3
##
## Non missing response frequency for each item
##      1 2 3 4 5 6 9 miss
## fbh_1 0.02 0.06 0.14 0.25 0.37 0.12 0.05 0
## fbh_2 0.00 0.01 0.13 0.14 0.38 0.31 0.03 0
## fbh_3 0.01 0.03 0.16 0.24 0.31 0.21 0.04 0
## fbh_4 0.01 0.05 0.12 0.19 0.35 0.22 0.05 0
## fbh_5 0.01 0.03 0.13 0.16 0.42 0.24 0.03 0
## fbh_6 0.02 0.04 0.13 0.18 0.35 0.26 0.03 0
## fbh_7 0.03 0.01 0.13 0.21 0.38 0.22 0.02 0
```

4.1.4 Reliability involvement

```
##
## Reliability analysis
## Call: alpha(x = subset(data, select = c("involvement_1", "involvement_2",
##   "involvement_3"))))
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##     0.81      0.81      0.76      0.59 4.4 0.027  5.5 1.2     0.63
##
##   lower alpha upper      95% confidence boundaries
## 0.76 0.81 0.86
##
## Reliability if an item is dropped:
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## involvement_1      0.77      0.77      0.63      0.63 3.4   0.037  NA  0.63
## involvement_2      0.65      0.65      0.48      0.48 1.8   0.057  NA  0.48
```

```
## involvement_3      0.81      0.81      0.68      0.68 4.2      0.031      NA      0.68
##
## Item statistics
##              n raw.r std.r r.cor r.drop mean  sd
## involvement_1 156  0.84  0.84  0.72   0.64  5.5 1.4
## involvement_2 156  0.89  0.90  0.84   0.76  5.4 1.4
## involvement_3 156  0.83  0.82  0.67   0.60  5.7 1.5
##
## Non missing response frequency for each item
##              1      2      3      4      5      6      7 miss
## involvement_1 0.02 0.04 0.04 0.08 0.19 0.38 0.26      0
## involvement_2 0.02 0.04 0.03 0.12 0.26 0.36 0.18      0
## involvement_3 0.02 0.04 0.03 0.08 0.15 0.33 0.35      0
```

4.2 Testing system usability using SUS

Before testing the hypotheses, the System Usability Scale score was used to determine whether the implementation of the conversational agent was sufficient: we determined that the average SUS score should at least fall in the “OK” category (a score of 50.9 with standard deviation of 13.8) as described by Bangor et al. (2009). The mean SUS is equal to 81.6 (SD 12.0), meaning that the score is closest to either “Good” or “Excellent”, showing that the implementation was sufficient.

```
## [1] 81.64557
## [1] 12.01552
```

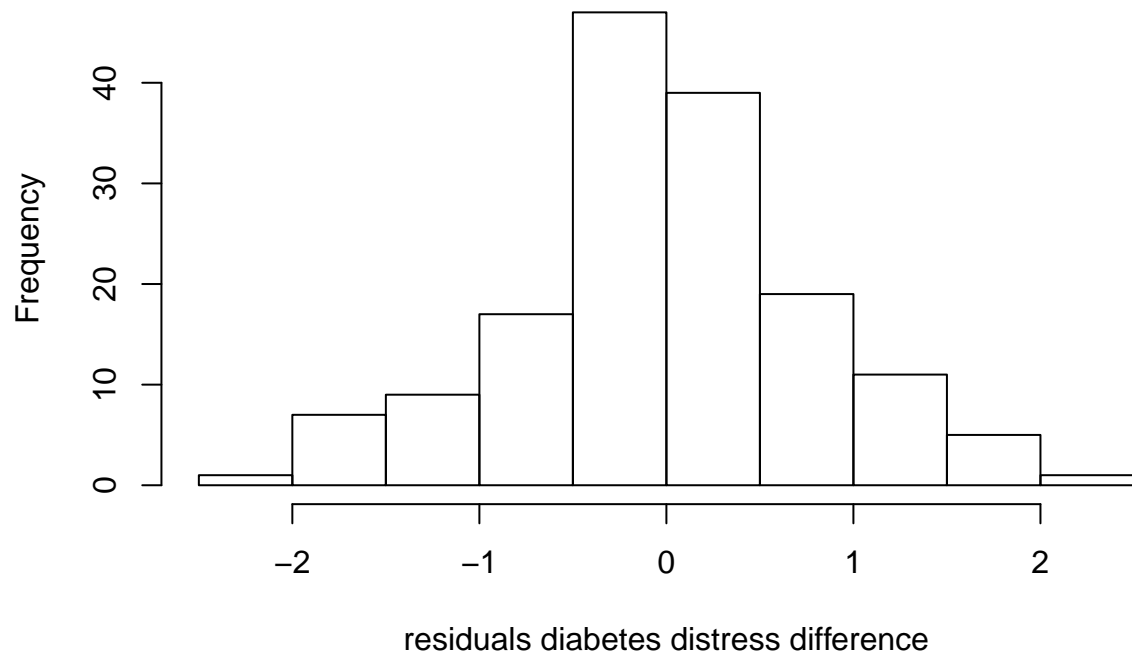
4.3 Testing hypothesis 1: People using the conversational agent will have a larger reduction in diabetes distress than the control group

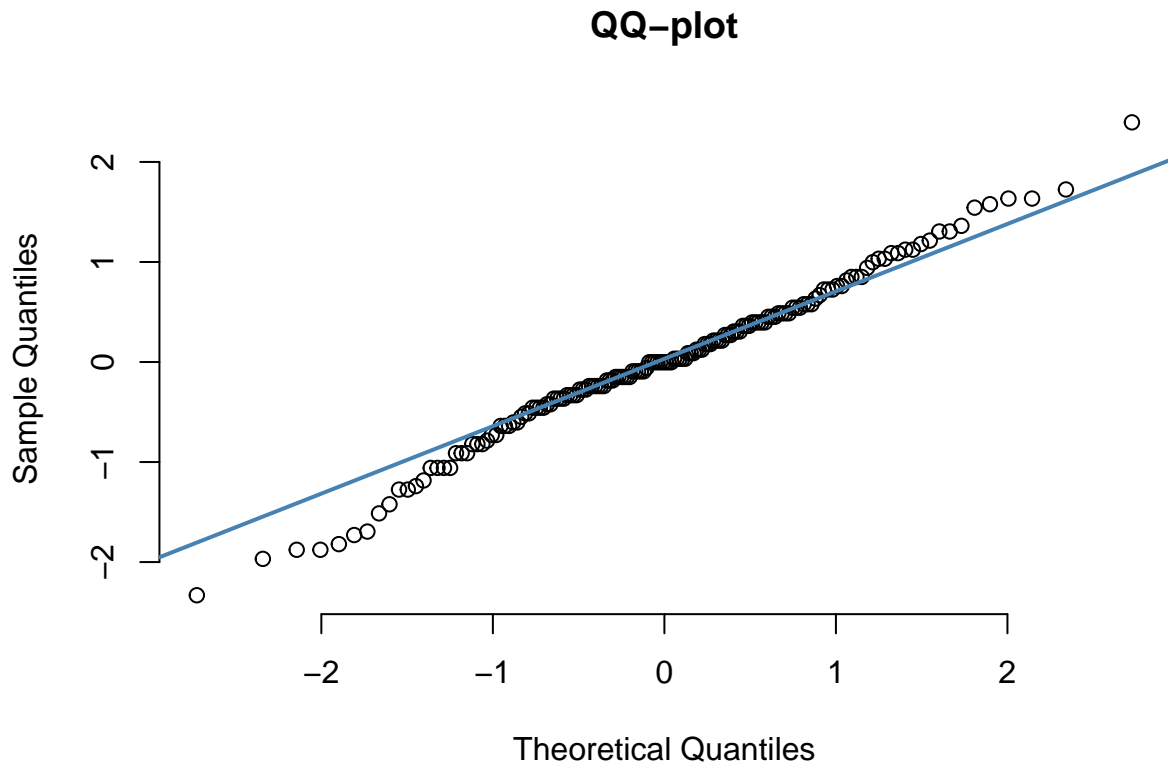
As explained in the thesis report, an independent t-test or Mann-Whitney U test was done depending on whether the assumptions for a independent t-test are met. First, the assumptions were all checked. The assumptions for an independent t-test with equal variances of groups were met.

4.3.1 Testing assumptions for H1

Below, we see that the data does not deviate from normality: the Shapiro-Wilk test, the QQ-plots and the histograms show that the data looks normal. In addition, the variances in both groups are equal.

Histogram of residuals





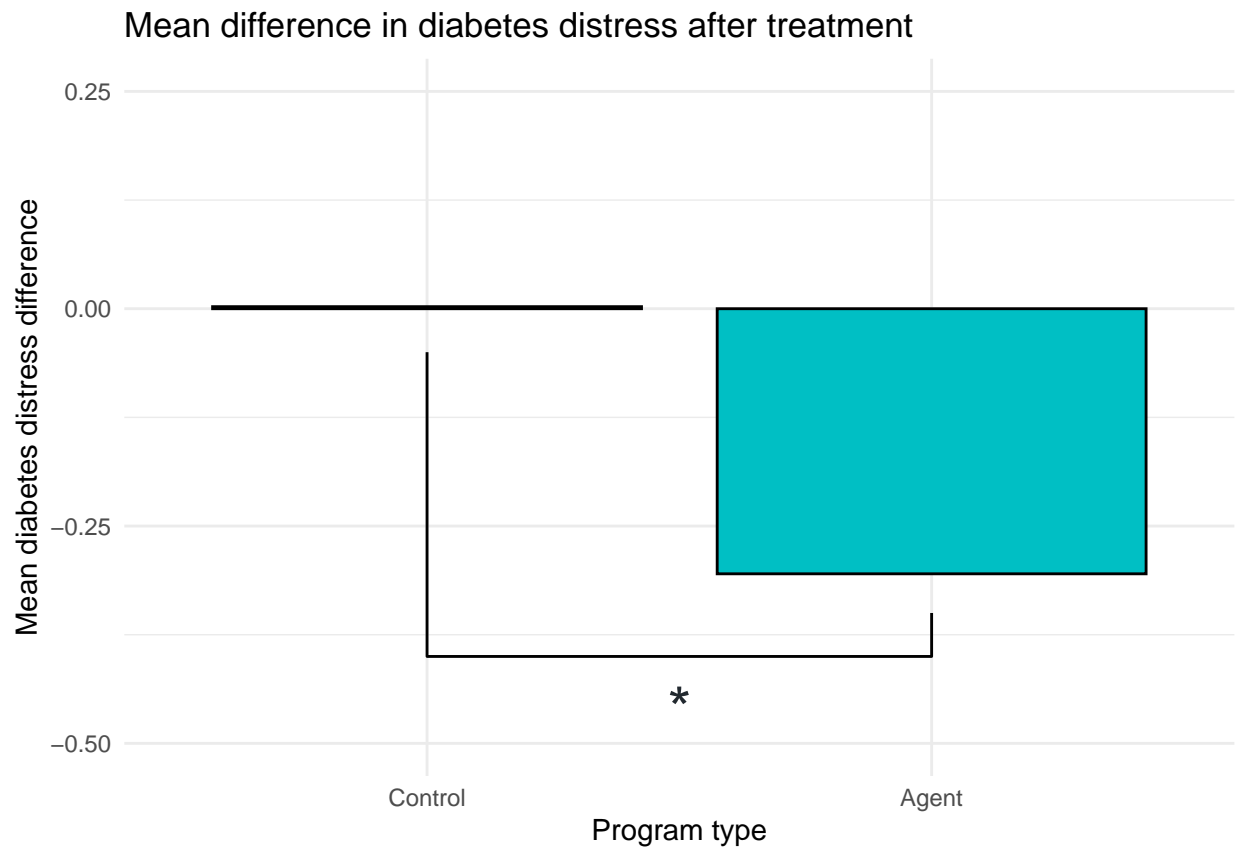
```
##
##  Shapiro-Wilk normality test
##
## data:  residuals_diabetes_distress_difference
## W = 0.98883, p-value = 0.2506
##
##  F test to compare two variances
##
## data:  diabetes_distress_diff by program_type
## F = 0.73805, num df = 76, denom df = 78, p-value = 0.1855
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4708759 1.1583540
## sample estimates:
## ratio of variances
##      0.7380485
```

4.3.2 Result for hypothesis 1

The results below show that the agent does cause a bigger reduction in diabetes distress than the group receiving a plain textual delivery of tips.

```
## # A tibble: 2 x 3
##   program_type mean_dd_diff sd_dd_diff
##   <fct>         <dbl>         <dbl>
## 1 Control      0.00236      0.743
```

```
## 2 Agent          -0.305      0.865
```



```
##
## Two Sample t-test
##
## data: diabetes_distress_diff by program_type
## t = 2.3771, df = 154, p-value = 0.01868
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.05191821 0.56270077
## sample estimates:
## mean in group Control    mean in group Agent
##      0.002361275         -0.304948216
##
## One Sample t-test
##
## data: subset(data, program_type == "Agent")$diabetes_distress_diff
## t = -3.1329, df = 78, p-value = 0.001219
## alternative hypothesis: true mean is less than 0
## 95 percent confidence interval:
##      -Inf -0.1429201
## sample estimates:
## mean of x
## -0.3049482
##
```

```
## One Sample t-test
##
## data: subset(data, program_type == "Control")$diabetes_distress_diff
## t = 0.027878, df = 76, p-value = 0.5111
## alternative hypothesis: true mean is less than 0
## 95 percent confidence interval:
##      -Inf 0.1434
## sample estimates:
## mean of x
## 0.002361275
## [1] 0.3806692
```

We observed an effect size of 0.38, which is between a small and medium effect according to Cohen (2013).

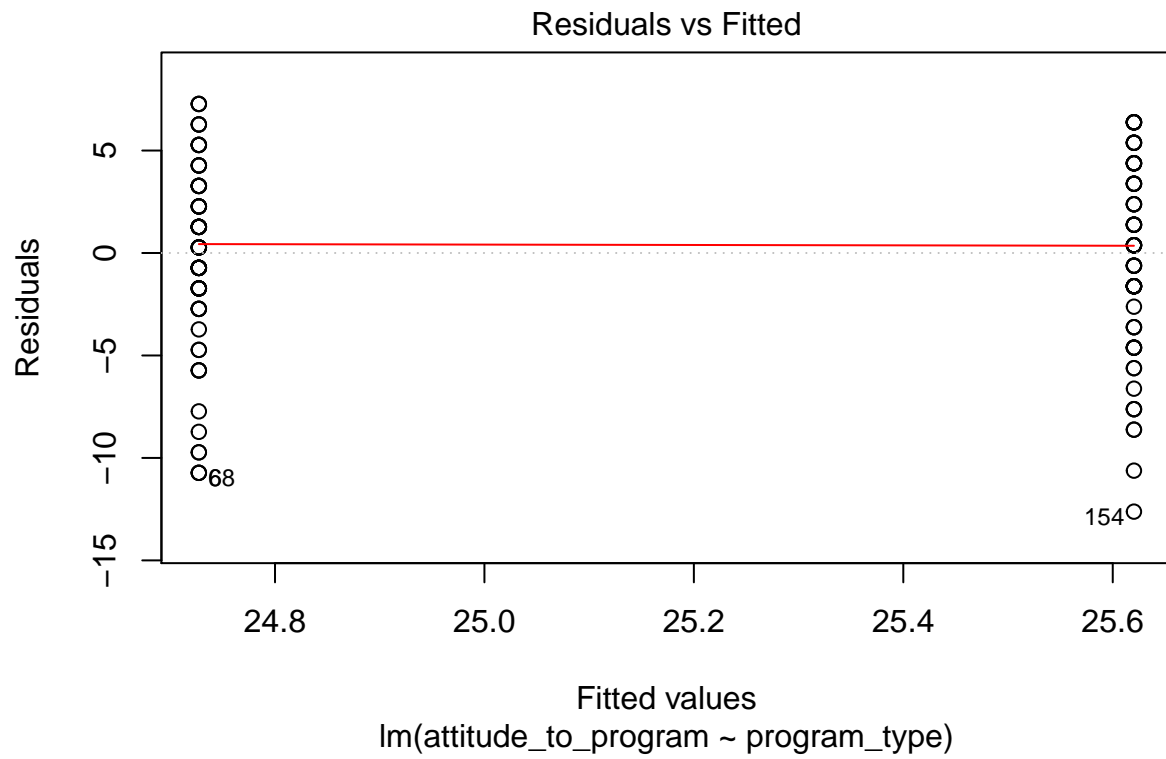
4.4 Testing hypothesis 2: The effect of the social help program type on the diabetes distress difference is mediated by the attitude towards the social help program

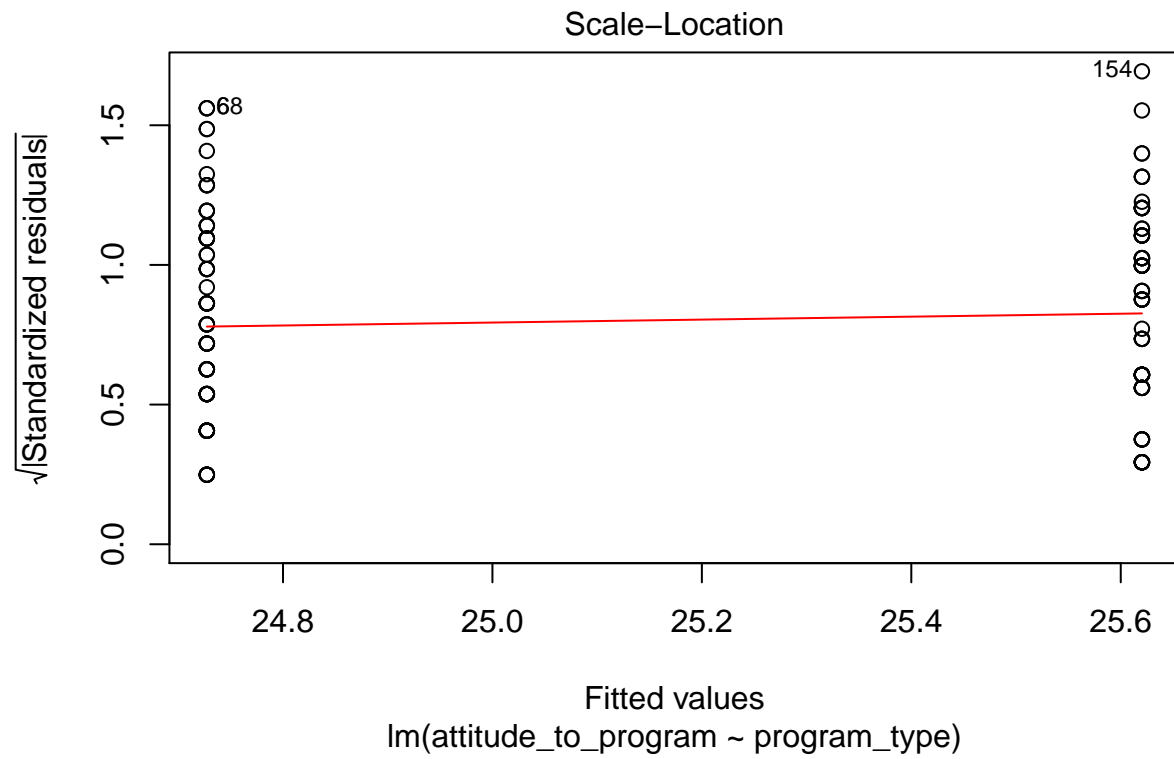
4.4.1 Testing assumptions for H2

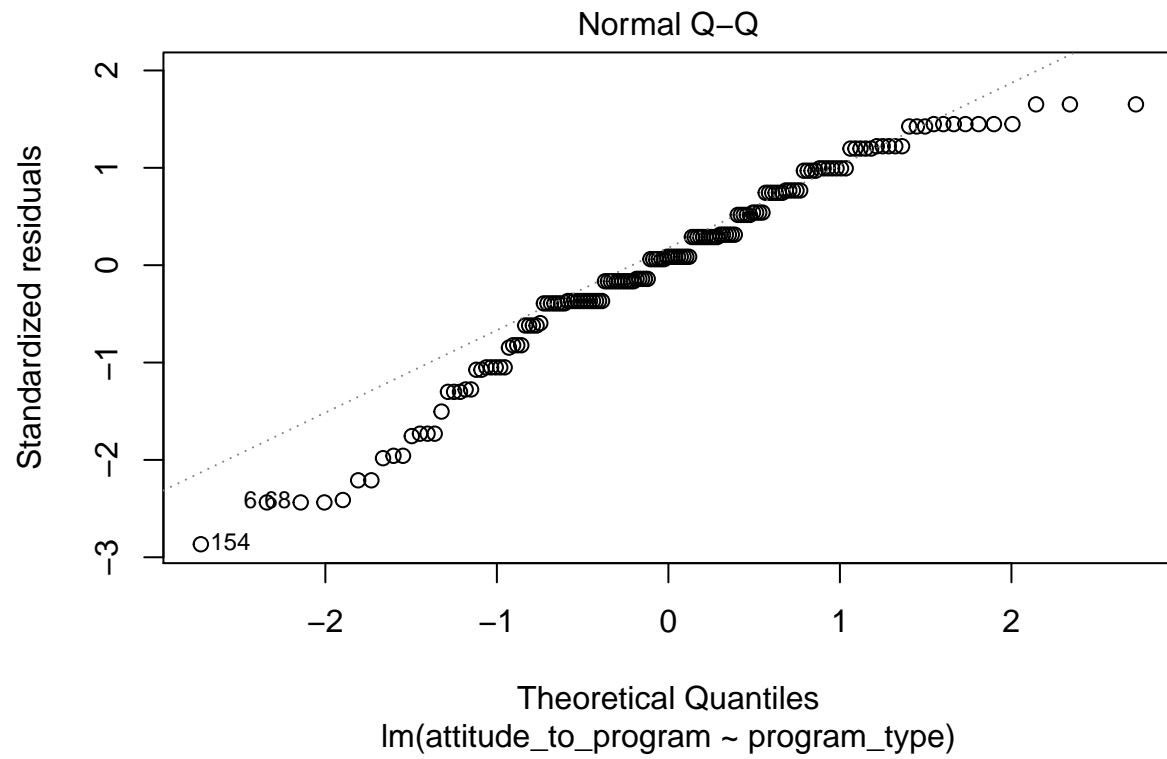
The QQ-plot of the model predicting the attitude towards the social help program shows the data is not normal. However, we do not change our setup since the bootstrapping by Preacher and Hayes does not assume normality of data.

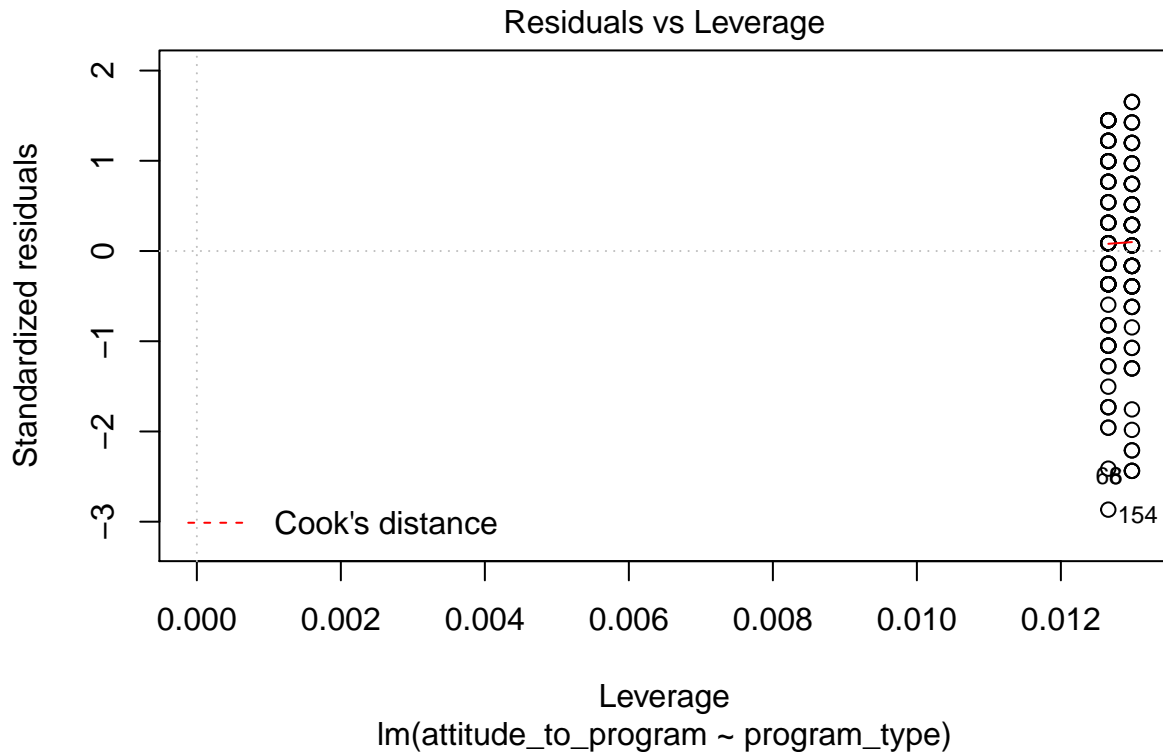
```
##
## Call:
## lm(formula = attitude_to_program ~ program_type, data = data)
##
## Coefficients:
##      (Intercept)  program_typeAgent
##           24.727           0.893
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance = 0.05
##
## Call:
## gvlma(x = M)
##
##
```

	Value	p-value	Decision
## Global Stat	9.700e+00	0.045801	Assumptions NOT satisfied!
## Skewness	9.646e+00	0.001898	Assumptions NOT satisfied!
## Kurtosis	7.285e-05	0.993190	Assumptions acceptable.
## Link Function	-1.146e-13	1.000000	Assumptions acceptable.
## Heteroscedasticity	5.359e-02	0.816935	Assumptions acceptable.

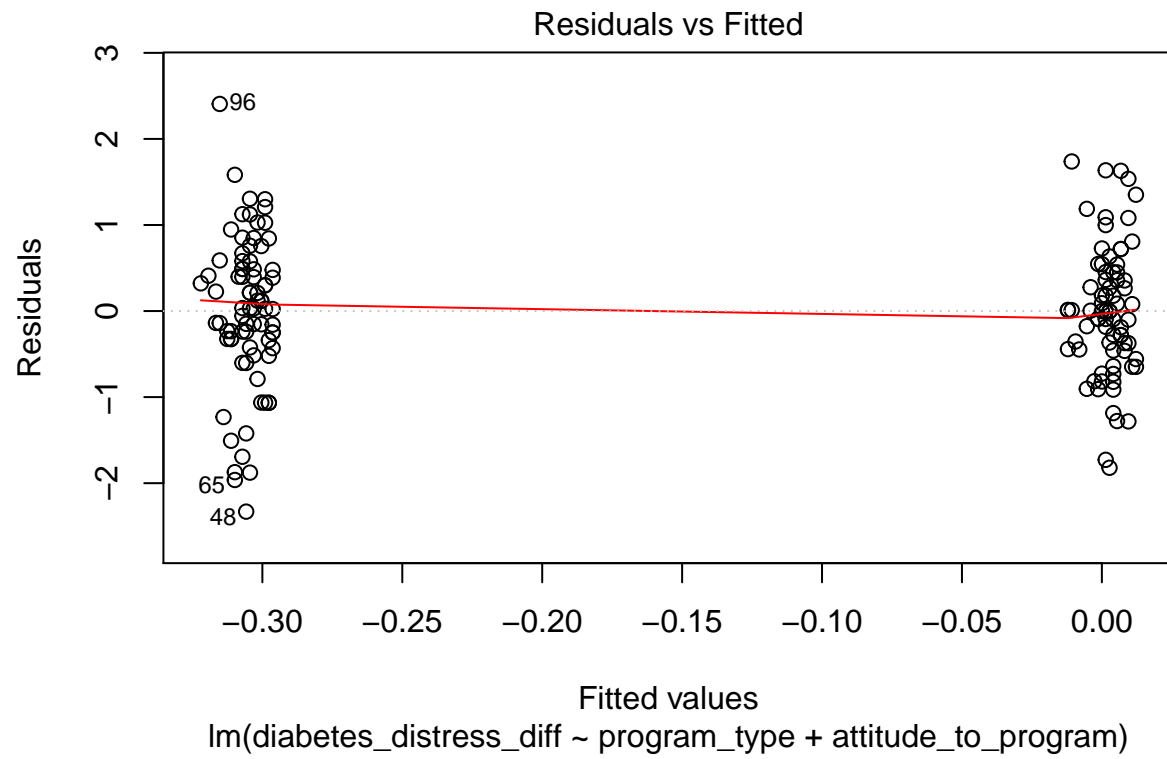


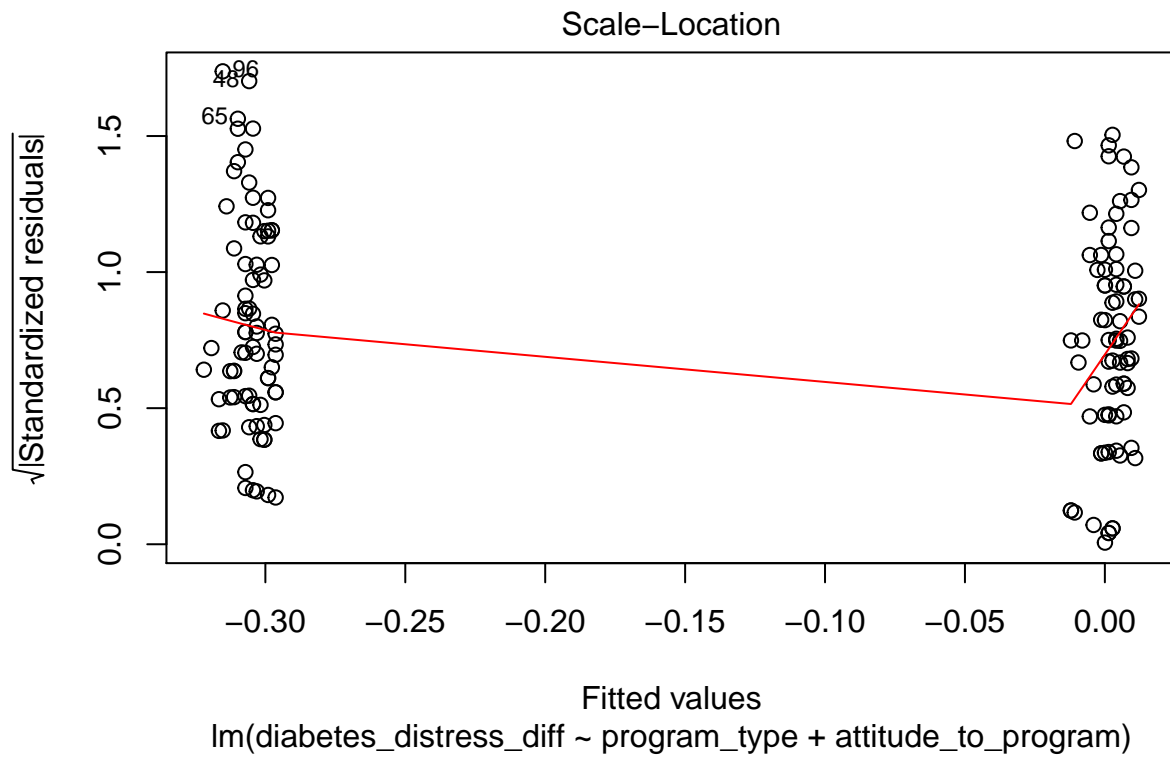


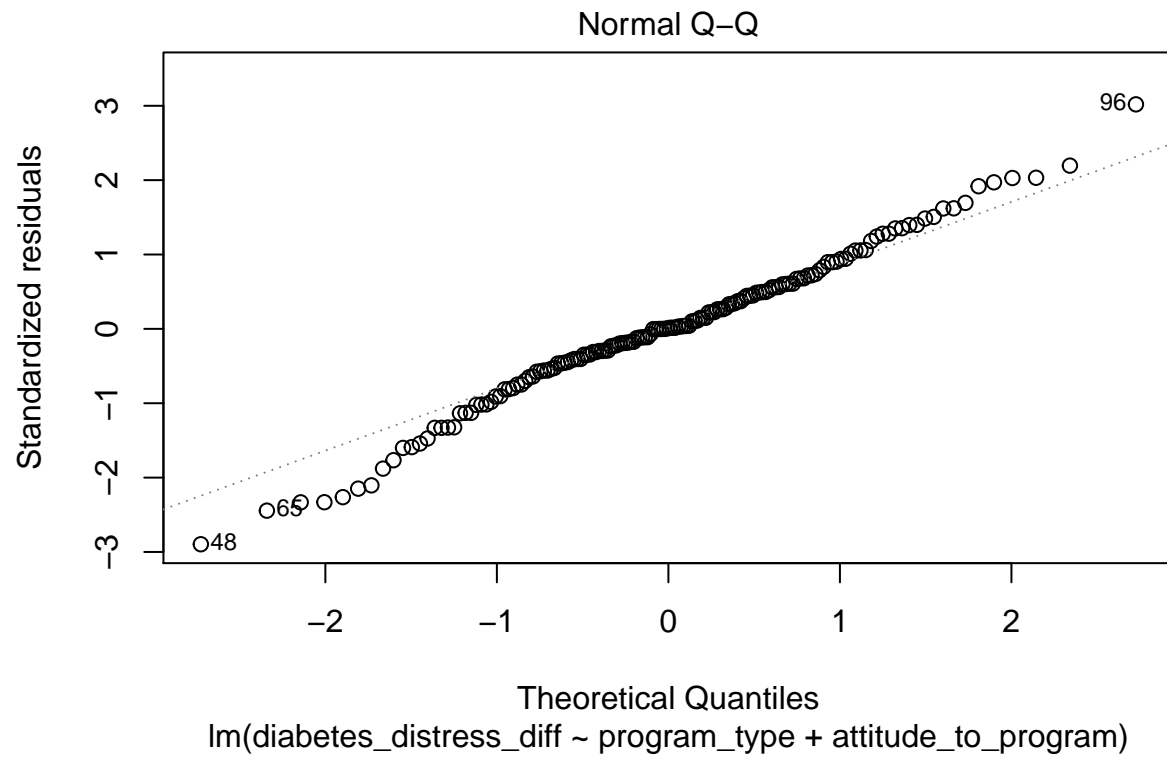


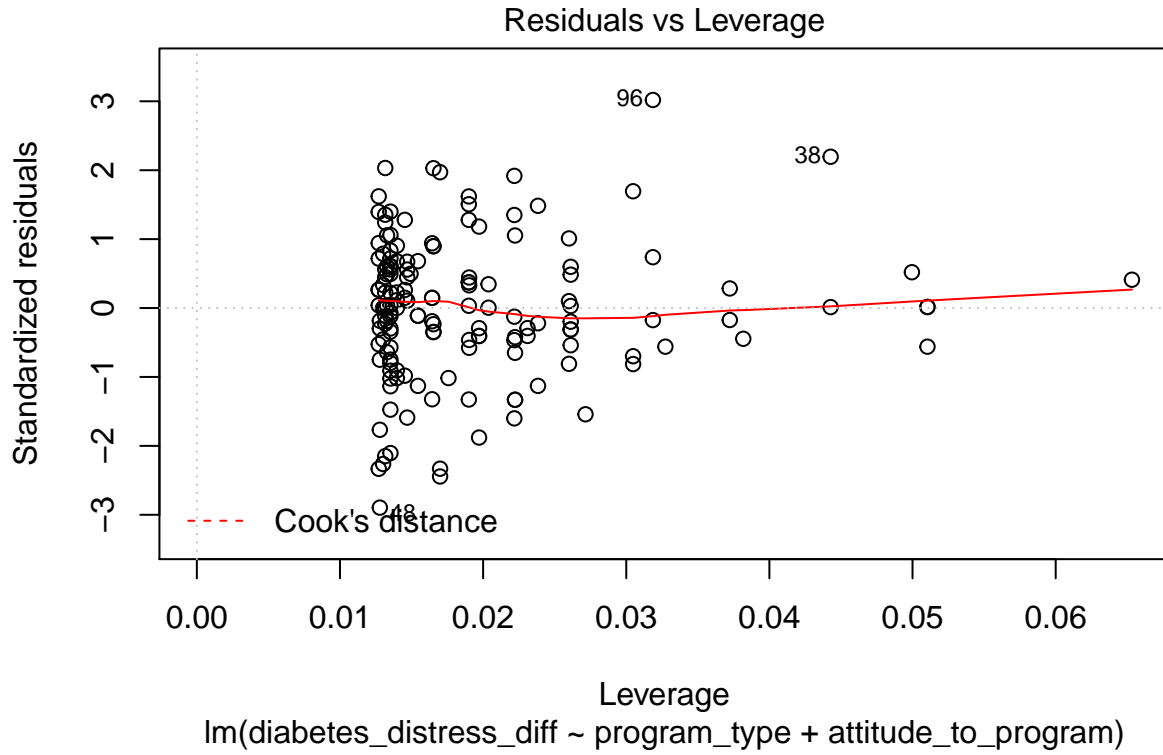


```
##
## Call:
## lm(formula = diabetes_distress_diff ~ program_type + attitude_to_program,
##     data = data)
##
## Coefficients:
##             (Intercept)      program_typeAgent  attitude_to_program
##             -0.03103          -0.30852              0.00135
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance = 0.05
##
## Call:
## gvlma(x = Y)
##
##              Value p-value          Decision
## Global Stat    3.24986  0.5169 Assumptions acceptable.
## Skewness       0.74684  0.3875 Assumptions acceptable.
## Kurtosis       1.41979  0.2334 Assumptions acceptable.
## Link Function   0.02165  0.8830 Assumptions acceptable.
## Heteroscedasticity 1.06158 0.3029 Assumptions acceptable.
```





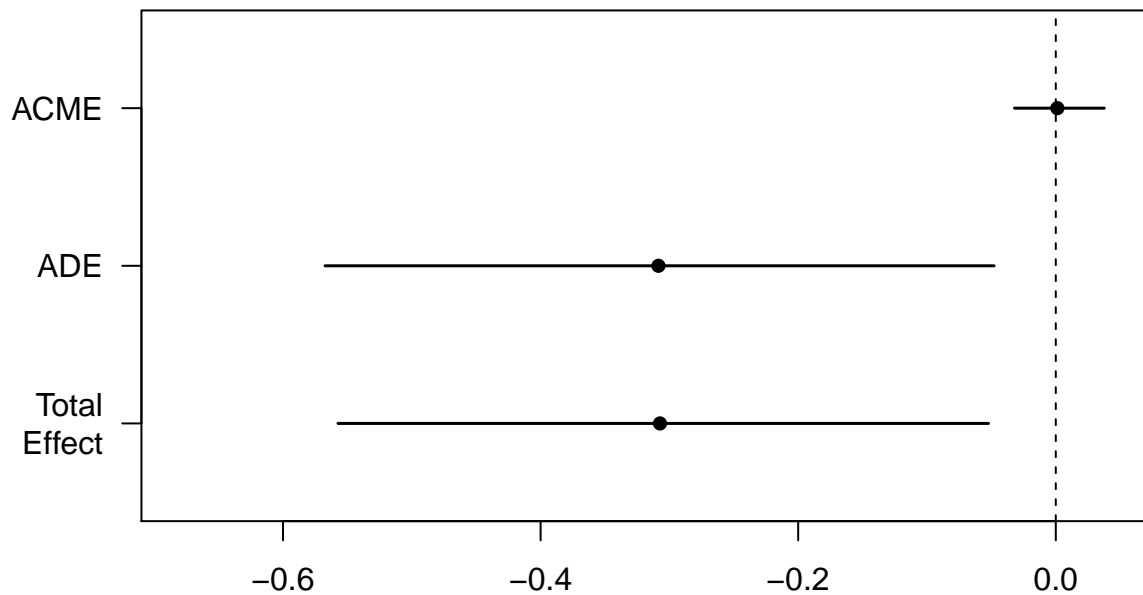




4.4.2 Results H2

The results show that our hypothesis is not supported, meaning mediation was not present in the data (note that the exact numbers can differ between simulation runs).

```
## Running nonparametric bootstrap
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##      Estimate 95% CI Lower 95% CI Upper p-value
## ACME      0.00121   -0.03205    0.04  0.920
## ADE      -0.30852   -0.56740   -0.05  0.020 *
## Total Effect -0.30731   -0.55731   -0.05  0.016 *
## Prop. Mediated -0.00392   -0.15332    0.18  0.918
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 156
##
##
## Simulations: 10000
```



4.5 Testing hypothesis 3: People who have higher initial diabetes distress will have a larger re-duction in diabetes distress than people with low initial diabetes distress

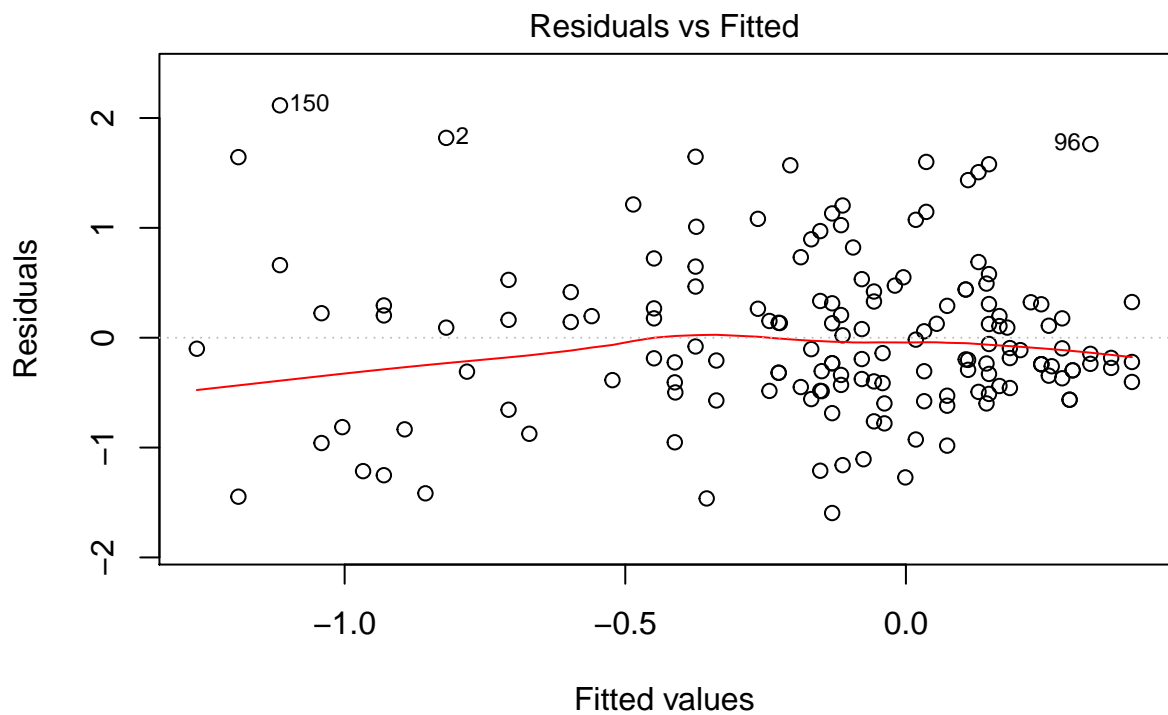
A classic moderation analysis was done to determine whether the pre-measurement diabetes distress is a moderator.

4.5.1 Testing assumptions for H3

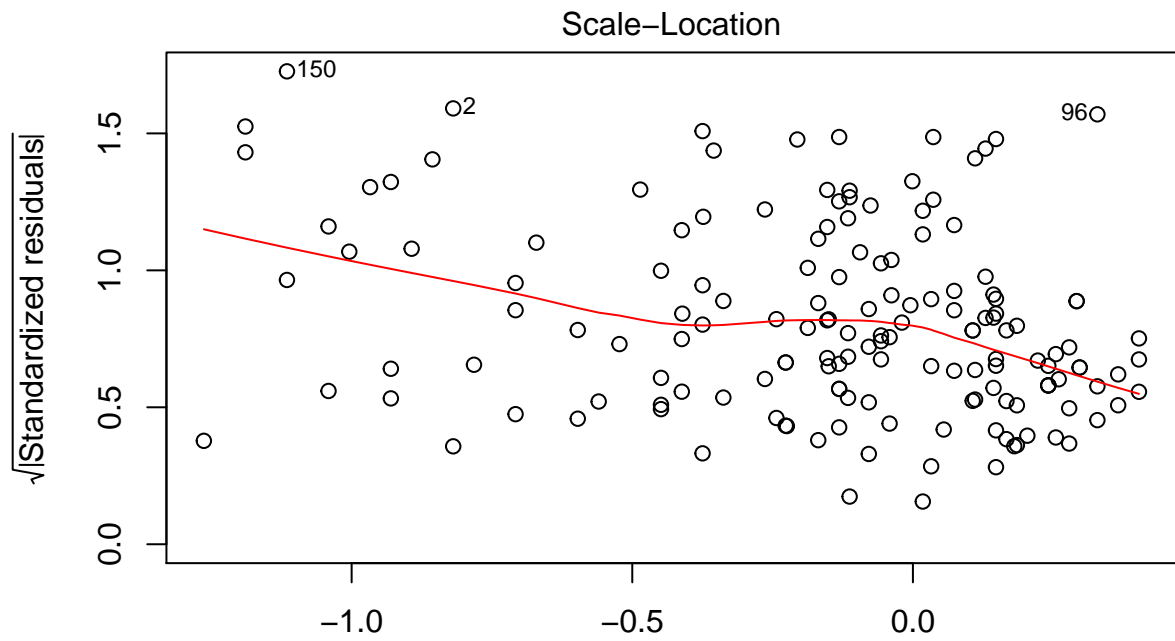
Below, we can see that the data is linear and normally distributed. There is homogeneity of variances and there are no extreme outliers.

```
##
## Call:
## lm(formula = data$diabetes_distress_diff ~ data$program_type +
##     data$pre_diabetes_distress + data$program_type:data$pre_diabetes_distress)
##
## Coefficients:
##                (Intercept)                0.5019
##                data$program_typeAgent                0.3080
##                data$pre_diabetes_distress                -0.2049
## data$program_typeAgent:data$pre_diabetes_distress                -0.2023
##
```

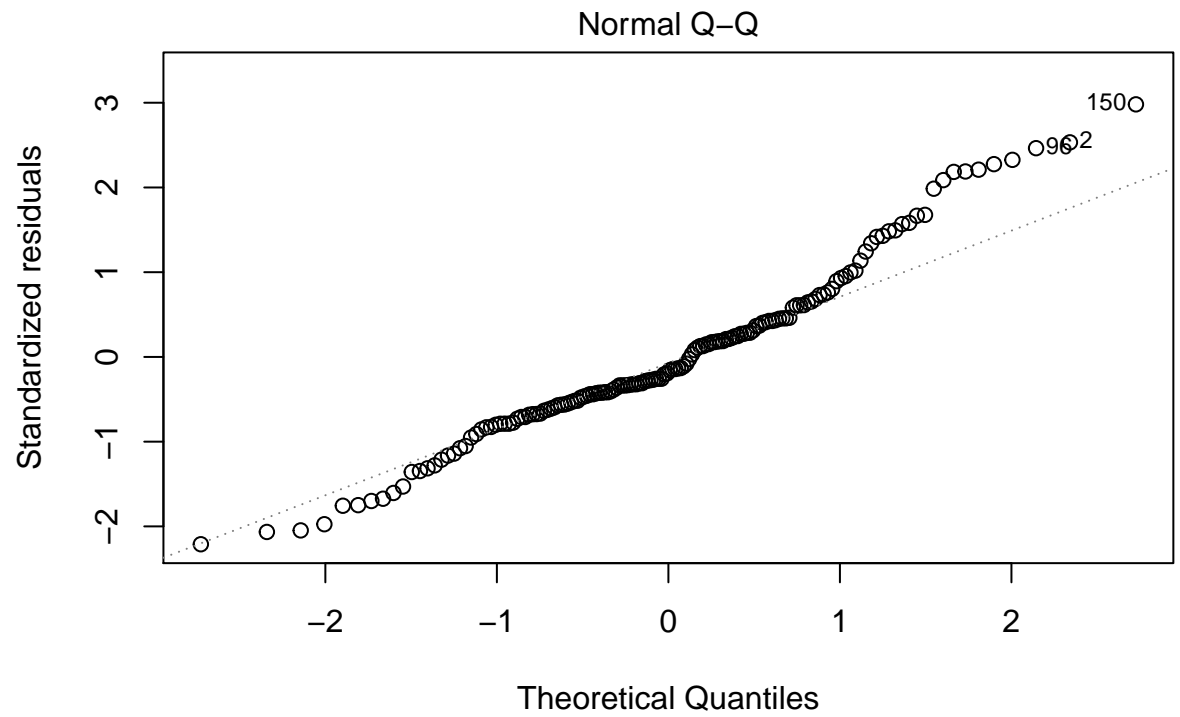
```
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance = 0.05
##
## Call:
## gvlma(x = moderation_model)
##
##              Value  p-value              Decision
## Global Stat      9.91040 0.041964 Assumptions NOT satisfied!
## Skewness         7.22763 0.007179 Assumptions NOT satisfied!
## Kurtosis         0.78208 0.376506 Assumptions acceptable.
## Link Function    0.05762 0.810296 Assumptions acceptable.
## Heteroscedasticity 1.84307 0.174592 Assumptions acceptable.
```



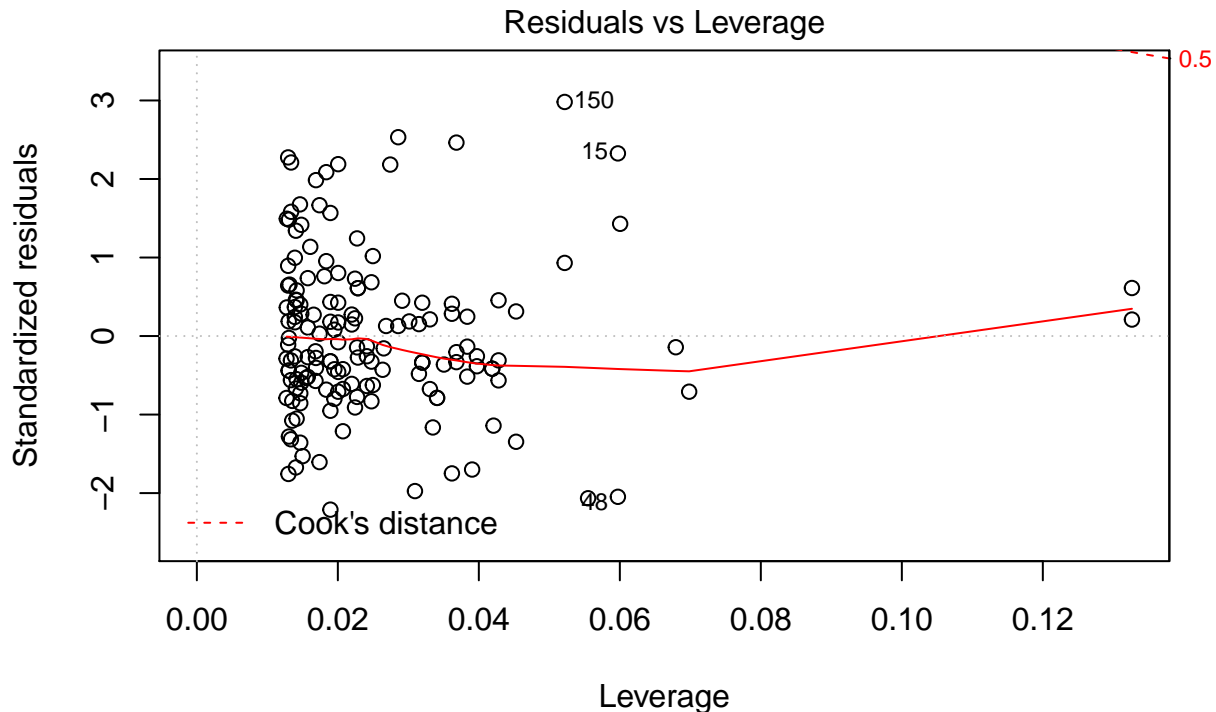
`lm(data$diabetes_distress_diff ~ data$program_type + data$pre_diabetes_dist ...`



Fitted values
`lm(data$diabetes_distress_diff ~ data$program_type + data$pre_diabetes_dist ...`



`lm(data$diabetes_distress_diff ~ data$program_type + data$pre_diabetes_dist ...`



lm(data\$diabetes_distress_diff ~ data\$program_type + data\$pre_diabetes_dist ...

4.5.2 Results H3

The interaction term is not significant, meaning that the hypothesis that the pre-measurement diabetes distress moderates the relationship between the social help program and the diabetes distress difference is not supported. We do see that the pre-measurement diabetes distress influences the diabetes distress difference.

```
##
## Call:
## lm(formula = data$diabetes_distress_diff ~ data$program_type +
##     data$pre_diabetes_distress + data$program_type:data$pre_diabetes_distress)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5958 -0.4322 -0.1266  0.3258  2.1153
##
## Coefficients:
##                                     Estimate Std. Error t value
## (Intercept)                        0.50195    0.22592   2.222
## data$program_typeAgent                0.30803    0.31225   0.986
## data$pre_diabetes_distress            -0.20492    0.08617  -2.378
## data$program_typeAgent:data$pre_diabetes_distress -0.20235    0.11282  -1.794
##                                     Pr(>|t|)
## (Intercept)                        0.0278 *
## data$program_typeAgent                0.3255
## data$pre_diabetes_distress            0.0187 *
## data$program_typeAgent:data$pre_diabetes_distress 0.0749 .
```

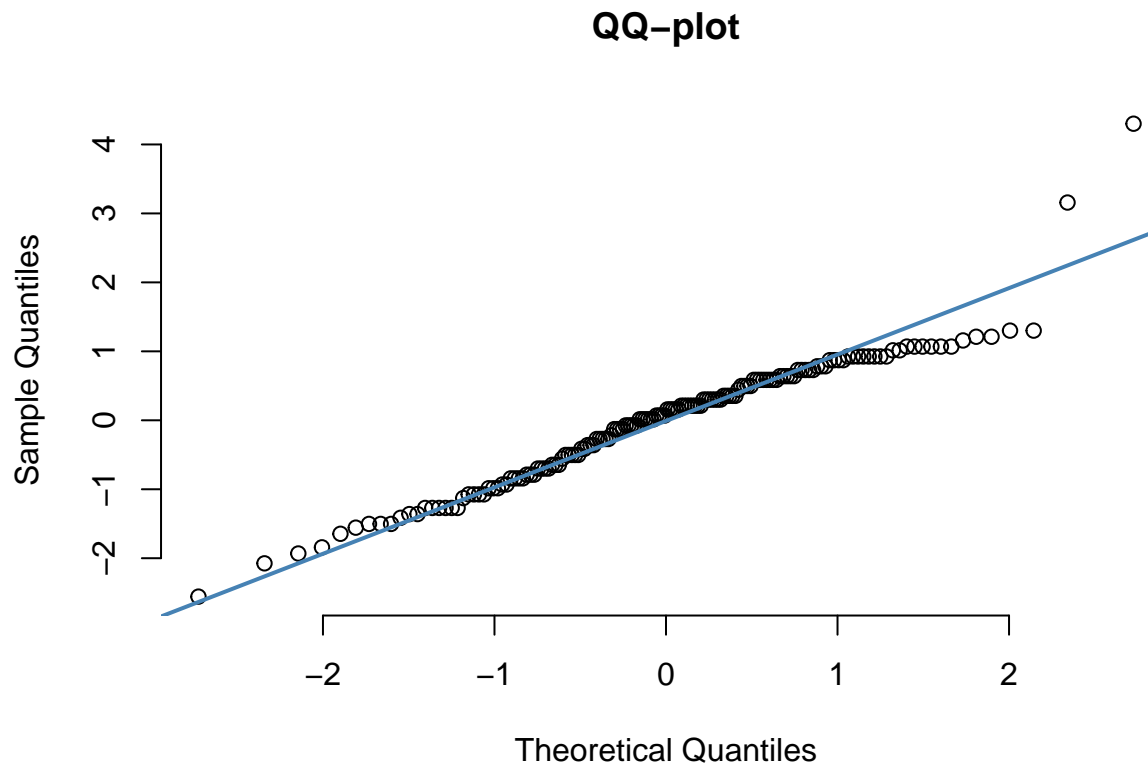
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7288 on 152 degrees of freedom
## Multiple R-squared:  0.224, Adjusted R-squared:  0.2086
## F-statistic: 14.62 on 3 and 152 DF,  p-value: 2.045e-08
```

4.6 Testing hypothesis 4: People who use the conversational agent have a higher feeling of being heard than the control group

As explained in the thesis report, an independent t-test or Mann-Whitney U test is done depending on whether the assumptions for a independent t-test are met.

4.6.1 Testing assumptions for H4

The data looks normal as we see in the QQ-plots, even though the Shapiro-Wilk test says differently. However,



we still assume normality.

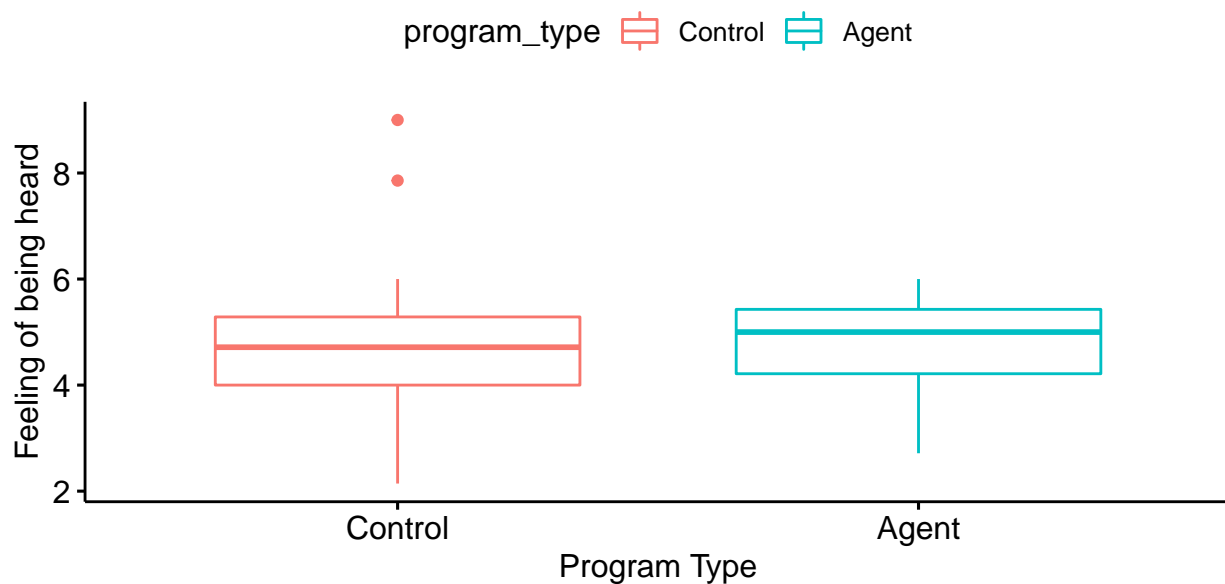
```
##
## Shapiro-Wilk normality test
##
## data: residuals_fbh
## W = 0.94912, p-value = 1.925e-05
##
## F test to compare two variances
##
## data: feeling_of_being_heard by program_type
## F = 1.4714, num df = 76, denom df = 78, p-value = 0.09176
```

```
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.9387851 2.3094099
## sample estimates:
## ratio of variances
##          1.471447
```

4.6.2 Result for hypothesis 4

The t-test shows there is no significant difference between the two groups in the means of the feeling of being heard.

```
## # A tibble: 2 x 4
##   program_type count  mean    sd
##   <fct>      <int> <dbl> <dbl>
## 1 Control         77  4.70 1.02
## 2 Agent          79  4.79 0.839
```



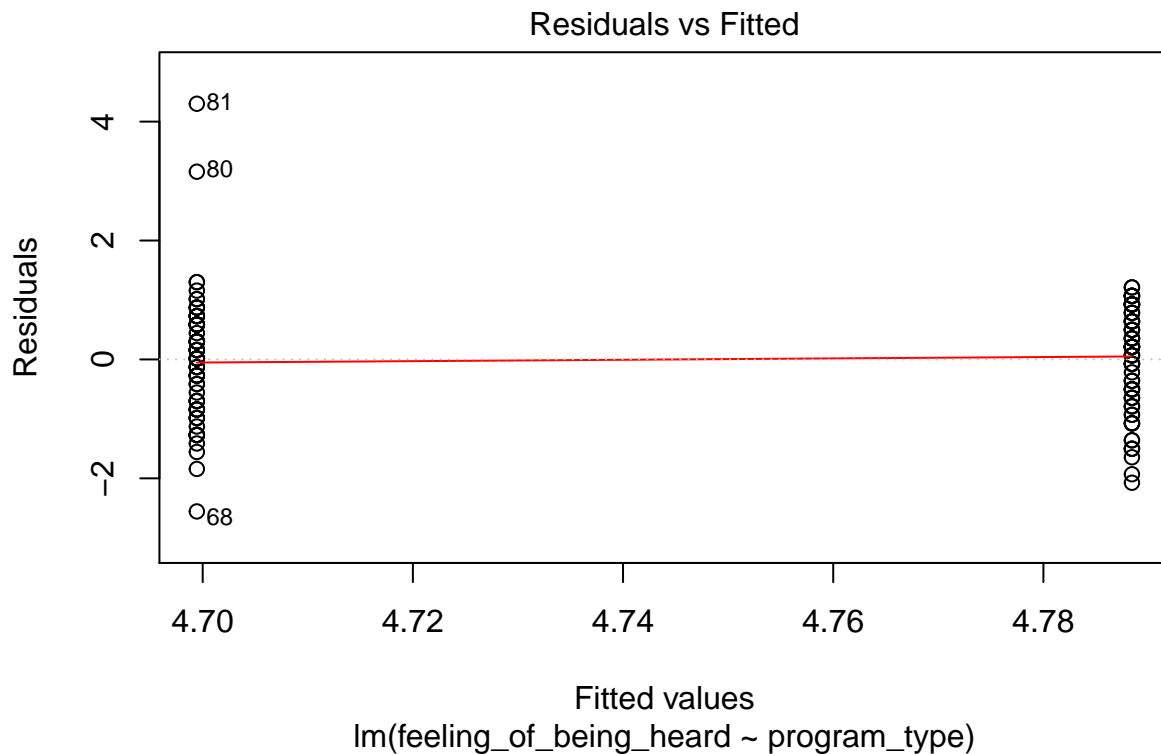
```
##
## Two Sample t-test
##
## data: feeling_of_being_heard by program_type
## t = -0.59637, df = 154, p-value = 0.5518
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3837440 0.2057773
## sample estimates:
## mean in group Control mean in group Agent
##          4.699443          4.788427
```

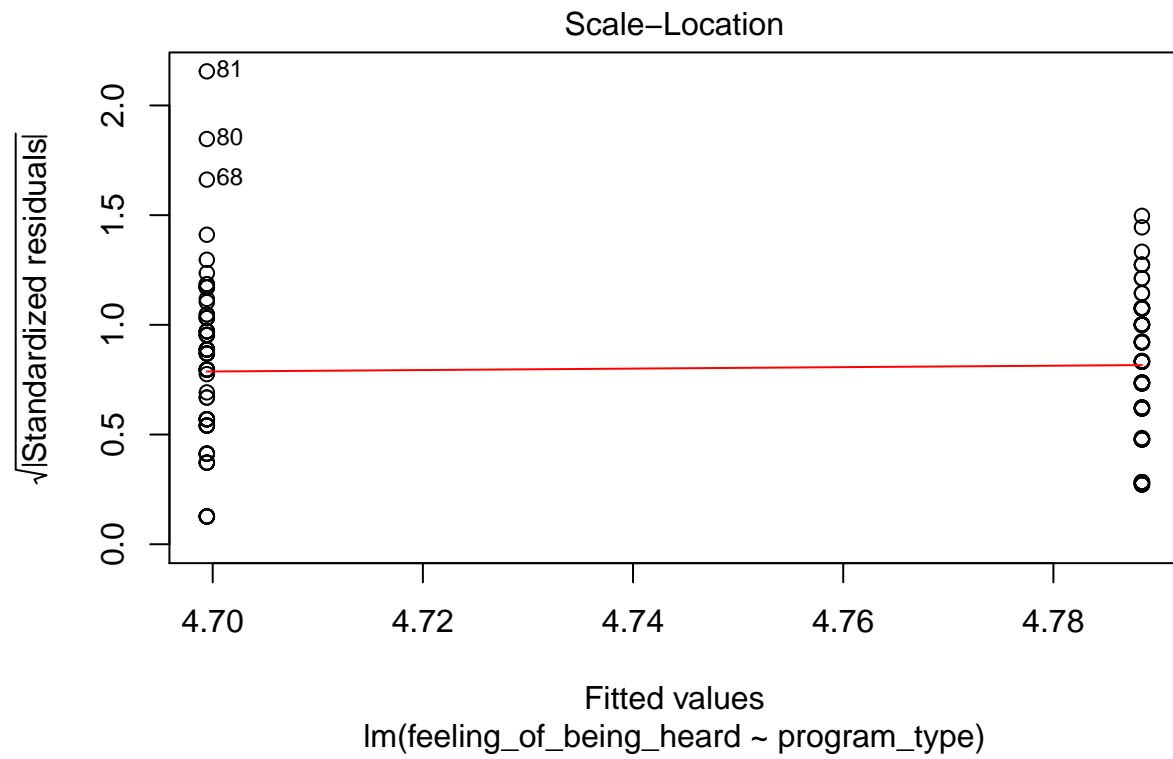
4.7 Exploratory research

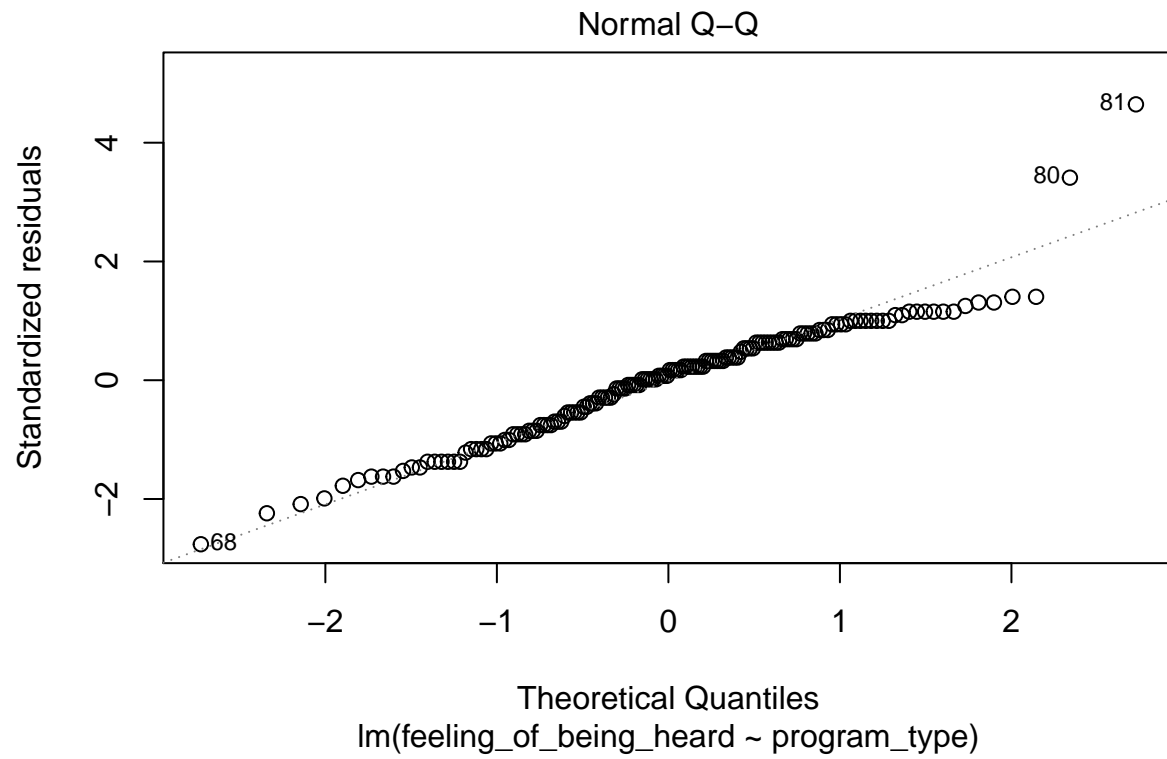
We considered the feeling of being heard as our mediator. A similar analysis is done as shown in H2.

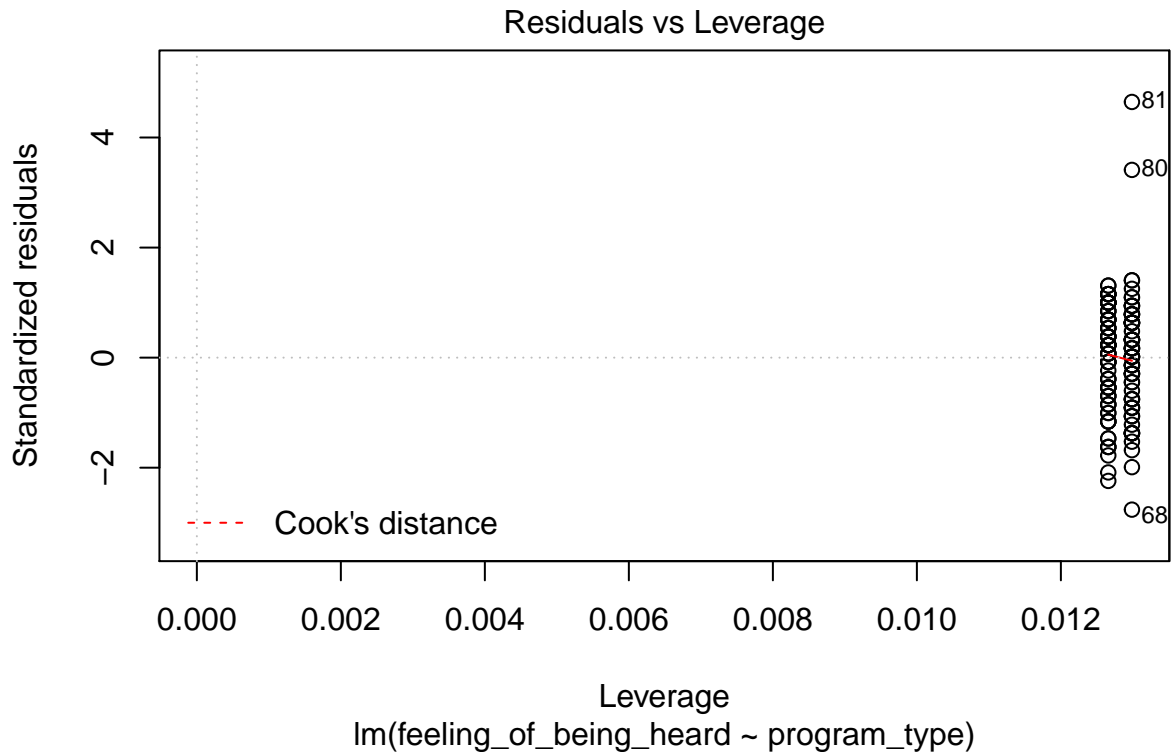
```
##
## Call:
```

```
## lm(formula = feeling_of_being_heard ~ program_type, data = data)
##
## Coefficients:
##      (Intercept)  program_typeAgent
##           4.69944           0.08898
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance = 0.05
##
## Call:
##  gvlma(x = M)
##
##              Value    p-value      Decision
## Global Stat    4.706e+01 1.479e-09 Assumptions NOT satisfied!
## Skewness       4.337e+00 3.729e-02 Assumptions NOT satisfied!
## Kurtosis       4.182e+01 1.001e-10 Assumptions NOT satisfied!
## Link Function  -1.250e-12 1.000e+00 Assumptions acceptable.
## Heteroscedasticity 9.071e-01 3.409e-01 Assumptions acceptable.
```

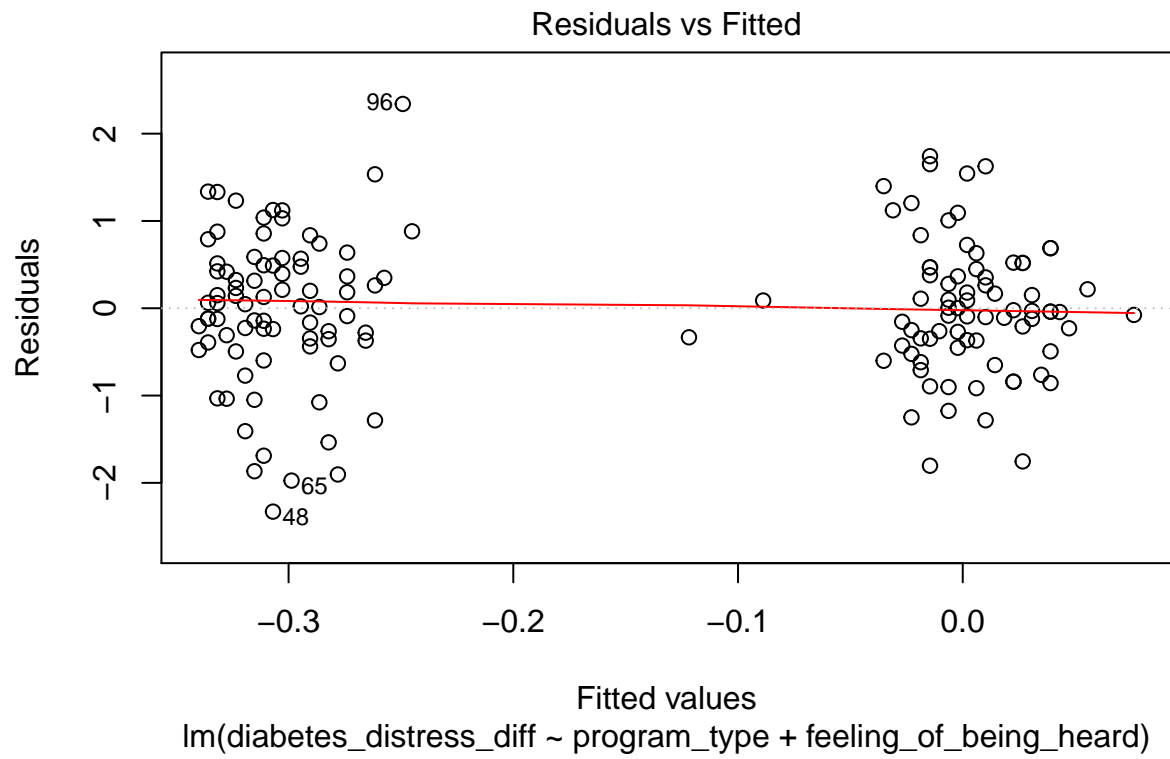


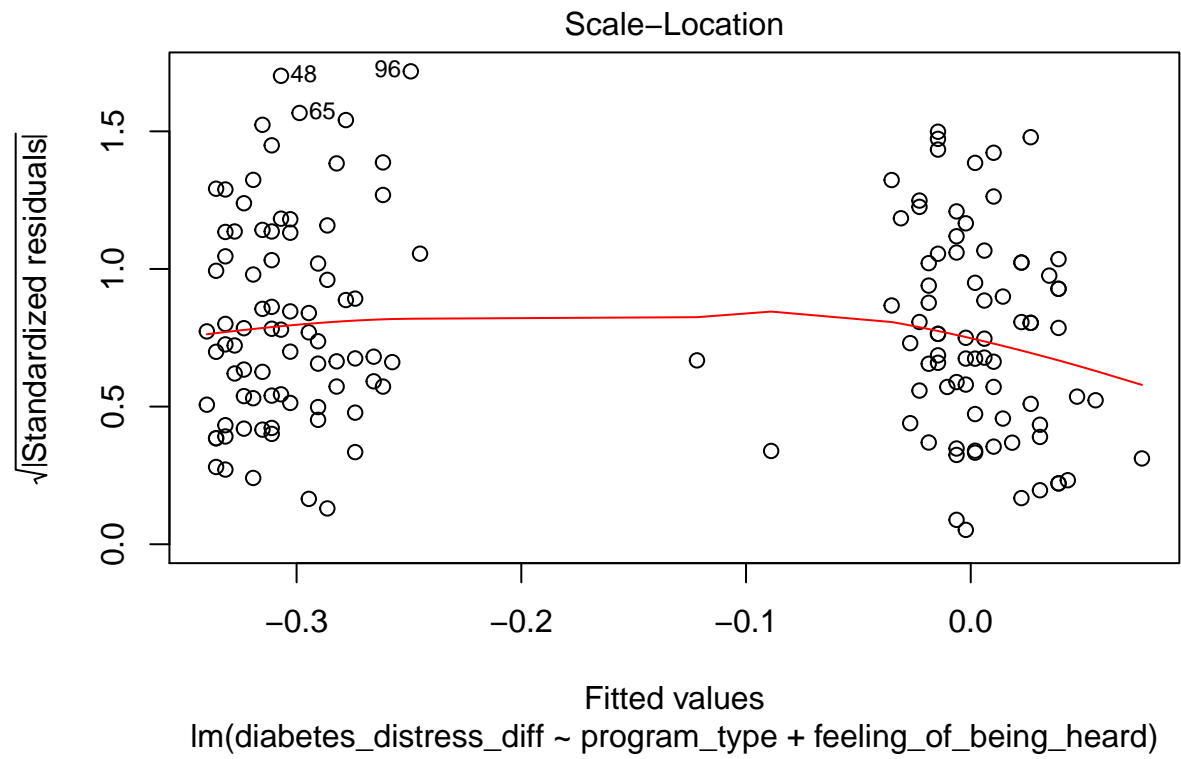


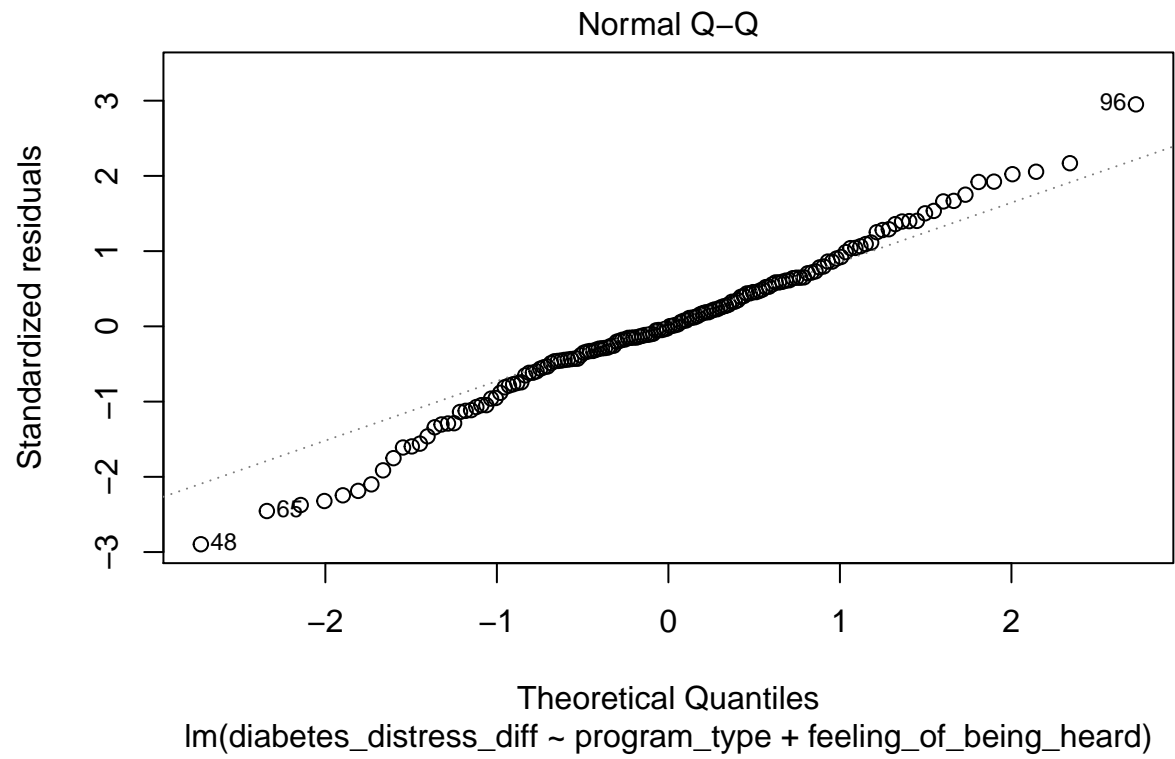


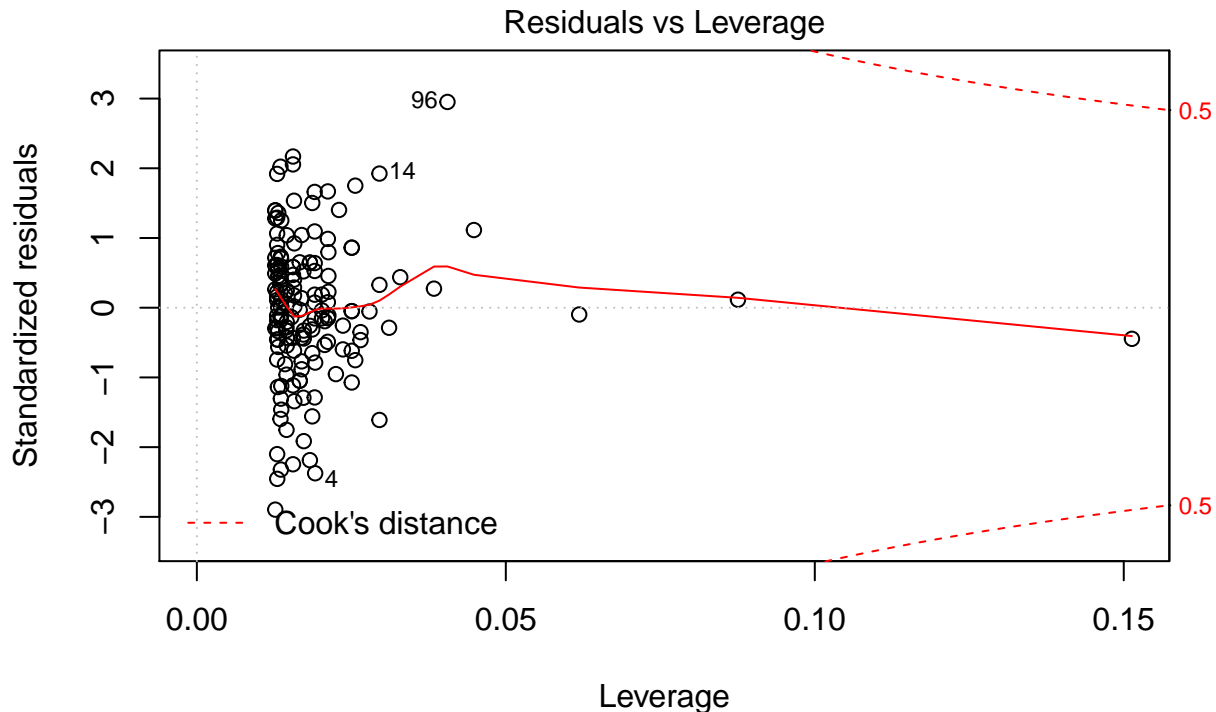


```
##
## Call:
## lm(formula = diabetes_distress_diff ~ program_type + feeling_of_being_heard,
##     data = data)
##
## Coefficients:
##             (Intercept)      program_typeAgent  feeling_of_being_heard
##                0.13806                -0.30474                 -0.02888
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance = 0.05
##
## Call:
## gvlma(x = Y)
##
##              Value p-value              Decision
## Global Stat    3.5965  0.4634 Assumptions acceptable.
## Skewness       0.8377  0.3601 Assumptions acceptable.
## Kurtosis       1.3378  0.2474 Assumptions acceptable.
## Link Function   0.3648  0.5459 Assumptions acceptable.
## Heteroscedasticity 1.0563 0.3041 Assumptions acceptable.
```





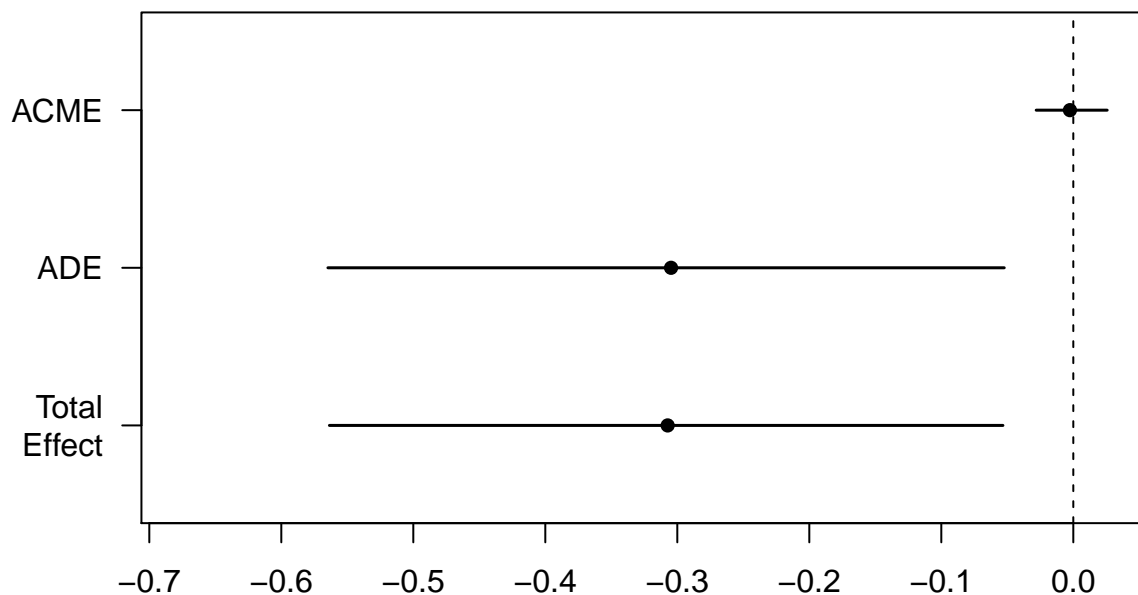




$\text{lm}(\text{diabetes_distress_diff} \sim \text{program_type} + \text{feeling_of_being_heard})$

The QQ-plot for the model predicting our mediator shows the residuals are not normally distributed. However, we do not change our setup since the bootstrapping by Preacher and Hayes does not assume normality of data. The results show that our hypothesis is not supported.

```
## Running nonparametric bootstrap
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           -0.00257   -0.02812    0.03  0.964
## ADE            -0.30474   -0.56468   -0.05  0.016 *
## Total Effect   -0.30731   -0.56356   -0.05  0.016 *
## Prop. Mediated  0.00836   -0.12162    0.13  0.976
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 156
##
##
## Simulations: 1000
```



5 R version information

During the above analysis, the following packages were used in R.

R version 3.6.0 (2019-04-26)

Platform: x86_64-w64-mingw32/x64 (64-bit)

locale: *LC_COLLATE=Dutch_Netherlands.1252, LC_CTYPE=Dutch_Netherlands.1252, LC_MONETARY=Dutch_Netherlands.1252, LC_NUMERIC=C and LC_TIME=Dutch_Netherlands.1252*

attached base packages: *stats, graphics, grDevices, utils, datasets, methods and base*

other attached packages: *pander(v.0.6.4), car(v.3.0-11), carData(v.3.0-4), psych(v.2.1.6), pwr(v.1.3-0), lsr(v.0.5), ggpubr(v.0.4.0), dplyr(v.1.0.7), ggplot2(v.3.3.5), stargazer(v.5.2.2), gulma(v.1.0.0.3), multi-level(v.2.6), nlme(v.3.1-139), rockchalk(v.1.8.144), mediation(v.4.5.0), sandwich(v.3.0-1), mvtnorm(v.1.1-2), Matrix(v.1.2-17) and MASS(v.7.3-51.4)*

loaded via a namespace (and not attached): *RColorBrewer(v.1.1-2), tools(v.3.6.0), backports(v.1.2.1), utf8(v.1.2.1), R6(v.2.5.0), rpart(v.4.1-15), Hmisc(v.4.5-0), DBI(v.1.1.1), colorspace(v.2.0-2), nnet(v.7.3-12), withr(v.2.4.2), mnormt(v.2.0.2), tidyselect(v.1.1.1), gridExtra(v.2.3), curl(v.4.3.2), compiler(v.3.6.0), cli(v.3.0.0), htmlTable(v.2.2.1), labeling(v.0.4.2), scales(v.1.1.1), checkmate(v.2.0.0), stringr(v.1.4.0), digest(v.0.6.27), foreign(v.0.8-71), minqa(v.1.2.4), rmarkdown(v.2.9), rio(v.0.5.27), base64enc(v.0.1-3), jpeg(v.0.1-8.1), pkgconfig(v.2.0.3), htmltools(v.0.5.1.1), lme4(v.1.1-27.1), highr(v.0.9), htmlwidgets(v.1.5.3), rlang(v.0.4.11), readxl(v.1.3.1), rstudioapi(v.0.13), farver(v.2.1.0), generics(v.0.1.0), zoo(v.1.8-9), zip(v.2.2.0), magrittr(v.2.0.1), Formula(v.1.2-4), Rcpp(v.1.0.6), munsell(v.0.5.0), fansi(v.0.5.0), abind(v.1.4-5), lifecycle(v.1.0.0), stringi(v.1.6.2), yaml(v.2.2.1), plyr(v.1.8.6), grid(v.3.6.0), parallel(v.3.6.0), forcats(v.0.5.1), crayon(v.1.4.1), lattice(v.0.20-38), kutils(v.1.70), haven(v.2.4.1), splines(v.3.6.0), hms(v.1.1.0), tmvnsim(v.1.0-*

2), *knitr*(v.1.33), *pillar*(v.1.6.1), *boot*(v.1.3-22), *ggsignif*(v.0.6.2), *lpSolve*(v.5.6.15), *glue*(v.1.4.2), *evaluate*(v.0.14), *latticeExtra*(v.0.6-29), *data.table*(v.1.14.0), *png*(v.0.1-7), *vctrs*(v.0.3.8), *nloptr*(v.1.2.2.2), *cellranger*(v.1.1.0), *gtable*(v.0.3.0), *purrr*(v.0.3.4), *tidyr*(v.1.1.3), *assertthat*(v.0.2.1), *xfun*(v.0.24), *openxlsx*(v.4.2.4), *xtable*(v.1.8-4), *broom*(v.0.7.8), *rstatix*(v.0.7.0), *survival*(v.3.2-11), *tibble*(v.3.1.2), *cluster*(v.2.0.8) and *ellipsis*(v.0.3.2)