

# **Automatisches Erstellen und Verschicken eines Benchmarking-Reports zur Qualitätssicherung dokumentierter Diabetesdaten mittels SAS 9.3**

Katharina Fink, Esther Molz, Andreas Hungele,  
Ramona Ranz, Matthias Grabert, Reinhard Holl  
Institut für Epidemiologie und med. Biometrie, ZIBMT  
Universität Ulm,  
Albert-Einstein-Allee 41  
89081 Ulm  
Katharina.Fink@uni-ulm.de

## **Zusammenfassung**

Eine große Datensammlung von Patienten mit Diabetes aus dem deutschsprachigen Raum wird mit der DPV-Software (Diabetes-Patienten-Verlaufsdokumentation) erhoben. Den teilnehmenden Zentren wird die DPV-Software kostenlos zur Verfügung gestellt. Dafür übermitteln sie zweimal jährlich ihre anonymisierten Daten nach Ulm.

Zur Qualitätssicherung wird für jedes Zentrum eine umfangreiche Vergleichsauswertung (ca. 100 Seiten), das DPV-Benchmarking, angefertigt. Die darin enthaltenen Grafiken dienen der Kontrolle der eigenen Dokumentation. Außerdem ist ein Vergleich mit anderen Zentren möglich.

Das Benchmarking wird durch drei Programme mit Hilfe von SAS 9.3 erstellt. Das erste Programm bereitet die Daten auf und berechnet die Qualitätsindikatoren für die Grafiken. Das Hauptprogramm legt das Benchmarking-Dokument an und ruft ein drittes Programm auf, welches Makros für die Erstellung unterschiedlicher Grafiken enthält. Im Benchmarking können unter anderem Querschnitts-, Trend-, und bivariate Grafiken dargestellt werden.

Durch die Angabe des Zentrums Kürzels im Hauptprogramm wird das Benchmarking für dieses Zentrum, in Form eines PDF-Dokumentes, automatisch ausgegeben und mit individueller Anrede per E-Mail an die Hauptansprechpartner dieses Zentrums geschickt.

Das Benchmarking kann in verschiedenen Varianten erstellt werden (jeweils getrennt nach pädiatrischen und internistischen Zentren): Zentrumsidentität anonym, Zentrumsidentität offen, Darstellung aller Zentren aus Deutschland, Darstellung aller Zentren aus Österreich, Darstellung für separate Anwendergruppen.

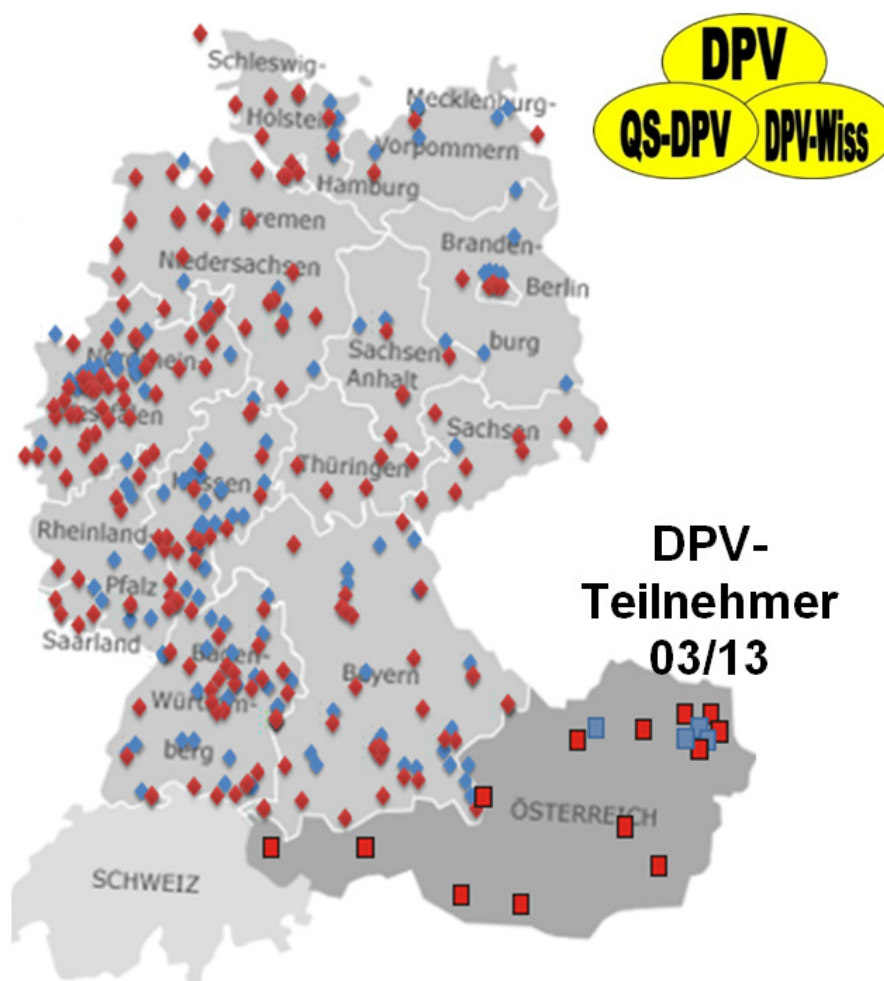
**Schlüsselwörter:** Benchmarking, Makro, FILE EMAIL, GPLOT-Prozedur, ODS PDF

## **1 Diabetes-Patienten-Verlaufsdokumentation**

Eine große Sammlung von Diabetesdaten wird mit der DPV-Software (Diabetes-Patienten-Verlaufsdokumentation) erhoben. Diese Software wird an der Universität Ulm, am Institut für Epidemiologie und medizinische Biometrie, ZIBMT, kontinuierlich weiterentwickelt. Sie ermöglicht eine standardisierte, strukturierte und prospektive Dokumentation von Diabetespatienten. Neben Untersuchungsbefunden, wie HbA1c, Blut-

druck oder Folgeerkrankungen an Herz, Niere oder Füßen, können auch Therapie-maßnahmen erfasst und abgebildet werden. Zusätzliche Funktionen, welche die DPV-Software bietet, sind die automatische Arztbriefschreibung, die Erinnerung an ausstehende Untersuchungen, die grafische und tabellarische Darstellung des Diabetesverlaufs und vieles mehr.

Die DPV-Software war ursprünglich für die Kinderdiabetologie gedacht. Inzwischen sind aber weit mehr Erwachsene als Kinder dokumentiert. Aktuell (Datenbestand 03/2013) liegen anonymisierte Verlaufsdaten von 308.049 Patienten (ca. 2,5 Millionen Kontakttermine) aus 393 Zentren vor. Die Karte zeigt DPV-Anwender in Deutschland und Österreich (siehe Abbildung 1).



**Abbildung 1:** Karte der DPV-Anwender

Die teilnehmenden Zentren bekommen die DPV-Software kostenlos zur Verfügung gestellt. Dafür sind sie bereit, zweimal jährlich ihre anonymisierten DPV-Daten an das Institut für Epidemiologie und medizinische Biometrie zu schicken. Das Institut verwendet die bearbeiteten Daten für wissenschaftliche Auswertungen. Die Regeln für den Umgang mit den Daten legt ein Gremium fest, das von den DPV-Anwendern gewählt wird.

Um möglichst fehlerfreie Daten für die Auswertungen zu erhalten, werden Inkonsistenzen mit Hilfe eines Korrekturlaufs ermittelt und zur Überprüfung an die Zentren zurückgeschickt. Nach der Bereinigung der Daten senden die Zentren ihre anonymisierten Archive erneut. Es findet kein Monitoring der Daten vor Ort statt.

Ein Visual FoxPro<sup>1</sup>-Programm und drei SAS-Programme setzen die Daten zusammen, benennen Variablen um, eliminieren doppelte Datensätze, erkennen Patienten, die das Zentrum gewechselt haben und vieles mehr.

## **2 Benchmarking-Report zur Qualitätssicherung**

Das DPV-Benchmarking ist eine umfangreiche Vergleichsauswertung im PDF-Format. Mit den darin enthaltenen Grafiken können die Zentren die Qualität ihrer Patientenbetreuung und ihrer eigenen Dokumentation überprüfen und sich mit anderen Zentren vergleichen. Die Auswertung wird zweimal jährlich per E-Mail an die Anwender verschickt, im Frühjahr für das gesamte Vorjahr und im Sommer für die erste Hälfte des aktuellen Jahres.

### **2.1 Wie wird das Benchmarking erstellt?**

Das Benchmarking wird durch drei Programme mit Hilfe von SAS 9.3 erstellt. Das erste Programm bereitet die Daten auf und berechnet die Qualitätsindikatoren für die Grafiken. Das Hauptprogramm legt das Benchmarking-Dokument an und ruft ein drittes Programm auf, welches Makros für die Erstellung unterschiedlicher Grafiken enthält. Das Benchmarking wird als PDF-Dokument mit Hilfe von ODS PDF erstellt. Im Benchmarking wird jedes Zentrum durch einen Balken repräsentiert. Die Zentrumsidentität ist anonym, jedes Zentrum erkennt nur seinen eigenen Balken, andere Zentren sind nicht identifizierbar. Aus diesem Grund wird das Benchmarking für jedes Zentrum einzeln erstellt.

### **2.2 Wie ist das Benchmarking aufgebaut?**

Das Benchmarking besteht aus folgenden Bereichen:

- Deckblatt
- Anmerkungen: Auf dieser Seite werden allgemeine Informationen und Neuerungen im Benchmarking aufgeführt.
- Inhaltsverzeichnis mit Link-funktion
- Liste der teilnehmenden Zentren
- Deutschland-/Österreich-karte: Die Karte wird auto-matisch mit SAS erstellt und enthält immer nur die aktuell teilnehmenden Einrichtungen, die durch einen Stern gekennzeichnet sind (siehe Abbildung 2).
- Inhalt: Der Inhalt besteht aus verschiedenen Blöcken, die jeweils durch Zwischen-seiten im PDF eingeleitet werden.

---

<sup>1</sup> Programmierumgebung von Microsoft

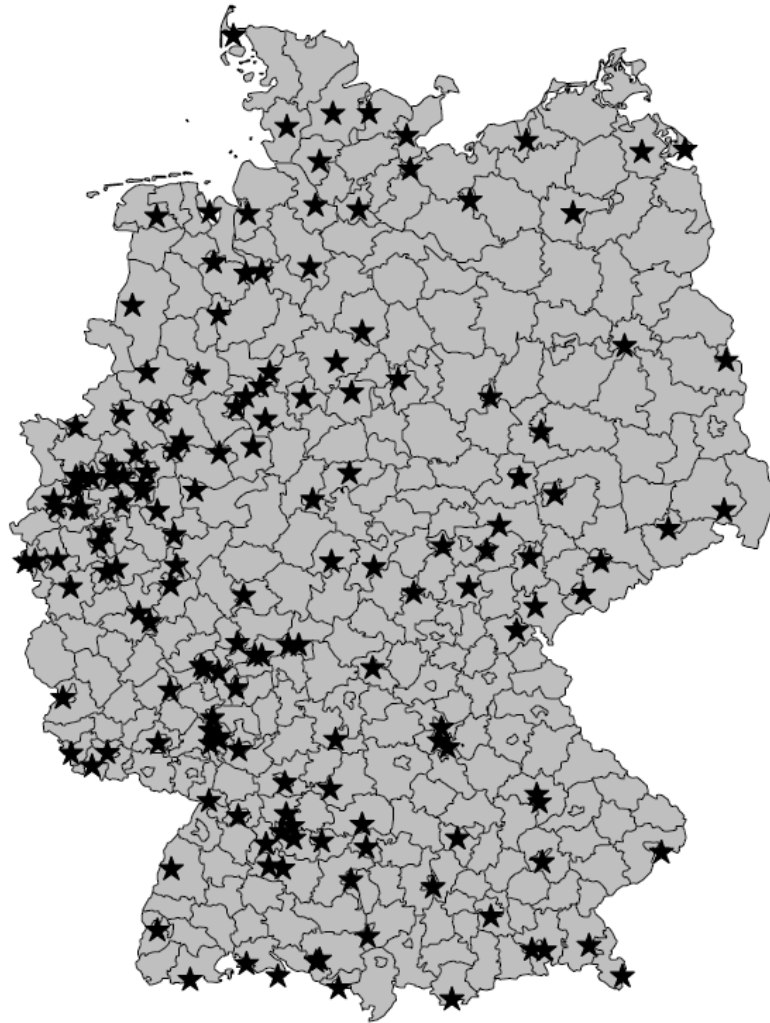


Abbildung 2: Deutschlandkarte

## 2.3 Welche Grafiken werden im Benchmarking verwendet?

Im Benchmarking werden verschiedene Grafiken dargestellt, unter anderem Querschnitts- und Trendgrafiken. Sie werden über Makros erstellt, damit sie variabel für verschiedene Parameter sind. Für diese beiden Grafiktypen wird die SAS-Prozedur GPLOT verwendet.

### 2.3.1 Trendgrafiken

In Abbildung 3 wird eine Trendgrafik für die Variable Alter dargestellt. Auf der x-Achse werden die Behandlungsjahre 1995-2012 eingezeichnet, die y-Achse enthält das Alter der Patienten in Jahren. Das mediane Alter der Patienten im eigenen Zentrum wird als schwarze Linie mit Punkten dargestellt, der Median aller Patienten als blaue Linie mit ausgefüllten Quadraten.

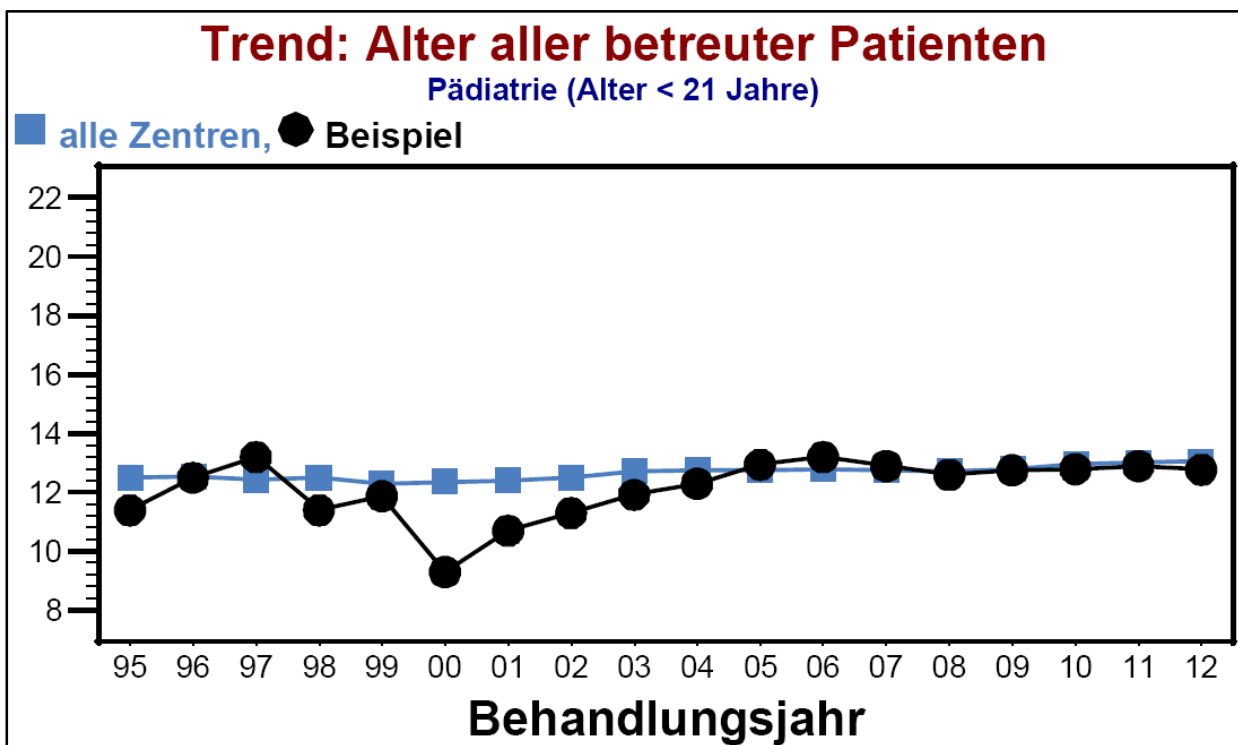


Abbildung 3: Trendgrafik für Alter in Jahren

Beim Aufruf des Makros im Hauptprogramm werden alle individuellen Parameter übergeben, das zu verwendende Data-Set, die Variable, die in der Grafik abgebildet werden soll, die Einheit für die y-Achse (PERCENT oder NORMAL) und der Titel, der oberhalb der Grafik ausgegeben werden soll.

```
%trend ( indata = dpva,
         _var    = alter,
         _unit   = NORMAL,
         _title  = "Trend: Alter aller betreuter Patienten");
```

Das eigentliche Grafikmakro befindet sich im dritten SAS-Programm. Hier werden die Variablen “\_vara“ und “\_vari“ erstellt. In “\_vara“ wird der mediane Wert aller Zentren gespeichert, in “\_vari“ der Wert des eigenen Zentrums.

```
%MACRO trend (_indata, _var, _unit, _title);

DATA _trenddat;
  SET &_indata.;
  IF zentrum = &eigenes_Zentrum. THEN _vari = &_var.;
  IF zentrum EQ 'all'           THEN _vara = &_var.;
RUN;
```

Im PLOT-Statement werden die Behandlungsjahre und die Variablen “\_vara“ und “\_vari“ angegeben. Damit nicht zwei einzelne Grafiken (“\_vara \* jahr“ und “\_vari \* jahr“) gezeichnet werden, können beide Grafiken mit der Option OVERLAY übereinan-

der ausgegeben werden. Zwei SYMBOL-Statements definieren die unterschiedlichen Farben und Symbole für die Linien.

```
SYMBOL1 i=j c=blue f=marker v='U' h=0.5 l=1 w=2.5;  
SYMBOL2 i=j c=black v=dot h=1 l=1 w=2.5;  
  
PROC GPLOT DATA=_trenddat;  
  PLOT (_vara _vari) * jahr / OVERLAY  
                                VAXIS = axis2  
                                HAXIS = axis3;  
  
RUN;  
QUIT;  
  
%MEND trend;
```

### 2.3.2 Querschnittsgrafiken

Abbildung 4 zeigt eine Querschnittsgrafik für die Vollständigkeit des Parameters “Zigarettenrauchen angesprochen“. Dieser Parameter muss mindestens einmal pro Jahr erfragt werden. Außerdem gehen nur Patienten ab 11 Jahren in die Grafik ein. Hier wird jedes Zentrum durch einen einzelnen Balken repräsentiert. Der schwarze Balken mit der Zahl darüber stellt den Balken des eigenen Zentrums dar. Alle anderen Balken sind in Ampelfarben unterteilt, um eine gute, mittlere und schlechte Vollständigkeit des Parameters zu symbolisieren. Die graue Linie stellt den Median aller Einrichtungen dar. Es werden die Daten des jeweiligen Beobachtungszeitraumes dargestellt, hier die erste Hälfte des Jahres 2012.

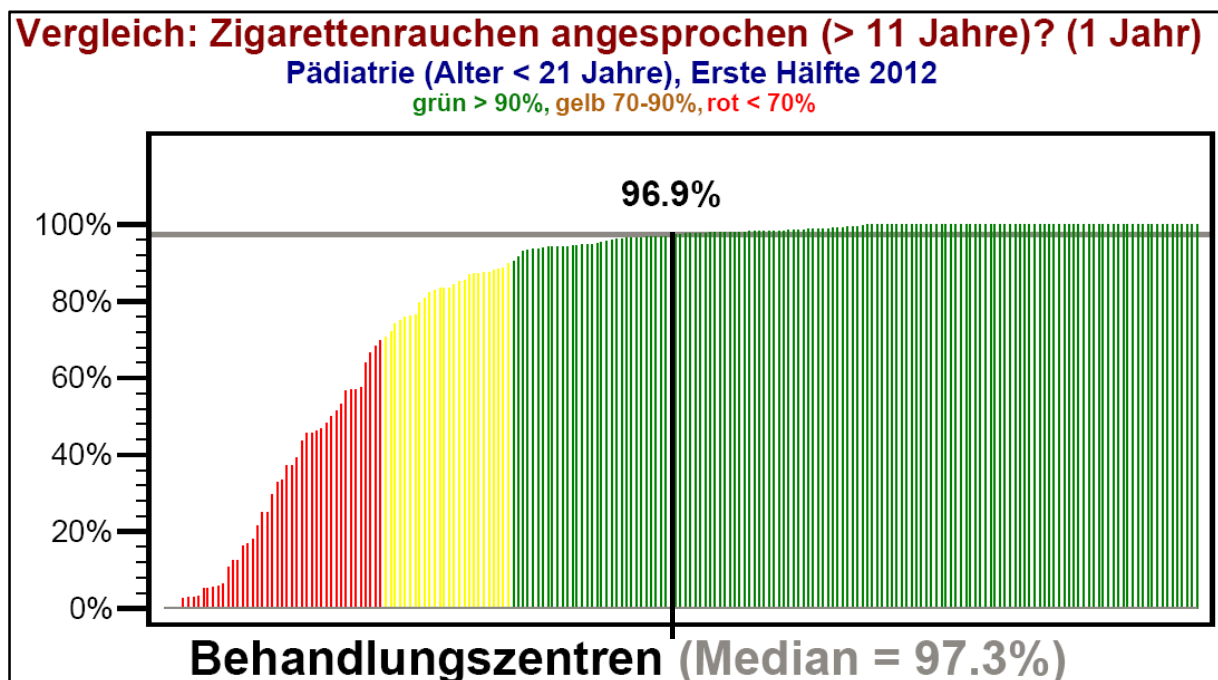


Abbildung 4: Querschnittsgrafik mit Ampelfarben

Der Makroaufruf für eine Querschnittsgrafik enthält die gleichen Parameter wie für eine Trendgrafik. Zusätzlich muss aber noch eine Variable für die Farbgestaltung (RYG, GYR oder BLUE) übergeben werden.

```
%benchmarking ( _indata = dpva_1_akt,
                _var      = rauchen,
                _unit     = PERCENT,
                _colour   = GYR);
```

Die Variable “\_var“ enthält für jedes Zentrum den Anteil Patienten, für den das Rauchverhalten nachgefragt wurde. Liegt die Vollständigkeit über 90% wird der Wert des Zentrums in die Variable “\_var1“ gespeichert, Werte zwischen 70 und 90% werden in “\_var2“ geschrieben und Werte <70% in “\_var3“. Zusätzlich wird der Wert des eigenen Zentrums in “sign“ geschrieben, um ihn oberhalb des Balkens anzuzeigen (Pointlabel).

```
%MACRO benchmarking ( _indata, _var, _unit, _colour);
```

```
DATA _benchdat;
  SET &_indata.;
```

```
* Ampelbereiche zuordnen für die Vollständigkeit: >90%,70-90%,<70%;
```

```
IF &_var. > 0.9 THEN _var1 = &_var.;
IF &_var. >= 0.7 and &_var. <= 0.9 THEN _var2 = &_var.;
IF &_var. < 0.7 THEN _var3 = &_var.;
IF zentrum = &eigenes_Zentrum. THEN _vari = &_var.;
```

```
* Wert des eigenen Zentrums;
```

```
IF _vari NE . THEN sign =
  LEFT(PUT(ROUND(_vari*100,.1),best.)||'%');
```

```
RUN;
```

Die Zentren werden auf der x-Achse abgetragen, die Vollständigkeit der Raumnachfrage auf der y-Achse. Somit würde man vier einzelne Grafiken erhalten, Zentren mit einer Vollständigkeit über 90%, Zentren mit einer Vollständigkeit zwischen 70 und 90%, Zentren mit einer Vollständigkeit unter 70% und das eigene Zentrum. Um alles in einer Grafik darzustellen, muss hier ebenfalls die Option OVERLAY angegeben werden. Die Optionen VREF, CVREF und WVREF im PLOT-Statement zeichnen die graue Referenzlinie, die den Median der Zentren darstellt.

Die SYMBOL-Statements definieren die unterschiedlichen Farben für die Balken. Die Reihenfolge der Farben wird über den Parameter “\_colour“ im Makroaufruf festgelegt. PROC GPLOT stellt eigentlich Punkte in einer Grafik dar, mit der Option I=NEEDLE werden vertikale Linien zur x-Achse gezeichnet.

```
*Symbol-Statements;  
SYMBOL1 i=needle c=green v=none l=1 w=1 pointlabel=none;  
SYMBOL2 i=needle c=yellow v=none l=1 w=1 pointlabel=none;  
SYMBOL3 i=needle c=red v=none l=1 w=1 pointlabel=none;  
SYMBOL4 i=needle c=black v=none l=1 w=1 pointlabel=(c=black  
h=0.6 '#sign' f="Arial/bold");  
PROC GGPLOT DATA=_benchdat;  
  PLOT (_var1 _var2 _var3 _vari) * zentren / overlay  
    vaxis = axis2  
    haxis = axis1  
    vref = &median.  
    cvref = grey  
    wvref = 2;  
RUN;  
QUIT;  
  
%MEND benchmarking;
```

### 2.3.3 Weitere Grafiken

Im Benchmarking werden noch weitere Grafik-Arten ausgegeben:

**Bivariate Grafiken:** Beispiel: Anzahl Hypos (y-Achse) \* HbA1c-Median (x-Achse)  
Jedes Quadrat repräsentiert eine Einrichtung. Das eigene Zentrum wird durch einen schwarzen Punkt dargestellt (siehe Abbildung 5).

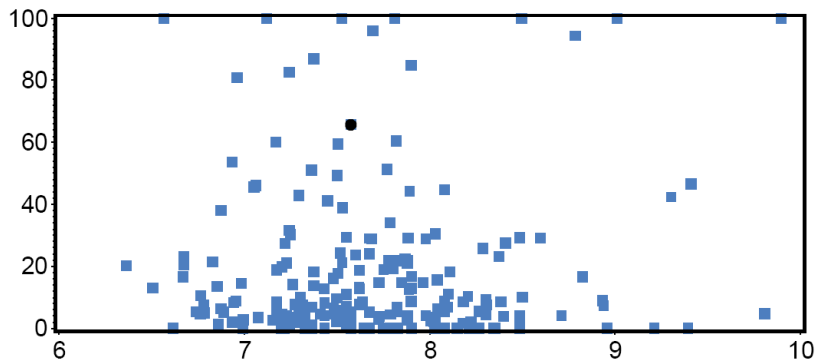


Abbildung 5: Bivariate Grafik



**Grafik:** Beispiel: HbA1c-Median \* Therapieart (CT, ICT, Pumpe)

Die schwarzen Balken (jeweils links) repräsentieren das eigene Zentrum, die blauen Balken (rechts) den Wert aller Zentren (siehe Abbildung 6).

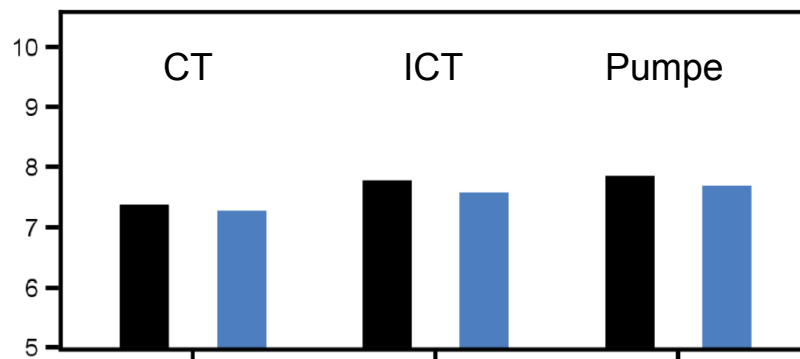


Abbildung 6: Grafik

**Kumulierte Grafik:** Beispiel: Alter in drei Kategorien

In dieser Grafik werden nur die Daten des eigenen Zentrums dargestellt. Die Balken zeigen die Verteilung der Patienten auf die Kategorien (siehe Abbildung 7).

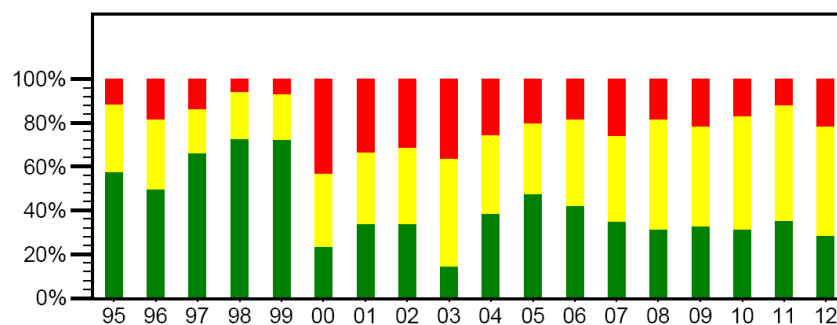


Abbildung 7: Kumulierte Grafik

## 2.4 Wie wird das Benchmarking verschickt?

Die Kontaktdaten der Ansprechpartner der Zentren sind in einer Adressdatenbank gespeichert. Rückmeldungen bezüglich Adress- oder E-Mail-Änderungen werden kontinuierlich eingearbeitet.

Diese Adressdatenbank wird im SAS-Hauptprogramm eingelesen. Die Kontaktdaten des gerade verarbeiteten Zentrums werden herausgesucht und in Makrovariablen gespeichert. Außerdem wird eine geschlechtsspezifische Anrede erstellt.

Der folgende Programmcode erstellt eine E-Mail mit Empfänger, Betreffzeile, Anhang und E-Mail-Text:

```
DATA emaildpv;
  SET emaildpv;
  CALL SYMPUT ('email', email);
  CALL SYMPUT ('ename', TRIM (name));
```

```
CALL SYMPUT ('eanrede', TRIM (anrede));
CALL SYMPUT ('eklinik', TRIM (name_b));
IF geschl = "Frau" THEN CALL SYMPUT ('eAnsprache', "Liebe Frau");
IF geschl = "Herr" THEN CALL SYMPUT ('eAnsprache', "Lieber Herr");
IF geschl NOT IN ("Herr", "Frau") THEN CALL SYMPUT ('eAnsprache',
"Liebe(r)");
RUN;
FILENAME message EMAIL
TO = ("%eemail")
SUBJECT = "DPV-Benchmarking erstes Halbjahr 2012 &eklinik."
ATTACH = ("D:\temp\qsdpv\DPV-Benchmarking-&benchart.-&jahr.-
&eins_sign.-&pdfoutput.pdf")
CC = ("katharina.fink@uni-ulm.de");
DATA _null_;
FILE message;
PUT
"Liebe Teilnehmer am Benchmarking, &eAnsprache &eanrede &ename!"//
"Heute schicken wir Ihnen die DPV-Auswertungen für das erste Halb-"/
"jahr 2012 zu. In der angehängten pdf-Datei finden Sie die"/
"jeweiligen Ergebnisse Ihrer Einrichtung in Bezug zu allen anderen"/
"Diabetes-Einrichtungen dargestellt. Die Auswertung erfolgt wie"/
"immer separat für pädiatrische und für internistische Zentren"//
"Für heute herzliche Grüße aus Ulm"//
"R. Holl - A. Hungele - K. Fink - E. Molz - R. Ranz - M. Grabert"//
"-----"/
"Institut für Epidemiologie und medizinische Biometrie, Uni Ulm"//
"Tel: 0731-502-5314"/
"FAX: 0731-502-5309"/
RUN;
```

Hat ein Zentrum mehrere Ansprechpartner, wird das Benchmarking mit Hilfe einer Schleife mehrfach verschickt.

Anhand von Optionen kann festgelegt werden, welche Absenderadresse verwendet werden soll.

## 2.5 Welche Varianten können erstellt werden?

Das Benchmarking kann in verschiedenen Varianten erstellt werden (jeweils getrennt nach pädiatrischen und internistischen Zentren):

- Zentrumsidentität offen: entscheidet sich eine Gruppe von Zentren für eine offene Auswertung, erscheint unter jedem Balken das jeweilige Kürzel der Zentren. Hierfür muss jedes Zentrum dieser Gruppe der Entanonymisierung schriftlich zugestimmt haben. Diese Variante wird zum Beispiel von regionalen Qualitätszirkeln genutzt.
- Darstellung aller Zentren aus Deutschland.
- Darstellung aller Zentren aus Österreich.
- Darstellung für separate Anwendergruppen (zum Beispiel pädiatrische Pumpen-AG).

Die Erstellung dieser verschiedenen Varianten wird über Makrovariablen, die am Anfang des Hauptprogrammes gesetzt werden, gesteuert. Im Folgenden sind ein paar Beispielmakrovariablen aufgeführt.

- Mit Hilfe der Makrovariablen “analyse“ kann festgelegt werden, ob das pädiatrische oder das internistische Benchmarking erstellt werden soll.  
`%LET analyse = paed;`
- Falls nur Zentren aus Österreich aufgeführt werden sollen, wird die Makrovariable “oest“ verwendet.  
`%LET oest = JA;`
- Für regionale Auswertungen setzt man die Makrovariable “regioqz“ auf “ja“. Zusätzlich muss der gewünschte Qualitätszirkel ausgewählt werden.  
`%LET regioqz = JA;`  
`%LET regional = 'Y111', 'Y234', 'X221', 'K223', 'L786';`

### 3 Schwierigkeiten und Weiterentwicklungen

Eine Schwierigkeit des Benchmarking-Programmes besteht darin, möglichst viele verschiedene Varianten innerhalb eines Programmes zu ermöglichen. Das kann zu Unübersichtlichkeit führen, allerdings besteht der Vorteil darin, dass nur ein Programm gepflegt werden muss.

Da sich mit jeder neuen Benchmarking-Runde die Datengrundlage verändert, enthalten die Grafikmakros eine automatische y-Achsenberechnung. In manchen Fällen kann es trotzdem nötig sein, die Achsen manuell einzuteilen.

Seit einer der letzten Versionen gibt es ein Inhaltsverzeichnis mit Linkfunktion. Über dies sind die Anwender sehr erfreut, da der Umfang des Benchmarking-Reports immer größer wird. Leider funktioniert der “Rücksprung“ zum Inhaltsverzeichnis noch nicht. Eine weitere Schwierigkeit ist, die Grafiken so verständlich wie möglich darzustellen, damit jeder Anwender einen Nutzen aus dem Benchmarking ziehen kann.

Wechselnde SAS-Versionen machen oftmals ebenfalls Probleme, wie zum Beispiel der Umstieg von SAS 9.2 auf 9.3. Hier musste an den Grafiken einiges angepasst werden.

Es ist wichtig, dass das Programm so automatisiert wie möglich abläuft. Dies soll in Zukunft noch weiter ausgebaut werden. Zum Beispiel soll das Programm automatisch für alle teilnehmenden Zentren durchlaufen werden, damit nicht mehr jedes Zentrum einzeln eingetragen werden muss. Denn bisher verschicken wir pro Jahr ca. 1000 Benchmarking-Reporte.

*K. Fink, E. Molz, A. Hungele, R. Ranz, M. Grabert, R. Holl*

Eine weitere Herausforderung ist es, in jeder Benchmarking-Runde neue Parameter oder Grafiktypen einzubauen, damit das Benchmarking weiterhin interessant gestaltet ist.

Im Institut für Epidemiologie und medizinische Biometrie, ZIBMT werden neben DPV noch weitere Software-Programme entwickelt, zum Beispiel für Adipositas. Für diese Dokumentationsprogramme wurde die Struktur des DPV-Benchmarkings übernommen.

Ein Beispielbenchmarking, eine Demoversion der Software und weitere Informationen finden sich auf unserer Homepage [www.d-p-v.eu](http://www.d-p-v.eu).