

Tabellendesign und Bericht: Mit SAS erzeugte Tabellen in formatierte Berichte einfügen

Renate Scheiner-Sparna
iOMEDICO AG
Ellen-Gottlieb-Straße 19
79106 Freiburg
Renate.Scheiner-
Sparna@iOMEDICO.com

Tanja Medinger
iOMEDICO AG
Ellen-Gottlieb-Straße 19
79106 Freiburg
Tanja.Medinger@iOMEDICO.com

Zusammenfassung

Die Berichte unserer Tumorregister enthalten Tabellen, Grafiken sowie Textabschnitte, in denen die Ergebnisse erläutert werden. Sie haben über die reinen Zahlen hinaus einen hohen Anspruch an intuitive Lesbarkeit, komprimierten Informationsgehalt und Design.

Mit SAS ist das Erstellen solcher Tabellen möglich. SAS kann auch Berichte generieren, jedoch werden unsere formatierten Berichte in einer Abteilung geschrieben, die keine SAS Kenntnisse hat. Außerdem ist SAS in der Textverarbeitung nicht komfortabel. Kopiervorgänge sind - auch mit Automatisierung - fehleranfällig und aufwändig.

Hier wird gezeigt, wie mit SAS erzeugte Tabellen in Microsoft Word Dokumente verknüpft werden. Formatierungen, Beschränkungen und Aspekte der Zusammenarbeit mit den verschiedenen Abteilungen, die den Bericht erstellen, werden betrachtet. Vor- und Nachteile zur Kopier-Methode werden aufgezeigt.

Schlüsselwörter: Bericht, Formatierung, ODS, TAGSETS.RTF

1 Ausgangssituation

Formatierte Berichte mit einem hohen Anspruch an intuitive Lesbarkeit und Design haben bei iOMEDICO eine lange Tradition. Sie werden von der Abteilung für Epidemiologie ausgearbeitet. Die Abteilung Statistik liefert die Auswertungen hierzu.

Über Jahre wurde Open Office als Textverarbeitungsprogramm verwendet. Die aufwändig formatierten Tabellen wurden mit Hilfe von Python-Skripten befüllt. Dieses Vorgehen hatte einige Vorteile, aber auch Nachteile. Es war ein sehr flexibles Verfahren, Tabellen konnten entworfen, geändert, erweitert, verkleinert, formatiert oder ausgeblendet werden, ohne dass die Tabellen der Statistikabteilung davon betroffen waren. Änderungen im Open Office Dokument wurden im Python-Skript abgebildet.

Die Python-Skripte entwickelten sich jedoch mit zunehmendem Umfang der Auswertungen zu unübersichtlichen Gebilden, die schwer zu warten waren. Kleine Änderungen erforderten höchste Konzentration, da die Zellenbereiche der Tabellen Zeilen- und Spaltenweise angesprochen wurden. Das Verfahren war zeitintensiv. Reviews wurden immer aufwändiger.

Hinzu kam, dass Open Office durch den wachsenden Umfang der Registerberichte an seine Grenzen stieß. Es wurde beschlossen, die verwendete Software und die Prozesse zu ändern.

2 Prozess der Tabellen für Registerberichte

Open Office wurde durch Microsoft Word abgelöst, da es mit den großen Dateien stabiler lief. Der Prozess wurde wie in Tabelle 1 dargestellt entwickelt:

Tabelle 1: Prozess der Tabellen für Registerberichte

Zuständigkeit	Prozessschritt
Epidemiologie / Statistik	Tabellendefinition
Epidemiologie	Aufsetzen von Teil-Dokumenten
Statistik	Programmierung und Ausgabe der formatierten Tabellen als separate RTF Files mit SAS
Epidemiologie	Einbinden/Verknüpfen der Tabellen, Einfügen von Überschriften, Fußnoten
Epidemiologie	Ausformulierung der Berichte, Reduktion der Tabellen nach Bedarf (Zeilen, Spalten)
Epidemiologie / Statistik	Review der Teil-Dokumente
Technical Writing	Zusammenstellung des Gesamt-Dokuments

Die Zusammenarbeit zwischen den beiden beteiligten Abteilungen erfordert es, dass umfangreiche Freiheiten für die Autoren des Berichts erhalten bleiben. Gleichzeitig soll die Tabellenprogrammierung einmalig erfolgen und von Anpassungen abgesehen werden.

3 Tabellendefinition der TAGSETS.RTF Tabellen

Zur Ausgabe aus SAS wird ODS TAGSETS.RTF verwendet, da dieses Format mit Microsoft Word kompatibel ist. Zwei Beispieltabellen sind in den Abbildungen 1 und 2 dargestellt. Im Original ist die erste Zeile in der Farbe des jeweiligen Registers hinterlegt.

Das allgemeine Layout der Tabellen wird durch einen ODS Style definiert, der im Projekt durchgängig angewendet wird. Zusätzlich zu den allgemeinen Layout-Vorgaben wie Schriftgröße, Rahmenlinien etc. müssen für die Verknüpfung ein paar Aspekte zusätzlich berücksichtigt werden.

Eine Tabelle darf keinen Seitenumbruch enthalten, da in diesem Fall nur die erste Seite der Tabelle im Word Dokument dargestellt wird. Dies ist eine Anforderung an die Tabellendefinition, die als Vorlage für die Programmierung dient, da SAS Tabellen, die nicht auf eine Seite passen, automatisch umbrechen.

	Total
Patients (N)	1132
Patients with surgery	
Yes n (%)	1166 (97.8%)
No n (%)	23 (1.9%)
Missing n (%)	3 (0.3%)
If surgery: Number of resected lymph nodes	
0 n (%)	418 (35.0%)
1 n (%)	311 (26.1%)
2 n (%)	113 (9.5%)
≥ 3 n (%)	306 (25.7%)
Unknown to site n (%)	15 (1.3%)
Missing n (%)	3 (0.3%)
If surgery: Number of metastatically affected lymph nodes	
0 n (%)	201 (16.7%)
1 n (%)	98 (8.1%)
2 n (%)	218 (18.3%)
≥ 3 n (%)	635 (53.3%)
Unknown to site n (%)	15 (1.3%)
Missing n (%)	1 (1.1%)

Abbildung 1: Beispieltabelle 1

	Aaaa	Ebbb	Cccc	Dddd	Eeee	Gesamt
Patienten (N)	76	48	21	20	16	92
Alter bei Therapiebeginn (Jahre)						
n	76	48	21	20	16	92
Median	75.5	76.3	68.7	76.7	76.2	76.8
ECOG						
ECOG 0 n (%)	18 (23.7%)	12 (25.0%)	7 (33.3%)	5 (25.0%)	5 (31.3%)	22 (23.9%)
ECOG ≥ 1 n (%)	29 (38.2%)	22 (45.8%)	5 (23.8%)	7 (35.0%)	5 (31.3%)	35 (38.0%)
Unbekannt n (%)	12 (15.8%)	6 (12.5%)	5 (23.8%)	3 (15.0%)	3 (18.8%)	14 (15.2%)
Ohne Angabe n (%)	17 (22.4%)	8 (16.7%)	4 (19.0%)	5 (25.0%)	3 (18.8%)	21 (22.8%)
Begleiterkrankungen vorliegend						
Ja n (%)	63 (82.9%)	40 (83.3%)	17 (81.0%)	17 (85.0%)	14 (87.5%)	77 (83.7%)
Nein n (%)	8 (10.5%)	4 (8.3%)	1 (4.8%)	2 (10.0%)	1 (6.3%)	9 (9.8%)
Ohne Angabe n (%)	5 (6.6%)	4 (8.3%)	3 (14.3%)	1 (5.0%)	1 (6.3%)	6 (6.5%)

Abbildung 2: Beispieltabelle 2

Die Breite der Tabelle auf der Seite ist definiert mit der ODS Option OUTPUTWIDTH = 17,0 cm (Hochformat) bzw. 27,1 cm (Querformat). Mit diesen Maßen können die vorgesehenen Ränder im Dokument eingehalten werden.

Die Fußnoten und Titel der Tabellen werden nicht mit SAS in das zu verknüpfende RTF File geschrieben, sondern im Word Dokument erstellt.

4 Details der Programmierung

Wenn Tabellen ohne Titel und Fußnoten generiert werden besteht die Schwierigkeit, ihre Herkunft nachvollziehbar zu halten. Aus diesem Grund wird parallel zur bereits beschriebenen Tabelle eine Tabelle identischen Inhalts und fast gleichen Namens mit Zusatzinformation erzeugt. Titel bzw. Fußnoten enthalten Zeitstempel und Datenstand. Die Namenskonventionen der Dateien geben weiteren Aufschluss über die Einordnung der Tabelle im Projekt.

Zur Standardisierung des Prozesses wurden Makros entwickelt. Das Makro `standard_out.sas` bündelt die Aktivitäten zur Ausgabe der Tabellen in den unterschiedlichen Formaten. Es wird aufgerufen:

```
%standard_out(inds, varlist, outname=&gprogramname., tabstatus=dev)
```

Die Makroparameter werden wie folgt belegt:

<code>inds</code>	Input Data Set, im Tabellenprogramm erzeugte SAS-Tabelle mit den darzustellenden Ergebnissen.
<code>varlist</code>	Liste der Variablen, die in der RTF Tabelle enthalten sein sollen, auch Sortier-Variablen.
<code>outname</code>	Namensstamm aller Outputs, vorbelegt mit dem Namen des SAS Programms.
<code>tabstatus</code>	Status der Programmierung, vorbelegt mit <code>dev</code> . <code>Tabstatus</code> wird auf <code>prod</code> gesetzt, wenn die Programmierung abgeschlossen ist. Dies wird über eine Makrovariable aus dem aufrufenden Programm gesteuert und bewirkt ein anderes Ausgabeverzeichnis sowie weitere Prozesse, die nur bei finalen Tabellen benötigt werden.

Der Ablauf des Makros enthält folgende Schritte, der SAS Code ist in Ausschnitten dargestellt:

- Zusammensetzen der Pfade für die Output Dateien in Filename Statements
- Festlegen der Tabellenbreite in Abhängigkeit von Hochformat / Querformat durch Auslesen aus den Dictionary Tables [1]

```
/* determine outputwidth depending on paper orientation;
proc sql noprint;
  select setting into: ori
  from dictionary.options
  where upcase(optname)='ORIENTATION';
quit;
%put Paper Orientation: &=ori;
%if %upcase(&ori.)=LANDSCAPE %then %let linkwidth=%str(27.1cm);
%if %upcase(&ori.)=PORTRAIT %then %let linkwidth=%str(17.0cm);
```

- Ausgabe eines SAS Data Sets mit den Ergebnissen und einigen zusätzlichen Spalten (z.B. Sortier-Variablen)
- Ausgabe einer TAGSETS.RTF Tabelle mit Titel und Fußnoten.
Diese Tabelle wird nur in der Statistik abteilungsintern verwendet, sie dient der Transparenz und Dokumentation.
Der Style ist projektübergreifend gültig.
Hier wird als Breite der Tabelle OUTPUTWIDTH=100% verwendet. Das geschachtelte Makro procreport enthält Standard- und tabellenspezifische Einstellungen für die Zeilen- und Spaltendefinition. Es wird im Tabellenprogramm jeweils definiert und vom standard_out Makro aufgerufen.

```
%*generate standard table with titles and footnotes;
ods tagsets.rtf file = tab style=<Standard Style>;
  %procreport (outputwidth=%nrbrquote(100%))
  run;
ods tagsets.rtf close;
```

- Speichern der Titles und Footnotes in einem Data Set, die aktuell aus dem Standard-Aufruf des Programms aktiv sind, durch Auslesen aus den Dictionary Tables

```
proc sql noprint;
  create table &ds1._tit as select *
  from dictionary.titles;
quit;
```

- Entfernen der Titles und Footnotes

```
title;
footnote;
```

- Ausgabe der zur Verknüpfung vorgesehenen TAGSETS.RTF Tabelle.
Hier wird ein adaptierter Style verwendet und über die Makrovariable LINKSTYLE eingesetzt, der die speziellen Projektbedürfnisse abbildet (z.B. Farbe, Linien). Die Option OUTPUTWIDTH wird entsprechend der Seitenausrichtung gesetzt.

```
ods tagsets.rtf file = linktab style=&linkstyle.;
  %procreport (outputwidth=%nrbrquote(&linkwidth.))
  <... additional syntax for grey lines inside table ...>
  run;
ods tagsets.rtf close;
```

- Titel und Fußnoten werden wieder mit den vorher gesicherten Inhalten gefüllt

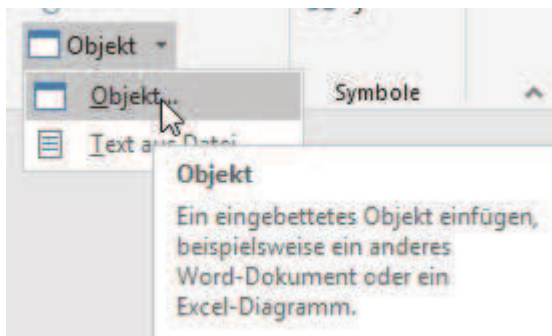
```
data _null_;
  set &ds1._tit;
  if type = 'T' then cmd = catt('title', number, " j=1 '",
trim(text), " '");
  else cmd = catt('footnote', number, " j=1 '", trim(text), " ' ');
  call execute(cmd);
run;
```

Ein Detail, das uns vor Herausforderungen stellte, war das automatische Erscheinen eines Zeilenumbruchs vor jeder Tabelle. Dieser führte zu Problemen im Dokument: Die Abstände zu den Titel-Zeilen wurden zu groß, manche Tabellen passten nicht mehr in den vorgesehenen Bereich. Der Umbruch musste anfangs für jede Tabelle von Hand entfernt werden. Wir lösen dieses Problem durch automatisierte Nachbearbeitung des .RTF Dokuments. Die beiden Zeilen, die den Umbruch erzeugen, werden nun gelöscht. Hierzu wird ein DATA _NULL_ Step mit INFILE und FILE Statement verwendet [2]:

```
data _null_;  
  infile "<pfad>\&outname._link0.rtf" lrecl = 500;  
  file "<pfad>\&outname._link.rtf" lrecl = 500;  
  input ;  
  if _infile_ = "{\*\bkmkstart IDX}{\*\bkmkend IDX}" then delete;  
  if _infile_ = "{\pard\par}" then delete;  
  put _infile_;  
run;
```

5 Details der Verknüpfung

Die Möglichkeit, Objekte in Microsoft Word Dokumente als Link einzufügen, bildet die Grundlage für dieses Verfahren. Über das Menü Einfügen lässt sich ein Objekt auswählen. Der Punkt „Verknüpfen“ muss markiert werden wie in Abb. 3 dargestellt, dann wird der Pfad im Dokument gespeichert. Wird dann die Tabelle an der abgelegten Stelle aktualisiert, wird diese Änderung auch im Word-File sichtbar.



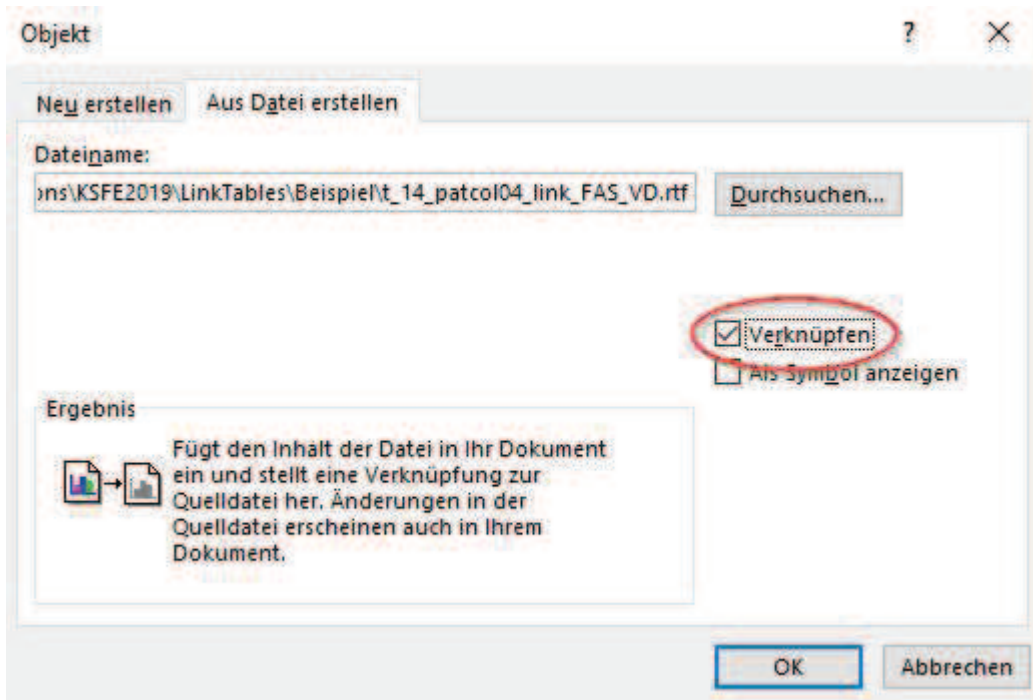
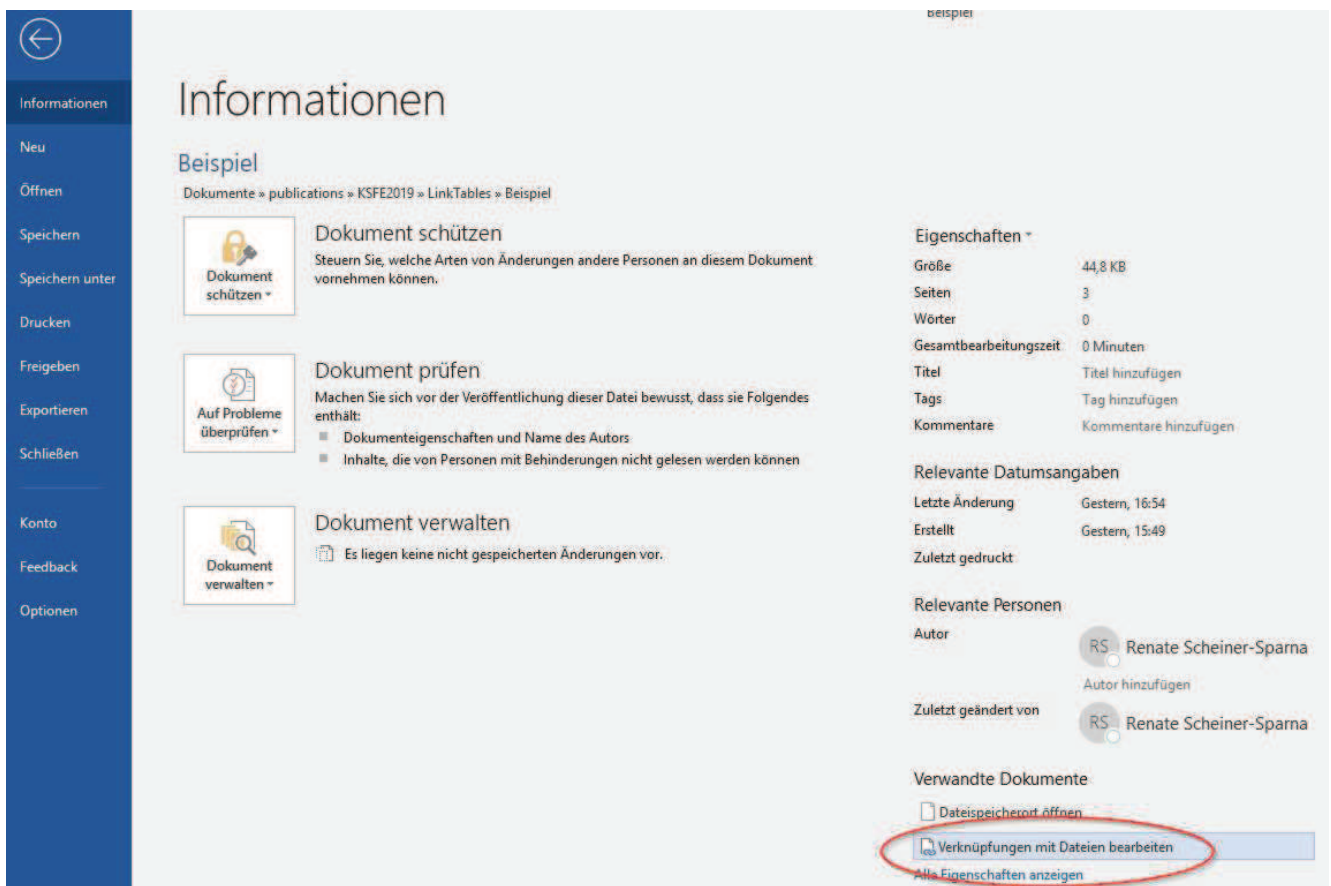


Abbildung 3: Objekt in Microsoft Word verknüpfen

Mit Doppelklick oder Rechtsklick – Verknüpftes Dokument-Objekt kann das .RTF File zum Bearbeiten direkt geöffnet werden. Word verfügt über ein Fenster, in dem die eingefügten Objekte mit ihrem Pfad sichtbar sind. Über das Menü Datei – Informationen – Verknüpfungen mit Dateien anzeigen kann es geöffnet werden (Abb. 4).



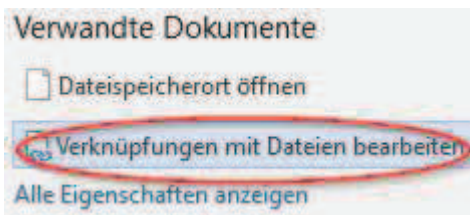


Abbildung 4: Verknüpfungen anzeigen und bearbeiten

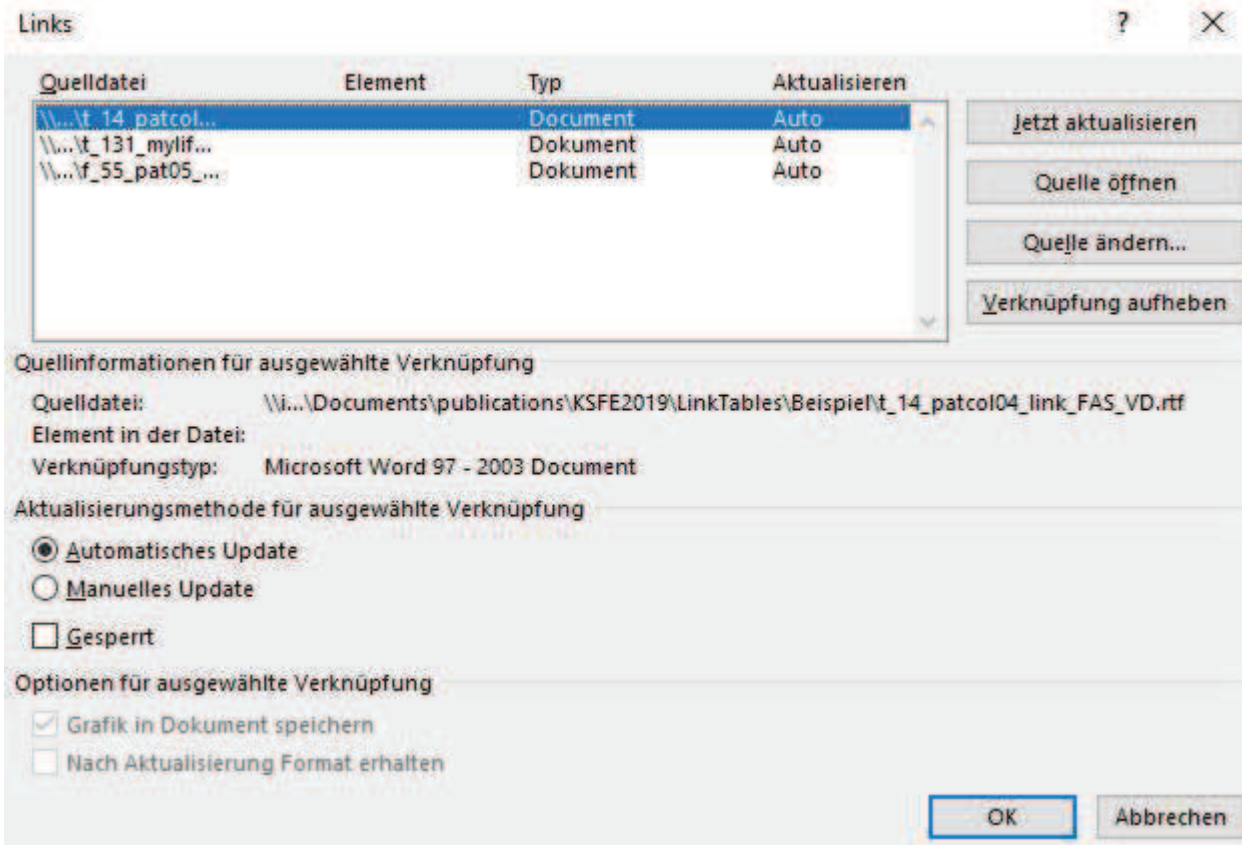


Abbildung 5: Verknüpfungen anzeigen und bearbeiten

Die Bearbeitung in diesem Fenster (Abb. 5) ist nicht komfortabel. Es ist nicht zu erkennen, welches Dokument an welcher Position eingefügt ist. Auch an der Position der verknüpften Tabelle gibt es keine direkte Information über ihren Ursprung.

Eine hilfreiche Funktion ist hier die Tastenkombination Alt+F9, mit der man zwischen der Anzeige der verknüpften Tabelle und der Anzeige des Links mit Pfad hin und her schalten kann. Dies ist die einzige Möglichkeit, den Pfad an Ort und Stelle sichtbar zu machen:

```
{ LINK Word.Document.8 \\\\iom-vs01\datenS\Home\scheiner-sparna\Documents\publications\KSFE2019\LinkTables\Beispiel\t_14_patcol04_link_FAS_VD.rtf \a \p \f 0 { EMBED Word.Document.8 \s }
```


6 Verknüpfung von Einzel- zu einem Master-Dokument

Bei der Erstellung von Berichten werden thematische Kapitel erstellt, in denen die Tabellen als Links enthalten sind. Diese zu einem Gesamtdokument (Master) zu kombinieren ist kein Problem, die Links werden von Microsoft Word durchgereicht, solange die Ablage der Tabellen sich nicht ändert. Es handelt sich um absolute Pfade, die beim Verschieben des Masterdokuments behalten werden. Zum Zweck der Archivierung, werden die Verknüpfungen entfernt und die Tabellen an sich eingebettet. In dem in Abb. 5 gezeigten Fenster kann man mehrere Verknüpfungen gleichzeitig markieren und mit dem Button „Verknüpfung aufheben“ die Tabellen einbetten.

7 Diskussion und Zusammenfassung

Die Verknüpfung von .RTF Tabellen mit einem Berichtsdocument ermöglicht es, ganz von Kopiervorgängen abzusehen und damit Fehleranfälligkeit zu reduzieren. Der Aufwand ist vertretbar, denn andere Prozesse wie z.B. Reviews werden erleichtert. Die Akzeptanz bei den beteiligten Mitarbeitern ist hoch, die Zusammenarbeit zwischen den Abteilungen funktioniert gut. Bei den Adressaten kommen die Berichte gut an.

Es erweist sich als sehr vorteilhaft, dass die RTF Tabellen als separates Objekt bearbeitet werden können. So können die Autoren der Berichte je nach Bedarf des Berichts in den Interimsauswertungen Spalten und Zeilen ausblenden, die noch nicht für den Bericht relevant sind, z.B. wegen zu geringer Fallzahlen. Fußnoten können ergänzt werden.

Die Übersicht über die Verknüpfungen in Microsoft Word ist unzureichend. Hier Transparenz zu erhalten erfordert zusätzliche Dokumentation, welche Tabelle an welcher Position im Dokument eingefügt ist. In Open Office war es möglich, über Rechtsklick – Beschreibung in einem Fenster ein Objekt übersichtlich zu kommentieren. Diese Funktion hält Microsoft Word nicht bereit, sie wäre in diesem Zusammenhang sehr hilfreich.

8 Danksagung

Herzlicher Dank geht an unsere Kolleginnen und Kollegen, die in ihrem jeweiligen Bereich wertvollen Input gegeben haben:

Renate Grugel, David Hamm, Kirsten Biedermann, Marco Chiabudini.

Literatur

- [1] R. Scheiner-Sparna: Dictionary Tables - Metadaten meines Workflows in SAS. In: Weyermann, Minkenber, Muche (Hrsg.): KSFE 2017, Proceedings der 21. Konferenz der SAS[®]-Anwender in Forschung und Entwicklung (KSFE). Shaker-Verlag Aachen 2017.
- [2] <https://support.sas.com/resources/papers/proceedings/pdfs/sgf2008/166-2008.pdf>
06.02.2019