

Metaanalyse und die Entwicklung eines dynamischen Forest-Plots mit Excel und SAS

Alexander Wagner Jun.
iAlternative München
Waldmeisterstraße 56
80935 München
E-Mail: a_wagner@gmx.de

Zusammenfassung

In der modernen medizinischen und gesundheitswissenschaftlichen Forschung spielt Metaanalyse eine große Rolle. Es gibt heutzutage eine ganze Reihe von kommerziellen und freien Software für die Metaanalyse und für die Erstellung eines Forest-Plots. Diese Grafiken sind in der Regel im PDF-Format [RevMan], MS Word, MS PowerPoint [CMA] oder als Picture-Format entwickelt und sind in der Regel statisch. Um diese Grafiken in einen Bericht einzufügen, muss man sie erst ausschneiden, kopieren und erst danach im Bericht platzieren. Oder man implementiert sie direkt als Bild-Objekt in den Bericht. Dies führt zu Schwierigkeiten bei Anpassung und Korrektur/Verbesserung, zum Beispiel bei der Änderung von Schriftgröße, Farbe, Überschriften, Labels und etc. Häufig muss man das System auch neu starten, um eine neue Grafik zu entwickeln, was zu Zeitverlust und evtl. Fehlern führt. MS Excel hat ein effektives Instrument für die Entwicklung der Grafiken verschiedener Art. Die Erzeugung eines dynamischen Forest-Plot mit Hilfe von MS Excel ist eine interessante und reizvolle Aufgabe.

Schlüsselwörter: Forest-Plot, Meta-Analyse, AMNOG, MS-Excel, AutoIt, VBS

1 Einführung

Zurzeit existiert keine vernünftige Lösung zur Erstellung eines Forest-Plot in Excel. Viele von vorgeschlagenen Lösungen werden entweder in Excel manuell entwickelt, was erheblichen Zeitaufwand bedeutet und auch sehr eingeschränkte visuelle Möglichkeiten für Grafiken einer solchen Art mit sich bringt.

Nun haben wir uns zur Aufgabe gestellt, eine Lösung auszuarbeiten, die Forest-Plots direkt und vollautomatisch in Excel erzeugt. Wir haben ein Programmtool für die automatische und dynamische (interaktive) Erzeugung der Grafiken entwickelt. Eine solche Lösung soll Zeit sparen, Fehler vermeiden und veränderbar sein. Als Hauptinstrument nahmen wir uns SAS mit entsprechenden Statistischen Prozeduren und Datenmanagement, Excel als Präsentationsgrundlage und AutoIt mit einem VBS ähnlicher Programmiersprache zur Grafenentwicklung und Datenaustausch.

Durch die Verwendung des Programmtools können Zeitaufwand, Betriebskosten und Fehlerrate bei der Entwicklung des AMNOG Dossiers reduziert werden.

Einige Beispiele für Forest-Plots in Excel

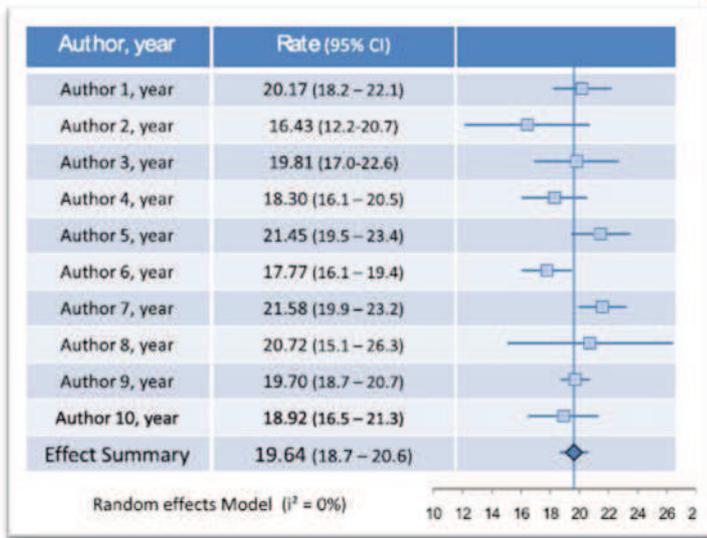


Abbildung 1: Beispiel Forest-Plot aus [14]

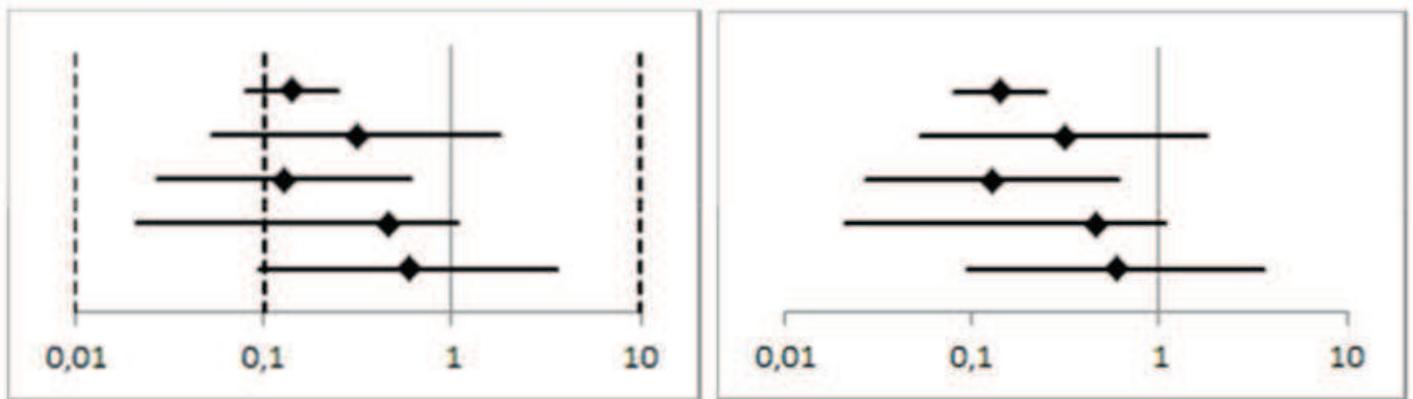


Abbildung 2: Beispiele Forest-Plots aus [12]

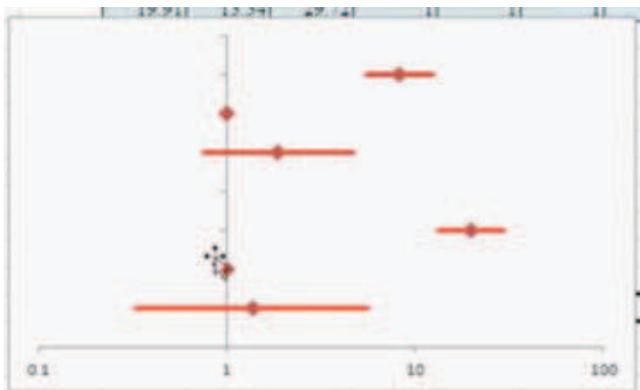


Abbildung 3: Beispiel Forest-Plot aus [13]

Einige Schwachstellen in existierenden Excel Forest-Plot Lösungen sind:

- in Abbildung 1 sind alle Quader gleich groß.
- in den Abbildungen 2 und 3 fehlen die Überschriften und alle Bezeichnungen, es fehlen die Daten und Diamond für den Gesamteffekt
- alle Grafen werden manuell erstellt.

2 Entwicklung der Excel Forest-Plot in Meta-Analyse

2.1 Durchführung der Meta-Analyse und Erzeugung einer SAS Output-Datei

Als Grundlage für die Erzeugung eines Forest-Plots wird eine speziell strukturierte SAS-Datei genommen. Diese Datei ist eine Output-Datei der Meta-Analyse und wird mit Hilfe von SAS Macros erzeugt. Die Meta-Output Datei ist in Abbildung 4 dargestellt. Diese Datei kann man zum Erzeugen von Grafen in MS Excel verwenden. Um diese Datei ins Excel zu übertragen, speichern wir sie zuerst mit dem SAS Befehl **clipbrd** in einer Zwischenablage. Der dazugehörige Programmblock befindet sich unten.

```
%let pfadprojekt=C:\Konferenz-KSFE-2015\Chart;
LIBNAME META "&pfadprojekt\SAS_DATEN";

FILENAME temp clipbrd;
data _null_;
LENGTH SUBGR $200;
RETAIN SUBGR;
file TEMP;
SET META.DATEN_SG end=eof;
PUT Dat1_1 @+1 Dat1_2 @+1 Dat1_3 @+1 Dat2_1 @+1 Dat2_2 @+1 Dat2_3 @+1
subgroup @+1 PR /;
RUN;
X "&pfadprojekt\PROGRAMME\ScatterChart.exe";
```

Studiennummer	Studienname	n/N	ES	95%KI	I ²	OR
1	Niedrigdosis	4/707	5/707	0.0	1.00	
2	Niedrigdosis	A 1997	4/707	0.4	0.00	
3	Niedrigdosis	B 2000	1/23164	1/23167	12.2	0.97
4	Niedrigdosis	C 2002	2293/15255	1256/9045	51.4	1.05
5	Niedrigdosis	H 2003	122/940	145/942	8.8	0.82
6	Niedrigdosis	Gesamt	2501/20066	1502/13554	0.0	1.00
7	Niedrigdosis	Heterogenität	Q=3.50, df=3, p=0.321, I ² =14.2%	0.0	0.0	1.00
8	Niedrigdosis	Gesamteffekt	Z Score=0.04, p=0.972, Tau=0.045	0.0	0.0	1.00
9	Hochdosis	Hochdosis	125/4297	120/4403	0.0	1.00
10	Hochdosis	D 1985	125/4297	120/4403	9.2	1.10
11	Hochdosis	E 1987	101/3272	95/3297	7.5	1.05
12	Hochdosis	F 1992	134/1051	157/1102	8.9	0.79
13	Hochdosis	G 1996	9/441	8/442	0.7	1.13
14	Hochdosis	Gesamt	372/9091	391/9244	0.0	0.97
15	Hochdosis	Heterogenität	Q=4.01, df=3, p=0.260, I ² =25.2%	0.0	0.0	1.00
16	Hochdosis	Gesamteffekt	Z Score=0.32, p=0.749, Tau=0.092	0.0	0.0	1.00
17	Alle	Alle	2873/29157	1973/33098	0.0	1.00
18	Alle	Gesamt	2873/29157	1973/33098	100.0	0.99
19	Alle	Heterogenität	Q=7.95, df=7, p=0.335, I ² =12.1%	0.0	0.0	1.00
20	Alle	Gesamteffekt	Z Score=0.17, p=0.868, Tau=0.043	0.0	0.0	1.00
21	Total	Novum vs. Kontrolle - Subgruppen nach Dosierung	1-Jahres-Monatsat	Modell mit zufälligen Effekten - DerSimonian und Laird	Heterogenität zwischen Subgruppen	Novum besser: Q=0.45, df=1, p=0.502, I ² =0%

Abbildung 4: Eine Meta-Output Datei in SAS zur Entwicklung eines Forest-Plots

2.2 Entwicklung Forest-Plot in MS Excel

Die Entwicklung eines Forest-Plots in MS Excel wird mit dem AutoIt erzeugt. Im ersten Schritt wird eine SAS Datei aus der Zwischenablage in ein AutoIt Array eingelesen. Folgende Befehle erzeugen dieses Array aus der SAS Output-Datei im Zwischenspeicher.

```
$content = ClipGet()
$array = StringSplit(StringStripCR($content), @LF)
$Base = _StringTo2DArray(ClipGet())
```

Das Array ist in Abbildung 5 dargestellt.

Row	Col 0	Col 1	Col 2	Col 3	Col 4	Col 5	Col 6	Col 7
[0]	Niedrigdosis	4/707	5/707	0.0	1.00	[1.00, 1.00]	Niedrigdosis	0
[1]	A 1997	4/707	5/707	0.4	0.80	[0.21, 2.99]	Niedrigdosis	1
[2]	B 2000	172/3164	176/3157	12.2	0.97	[0.78, 1.21]	Niedrigdosis	1
[3]	C 2002	2203/15255	1256/9048	51.4	1.05	[0.97, 1.13]	Niedrigdosis	1
[4]	H 2003	122/940	145/942	8.8	0.82	[0.63, 1.06]	Niedrigdosis	1
[5]	Gesamt	2501/20066	1582/13854	0.0	1.00	[0.91, 1.10]	Niedrigdosis	1
[6]	Heterogenität: Q=3...	0/0	0/0	0.0	1.00	[1.00, 1.00]	Niedrigdosis	2
[7]	Gesamteffekt: Z Sc...	0/0	0/0	0.0	1.00	[1.00, 1.00]	Niedrigdosis	2
[8]	Hochdosis	128/4297	120/4403	0.0	1.00	[1.00, 1.00]	Hochdosis	0
[9]	D 1985	128/4297	120/4403	9.2	1.10	[0.85, 1.41]	Hochdosis	1
[10]	E 1987	101/3272	96/3297	7.5	1.06	[0.80, 1.41]	Hochdosis	1
[11]	F 1992	134/1081	167/1102	9.8	0.79	[0.62, 1.01]	Hochdosis	1
[12]	G 1996	9/441	8/442	0.7	1.13	[0.43, 2.96]	Hochdosis	1
[13]	Gesamt	372/9091	391/9244	0.0	0.97	[0.81, 1.16]	Hochdosis	1
[14]	Heterogenität: Q=4...	0/0	0/0	0.0	1.00	[1.00, 1.00]	Hochdosis	2

Abbildung 5: AutoIt Array

Für Entwicklung des Forest-Plots wird ein Scatter Plot von MS Excel genutzt. Jede Studie „Niedrigdosis“ oder „Hochdosis“ besteht aus mehreren Serien „Row“ in AutoIt Array, zum Beispiel: „A 1997“, „B 2000“, usw. Einer Serie entspricht eine horizontale Linie. Jede diese Linie hat eine bestimmte Länge. Im Forest-Plot Graphen sind es die x-Koordinaten für den Anfangs- und den Endwert dieser Linie, die die Länge der Linien definieren.

Jede Plot-Serie „Row [1]“, „Row [2]“ usw. wird aus zwei Punkten „LowCI“ und „UpperCI“. Diese Punkte lesen wir aus der Spalte „Col 5“ des AutoIt-Arrays (Trennzeichen ist ","). Zum Beispiel, in der Zeile 2 „Row [1]“ und Spalte 5 „Col 5“ befindet sich ein Text "[0.21, 2.99]", es bedeutet für Serie „A 1997“: X1=0.21 und X2=2.99. Und so weiter für alle Zeilen in Array, z.B. Serie „Row [13]“: X1=0.81 und X2=1.16.

Am Ende bilden wir noch eine Serie für die Effektschätzungen aller Studie, diese Serie besteht aus N Punkten, wo N ist Anzahl der Studie ist.

Hier ist der Autoit-Programmcode für die Serienbildung:

```
For $SER=0 To $Nrow-1
    .SeriesCollection.NewSeries
    .SeriesCollection($SER+1).XValues=$X
    .SeriesCollection($SER+1).Values=$Y
    .SeriesCollection($SER+1).FORMAT.Line.Weight = 1
    .SeriesCollection($SER+1).Format.Line.FOREColor.RGB = 0
    .SeriesCollection($SER+1).MarkerStyle = -4142
Next
```

Die temporäre \$X, \$Y Autoit-Arrays sind Koordinaten der linken und rechten Punkte für die Linie jeder einzelnen Serie (Studie). Der Befehl `.MarkerStyle = -4142` setzt Leerzeichen am Anfang/Ende (links und rechts) der einer Linie jeder Serie.

Setzung der Quadrate für die Effektschätzung und der Labels für die Serien:

Die Punktschätzer jeder Studie nehmen wir aus Spalte 4 des Arrays. Zu diesem Zweck benötigen wir die Anzahl der Points. Mit folgendem Befehl definieren wir die Anzahl der Points (`$Npoints`) in der Serie der Effektschätzer:

```
$Npoints = .SeriesCollection($NStudy).Points.Count
```

wobei `$NStudy` eine Nummer der Serie für Effektschätzung ist.

Für jeden Point setzen wir Marken in Form eines schwarzen Quadrats

(`$xlMarkerStyleSquare`).

```
.SeriesCollection($Nrow).Points($I).MarkerStyle = $xlMarkerStyleSquare
.SeriesCollection($Nrow).Points($I).MarkerSize = $Size
```

wobei **\$Size** die Größe der Effektschätzer und gleich der Wert aus der Spalte 4 ist. Die **\$Top** Variable (Y-Koordinate der Effektschätzungen) verwenden wir für die Bestimmung der Positionen der Labels des Forest-Plots.

```
For $i = 1 To $Npoints
    .SeriesCollection($Nrow).Points($I).MarkerStyle = $xlMarkerStyleSquare
    .SeriesCollection($Nrow).Points($I).MarkerBackgroundColor = $nColor
    .SeriesCollection($Nrow).Points($I).MarkerForegroundColor = $nColor
    .SeriesCollection($Nrow).Points($I).hasDataLabel = True
    $TOP=.SeriesCollection($Nrow).Points($I).DataLabel.Top
Next
```

Labels setzen:

Für jede Serie setzen wir die Labels aus dem Autoit-Array. Die sind folgende:

- Studienname
- n1/N1
- n2/N2
- Effektmaß
- Gewicht
- Koordinaten.

`$SYi=$Top` für `$Points i`

A. Wagner Jun.

$\$SXi$ definieren wir für jede Labels-Typ in Forest-Plot, z.B für die Spalte "Studienname" $\$SXi=0$, für die Spalte "Gewicht" $\$SXi=250$ u.s.w.
Alle Labels befinden sich in der Spalte "Col 0" des Autoit-Array.

Achsen und Beschriftung

Die horizontalen und vertikalen Achsen definiert man mit folgenden Autoit-Programmcode:

Horizontal-Achse

```
.Axes(1).HasTitle = False  
.Axes(1).Crosses = $xlAutomatic  
.Axes(1).MajorGridlines.Format.Line.Visible = False  
.Axes(1).ScaleType = $xlLogarithmic
```

Variable $\$xlLogarithmic$ zeigt, das die X-Scala logarithmisch ist.

Vertikal-Achse

```
.Axes(2).Crosses = $xlAutomatic  
.Axes(2).MinimumScale = 0  
.Axes(2).MaximumScale = $Nrow+1 (Anzahl Serien)  
.Axes(2).MajorTickMark = $xlNone  
.Axes(2).TickLabelPosition = $xlNone  
.Axes(2).MajorGridlines.Format.Line.Visible = False
```

Beschriftung der Achsen:

Die Vertikale Achse hat keine Beschriftung.

Die horizontale Achse hat folgende Beschriftung:

"<Test-Medikament> besser <Kontrolle-Medikament> besser"

Title, Area und Plot Bereiche der Forest-Plots:

Hier sind die Befehle für die Definition der Höhe und Breite des Grafik-Bereiches, wobei $\$Nrow$ ist die Anzahl Zeilen in Autoit-Array, 15 – Höhe der Serien Zeile im Forest-Plot.

```
.ChartArea.Height = 150 + $Nrow*15  
.ChartArea.Width = 540
```

Hier sind die Befehle für die Definition der Höhe und Breite des Forest-Plot Bereiches.

```
.PlotArea.Height = 20 + $Nrow * 15  
.PlotArea.Width = 260  
.PlotArea.Top = 70  
.PlotArea.Left = 180
```

Hier sind die Befehle für die Definition der Überschriften des Forest-Plot.

Es konnte man mehrere Zeilen sein durch den Trennzeichen "A10x" getrennt.

Wo $\$Title$ Überschriftvariable ist.

```
.HasTitle = True  
.ChartTitle.Characters.Text = $Title  
.ChartTitle.LEFT=-4  
.ChartTitle.HorizontalAlignment=1  
$oTitle = $oChart.ChartTitle  
$oObject = $oTitle.Characters(1, 400)
```


4. Das Programm ist leicht installierbar, hat ein benutzerfreundliches Menu und ist einfach zu verwenden.
5. Das Programm ist universell und kann Forest-Plot unterschiedlicher Art erzeugen.

Literatur

- [1] http://support.sas.com/documentation/cdl_main/93/docindex.html
(Zugriff am 15.10.2013)
- [2] <http://www.g-ba.de/institution/themenschwerpunkte/arzneimittel/nutzenbewertung35a/anlagen/> (Zugriff am 15.10.2013)
- [3] AutoIt v3.3.8 released (Jul 30th, 2012).
<http://www.autoitscript.com/autoit3> (Zugriff am 15.10.2013)
- [4] Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011. Review Manager (RevMan) [Computer program]. Version 5.1
<http://ims.cochrane.org/revman>. (Zugriff am 15.10.2013)
- [5] Borenstein M, Hedges L, Higgins J, Rothstein H: Comprehensive Metaanalysis Version 2. Biostat, Englewood NJ; 2005, <http://www.meta-analysis.com>.
- [6] Guido Skipka. Tutorial „Software zur Durchführung von Metaanalysen Metaanalysen mit SAS, 55. GMDS-Jahrestagung, Mannheim, 2010,
<http://www.gmds2010.de/cms/wordpress/wp-content/uploads/fohlen/tu1-skipka.pdf>
- [7] William E Benjamin Jr. Give the Power of SAS to Excel Users Without Making Them Write SAS code. Proceedings of the SAS Global Forum 2013 Conference. Cary, NC: SAS Institute Inc. Available at.
<http://support.sas.com/resources/papers/proceedings13/010-2013.pdf>
- [8] Arthur S. Tabachneck. Copy and Paste from Excel to SAS. SAS Conference Proceedings: Midwest SAS User Group 2013,
<http://www.mwsug.org/proceedings/2013/RF/MWSUG-2013-RF04.pdf>
- [9] Michiel Hagendoorn at all, Save Those Eyes: A Quality-Control Utility for Checking RTF Output Immediately And Accurately. SAS Institute Inc. 2006. Proceedings of the Thirty-first Annual SAS Users Group International Conference. Cary, NC: SAS Institute Inc.
<http://www2.sas.com/proceedings/sugi31/066-31.pdf>
- [10] Donsung Cao, Automation of Comparing ODS RTF Outputs in Batch using VBA and SAS. PharmaSUG Conference Proceedings: PharmaSUG 2012,
<http://www.pharmasug.org/proceedings/2012/CC/PharmaSUG-2012-CC08.pdf>
- [11] Alexander Wagner: SAS-Tabellen und -Graphiken als MS PowerPoint-Applikation. Alte Probleme und neue Möglichkeiten. Ralf-Dieter Hilgers, Nicole Heussen, Wolfgang Herff, Carina Ortseifen (Hrsg.): KSFE 2008. Proceedings der 12. Konfe-

renz der SAS-Anwender in Forschung und Entwicklung (KSFE). Shaker Verlag, Aachen 2008.

- [12] Sylvia Büttner, Moodle-Kurs Excel 2010, Forest-Plot erstellen. Medizinische Fakultät Mannheim der Universität Heidelberg. Med. Statistik, Biomathematik und Informationsverarbeitung
http://www.umm.uniheidelberg.de/inst/biom/edv_kurse/skripten/Excel/ForestPlot.pdf
- [13] Forestplotexcel von Truong Hoang. 14.04.2013.
<https://www.youtube.com/watch?v=ZXXKysUK4BZI>
- [14] Jeruza L Neyeloff, Sandra C Fuchs and Leila B Moreira. Meta-analyses and Forest plots using a Microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. Neyeloff et al. BMC Research Notes 2012.
<http://www.biomedcentral.com/1756-0500/5/52>

Anhang:

1. SAS Macro für die Meta-Analyse

SAS Macro von Dr. Skipka

```
%ma_sg(Beispiel_Studiendaten.xls, %str(vergleich="Novum vs  
Kontrolle" and subgruppe=""), binaer,  
  %nrstr(Novum), %nrstr(Kontrolle), %nrstr(Nov),  
%nrstr(Kontrolle), %nrstr(Subgruppen nach Dosierung),  
  %nrstr(), %nrstr(), %nrstr(),  
  %nrstr(1-Jahres-Mortalität), mortalitaet, or, n, dsl, 2, 1, 1,  
1, %str(-2,-1,0,1,2,3),  
  0, 1, 0.7, %str(0,26,37,41,81,87,97), 0, d, besser,  
  Studienspezifikum,  
  %str(Niedrigdosis,Hochdosis),2,2,0,1);
```

Speichern Meta Output-Dazei in Zwischenspeicher

```
FILENAME temp clipbrd;  
data _null_;  
LENGTH SUBGR $200;  
RETAIN SUBGR;  
file TEMP;  
SET META.DATEN_SG end=eof;  
PUT Dat1_1 @+1 Dat1_2 @+1 Dat1_3 @+1 Dat2_1 @+1 Dat2_2 @+1 Dat2_3  
@+1 subgroup @+1 PR /;  
RUN;  
options noxwait;  
Start Autoit Scatter Plot Programm  
X "&pfadprojekt\PROGRAMME\ScatterChart12_SG.exe";
```

2. Autoit Programm für die Erzeugung des Forest-Plot in einem Word-Dokument kann vom Autor angefordert werden.