IFT · Institut für
Therapieforschung
München

Ludwig Kraus
Martin Steppan
Daniela Piontek

Schätzung der Prävalenz substanzbezogener Störungen in Berlin: Opioide, Kokain und Stimulanzien

17. April 2015

Titel und Verantwortliche

Projekttitel: Schätzung der Prävalenz substanzbezogener Störungen in Berlin: Opioide, Kokain und Stimulanzien

Förderung: Senatsverwaltung für Gesundheit und Soziales Berlin

Leitung: Prof. Dr. Ludwig Kraus

Mitarbeiter: Dr. Martin Steppan
Dr. Daniela Piontek

Kontaktdaten: IFT Institut für Therapieforschung
Herr Prof. Dr. Ludwig Kraus
Parzivalstraße 25
80804 München
Tel.: 089-360804-30
Fax: 089-360804-19
kraus@ift.de

Inhaltsverzeichnis

Inhaltsverzeichnis 5
Abbildungsverzeichnis 7
Tabellenverzeichnis 7
Zusammenfassung der Ergebnisse 9
1 Einleitung 11
2 Methoden 13
2.1 Capture-recapture Methode 13
2.2 Datenquellen 14
2.3 Datenschutz und Kryptographierung 15
2.3.1 Rechtlicher Hintergrund 15
2.3.2 Technische Umsetzung und Kryptographierung 15
3 Statistische Auswertung 19
3.1 Fehleranfälligkeit des HIV-Codes 19
3.2 Datenbasis und Datenbereinigung 19
3.3 Schätzung der Populationsgrößen 19
3.3.1 Log-lineare Regression 20
3.3.2 Probabilistische Methode 21
3.3.3 Schätzung intravenöser Konsumenten 21
3.4 Plausibilitätsprüfung der Schätzung und Übereinstimmung der Schätzverfahren 22
3.5 Vergleich mit anderen Großstädten 22
3.6 Verlinkung verschiedener Datenquellen 22
4 Ergebnisse 23
4.1 Deskriptive Statistiken und Stichprobengrößen 23
4.2 Fehleranfälligkeit des HIV-Codes 24
4.3 Überschneidung zwischen verschiedenen Substanzen 24
4.4 Schätzung der Populationsgrößen 25
4.4.1 Log-lineare Regression 25
4.4.3 Probabilistische Methode 28
4.4.4 Schätzung intravenöser Konsumenten 30
4.4.5 Schätzung polyvalenter Störungen 31

4.5 Vergleich mit anderen Großstädten 32
4.6 Plausibilitätsprüfung der Schätzung und Übereinstimmung der Schätzverfahren 35
4.6.1 Übereinstimmung beider Methoden 35
4.6.2 Summationsüberprüfung 35
4.7 Verlinkung verschiedener Datenquellen 36
5 Diskussion 39
6 Danksagung 43
7 Literatur 45
Anhang I 47
8 Schätzverfahren 49
9 Tabellen und Abbildungen der Schätzungen für das Jahr 2011 55
Anhang II 61

Abbildungsverzeichnis

Abbildung 1. Altersverteilung der drei Substanzen und des intravenösen Konsums (Datenjahr 2010) 24

Abbildung 2. Substanzkonsum der registrierten Drogenkonsumenten (Datenjahr 2010) 25

Abbildung 3. Zusammenfassung der beiden Schätzmethoden nach Substanzgruppen 30

Abbildung 4. Schätzung des Anteils mono- und polyvalenter Störungen 31

Abbildung 5. Europäischer Trend des Rückgangs der Prävalenz des Opioidkonsums auf Basis der Datenbank der EMCDDA 32

Abbildung 6. Graphischer Vergleich der Berliner Schätzwerte mit den Schätzungen anderer Europäischer Großstädte 34

Abbildung 7. Übereinstimmung der Schätzungen nach Substanz, Altersgruppe und Geschlecht aus den Tabellen 2 und 3 35

Abbildung 8. Abweichung der Summe der Einzelschätzungen von den Gesamtschätzungen 36

Tabellenverzeichnis

Tabelle 1. Stichprobengrößen nach Datenquelle, Datenjahr, Geschlecht und Substanz 23

Tabelle 2. Stichprobengrößen nach Altersgruppe, Substanz, Geschlecht und i.v. Konsum (Datenjahr 2010) 23

Tabelle 3. Populationsschätzung anhand log-linearer Regression nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010) 26

Tabelle 4. Populationsschätzung anhand der probabilistischen Methode nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010) 28

Tabelle 5. Vergleich der Berliner Schätzwerte mit den Schätzungen anderer Europäischen Großstädte (Opioide) 33

Tabelle 6. Überzufälligkeit des Overlaps verschiedener Datenquellen 37

Zusammenfassung der Ergebnisse

Ziel und Hintergrund. Die Abschätzung der Populationsgröße von Drogenabhängigen stellt eine wichtige Voraussetzung für gesellschaftliche, gesundheitliche und politische Erwägungen dar. Die vorliegende Untersuchung verfolgte das Ziel, möglichst akkurate und aktuelle Schätzungen der Größe der Population von Abhängigen harter Drogen (Opioide, Kokain, Stimulanzien) in Berlin vorzunehmen. *Methode.* Für die Schätzungen wurden fünf Datenquellen aus den Jahren 2010 und 2011 verwendet: (1) Krankenhausstatistik, (2) Kriminalitätsstatistik, (3) Suchthilfe ambulant, (4) Suchthilfe stationär und (5) Drogentodesstatistik. Zur Identifizierung von Fällen über die verschiedenen Datenquellen hinweg wurden die personenbezogenen Daten von den Datenhaltern verschlüsselt (kryptographiert). Die technische Durchführung der Kryptographierung erfolgte über eine eigens dafür entwickelte Webseite. Unter Einbeziehung verschiedener Interaktionen zwischen den Datenquellen wurden mehrere log-lineare Modelle an die Daten angepasst (Log-lineares Modell). Zusätzlich wurden Schätzungen mit dem für *n* Datenquellen verallgemeinerten einfachen Capture-recapture Modell (zwei Datenquellen) vorgenommen (Probabilistisches Modell). Die Definition der Konsumenten harter Drogen orientierte sich weitgehend an der Definition psychischer und Verhaltensstörungen durch psychotrope Substanzen gemäß der internationalen Klassifikation von Krankheiten (ICD-10). Mit Ausnahme der vom Landeskriminalamt Berlin im Zusammenhang mit Drogenkonsum registrierten Personen und der Drogentodesstatistik wurden Diagnosen verwendet. *Ergebnisse.* Schätzungen substanzbezogener Störungen bei 15- bis 64-jährigen Berlinern für das Jahr 2010 (jeweils Mittelwerte): opioidbezogene Störung (11.300-16.700 oder 5-7 von Tausend), kokainbezogene Störungen (5.300-9.500 oder 2-4 von Tausend) und stimulanzienbezogene Störungen (6.500-9.400 oder 3-4 von Tausend). Störungen im Zusammenhang mit Opioiden, Kokain oder Stimulanzien wurden auf 9-12 pro Tausend geschätzt (21.700-27.300). Bei nahezu einem Drittel der Drogenkonsumenten (ca. 8.000) lag eine Störung im Zusammenhang mit mindestens zwei Substanzen vor. Schätzungen der Anzahl von Personen mit Störungen im Zusammenhang mit i.v. Opioidkonsum auf der Grundlage der ambulanten Suchthilfestatistik ergaben ein N von 8.400-10.100 (Log-lineares Modell) bzw. von 6.000-7.000 (Probabilistisches Modell) oder 2.5-4.25 pro Tausend. *Diskussion*. Gegenüber der Schätzung für 1992 von 6.000-8.000 Personen mit einer i.v. Heroinabhängigkeit (2.5-3.3 pro Tausend) ist von einer leichten Zunahme der Prävalenz von Personen mit einer opioidbezogenen Störung, die i.v. konsumieren, auszugehen. Vergleiche mit den Prävalenzschätzungen von Störungen im Zusammenhang mit Opioidkonsum aus anderen Großstädten Europas zeigen, dass sich die Schätzungen für Berlin bezüglich des Umfangs des Opioidproblems im mittleren bis unteren Bereich bewegen. Obwohl keine Vergleichszahlen vorliegen, weisen die aktuellen Schätzungen zu kokain- und stimulanzienbezogenen Störungen auf eine Diversifizierung des Drogenmarktes in Berlin vergleichbar mit anderen europäischen Großstädten hin. Damit ist trotz einer hohen Überschneidung mit opioidbezogenen Störungen (polyvalenter Konsum) von einer Zunahme der Prävalenz drogenbezogener Störungen in Berlin auszugehen. Hinweise eines Rückgangs der Inzidenz der Opioidabhängigkeit finden sich in verschiedenen europäischen Städten. Während die Inzidenz zurückgeht, nimmt die Prävalenz aufgrund des flächendeckenden Ausbaus der Substitution mit der Folge einer Überlebenssicherung jedoch nur langfristig ab.

1 Einleitung

Die Abschätzung der Populationsgröße von Abhängigen sogenannter „harter Drogen“ stellt eine wichtige Voraussetzung für gesellschaftliche, gesundheitliche und politische Erwägungen dar (Babor et al., 2010). Basierend auf der Annahme, dass mit dem Konsum dieser Substanzen eine starke körperliche und seelische Beeinträchtigung der betroffenen Personen verbunden ist (Degenhardt & Hall, 2012), entstehen durch eine hohe Anzahl von Personen mit einem solchen Problem besondere Notwendigkeiten in der medizinischen und psychologischen Versorgung. Die Abschätzung der Größe dieser Personengruppe für ein geographisches Gebiet, im Besonderen für den urbanen Raum, stellt eine wesentliche Voraussetzung dar, um die Problematik zu quantifizieren und entsprechende Entscheidungen treffen zu können.

Mit „harten Drogen“ sind in diesem Zusammenhang vor allem die Substanzgruppen Opioide (Derivate des Schlafmohns), Kokain (Pulver, Crack-Kokain) und Stimulanzien (z.B. Ecstacy, Amphetamin, Ritalin, MDMA, „Chrystal Meth“ und ähnliche) gemeint (Laux & Möller, 2011). Der Begriff harte Drogen ist in diesem Fall insofern irreführend, als man annehmen könnte, dass andere psychotrope Substanzen wie Alkohol, Cannabis oder Nikotin vergleichsweise „weich“, d.h. unproblematischer sind. Dies ist nicht der Fall, da auch bei diesen Substanzen starke Abhängigkeitserscheinungen entstehen und gravierende psychosoziale und medizinische Beeinträchtigungen die Folge sein können (Nutt, King & Phillips, 2010). Dennoch gibt es gewisse Merkmale „harter“ Drogen, die diese von Seiten der Öffentlichkeit besonders relevant machen. Erstens weisen insbesondere Opioide und Kokain sehr starke körperliche und pharmakologisch ableitbare Abhängigkeitsrisiken auf (Schäfer, 2011). Zweitens geht von allen drei Substanzen ein nicht unerhebliches Risiko aus, direkt im Zusammenhang mit dem Konsum oder den kurzfristigen Folgen zu versterben (Arendt, Munk-Jorgensen, Sher & Jensen, 2011). Drittens ist diesen drei Substanzgruppen gemeinsam, dass der Konsum auch intravenös erfolgen kann, wodurch ein erhöhtes Risiko für Infektionskrankheiten - insbesondere HIV, Hepatitis und ähnliche - gegeben ist, sofern Spritzen von Abhängigen gemeinsam verwendet werden (Strathdee et al., 2010). Schließlich scheint vor allem der Konsum von Opioiden zumindest in der öffentlichen Wahrnehmung mit einem erhöhten Risiko der „Verelendung“, d.h. des Abdriftens in prekäre Lebensumstände assoziiert zu sein (Bernard, 2013). Ebenso wird häufig davon ausgegangen, dass mit dem Konsum „harter Drogen“ auch eine erhöhte Wahrscheinlichkeit für kriminelle Handlungen im Sinne von „Beschaffungskriminalität“ verbunden ist (Groenemeyer, 2012). Es gibt somit einige Gründe, der Population Abhängiger von Opioiden, Kokain und Stimulanzien besonderes Augenmerk zu widmen. Die vorliegende Untersuchung beschäftigt sich daher mit diesen drei Substanzgruppen.

Aufgrund der dargestellten Problematik und Bedeutung „harter Drogen“ werden Daten zu dieser Konsumentengruppe in den meisten Ländern der Welt von Gesundheits- und Strafverfolgungsbehörden sowie von medizinischen und sozialwissenschaftlichen Instituten gesammelt und erforscht. Die Europäische Union (EU) hat für diesen Problembereich eine eigene Agentur eingerichtet (European Monitoring Centre for Drugs and Drug Addiction, EMCDDA), die Daten in den Mitgliedsländern sammelt und überwacht. Dabei wurde eine eigene Referenzkategorie gebildet, die harte Drogen und intravenösen Konsum umfasst und als „prob-

lem drug use“ (PDU) bzw. seit kurzem „high risk drug use“ (HRU) bezeichnet wird. Die Prävalenzschätzungen in diesem Bericht erfolgen neben den einzelnen Substanzen gemäß der Definition der EMCDDA für PDU/HRU, um hier eine Vergleichsmöglichkeit mit anderen Ländern und Städten in Europa gewährleisten zu können (vgl. Methoden). International wird der Konsum harter Drogen auch von der Weltgesundheitsorganisation (WHO) im Rahmen des World Drug Report (United Nations Office on Drugs and Crime, World Drug Report 2014) beobachtet. In Deutschland werden Daten zu diesen Konsumentengruppen vom Bundesministerium für Gesundheit (BMG), von den Landeskriminalämtern, von Kranken- und Rentenversicherungen, zunehmend aber auch von Ländern und Gemeinden und von universitärer Seite verarbeitet. Für das offizielle Berichtswesen zuständig ist die Deutsche Beobachtungsstelle für Drogen und Drogensucht (DBDD), die die Daten aus Deutschland an die EMCDDA und das BMG berichtet.

Aufschlussreich sind Beobachtungen von Drogenszenen im urbanen Raum (Korf, 2004; Tretter & Kraus, 2004). Ballungsräume sind häufig auch Zentren des Drogenkonsums. Höhere Anonymität, bessere Versorgungslage mit Drogen, geringerer Druck seitens der Exekutive, geringere Stigmatisierung, entsprechende Infrastruktur (z.B. Drogenkonsumräume) und Hilfsangebote könnten Gründe sein, warum vor allem Großstädte mit einer erhöhten Drogenproblematik rechnen müssen. Aus diesem Grund gibt es eine Reihe von Untersuchungen zur Größe dieser Szenen in verschiedenen europäischen Großstädten. Besonders hohe Zahlen liegen dabei für Städte in Großbritannien vor, z.B. Manchester, Birmingham und Liverpool, wo die Populationsgröße jeweils über 19 pro tausend Einwohner liegt. Ebenfalls hohe Anteile weisen Dublin, Kopenhagen und Sofia (13 pro T.) auf. Die höchsten Zahlen in Mitteleuropa werden für Wien (11 pro T.) und Prag (9 pro T.) berichtet. Vergleichsweise geringe Zahlen liegen für Budapest und Warschau vor (jeweils 2 pro T.). Für Berlin liegen Schätzungen zum Umfang der Drogenabhängigkeit aus den Jahren 1979 (Skarabis & Patzak, 1981) und 1992 (Kirschner et al., 1994) vor. Die Schätzungen aus dem Jahr 1979 beziehen sich auf Westberlin und die Zielgruppe der Heroinabhängigen. Mithilfe verschiedener Ansätze der Capture-recapture Methode wurde die Prävalenz Heroinabhängiger in Westberlin auf 6.000 (95% KI 5.820-6.280) geschätzt. Zielgruppe der Schätzungen für Gesamtberlin aus dem Jahr 1992 waren i.v. Heroinabhängige. Neben der Capture-recapture Methode unter Verwendung vergleichbarer Datenquellen wie im Jahr 1979 führten die Autoren ein anonymes Monitoring in Praxen niedergelassener Ärzte (ANOMO) durch. Auf der Basis dieser beiden Methoden wurde die Prävalenz i.v. Heroinabhängiger in Gesamtberlin auf 6.000-8.000 Personen geschätzt. Aufgrund der Unterschiede hinsichtlich Zielpopulation und Einzugsbereich (Berlin West und Berlin Gesamt) ist die Vergleichbarkeit der Schätzungen erheblich eingeschränkt.

Die vorliegende Untersuchung verfolgt das Ziel, möglichst akkurate und aktuelle Schätzungen der Größe der Population von Abhängigen harter Drogen (Opioide, Kokain, Stimulanzien) in Berlin vorzunehmen. Die Schätzungen erfolgen analog der Richtlinien der EMCDDA, um Vergleiche mit anderen Metropolen vornehmen zu können.

2 Methoden

2.1 Capture-recapture Methode

„Möchte jemand die Anzahl von Fischen in einem See abschätzen, so empfiehlt es sich, hundert Fische zu fangen (capture), diese zu markieren (z.B. rot anzumalen) und wieder freizulassen. Anderntags möge man wieder z.B. zweihundert Fische fangen (recapture). Findet man zwei rote Fische wieder, so ergibt sich die Schätzung, zwanzigtausend (100 mal 200) geteilt durch die Anzahl der wiedergefangenen (zwei), d.h. tausend Fische.“ So könnte eine vereinfachende Anleitung zur Capture-recapture Methode lauten.

Anhand dieser simplifizierenden Beschreibung soll die Grundidee des hier angewandten Ansatzes beschrieben werden. Die Methodik stammt ursprünglich aus der Biologie, wo die Abschätzung einer Population für eine Reihe von zoologischen Fragestellungen relevant ist. Die erste Anwendung dieses Prinzips auf die Schätzung von Drogenabhängigen stammt aus den Siebzigerjahren des 20. Jahrhunderts (Fienberg, 1972; Wittes, 1974). Der Vorteil der Anwendung dieser Methode beim Menschen besteht darin, dass eine Markierung meistens nicht notwendig ist, da Menschen bereits durch Namen relativ eindeutig „markiert“ sind. Fügt man weitere Informationen hinzu (z.B. Geburtsdatum, Geschlecht, Wohnort), so ist die Markierung in den allermeisten Fällen fast perfekt. Lässt man Probleme des Datenschutzes außer Acht (vgl. Datenschutz in Anhang II), so lässt sich anhand zweier Datenbanken (z.B. Krankenhaus und Polizei) bereits eine Capture-recapture Untersuchung durchführen, in dem man einfach einen Namensabgleich und dann die entsprechende Berechnung wie eingangs dargestellt durchführt.

Dasselbe Prinzip kann angewandt werden, wenn mehrere Rückfänge (recaptures) gezogen werden. Zieht man beispielsweise nicht nur einen Rückfang wie im obigen Beispiel, sondern fünf, so ergeben sich fünf Schätzungen, die zusammen eine adäquatere Schätzung der Gesamtzahl liefern, da sich eventuelle zufallsbedingte Ausreißer von selbst bereinigen. Wenn bei jedem Rückfang zusätzlich noch eine weitere Markierung erfolgt, was, um beim Beispiel Mensch zu bleiben, nicht notwendig ist, sofern ein Identifikator (z.B. Name oder Code) vorliegt, so ergeben sich bei sechs Namenslisten, „sechs über zwei“ Kombinationen[1], d.h. 15 Zweierkombinationen, die jeweils eine Schätzung ermöglichen. Zusätzlich gibt es 20 Dreierkombinationen („sechs über drei“), 15 Viererkombinationen („sechs über vier“), 6 Fünferkombinationen („sechs über fünf“) und eine Sechserkombination („sechs über sechs“). Insgesamt lassen sich 57 Schätzungen aus dem Abgleich aller Namenslisten ableiten. Je höher die Anzahl der Rückfänge, d.h. die Anzahl der verfügbaren Personenlisten, desto adäquater die Schätzung der Population.

---

[1] $\binom{n}{k}$ [$sprich$ n über k"]= $\frac{\text{n!}}{\text{k!}\cdot(\text{n-k})!}$ „sechs über zwei“ = $\frac{6!}{2!\cdot 4!} = 15$

2.2 Datenquellen

Für die vorliegende Untersuchung wurden fünf Datenquellen genutzt (Krankenhausstatistik, Kriminalitätsstatistik, Suchthilfe ambulant, Suchthilfe stationär, Drogentodesstatistik). Es können somit zehn Zweierkombinationen, zehn Dreierkombinationen, fünf Viererkombinationen und eine Kombination aus allen Datenquellen gebildet werden. Es gibt somit 26 mögliche Abgleiche der Personenlisten, die zu einer Schätzung der Gesamtpopulation führen.

Für die Schätzung wurden folgende Datenquellen in Berlin für die Jahre 2010 und 2011 herangezogen: (1) Krankenhausstatistiken zur Drogenentgiftung; (2) Statistiken des Landeskriminalamts zu Verstößen gegen das Betäubungsmittelgesetz (BtMG); Daten der Deutschen Suchthilfestatistik (DSHS) (3) zu Einrichtungen der ambulanten Suchthilfe und (4) zu stationären Rehabilitationseinrichtungen; und (5) Drogentodesstatistiken.

Krankenhausstatistik

Erfasst wurden alle Neuzugänge der Entgiftungsstationen der Vivantes Kliniken. Auswahlkriterien waren die Diagnosen Psychische und Verhaltensstörungen durch Opioide (F11), Kokain (F14) oder Stimulanzien (F15). Die Vivantes Kliniken decken etwa 60% der Entgiftungen in Berlin ab.

Verstöße gegen das Betäubungsmittelgesetz (BtMG)

Dies betraf alle Personen, die im Zusammenhang mit Verstößen gegen das Betäubungsmittelgesetz (BtMG) vom Landeskriminalamt registriert wurden Die Einträge umfassten für Opioide: Heroin, Morphin, Methadon, Subutex, Tilidin, Codein, Dehydrocodein oder Fentanyl; für Kokain: Kokain oder Crack; und für Stimulanzien: Amphetamin/Metamphetamin und deren Derivate in Pulver- oder flüssiger Form; oder Amphetamin/Metamphetamin und deren Derivate in Tabletten- bzw. Kapselform; oder Khat.

Ambulante Suchthilfeeinrichtungen

In den Daten der DSHS wurden alle Personen identifiziert, die in ambulanten Beratungsstellen aufgrund der Diagnosen schädlicher Gebrauch oder Abhängigkeit von Opioiden (F11.1 oder F11.2), Kokain (F14.1 oder F14.2) oder Stimulanzien (F15.1 oder F15.2) betreut wurden.

Stationäre Suchthilfeeinrichtungen

Die Daten wurden ebenfalls der DSHS entnommen. Sie betrafen alle Personen, die in stationären Rehabilitationseinrichtung mit den Diagnosen schädlicher Gebrauch oder Abhängigkeit von Opioiden (F11.1 oder F11.2), Kokain (F14.1 oder F14.2) oder Stimulanzien (F15.1 odeF15.2) behandelt wurden.

Drogentodesfälle

Es besteht eine Meldepflicht für Todesfälle, die in einem kausalen Zusammenhang mit dem missbräuchlichen Konsum von Betäubungs- oder Ausweichmitteln stehen. Darunter fallen (1) Todesfälle infolge beabsichtigter oder unbeabsichtigter Überdosierung (Organversagen aufgrund einer akuten Vergiftung), (2) Todesfälle infolge langzeitigen Missbrauchs (Langzeit-

schäden, drogentoxische Schädigungen durch konsumierte Substanzen, verminderte körperliche Abwehrkräfte, Schädigungen durch Streckungsmittel, Erkrankungen und Infektionen durch i. v. Applikation, Hepatitis oder AIDS), (3) Selbsttötung aus Verzweiflung über die Lebensumstände oder unter Einwirkung von Entzugserscheinungen außer durch Überdosierung und (4) tödliche Unfälle von unter Drogeneinfluss stehenden Personen (v. a. im Straßenverkehr). Um die Auswahlwahrscheinlichkeit nicht zu verzerren, wurden die Daten des jeweils folgenden Jahres erhoben. Damit wurde sichergestellt, dass die Person vor ihrem Tod über den gesamten Beobachtungszeitraum von jeweils einem Jahr hinweg in den anderen Datenquellen registriert werden konnte. Die Berücksichtigung von Drogentoten aus dem gleichen Beobachtungszeitraum wie die anderen Datenbestände führt, wie Uhl und Seidler (2000) zeigen konnten, zu einer systematischen Verzerrung der Ergebnisse. Berücksichtigt wurden alle Fälle, bei denen ein Bezug zum Konsum von Opioiden, Kokain oder Stimulanzien hergestellt werden konnte.

2.3 Datenschutz und Kryptographierung

2.3.1 Rechtlicher Hintergrund

Die Durchführung der Untersuchung erfolgte unter Bezugnahme auf eine datenschutzrechtliche Einschätzung des Berliner Beauftragten für Datenschutz und Informationssicherheit (Hauptausschuss 0645). In einem mehrstufigen Prozess wurden Vorschläge für eine sichere Datenschutzarchitektur erarbeitet und auf die Besonderheiten der vorliegenden Situation und der einzelnen Datenhalter abgestimmt. Das Datenschutzdesign wurde auch in Zusammenarbeit mit den jeweiligen Datenschutzbeauftragten der einzelnen Datenhalter erarbeitet und abgestimmt. Daten der Kassenärztlichen Vereinigung (KV) konnten für diese Studie entgegen der ursprünglichen Intention nicht verwendet werden, da in diesem Fall eine Weitergabe von personenbezogenen Daten (auch in pseudonymisierter Form) gemäß § 287 SGB V nicht möglich war (vgl. Anhang II, A, B und C).

2.3.2 Technische Umsetzung und Kryptographierung

Für die Abschätzung der Populationsgrößen substanzbezogener Störungsgruppen ist es notwendig, Fälle über verschiedene Datenquellen hinweg zu identifizieren, d.h. zu „matchen“. Um dies bewerkstelligen zu können, ist eine ein-eindeutige Zuordnung eines Codes zum äquivalenten Code in den weiteren Datensätzen erforderlich, d.h. ein „unique identifier“. Der Klarname einer Person ist als „unique identifier“ ungeeignet, zum einen aus Gründen des Datenschutzes, zum anderen sind manche Klarnamen nicht eindeutig genug (z.B. Stefan Maier). Es wurden aus diesem Grund personenbezogene Daten von den Datenhaltern benötigt, die in zwei Schritten verschlüsselt (kryptographiert) wurden, sodass eine Rückführung auf den ursprünglichen Klarnamen unmöglich war: Vorname, Nachname, Geschlecht, Geburtsjahr. Zusätzlich wurden Angaben zum Drogenkonsum (Opioide, Kokain, Stimulanzien) bzw. den Diagnosen (F11, F14, F15), der Art des Opioidkonsums (intravenös) und dem Jahr der Datenerhebung (2010 oder 2011) erhoben. Die Daten drogenbedingter Todesfälle wurden jeweils für das darauffolgende Jahr, d.h. 2011 und 2012 erhoben.

Die Verschlüsselung der ersten vier Variablen (Vorname, Nachname, Geschlecht, Geburtsjahr) erfolgte auf Basis des sogenannten HIV-Codes (einer Verschlüsselungsvorschrift, die aus dem dritten Buchstaben des Vornamens, der Länge des Vornamens, dem dritten Buchstaben des Nachnamens, der Länge des Nachnamens, dem Geschlecht [M oder W] und dem Geburtsjahr besteht). Außerdem wurden noch vier weitere HIV-Codes mit ähnlicher Vorschrift gebildet, um die Fehleranfälligkeit und Eindeutigkeit dieses Codes zu überprüfen. Im Anschluss wurden diese verschiedenen Codes auf Basis eines H-MAC Algorithmus[2] und eines 12-stelligen Schlüssels weiter kryptographiert. Zwei der vier für die Verschlüsselung benötigten Variablen (Vorname, Nachname) wurden nicht in den finalen Datensatz übernommen. Geschlecht und Geburtsjahr wurden für die Gruppenbildung nach Geschlecht und Altersgruppen benötigt. Der finale Datensatz bestand somit aus den Kryptogrammen, Geschlecht, Geburtsjahr und den Variablen Opioide (Opioide bzw. F11), Kokain (Kokain bzw. F14), Stimulanzien (Stimulanzien bzw. F15), Intravenoes (intravenöser Opioidkonsum) und Datenjahr (Datenerhebung 2010 oder 2011).

Die technische Durchführung der Kryptographierung erfolgte über eine eigens dafür entwickelte Webseite. Die Webseite erlaubte es, von den Rohdaten ausgehend vier Schritte durchzuführen: (1) Trimmen und Normalisieren der Daten, d.h. es wurden überflüssige Leer- und Sonderzeichen entfernt bzw. entsprechend der Codierungsregeln umgewandelt; (2) Bildung der HIV-Codes (wie oben beschrieben); (3) Umwandlung anhand eines H-MAC Algorithmus und eines 12-stelligen Schlüssels. Der Schlüssel wurde von Seiten der Berliner Senatsverwaltung für Gesundheit und Soziales gebildet, und an die einzelnen Datenhalter weitergegeben. Die Instruktion für die Datenhalter bestand weiterhin darin, den Schlüssel nach Durchführung der Kryptographierung sofort zu löschen. (4) Lokale Speicherung des kryptographierten Datensatzes. Der kryptographierte Datensatz wurde bei den einzelnen Datenhaltern lokal gespeichert, per E-Mail an das IFT Institut für Therapieforschung versandt und im Anschluss beim Datenhalter gelöscht (vgl. Anhang II D).

Bei der Umwandlung der Daten, die im Rahmen der Berliner Suchthilfestatistik in ambulanten und stationären Einrichtungen entstehen, war eine Sonderbearbeitung notwendig, da diese Daten an keiner Stelle außer den Einrichtungen selbst Klarnamen enthält, sondern lediglich einen HIV-Code (wie oben beschrieben). Die Bildung weiterer HIV-Codes war daher nicht möglich. Da dem IFT diese HIV-Codes außerdem für jährliche Auswertungen vorliegen, musste eine dritte Instanz („third trusted party“) gefunden werden, die diese HIV-Codes nach den beschriebenen Schritten (2-4) verschlüsselt, damit die auswertende Einrichtung (IFT) an keiner Stelle mit dem 12-stelligen Schlüssel in Kontakt kommt. Dieser Schritt wurde vom Sozialpädagogischen Institut Berlin (SPI) durchgeführt. Die kryptographierten HIV-Codes wurden im Anschluss an die Gesellschaft für Standarddokumentation und Datenauswertung (GSDA) weitergeleitet, die weitere Daten aus der Berliner Suchthilfestatistik (z.B. die Hauptdiagnose) zu den verschlüsselten Daten zuspielen konnte. Dieses „matching“ erfolgte über die Variablen Einrichtungscode, Fallnummer und Betreuungsnummer, welche der GSDA vorliegen (jedoch nicht die HIV-Codes). Der finale Datensatz einschließlich der zugespielten

[2] Keyed-Hash Message Authentication Code

Daten wurde schließlich an das IFT übermittelt (ohne die letztgenannten drei Variablen und ohne den eigentlichen HIV-Code, sondern dessen kryptographiertes Äquivalent). So konnte sichergestellt werden, dass seitens des IFT kein Rückschluss von den Kryptogrammen auf die ursprünglichen HIV-Codes gezogen werden kann.

3 Statistische Auswertung

3.1 Fehleranfälligkeit des HIV-Codes

Vor der Durchführung der Analysen wurde die Fehleranfälligkeit des verwendeten Personenidentifikators (kryptographierter HIV-Code) überprüft. Da im Rahmen der Untersuchung mehrere HIV-Codes gebildet wurden, konnte geprüft werden, in wie vielen Fällen in einem naturalistischen Datensatz ein HIV-Code identisch ist, obwohl es sich um unterschiedliche Personen handelt (falsch positive Übereinstimmung). Diese Information gibt Aufschluss darüber, zu welchem Anteil die Überschneidung zweier Datenquellen überschätzt wird. Die Fehleranfälligkeit des verwendeten Kryptogramms stellt eine wesentliche Information für die Bewertung der Güte der Schätzung dar. Zur Bestimmung der Fehleranfälligkeit des HIV-Codes wurde geprüft, wie stark sich die Überschneidung von Datenquellen verändert, wenn schrittweise ein weiterer HIV-Code hinzugefügt wird, sodass eine längere Zeichenkette als Identifikator zur Verfügung steht und der Informationsgehalt erhöht wird. Würde man die gleichen Überschneidungen beobachten, hätte der jeweilige HIV-Code eine Fehleranfälligkeit von Null. Ermittelt man eine Fehleranfälligkeit des HIV-Codes, die größer als Null ist, sollte sich durch Hinzufügen jedes weiteren HIV-Codes die Fehleranfälligkeit exponentiell reduzieren. Auf diese Weise kann ein Modell gebildet werden, das mit den beobachteten Daten verglichen werden kann. Eine Fehleranfälligkeit des HIV-Codes, die darin besteht, dass die zwei verschiedenen HIV-Codes ein und dieselbe Person repräsentieren (falsch negativ), die auf Fehlanwendung der Kodierungsvorschrift oder Namensänderung zurückzuführen sind, kann auf Basis dieser Daten nicht eingeschätzt werden. Für diesen Fehler fehlt ein objektives Außenkriterium, das die Identität zweier Codes belegen würde. Es muss aber auch von Fehlkodierungen ausgegangen werden, die die Anzahl von Überschneidungen künstlich reduzieren.

3.2 Datenbasis und Datenbereinigung

Die durchgeführten Analysen beziehen sich auf die Altersgruppe der 15- bis 64-Jährigen, was die Vergleichbarkeit mit internationalen Studien gemäß den Standards der EMCDDA für problem drug use (PDU) bzw. high risk drug use (HRDU) gewährleistet.

3.3 Schätzung der Populationsgrößen

Die Schätzung der Gesamtpopulation erfolgte auf zwei Arten: (A) log-lineare Regression, und (B) probabilistische Methode. Beide Methoden weisen zu erwartende gegenläufige Verzerrungen der Gesamtschätzung auf. Um diese Verzerrungen auszugleichen, wurden beide Methoden angewandt, sodass grobe Über- oder Unterschätzungen vermieden wurden. Die Schätzungen wurden stratifiziert nach Geschlecht und drei Altersgruppen (15-24, 25-34, 35-64 Jahre) durchgeführt, d.h. es gibt Schätzungen für die drei Altersgruppen und über alle Altersgruppen für Männer, für die drei Altersgruppen und über alle Altersgruppen für Frauen und für die drei Altersgruppen und über alle Altersgruppen für Männer und Frauen. Das ergibt insgesamt 12 Einzelmodelle pro Methode.

3.3.1 Log-lineare Regression

Im Vergleich zur linearen Regression wird in der log-linearen Regression davon ausgegangen, dass die Häufigkeit des Vorkommens einer Person in verschiedenen Listen poissonverteilt ist. Bei dieser Verteilung handelt es sich um eine „Sonderform“ der Gauß'schen Normalverteilung, die in der Natur insbesondere bei seltenen Ereignissen vorkommt, z.B. die Anzahl der Blitzschläge auf einem Quadratkilometer Land pro Jahr, die Anzahl der Personen, die ein spezifisches Schokoladengeschäft in Berlin pro Minute betreten etc. Die Poisson-Verteilung hat die Eigenschaft, dass sie sich mit zunehmender Häufigkeit des Ereignisses (Lambda) der Gauß'schen Normalverteilung (Glockenkurve) annähert. Im vorliegenden Beispiel wird davon ausgegangen, dass eine Person nicht in jeder Datenquelle erfasst wurde, d.h. dass das durchschnittliche Lambda geringer als fünf ist. In diesem Fall ist die Poisson-Verteilung „linkssteil“. Auf Basis der log-linearen Regression kann eine Poisson-Verteilung mit passendem Lambda gefunden werden. Auf Basis dieser Verteilung wird dann abgeschätzt, wie viele Personen in keiner der Datenquellen vorkommen.

Der Vorteil der log-linearen Regression besteht darin, dass die Auswertung mit einem Statistikprogramm (z.B. SPSS) durchgeführt werden kann, und dass die Methode erprobt ist. Die Methode hat außerdem den Vorteil, dass Interaktionen zwischen Datenquellen berücksichtigt werden können. Dies kann der Fall sein, wenn z.B. Personen aus dem ambulanten Bereich mit einer gewissen Regelhaftigkeit in eine andere Datenquelle gelangen, und dies mit einer höheren Wahrscheinlichkeit als zufallsbedingt zu erwarten wäre. Das könnte dann der Fall sein, wenn Personen, die mit der Polizei und in der Folge mit der Justiz in Kontakt kommen, auch häufiger Suchthilfeeinrichtungen aufsuchen, weil sie eine Auflage von Seiten des Gerichts („Therapie statt Strafe“) mitbringen. In einem solchen Fall würde eine Nichtberücksichtigung dieser Interaktion dazu führen, dass die Überschneidung zwischen den beiden Datenquellen höher als erwartet ausfällt, wodurch das Lambda „künstlich“ erhöht wird, was zu einer geringeren Schätzung der Gesamtpopulation führen würde. Ungünstig auf die Methode wirkt sich aus, wenn die Datenquellen stark unterschiedliche Stichprobengrößen aufweisen. Diese Verzerrung besteht auch in der vorliegenden Untersuchung, da die Datenquelle „ambulante Einrichtungen“ eine sehr hohe Fallzahl aufweist. Somit ist die Wahrscheinlichkeit, dass ein Fall lediglich in dieser Datenquelle vorkommt, hoch. Damit kommt es zu einer Annäherung des geschätzten Lambdas zum Wert 1. Liegt der „wahre“ Wert von Lambda höher als 1, führt dies zu einer künstlichen Reduktion, und damit zu einer zu hohen Gesamtschätzung der Populationsgröße. Liegt der „wahre“ Wert von Lambda unter 1, kommt es zu einer künstlichen Erhöhung von Lambda, d.h. zu einer geringeren Schätzung der Gesamtpopulation. Ein weiterer Nachteil dieser Methode besteht darin, dass eine zunehmend höhere Anzahl von Datenquellen das technische Vorgehen deutlich aufwändiger macht. Aus diesem Grund wurden für die log-lineare Regression nur vier Datenquellen verwendet und die verhältnismäßig kleine Datenquelle „Mortalität“ aus der Analyse ausgeschlossen (zum statistischen Verfahren der Capture-recapture Methode und zur Modellauswahl vgl. Anhang I, Kapitel 8.1).

3.3.2 Probabilistische Methode

Aufgrund der dargestellten Problematik, die die log-lineare Regression mit sich bringt, wurde in der vorliegenden Untersuchung ein weiterer methodischer Ansatz verfolgt. Auf Basis jeder einzelnen Überschneidung zwischen den Datenquellen, kann ein einfacher wahrscheinlichkeitstheoretischer Dreisatz angewandt werden (Steppan, 2013). Der Vorteil dieser Methode besteht darin, dass sie nachvollziehbare Resultate ermöglicht und dass sie robust ist gegenüber unterschiedlichen Stichprobengrößen der Datenquellen. Als Nachteil dieser Methode ist zu nennen, dass keine Interaktionen zwischen den Datenquellen berücksichtigt werden. Ist, wie am Beispiel „Therapie statt Strafe“ dargestellt, die Wahrscheinlichkeit einer Überschneidung zwischen den Datenquellen höher als zufallsbedingt zu erwarten wäre, würde dies zu einer Unterschätzung der Populationsgröße führen. Um die Überschätzungstendenz der log-linearen Regression zu kompensieren, wurde daher die probabilistische Methode angewandt. Ein weiterer Vorteil der Methode ist, dass eine beliebige Anzahl von Datenquellen für die Analyse herangezogen werden kann. In die Schätzungen mit der probabilistischen Methode flossen daher alle fünf Datenquellen ein (zum statistischen Verfahren der probabilistischen Methode vgl. Anhang I, Kapitel 8.2)

3.3.3 Schätzung intravenöser Konsumenten

Für die Population intravenöser Konsumenten lagen nur unzureichende Daten vor. Aus diesem Grund musste ein anderer Weg gefunden werden, um eine reliable Abschätzung der Anzahl von injizierenden Personen zu erhalten. Herangezogen wurden die Daten der Deutschen Suchthilfestatistik für Berlin aus dem ambulanten und dem stationären Bereich. In den Suchthilfeeinrichtungen wird auf Basis des Kerndatensatzes (KDS) dokumentiert, welcher Anteil der wegen der jeweiligen Substanz behandelten Personen auch intravenös konsumiert. Das Problem dieser Daten liegt darin, dass jene Personen, die Hilfe in Anspruch nehmen, vermutlich häufiger injizieren als jene Personen, die in keiner der Datenquellen auftauchen und sich somit im Dunkelfeld befinden, wodurch die Daten aus den Einrichtungen der Suchthilfe den Anteil intravenöser Konsumenten überschätzen. Es musste daher ein Weg gefunden werden, wie die Daten der Suchthilfe „kalibriert“, d.h. korrigiert werden können. Als Hilfskonstruktion wurden für diese Kalibrierung die Daten der Drogentodesfälle herangezogen. Es kann mit hoher Wahrscheinlichkeit davon ausgegangen werden, dass in dieser Quelle nur wenige Personen enthalten sind, die nicht intravenös konsumiert haben, da das Risiko für eine Überdosierung durch Inhalieren oder orale Einnahme deutlich geringer ist. Eine Untersuchung von Darke und Ross (2000) gibt einen Anteil von 1% an. Somit kann angenommen werden, dass nahezu alle heroinbedingten Todesfälle im Zusammenhang mit intravenösem Konsum stehen. Basierend auf der Analyse der Überzufälligkeit (vgl. Verlinkung von Datenquellen), sowohl in den ambulanten Daten als auch in der Drogentodesstatistik aufzutauchen, kann auf relativ einfache Art und Weise dieser Selektionsbias in den ambulanten Daten herausgerechnet werden.

3.4 Plausibilitätsprüfung der Schätzung und Übereinstimmung der Schätzverfahren

Um die Ergebnisse der beiden angewandten Methoden auf Plausibilität zu überprüfen, wurden zwei Tests durchgeführt. (1) Übereinstimmung beider Methoden: Hierbei wurden die Ergebnisse beider Methoden für die verschiedenen Altersgruppen und getrennt nach Geschlecht miteinander verglichen. Im günstigsten Fall kommen zwei unabhängige Methoden zu einer hohen Übereinstimmung der Resultate. Um diese zu überprüfen wurde die Korrelation beider Methoden über die verschiedenen Gruppen berechnet. (2) Summationsüberprüfung: Im günstigsten Fall ergibt die Summe der Einzelschätzungen für verschiedene Altersgruppen und Geschlechter einen ähnlichen Wert wie die Gesamtschätzung über alle Gruppen hinweg. Berechnet wurde die Abweichung der aufsummierten Einzelschätzungen von den Gesamtschätzungen.

3.5 Vergleich mit anderen Großstädten

Die Schätzung der reinen Populationsgröße ist für gesundheitspolitische Fragestellungen von hoher Relevanz. Von Interesse ist aber auch der Vergleich der Schätzungen mit denen anderer europäischer Großstädte. Dieser Vergleich ermöglicht eine Bewertung des Ausmaßes der Problematik in Berlin im Verhältnis zu anderen großen Städten in Europa. Das Problem von derartigen Vergleichen ist, dass sich die Schätzungen auf unterschiedliche Zeiträume beziehen. Insgesamt ist in vielen Städten ein Rückgang der Prävalenz des Konsums harter Drogen zu beobachten. Um die in dieser Untersuchung gewonnenen Zahlen in einen sinnvollen Vergleich zu anderen Prävalenzschätzungen zu setzen, wurde anhand der öffentlich zugänglichen Datensammlung der EMCDDA ein Trend berechnet und auf die Originalschätzungen angewendet. In der Datenbank der EMCDDA liegen Prävalenzschätzungen für problematischen Opioidkonsum für N=568 Regionen (Großstädte) vor. Als Prädiktor wurde das Studienjahr (Datenjahr) herangezogen, und als abhängige Variable die zentrale Schätzung pro tausend Einwohner. Dieses Vorgehen ist nicht ohne Einschränkungen. Für einige Metropolen in Europa mag dieser Trend ungültig sein. Es wurde daher auch der Vergleich mit den Originaldaten herangezogen.

3.6 Verlinkung verschiedener Datenquellen

Ein großer Vorteil der log-linearen Methode besteht darin, dass statistische Interaktionen zwischen Datenquellen berücksichtigt werden können, sodass es möglich ist, zu kontrollieren, ob die Wahrscheinlichkeit der Überschneidung zweier Datenquellen höher ist, als zufallsbedingt zu erwarten wäre. Die probabilistische Methode geht dagegen davon aus, dass die Datenquellen voneinander unabhängig sind, d.h. die Überschneidung nur durch den Zufall bedingt ist. Auf Basis der endgültigen Schätzungen kann anhand der probabilistischen Methode rückgerechnet werden, um welchen Faktor eine Datenquelle mit einer anderen assoziiert ist. Diese Analyse liefert Informationen darüber, welche Datenquellen für eine Schätzung ohne Berücksichtigung von Interaktionen ungeeignet sind. Zudem zeigt die Analyse typische Wege, die von Drogenkonsumenten beschritten werden.

4 Ergebnisse

4.1 Deskriptive Statistiken und Stichprobengrößen

Tabelle 1 zeigt die Fallzahlen der fünf Datenquellen nach Substanz, Geschlecht und intravenösem Konsum. Die höchsten Fallzahlen stellen die ambulanten Einrichtungen mit 7.657 Fällen im Jahr 2010, gefolgt von den stationären Einrichtungen und der Polizei. Erwartungsgemäß finden sich in der Drogentodesstatistik die wenigsten Fälle. Das Größenverhältnis männlich/weiblich liegt im Durchschnitt bei 3 bis 4 zu 1.

Tabelle 1. Stichprobengrößen nach Datenquelle, Datenjahr, Geschlecht und Substanz (Datenjahr 2010)

		Substanz			Geschlecht			
	Datenjahr	Opioide	Kokain	Stimulanzien	männlich	weiblich	Gesamt	i.v.**
Entgiftung	2010	665	241	182	625	222	847	136
Landeskriminalamt (LKA)	2010	806	626	660	1.430	274	1.704	232
Ambulante Einrichtungen	2010	5.632	3.212	2.195	5.841	1.816	7.657	3.470
Stationäre Einrichtungen	2010	1.042	1.081	1.017	1.415	306	1.721	717
Drogentodesfälle	2011*	89	26	18	86	21	107	26

Anmerkung. Gesamtzahl als Fallzahlen (N).
* Drogentodesfälle wurden jeweils ein Jahr später erfasst, da dieses Ereignis die Wahrscheinlichkeit für die Registrierung in einer anderen Datenquelle beeinflusst.
** i.v.=intravenöser Konsum (Daten liegen nur für die Einrichtungen der Suchthilfe vor).

Tabelle 2 zeigt die Fallzahlen nach Altersgruppe, Substanz, Geschlecht und intravenösem Konsum für das Datenjahr 2010. Die größte Altersgruppe stellen die 35- bis 64-jährigen Personen mit 4.339 Fällen.

Tabelle 2. Stichprobengrößen nach Altersgruppe, Substanz, Geschlecht und i.v. Konsum (Datenjahr 2010)

	Substanz			**Geschlecht**			
	Opioide	Kokain	Stimulanzien	männlich	weiblich	Gesamt	i.v.*
15-24 Jahre	983	681	991	1451	548	1999	384
25-34 Jahre	2355	1646	1310	3038	887	3925	1317
35-64 Jahre	3351	1440	607	3429	910	4339	2064

Anmerkung. Gesamtzahl als Fallzahlen.
* i.v.=intravenöser Konsum (Daten liegen nur für die Einrichtungen der Suchthilfe vor).

Abbildung 1 zeigt die Altersverteilung für die drei Substanzgruppen und für intravenösen Konsum. Die im Durchschnitt älteste Konsumentengruppe ist jene, die intravenös konsumiert (Mittelwert 36,3 Jahre), vor den Opioid- (35,6 Jahre) und den Kokainkonsumenten (32,7 Jahre). Am jüngsten sind Stimulanzienkonsumenten (28,8 Jahre). Bei Opioiden, Kokain und intravenösem Konsum zeigen sich lokale Maxima um das Alter von 30 Jahren und das Alter

von 41 Jahren. Opioid- und Kokainkonsumenten wurden am häufigsten im Alter von 30 Jahren registriert. Die Altersverteilung fällt bis zum Alter von 50 Jahren nur langsam ab. Stimulanzienkonsumenten wurden dagegen am häufigsten im Alter zwischen 20 und 30 Jahren von der Polizei erfasst oder nahmen Hilfe in Anspruch. Im Vergleich zu Opioid- und Kokainkonsumenten fällt der Anteil der Registrierungen mit dem Alter deutlich ab.

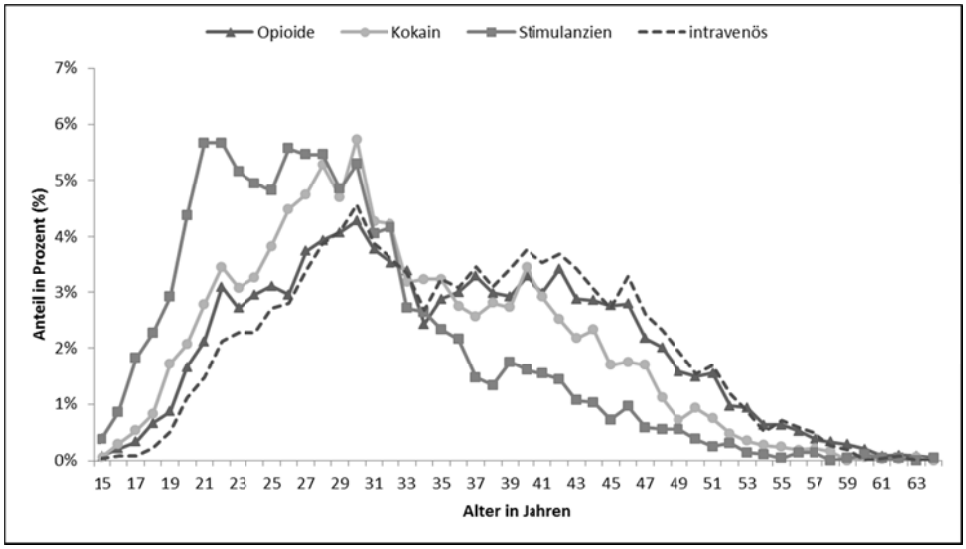


Abbildung 1. Altersverteilung der drei Substanzen und des intravenösen Konsums (Datenjahr 2010)

4.2 Fehleranfälligkeit des HIV-Codes

Eine Analyse der Abnahme der Überschneidung zwischen den Datenquellen, wenn zusätzliche HIV-Codes als Verbesserung der Information hinzugezogen werden, zeigt, dass bei etwa 7,5% der Fälle von unterschiedlichen Personen auszugehen ist, deren Identität von nur einem einzigen HIV-Code als identisch beurteilt wurde. Diese Fehleranfälligkeit kann dadurch bedingt sein, dass gewisse Buchstaben- oder Namenskombinationen gehäuft auftreten. Diese Rate falsch-positiver Identifikatoren führt dazu, dass die Überschneidung zwischen den Datenquellen zu hoch ausfällt. Die Schätzung der Gesamtpopulation würde in diesem Falle geringer ausfallen als ohne diesen Fehler. In der finalen Schätzung der Populationsgrößen wurde die Unterschätzung berücksichtigt.

4.3 Überschneidung zwischen verschiedenen Substanzen

Abbildung 2 zeigt die Verteilung der in den verschiedenen Datenquellen registrierten Drogenkonsumenten nach Substanzkonsum für das Datenjahr 2010. Der weitaus größte Teil sind Personen, die ausschließlich wegen ihres Opioidkonsums erfasst wurden (n=4779; 46,8%). Die zweitgrößte Gruppe sind Personen, die ausschließlich der Kategorie „Kokain“ zugeordnet wurden (n=1402; 13,7%). Die häufigste Kombination polyvalenten Substanzkonsums stellen Opioide und Kokain dar (n=1110; 10,9%), gefolgt von Kokain und Stimulanzien

(n=682; 7%) und allen drei Substanzen (n=574; 5,6%). Am seltensten wurden Personen im Zusammenhang mit Opioiden und Stimulanzien registriert (n=278; 2,7%). Damit wird deutlich, dass es sich bei den untersuchten Substanzgruppen um stark überlappende Problembereiche handelt.

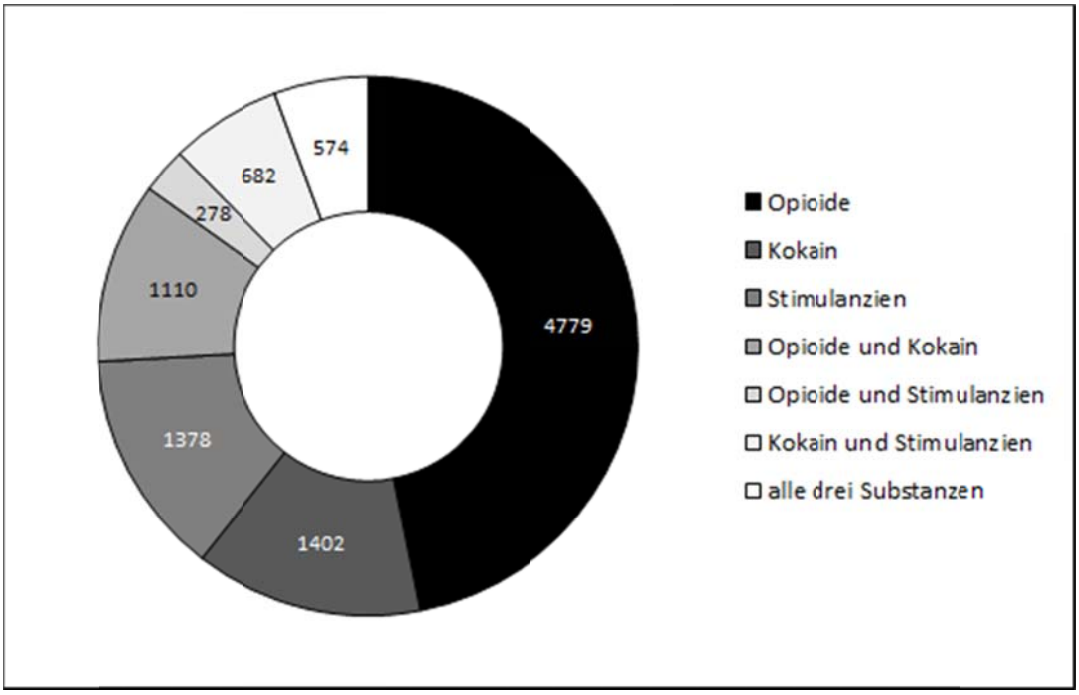


Abbildung 2. Substanzkonsum der registrierten Drogenkonsumenten (Datenjahr 2010).

4.4 Schätzung der Populationsgrößen

4.4.1 Log-lineare Regression

Tabelle 3 zeigt die Schätzungen des log-linearen Modells für das Jahr 2010. Die meisten Drogenkonsumenten pro Tausend der Berliner Bevölkerung finden sich in allen drei Substanzklassen in der Altersgruppe der 15- bis 24-jährigen Männer: Stimulanzien (22 pro T.), Opioide (17 pro T.), Kokain (15 pro T.). Über alle drei Substanzen hinweg sind in dieser Altersgruppe 49 von tausend Berliner Männern betroffen, was einem Anteil von etwa 5% der 15- bis 24-jährigen männlichen Berliner entspricht. Über alle Altersgruppen und beide Geschlechter hinweg wird die Anzahl von Konsumenten harter Drogen auf etwa 12 pro tausend Einwohner geschätzt, was etwas mehr als einem Prozent (1,15%) der 15- bis 64-jährigen Berliner Bevölkerung entspricht.

Tabelle 3. Populationsschätzung anhand log-linearer Regression nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010)

OPIOIDE

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	2.221	1.587	3.328	S1xS2, S2xS3	186.544	11,9
weiblich	25 bis 34	728	625	872	S3xS4, S2xS4	269.852	2,7
weiblich	35 bis 64	1.353	1.038	1.849	S1xS2, S3xS4	721.457	1,9
weiblich	15 bis 64	5.229	4.235	6.623	S1xS2, S2xS3	1.177.853	4,4
männlich	15 bis 24	3.142	2.683	3.762	S1xS2, S2xS3	182.765	17,2
männlich	25 bis 34	3.651	3.106	4.381	S1xS2, S2xS3	271.991	13,4
männlich	35 bis 64	5.086	4.253	6.205	S1xS2, S2xS3	741.535	6,9
männlich	15 bis 64	11.772	10.655	13.096	S1xS2, S2xS3	1.196.291	9,8
beide	15 bis 24	5.150	4.429	6.089	S1xS2, S2xS3	369.309	13,9
beide	25 bis 34	5.039	4.332	5.959	S1xS2, S2xS3	541.843	9,3
beide	35 bis 64	6.662	5.659	7.971	S1xS2, S2xS3	1.462.992	4,6
GESAMT	15 bis 64	16.773	15.300	18.486	S1xS2, S2xS3	2.374.144	7,1

KOKAIN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	1.167	796	1.872	S1xS2	186.544	6,3
weiblich	25 bis 34	765	451	1.534	S1xS2, S1xS4	269.852	2,8
weiblich	35 bis 64	206	176	257	S2xS3, S1xS3	721.457	0,3
weiblich	15 bis 64	2.938	2.053	4.512	S1xS2, S2xS4	1.177.853	2,5
männlich	15 bis 24	2.642	2.262	3.141	S1xS2, S2xS5	182.765	14,5
männlich	25 bis 34	2.797	2.370	3.362	S1xS2, S1xS4	271.991	10,3
männlich	35 bis 64	2.138	1.721	2.733	S1xS2, S1xS4	741.535	2,9
männlich	15 bis 64	7.446	6.742	8.277	S1xS2, S1xS4	1.196.291	6,2
beide	15 bis 24	3.712	3.185	4.396	S1xS2, S2xS5	369.309	10,1
beide	25 bis 34	3.467	2.950	4.142	S1xS2, S1xS4	541.843	6,4
beide	35 bis 64	2.546	2.088	3.181	S1xS2, S1xS4	1.462.992	1,7
GESAMT	15 bis 64	9.539	8.679	10.543	S1xS2, S1xS4	2.374.144	4,0

Tabelle 3 (Fortsetzung). Populationsschätzung anhand log-linearer Regression nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010)

STIMULANZIEN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	508	293	1.069	S1xS2	186.544	2,7
weiblich	25 bis 34	803	535	1.349	S1xS2, S2xS4	269.852	3,0
weiblich	35 bis 64	174	107	345	S1xS2	721.457	0,2
weiblich	15 bis 64	3.005	2.135	4.495	S1xS2, S1xS4	1.177.853	2,6
männlich	15 bis 24	4.081	3.470	4.872	S1xS2, S2xS4	182.765	22,3
männlich	25 bis 34	5.529	2.486	17.142	S1xS2, S2xS4	271.991	20,3
männlich	35 bis 64	1.841	1.502	2.317	S1xS2, S1xS4	741.535	2,5
männlich	15 bis 64	6.726	5.935	7.691	S1xS2, S2xS4	1.196.291	5,6
beide	15 bis 24	6.180	5.246	7.385	S1xS2, S2xS4	369.309	16,7
beide	25 bis 34	2.294	1.889	2.849	S1xS2, S1xS4	541.843	4,2
beide	35 bis 64	953	678	1.453	S1xS2, S2xS4	1.462.992	0,7
GESAMT	15 bis 64	9.407	8.322	10.720	S1xS2, S2xS4	2.374.144	4,0

ALLE SUBSTANZEN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	4.100	3.264	5.292	S1xS2, S3xS4	186.544	22,0
weiblich	25 bis 34	1.973	1.555	2.590	S1xS2, S3xS4	269.852	7,3
weiblich	35 bis 64	1.783	1.387	2.381	S1xS2, S3xS4	721.457	2,5
weiblich	15 bis 64	7.737	6.712	9.015	S1xS2, S3xS4	1.177.853	6,6
männlich	15 bis 24	8.933	7.574	10.723	S1xS2, S2xS3	182.765	48,9
männlich	25 bis 34	6.392	5.787	7.103	S1xS2, S1xS4	271.991	23,5
männlich	35 bis 64	6.672	5.874	7.647	S1xS2, S1xS4	741.535	9,0
männlich	15 bis 64	20.101	18.942	21.375	S1xS2, S3xS4	1.196.291	16,8
beide	15 bis 24	11.317	10.365	12.411	S1xS2, S3xS4	369.309	30,6
beide	25 bis 34	7.981	7.299	8.768	S1xS2, S3xS4	541.843	14,7
beide	35 bis 64	8.126	7.274	9.143	S1xS2, S3xS4	1.462.992	5,6
GESAMT	15 bis 64	27.335	25.866	28.939	S1xS2, S3xS4	2.374.144	11,5

Anmerkung. Schätzungen gerundet auf ganze Zahlen. Konfidenzintervall entspricht den mittleren 95% und liegt asymmetrisch um die zentrale Schätzung (links näher als rechts). Modell enthält die dargestellten Interaktionen zwischen Datenquellen (S1=ambulant, S2=ambulant, S3=Polizei, S4=Entgiftung). Bevölkerungszahlen der jeweiligen Altersgruppe gemäß des Bevölkerungsstandes Berlin zum Stichtag 31.12.2010 (Statistisches Bundesamt, 2014). Prävalenz p.T.(„pro tausend Einwohner“) der jeweiligen Altersgruppe.

4.4.3 Probabilistische Methode

Tabelle 4 zeigt die Schätzungen der probabilistischen Methode. Die höchsten Prävalenzwerte finden sich gemäß dieser Methode in der Gruppe der 25- bis 34-jährigen Männer: Opioide (11 pro T.), Stimulanzien (9 pro T.), Kokain (8 pro T.). Für alle drei Substanzen wird die Anzahl harter Drogenkonsumenten in dieser Gruppe auf 24 pro tausend Berliner Einwohner (25- bis 34-jährige Männer) geschätzt. Dies entspricht einer Prävalenz von etwa 2,4%. Insgesamt schätzt die probabilistische Methode die Anzahl harter Drogenkonsumenten (Opioide, Kokain oder Stimulanzien) in der Berliner Bevölkerung im Alter von 15 bis 64 Jahren auf 9 pro Tausend und liegt damit etwas niedriger als die Schätzung der log-linearen Regression unter Berücksichtigung der Interaktionen.

Tabelle 4. Populationsschätzung anhand der probabilistischen Methode nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010)

OPIOIDE

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	586**	>642	>642	1.424	186.544	3,1
weiblich	25 bis 34	854	854	468	4.759	269.852	3,2
weiblich	35 bis 64	1.336	1336	509	3.465	721.457	1,9
weiblich	15 bis 64	3.008	3008	1619	4.547	1.177.853	2,6
männlich	15 bis 24	690**	>1524	>1524	8.818	182.765	3,8
männlich	25 bis 34	3.052	3052	1691	6.700	271.991	11,2
männlich	35 bis 64	4.006	4006	2279	5.733	741.535	5,4
männlich	15 bis 64	7.768	7768	5022	15.515	1.196.291	6,5
beide	15 bis 24	1.157**	>2166	>2166	8.722	369.309	3,1
beide	25 bis 34	4.196	4196	2159	9.504	541.843	7,7
beide	35 bis 64	5.428	5428	3261	7.594	1.462.992	3,7
GESAMT	15 bis 64	11.365	11365	6641	19.407	2.374.144	4,8

KOKAIN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	338**	>354	>354	3.421	186.544	1,8
weiblich	25 bis 34	594	594	185	1.338	269.852	2,2
weiblich	35 bis 64	388	388	162	2.936	721.457	0,5
weiblich	15 bis 64	1.527	1527	701	3.306	1.177.853	1,3
männlich	15 bis 24	459**	>1195	>1195	1.195	182.765	2,5
männlich	25 bis 34	2.138	2138	1120	5.227	271.991	7,9
männlich	35 bis 64	1.827	1827	742	3.086	741.535	2,5
männlich	15 bis 64	3.939	3939	3057	11.861	1.196.291	3,3
beide	15 bis 24	633**	>1549	>1549	1.549	369.309	1,7
beide	25 bis 34	2.767	2767	1305	7.484	541.843	5,1
beide	35 bis 64	2.303	2303	904	3.917	1.462.992	1,6
GESAMT	15 bis 64	5.297	5297	3758	12.964	2.374.144	2,2

Tabelle 4 (Fortsetzung). Populationsschätzung anhand der probabilistischen Methode nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2010)

STIMULANZIEN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	541	541	467	616	186.544	2,9
weiblich	25 bis 34	844	844	771	916	269.852	3,1
weiblich	35 bis 64	78	78	71	181	721.457	0,1
weiblich	15 bis 64	1.617	1.617	632	3.605	1.177.853	1,4
männlich	15 bis 24	1.224**	>1.325	>1.325	11.306	182.765	6,7
männlich	25 bis 34	2.485	2.485	426	5.066	271.991	9,1
männlich	35 bis 64	424	674	674	2.230	741.535	0,6
männlich	15 bis 64	4.223	4.223	2.270	7.249	1.196.291	3,5
beide	15 bis 24	1.980	1.980	1.751	3.208	369.309	5,4
beide	25 bis 34	3.911	3.911	809	8.088	541.843	7,2
beide	35 bis 64	550	550	342	3.466	1.462.992	0,4
GESAMT	15 bis 64	6.462	6.462	2.902	11.356	2.374.144	2,7

ALLE SUBSTANZEN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	1.857	1.857	1.063	6.084	186.544	10,0
weiblich	25 bis 34	1.800	1.800	644	8.823	269.852	6,7
weiblich	35 bis 64	1.855	1.855	605	4.921	721.457	2,6
weiblich	15 bis 64	5.538	5.538	2.704	8.372	1.177.853	4,7
männlich	15 bis 24	2.263**	>2.927	>2.927	6.635	182.765	12,4
männlich	25 bis 34	6.531	6.531	2.595	10.933	271.991	24,0
männlich	35 bis 64	6.503	6.503	3.823	9.183	741.535	8,8
männlich	15 bis 64	15.988	15.988	7.772	25.222	1.196.291	13,4
beide	15 bis 24	3.549**	>3.990	>3.990	10.676	369.309	9,6
beide	25 bis 34	8.811	8.811	3.239	15.779	541.843	16,3
beide	35 bis 64	8.687	8.687	5.342	12.032	1.462.992	5,9
GESAMT	15 bis 64	21.677	21.677	12.380	30.973	2.374.144	9,1

Anmerkung. Schätzungen gerundet auf ganze Zahlen. Konfidenzintervall entspricht den mittleren 95% und liegt teilweise asymmetrisch um die zentrale Schätzung. Bevölkerungszahlen der jeweiligen Altersgruppe gemäß des Bevölkerungsstandes Berlin am Stichtag 31.12.2010 (Statistisches Bundesamt, 2014). Prävalenz p.T. („pro tausend Einwohner“) der jeweiligen Altersgruppe.
* Sofern die Schätzung unterhalb der registrierten Anzahl von Personen („Hellfeld“) liegt, wurde diese Zahl als das korrigierte Schätzminimum und als unteres Minimum des Konfidenzintervalls angenommen.
** Die Schätzung liegt unterhalb der registrierten Anzahl von Personen („Hellfeld“). Die registrierte Anzahl von Personen stellt somit das empirische Minimum der Schätzung dar (>). Bei mit ** und > markierten Zeilen kann somit die untere Grenze des Konfidenzintervalls nicht eindeutig bestimmt werden.

Abbildung 3 zeigt einen Überblick der Ergebnisse beider Schätzmethoden getrennt nach Substanz und für alle drei Substanzen gemeinsam. Die Abbildung macht ersichtlich, dass (1) die probabilistische Methode jeweils geringere Schätzungen produziert; (2) Die Konfidenzintervalle der log-linearen Regression sind deutlich enger; (3) Die Konfidenzintervalle der log-linearen Regression liegen jeweils innerhalb der Vertrauensintervalle der probabilistischen Methode.

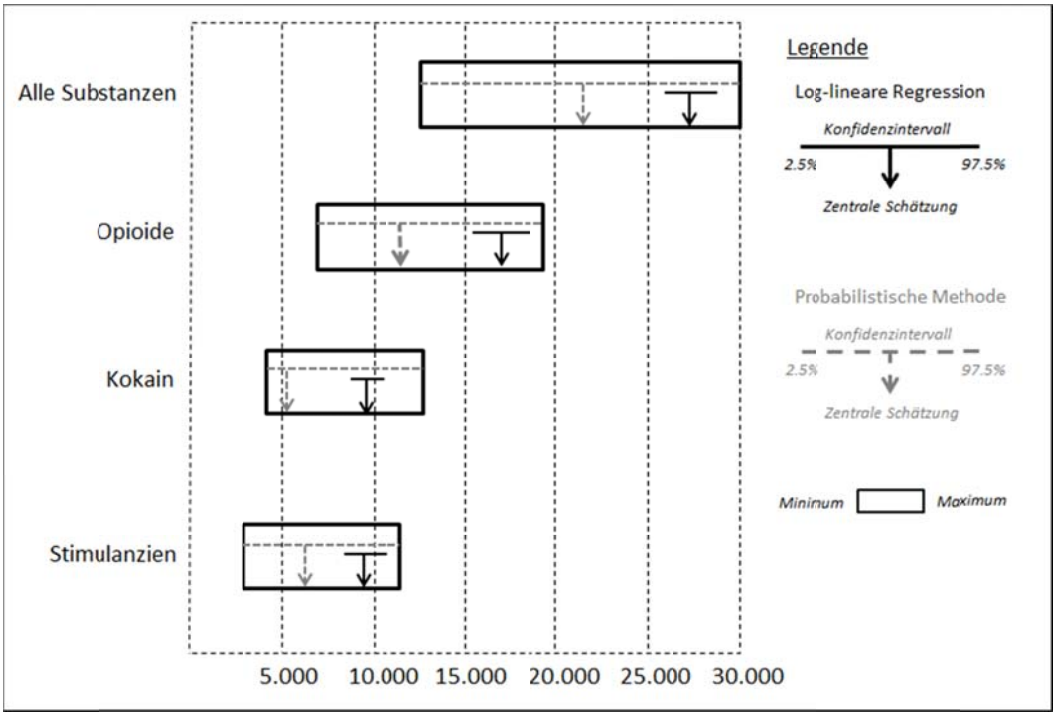


Abbildung 3. Zusammenfassung der beiden Schätzmethoden nach Substanzgruppen (Datenjahr 2010)

4.4.4 Schätzung intravenöser Konsumenten

Aus den Daten der Berliner Suchthilfestatistik für das Jahr 2012 (Pfeiffer-Gerschel, Steppan & Brand, 2013) geht hervor, dass 31,7% der wegen Opioiden behandelten Personen angeben, nie intravenös konsumiert zu haben. Dieser Anteil dürfte außerhalb des ambulanten Settings, d.h. bei Personen, die keine Hilfe in Anspruch nehmen, geringer sein. Dies kann durch eine Überzufälligkeit des Auftretens von Fällen mit i.v. Konsum in der Drogentodesstatistik und in den ambulanten Daten bedingt werden (1,38-fach; vgl. Verlinkung der Datenquellen). Geht man davon aus, dass bei nicht-intravenösem Konsum das Mortalitätsrisiko gering ist (1% der opioidbedingten Todesfälle; Darke & Ross, 2000), so kann der inverse Anteil von 100% - 31,7% = 68,3% durch den Divisor 1,38 nach unten korrigiert werden. Damit ergibt sich ein geschätzter Anteil von exakt 50,0% der Opioidkonsumenten, die intravenös konsumierten. Für die Diagnose Kokain (4,8%) und Stimulanzien (4,6%) ergeben sich auf dieselbe Weise entsprechend geringere Anteile intravenösen Konsums.

4.4.5 Schätzung polyvalenter Störungen

Drogenkonsumenten nehmen in der Regel nicht nur eine Substanz, sondern verschiedene Substanzen, sei es zu unterschiedlichen Zeitpunkten oder gleichzeitig, um bestimmte Wirkungen zu erzielen. Die hier vorgenommen Schätzungen der Mehrfachabhängigkeit von den Substanzen Opioide, Kokain oder Stimulanzien sind aber nicht mit polyvalenten Gebrauchsmustern zu verwechseln; hier geht es ausschließlich um die Schätzung der Mehrfachabhängigkeit bzw. Mehrfachstörung, d.h. von jeweils zwei der drei Substanzen oder von allen drei Substanzen. Für die Schätzung des Anteils der Mehrfachstörung wurde die probabilistische Methode verwendet. Da es hier um die Abschätzung von Anteilen an der Gesamtpopulation geht, spielt die absolute Höhe eine nur untergeordnete Rolle.

Abbildung 4 zeigt die anhand der probabilistischen Methode geschätzten Anteile mono- und polyvalenter Substanzstörung bezogen auf die Gesamtschätzung der drei Störungsgruppen (Opioide, Kokain und/oder Stimulanzien). Mit 47% ist die Gruppe mit einer ausschließlichen Störung im Zusammenhang mit Opioiden am größten, gefolgt von der Gruppe, die ausschließlich eine stimulanzienbezogene Störung (15%) aufweist und der Gruppe mit einer ausschließlichen Kokainstörung. Die Gruppe mit einer Störung im Zusammenhang mit Opioiden und Kokain wird auf 6% geschätzt, Störungen im Zusammenhang mit Kokain und Stimulanzien auf 3% und 1% weisen eine Störung im Zusammenhang mit den Opioiden und Stimulanzien auf. Der Anteil der Personen mit einer opioid-, kokain- und stimulanzienbezogenen Störung an der Gesamtschätzung wird auf 3% geschätzt.

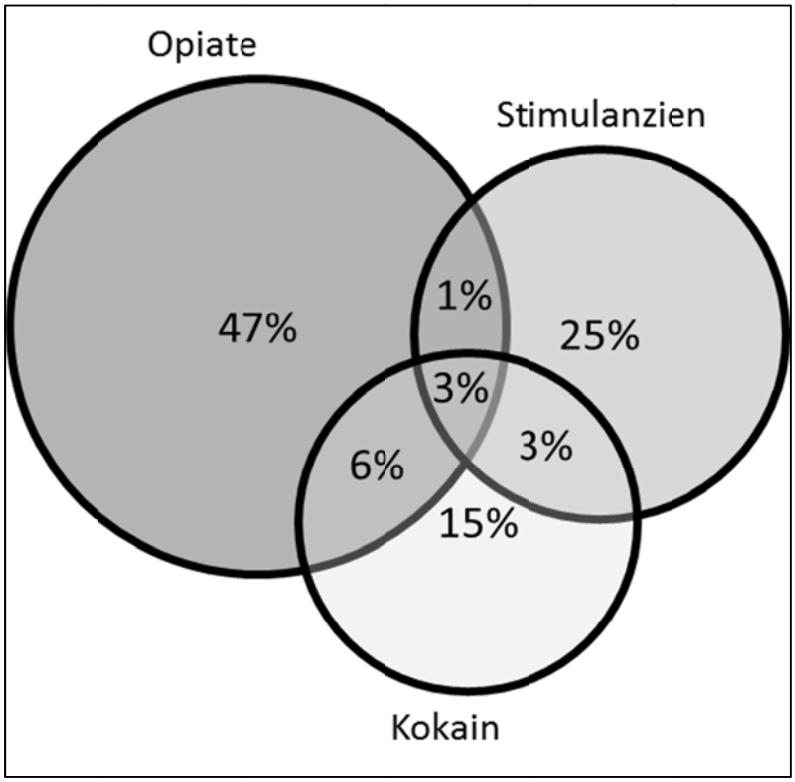


Abbildung 4. Schätzung des Anteils mono- und polyvalenter Störungen (Datenjahr 2010)

Geht man bei der Hochrechnung davon aus, dass die mit der probabilistischen Methode ermittelten Anteile polyvalenter Störungsbilder gleichfalls für die Gesamtschätzung der Personen mit einer Substanzstörung im Zusammenhang mit Opioiden, Kokain oder Stimulanzien gilt, ergeben sich folgende Schätzungen: Opioide & Stimulanzien: 200-300 Personen; Opioide & Kokain: 1.300-1.600 Personen; Kokain & Stimulanzien 600-800 Personen; unter den insgesamt hochgerechneten 2.100-2.700 Personen mit einer dualen Störung ist mit etwa 600-800 Personen mit einer Störung hinsichtlich aller drei Substanzen zu rechen.

4.5 Vergleich mit anderen Großstädten

Abbildung 5 zeigt ein Streudiagramm für N=568 regionale Schätzungen aus der Datenbank der EMCDDA zu Opioidkonsum. Die Linie stellt die Regressionsgerade dar, deren Gleichung oben dargestellt ist. Es ergibt sich im Zeitraum zwischen 1995 und 2010 ein Rückgang der Prävalenz von 0.15 pro tausend Fälle pro Jahr. Liegt die Schätzung zehn Jahre zurück, ist diesem Modell zufolge der wahre Wert heute um 1.5 pro tausend Einwohner niedriger.

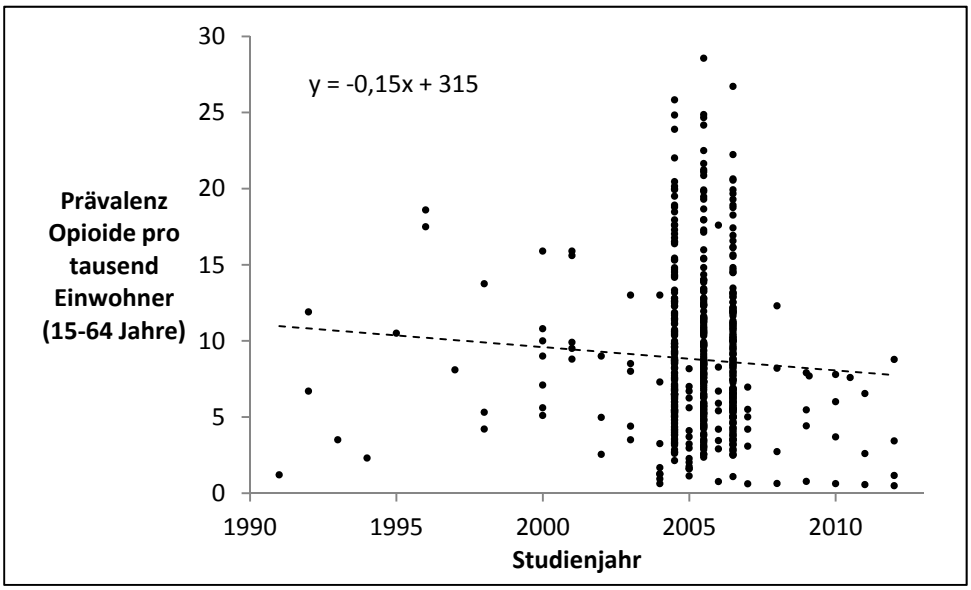


Abbildung 5. Europäischer Trend des Rückgangs der Prävalenz des Opioidkonsums auf Basis der Datenbank der EMCDDA

Tabelle 5. Vergleich der Berliner Schätzwerte mit den Schätzungen anderer europäischer Großstädte (Opioide)

Region	Studienjahr	Originalschätzung	Extrapolation 2010
Manchester	2006	19,8	19,1
Liverpool	2007	19,7	19,1
Birmingham	2006	19,3	18,6
Dublin	2006	17,6	17,0
Kopenhagen	1996	17,5	15,3
Sofia	2004	13,0	12,1
Riga	2008	12,3	12,0
Ljubljana	2003	13,0	11,9
London	2007	10,0	9,5
Madrid	1992	11,9	9,1
Prag	2012	8,8	9,1
Genua	2001	9,9	8,5
Wien	1995	10,5	8,2
Berlin*	2010	7,0	7,0
Mailand	2005	7,0	6,2
Enschede	2010	6,0	6,0
Verona	2000	7,1	5,6
Amsterdam	2009	5,5	5,3
Vilnius	2006	5,4	4,8
Berlin**	2010	4,7	4,7
Lazium (Rom)	1992	6,7	3,9
Venedig	2000	5,1	3,6
Athen	2004	3,3	2,3
Helsinki	2005	3,0	2,2
Budapest	2005	2,3	1,5
Thessaloniki	2002	2,5	1,3
Utrecht	1993	3,5	0,9

Anmerkung. Originalschätzungen in Fälle pro tausend Einwohner (Opioidkonsumenten). Extrapolation gemäß einer Abnahme von 0.1529 pro tausend Fälle pro Jahr.
* entspricht den Schätzungen der log-linearen Regression.
** entspricht den Schätzungen der probabilistischen Methode.

Tabelle 5 zeigt den Vergleich der für Berlin ermittelten Anzahl von Opioidkonsumenten pro tausend Einwohner für das Jahr 2010 mit den Schätzungen anderer europäischer Großstädte (aus der Datenbank der EMCDDA). Neben der Originalschätzung enthält die Tabelle den geschätzten Wert für 2010 gemäß der oben angeführten Korrektur. Für Berlin sind die beiden Schätzungen in der Tabelle eingetragen. Die Schätzungen liegen in der Mitte bzw. in der unteren Hälfte der Tabelle. Im Vergleich zur stark ausgeprägten Problematik in englischen Städten, ist die Problematik in Berlin um das Drei- bis Vierfache geringer. Berlin liegt hinsichtlich des Umfangs von Opioidkonsumenten im Bereich norditalienischer und niederländischer Städte. Etwas höhere Prävalenzwerte finden sich auch in Wien, Prag und Riga, die geographisch eine relativ geringe Distanz zu Berlin aufweisen. Aus der Datenbank der EMCDDA wurden für diese Tabelle die aktuellsten Schätzungen verwendet.

Abbildung 6 zeigt den Vergleich der Berliner Schätzungen mit den Schätzungen anderer europäischer Großstädte. Die horizontalen Linien repräsentieren die 95%-Konfidenzintervalle für die Berliner Schätzungen. Das graue Viereck zeigt den 95%-Vertrauensbereich an. Gemäß der log-linearen Regression ist die Prävalenz des Opioidkonsums in Berlin mit mehr als 95%iger Wahrscheinlichkeit niedriger als die von Wien und höher als die von Mailand. Gemäß der probabilistischen Methode unterscheidet sich die Prävalenz in Berlin von den Prävalenzen in Wien und Athen.

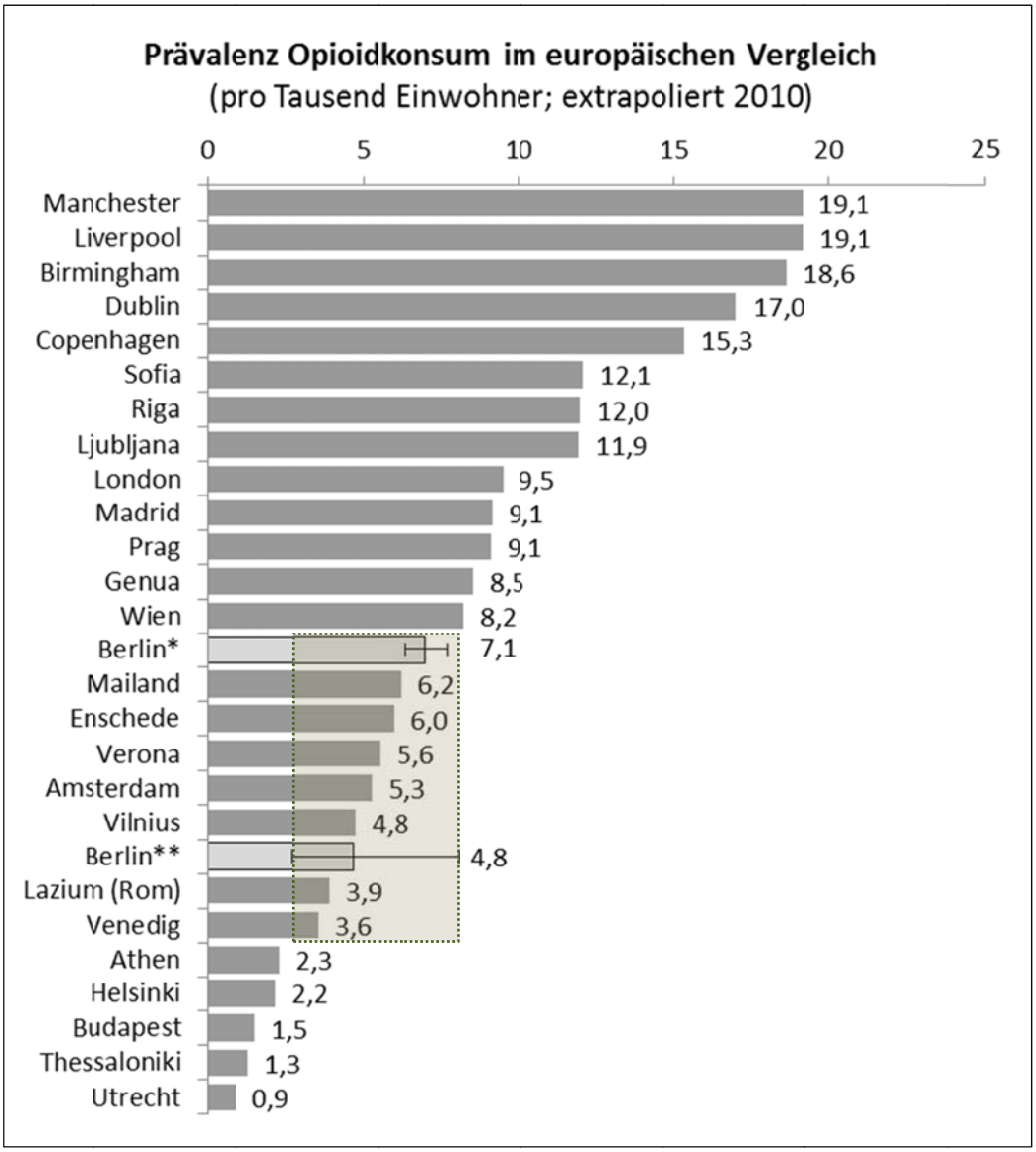


Abbildung 6. Graphischer Vergleich der Berliner Schätzwerte mit den Schätzungen anderer Europäischer Großstädte

4.6 Plausibilitätsprüfung der Schätzung und Übereinstimmung der Schätzverfahren

4.6.1 Übereinstimmung beider Methoden

Abbildung 7 zeigt den hoch linearen Zusammenhang zwischen den beiden Schätzmethoden. Im Diagramm dargestellt sind sowohl die Regressionsgleichung als auch das Quadrat der Korrelation (Bestimmtheitsmaß). Beide Methoden führen zu ähnlichen Ergebnissen in Bezug auf die relative Höhe der Prävalenzwerte. Die Korrelation der beiden Schätzmethoden beträgt r=.93, sodass sich ein R^2 von 0.87 ergibt. In absoluten Zahlen sind die Schätzwerte der probabilistischen Methode um den Faktor 0,7394 niedriger. In der Abbildung sind alle Schätzungen pro Substanz nach Altersgruppe und Geschlecht aus den Tabellen 3 und 4 dargestellt (48 Datenpunkte).

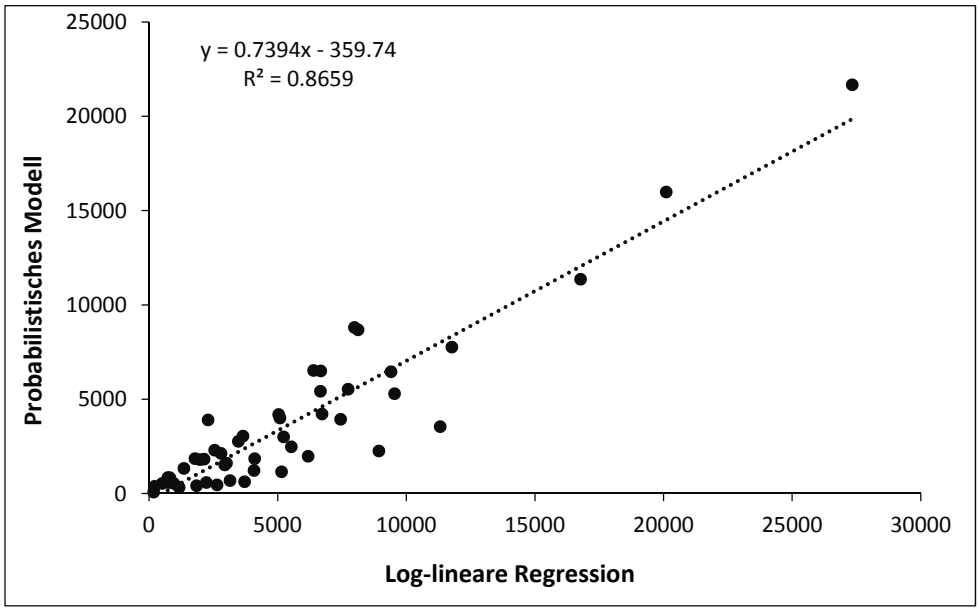


Abbildung 7. Übereinstimmung der Schätzungen nach Substanz, Altersgruppe und Geschlecht aus den Tabellen 2 und 3.

4.6.2 Summationsüberprüfung

Abbildung 8 zeigt die Abweichungen der über die Altersgruppen aufsummierten Prävalenzschätzungen pro Substanz für Männer, Frauen und beide Geschlechter vom jeweiligen Gesamtwert (4 mal 3 = 12 Summenbildungen pro Methode, vgl. Tabelle 3 und 4). Im günstigsten Fall liegt die Abweichung nahe Null, d.h. die Aufsummierung deckt sich mit der Gesamtschätzung. Die Abbildung zeigt, dass bei der log-linearen Regression höhere Abweichungen von Null für einzelne Teilschätzungen zu beobachten sind. Die Abweichungen der Teilschätzungen von den Gesamtschätzungen der probabilistischen Methode liegen deutlich näher bei null, was für die Robustheit dieser Methode spricht. Werden die Abweichungen gemittelt, weisen die Schätzungen der log-linearen Regression im Durchschnitt eine Abweichung von 0,7% und die der probabilistischen Methode auf 2% auf. Bei beiden Methoden

führt die Summation der Einzelschätzungen zu einer sehr ähnlichen Schätzung wie die Gesamtschätzung, was ebenfalls für deren Plausibilität spricht.

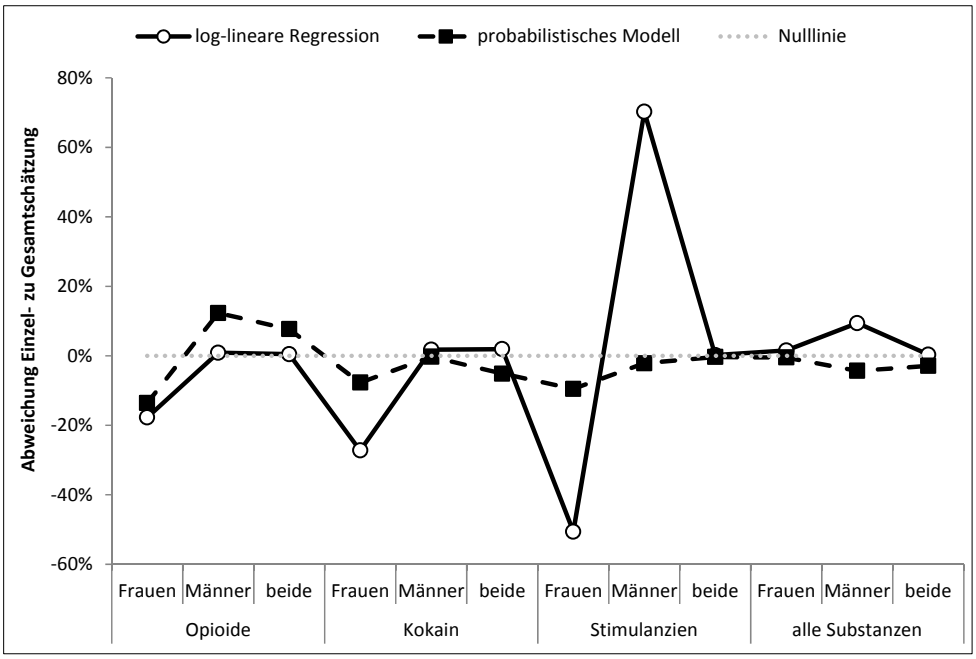


Abbildung 8. Abweichung der Summe der Einzelschätzungen von den Gesamtschätzungen

4.7 Verlinkung verschiedener Datenquellen

Tabelle 6 zeigt den jeweiligen Faktor, mit dem man die zufallsbedingten Überschneidungen zwischen den jeweiligen Datenquellen multiplizieren müsste, damit man die beobachteten Überschneidungen erhält. Die Tabelle zeigt, dass die höchste Verlinkung zwischen dem ambulanten und dem stationären Setting besteht. Personen, die in der ambulanten Datenquelle auftauchen, sind um das 2,14-fache häufiger in den stationären Daten enthalten, als dies bei völliger Unabhängigkeit der Datenquellen der Fall wäre. Ebenso weisen Mortalität und das ambulante Setting eine erhöhte Assoziation (1,38) auf. Umgekehrt liegt eine negative Assoziation zwischen Entgiftungsbehandlung und der Polizeistatistik vor (0,45). Personen, die sich einer Entgiftungsbehandlung unterzogen haben, werden deutlich seltener von der Polizei als drogenauffällig registriert, als zufallsbedingt zu erwarten wäre.

Tabelle 6. Überzufälligkeit des Overlaps verschiedener Datenquellen

	Entgiftung	**LKA**	**ambulant**	**stationär**
Entgiftung	-	-	-	-
Landeskriminalamt (LKA)	0,45	-	-	-
Ambulante Einrichtungen	1,03	1,04	-	-
Stationäre Einrichtungen	1,21	1,06	2,14	-
Mortalität	0,9	0,9	1,38	0,72

Anmerkung. Faktorielle (zu multiplizierende) Überzufälligkeit der Überschneidung der verschiedenen Datenquellen für alle Substanzen, basierend auf dem Mittelwert der beiden Schätzmethoden.

5 Diskussion

Die vorliegenden Prävalenzschätzungen des Konsums harter Drogen in Berlin kommen für die 15- bis 64-jährige Bevölkerung zu folgenden Ergebnissen: opioidbezogene Störung (11.300-16.700 oder 5-7 von Tausend), kokainbezogene Störungen (5.300-9.500 oder 2-4 von Tausend) und stimulanzienbezogene Störungen (6.500-9.400 oder 3-4 von Tausend). Störungen im Zusammenhang mit Opioiden, Kokain oder Stimulanzien wurden auf 9-12 pro Tausend geschätzt (21.700-27.300). Die Definition substanzbezogener Störungen orientiert sich weitgehend an der Definition psychischer und Verhaltensstörungen durch psychotrope Substanzen gemäß der internationalen Klassifikation von Krankheiten (ICD-10). Mit Ausnahme der vom Landeskriminalamt Berlin im Zusammenhang mit Drogenkonsum registrierten Personen und der Drogentodesstatistik wurden Diagnosen verwendet. Im Falle eines im Zusammenhang mit harten Drogen registrierten Drogentodes liegt jedoch mit hoher Wahrscheinlichkeit ein ursächlicher Zusammenhang vor, der auf eine den Drogentod vorausgehende Substanzstörung der Betroffenen hinweist. Mit wenigen Ausnahmen, z.B. Dealern, die selbst keine Drogen nehmen, kann auch bei Personen, die wegen BtMG Verstößen auffallen, mit hoher Wahrscheinlichkeit von einer substanzbezogenen Störung ausgegangen werden. Für die Entgiftung im Krankenhaus wurden die Diagnosen psychische und Verhaltensstörungen durch Opioide (F11), Kokain (F14) oder Stimulanzien (F15) und für ambulante und stationäre Beratung/Behandlung die Diagnosen schädlicher Gebrauch oder Abhängigkeit von Opioiden (F11.1 oder F11.2), Kokain (F14.1 oder F14.2) oder Stimulanzien (F15.1 oder F15.2) herangezogen.

Vergleiche mit früheren Schätzungen zum Umfang des Konsums harter Drogen in Berlin sind wegen der Verwendung unterschiedlicher Methoden und Definitionen sowie aufgrund historischer Entwicklungen nur bedingt möglich. So bezieht sich die Schätzung für das Jahr 1979 auf Berlin West und „Heroinabhängige“ (Skarabis and Patzak, 1981). Die Schätzung aus dem Jahr 1992 umfasste zwar Berlin als gesamte Region, schränkte aber die Schätzung auf intravenös konsumierende Heroinabhängige ein (Kirschner et al., 1994). Während in den 1970er und 1980er Jahren Heroin die Droge der Wahl war und die intravenöse Applikation die überwiegende Form der Einnahme darstellte, hat sich der Drogenkonsum in den letzten 20 bis 25 Jahren insbesondere in den Großstädten Europas weitgehend diversifiziert. Neben Heroin haben sich Kokain, Crack und insbesondere Stimulanzien auf dem Drogenmarkt etabliert. Polyvalenter Konsum stellt bei Drogenkonsumenten die Regel und nicht die Ausnahme dar (Coffin et al., 2003; Darke & Hall, 1995; Leri, Bruneau & Stewart, 2003). Die Ergebnisse der DSHS weisen darauf hin, dass neben Heroin andere opioidhaltige Substanzen zum Störungsbild beitragen. Bei den Hauptdiagnosen entfallen knapp zwei Drittel auf Heroin, d.h. Heroin ist zwar nach wie vor das am häufigsten konsumierte Opiat, jedoch entfällt ein nicht zu vernachlässigender Anteil der behandlungsleitenden Diagnosen auf Störungen im Zusammenhang mit anderen Opioiden (Pfeiffer-Gerschel, Steppan & Brand, 2013). Die Hochrechnungen polyvalenter Störungen ergeben einen Anteil von etwa 10% von Personen mit einer dualen Störung an der Gesamtschätzung. Damit weisen von den geschätzten 21.700-27.300 Personen mit einer Störung im Zusammenhang mit Opioiden, Kokain oder

Stimulanzien etwa 2000-3000 eine duale Störung mit jeweils zwei der drei Substanzen auf. In dieser Gruppe sind etwa 600-800 Personen mit einer Dreifachstörung (Opioiden, Kokain und Stimulanzien) belastet.

Zieht man zur Schätzung des Anteils des intravenösen Konsums (i.v.) auch die Daten der Suchthilfestatistik heran, findet man einen Anteil von knapp 60% der Heroinkonsumenten in ambulanter Betreuung, die angaben, Heroin im letzten Jahr injiziert zu haben und 30%, die angaben, Opioide noch nie injiziert zu haben (Pfeiffer-Gerschel, Steppan & Brand, 2013). Die Schätzungen des i.v. Konsums von Opioiden aus den vorliegenden Daten für Berlin kommen auf einen Anteil von 50%. Berechnet man auf der Grundlage eines Anteils von 40-50% von nicht i.v. Konsumenten die Prävalenz der Personen mit einer opioidbezogenen Störung der vorliegenden Schätzung für das Jahr 2010, kommt man auf ein N von 8.400-10.100 (Log-lineares Modell) bzw. von 6.000-7.000 (Probabilistisches Modell) i.v. Opioidkonsumenten (2.5-4.25 pro Tausend). Verglichen mit der Schätzung für 1992 von 6.000-8.000 i.v. konsumierenden Heroinabhängigen (2.5-3.3 pro Tausend) ist demnach von einer leichten Zunahme der Prävalenz von Personen mit einer opioidbezogenen Störung, die i.v. konsumieren, auszugehen.

Auf der Grundlage der Meldungen an das Substitutionsregister (§5a Betäubungsmittel-Verschreibungsverordnung) der Kassenärztlichen Vereinigung (KV) waren in Berlin im Jahr 2010 5.082 Heroinabhängige in Substitutionsbehandlung (Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Stand Januar 2011). Da die Schätzungen opioidbezogener Störungen ohne das Substitutionsregister durchgeführt werden musste, sind nur diejenigen Substituierten, die in einer der verwendeten Quellen registriert wurden, in die Schätzungen eingeflossen. Mit Ausnahme der psychosozialen Betreuung in der ambulanten Suchthilfe liegen aber keine Informationen zum Anteil Substituierter in den Datenquellen vor. In der ambulanten Suchthilfe wurden 2010 2.455 Personen als Substituiert gemeldet, so dass von einem Anteil von 50% im ambulanten Setting betreuten substituierten Opioidabhängigen ausgegangen werden kann. Damit waren von den geschätzten 6.500-9.250 i.v. konsumierenden Opioidabhängigen 2.500 Abhängige in Substitutionsbehandlung. Andererseits bedeutet dies, dass zur vorliegenden Schätzung von 11.300-16.700 Personen mit einer opioidbezogenen Störung 50% (2.500) Substituierte addiert werden müssten und man insgesamt von 16.500 (13.800-19.200) Personen mit einer opioidbezogenen Störung ausgehen müsste.

Die adaptierten Hochrechnungen vergleichbarerer Schätzungen der Prävalenz opioidbezogener Störungen aus anderen Großstädten Europas für das Jahr 2010 zeigen, dass sich die Schätzungen für Berlin bezüglich des Umfangs des Opioidproblems im mittleren bis unteren Bereich der verfügbaren Schätzungen bewegen. Bezogen auf die log-lineare Schätzung von 7 pro tausend Personen mit einer opioidbezogenen Störung hat das Drogenproblem in Berlin eine mit Wien und Mailand vergleichbare Größenordnung. Ein höherer Problemumfang findet sich in Madrid, Prag, Riga, Kopenhagen, London und Dublin. Im Vergleich zu den britischen Städten Birmingham, Liverpool und Manchester ist das Opioidproblem in Berlin um das Dreifache geringer. Niedrigere Prävalenzwerte werden für Amsterdam, Rom, Athen und auch Helsinki berichtet.

Die Schätzungen der beiden Modelle für das Jahr 2011 (vgl. Anhang I Kapitel 9) fallen für alle Substanzen etwas höher aus, weisen aber insgesamt auf die Robustheit der Schätzungen hin. Auch fallen die Schätzungen des probabilistischen Modells systematisch niedriger aus als die der log-linearen Regression und die Konfidenzintervalle sind weiter. Trotz systematisch niedriger Werte unterscheiden sich die Schätzungen des probabilistischen Modells statistisch nicht von den Schätzungen des etablierten Verfahrens der Capture-recapture Methode mit der log-linearen Regression. Als Gründe für die Unterschiede in den Modellschätzungen sind die Berücksichtigung der Abhängigkeit zwischen den Datenquellen im log-linearen Modell sowie die für die Schätzung verwendeten unterschiedlichen Datenquellen zu nennen. Methodenkritisch sei auf eine neuere Studie zur Validität der Standard Capture-recapture Methode verweisen. Jones und Kollegen (2014) kommen in ihren Analysen zu dem Schluss, dass Standard log-lineare Modelle die Population überschätzen, insbesondere wenn die beobachtete Population klein ist ["We have demonstrated that, in the presence of referral mechanisms between sources, the standard log-linear modeling approach to CRC is too restrictive and can lead to grossly misleading prevalence estimates. The absolute bias will be largest when the proportion of the population observed is low." (Jones et al., 2014, S. 6)].

Für Berlin liegen mit dieser Studie erstmals Schätzungen des Umfangs der Suchtbelastung durch Kokain (5.300-9.500) und Stimulanzien (6.500-9.400) vor. Prävalenzschätzungen auf der Grundlage von Bevölkerungsbefragungen berichten, dass Kokain und Stimulanzien nach Cannabis die am weitesten verbreiteten Drogen in Deutschland sind. Die Prävalenz kokain- und stimulanzienbezogener Störungen nach DSM-IV wird in der Allgemeinbevölkerung auf 0,2% bzw. 0,3% geschätzt (Pabst, Kraus, Gomes de Matos & Piontek, 2013). Die für die größte deutsche Stadt geschätzte Suchtbelastung durch Kokain und Stimulanzien ist damit mit jeweils 0,4% etwas höher als in Gesamtdeutschland. Insgesamt weisen 9-12 von Tausend 15- bis 64-jährigen Berlinern eine substanzbezogene Störung im Zusammenhang mit Opioide, Kokain oder Stimulanzien auf. Bei nahezu einem Drittel der Drogenkonsumenten (ca. 8.000) liegt eine Störung im Zusammenhang mit mindesten zwei Substanzen vor.

Zur Beurteilung der Güte unserer Schätzungen müssen mögliche Verletzungen der Grundannahmen des Capture-recapture Modells bewertet werden. Die eindeutige Identifizierbarkeit der relevanten Fälle halten wir bei allen Datenquellen mit Ausnahme der polizeiauffälligen Personen für gegeben. Inwieweit hier Fehlzuordnungen vorliegen, also beispielsweise Dealer als Konsumenten identifiziert wurden, muss offen bleiben. Die eindeutige Identifizierbarkeit identischer Fälle ist auch bei der hier vorgenommenen Kodierung über den HIV-Kode nicht vollständig gewährleistet. Allerdings muss auch bei einem scheinbar guten Kode mit Fehlern gerechnet werden (Uhl & Seidler, 2000). Die getrennten Schätzungen für die Jahre 2010 und 2011 führen gegenüber einer einzigen Schätzung über einen Beobachtungszeitraum von zwei Jahren zu geringeren Verzerrungen durch mögliche Fluktuation in der Population (Zuzug, Abwanderung etc.). Auf eine Schätzung über den gesamten Zwei-Jahres-Zeitraum wurde daher verzichtet.

Die Möglichkeiten, Verletzungen der Annahme der Homogenität, zu kompensieren sind in allen Capture-recapture Studien begrenzt. Die Wahrscheinlichkeit, Einrichtungen der Drogenhilfe aufzusuchen, ist für Schwerstabhängige sicherlich höher als bei Personen mit einer

weniger ausgeprägten Suchtproblematik. Zur Erhöhung der Homogenität wurden die Schätzungen stratifiziert nach Altersklassen und Geschlecht vorgenommen.

Verletzungen der Annahme der Unabhängigkeit wurden im Falle der log-linearen Regression durch Einbeziehung von Interaktionseffekten und der Wahl des korrekten Modells statistisch weitgehend kompensiert. Unsere Schätzungen beruhen auf den statistisch besten Modellen, was nicht gleichbedeutend ist mit der Frage, ob die ausgewählten Modelle die Realität widerspiegeln. Die vorgenommen Schätzungen aller möglichen Modelle variieren deutlich. Das Ausmaß einer Fehlschätzung durch eine falsche Modellauswahl darf nicht unterschätzt werden. Zur Verminderung von Verletzungen der Unabhängigkeitsannahme wurden Personen, die im Beobachtungsjahr an den Folgen des Drogenkonsums verstarben, nicht berücksichtigt, weil im Todesfall die Wahrscheinlichkeit, später noch in einer der anderen Datenquellen-registriert zu werden, gleich null und somit nicht unabhängig von diesem Ereignis ist. Zur Vermeidung dieses Fehlers wurden die Drogentodesfälle des jeweiligen Folgejahres herangezogen. Diese Vorgangsweise verlängert allerdings den Gesamtbeobachtungszeitraum, was wiederum die Geschlossenheit der Population verringert. Beispielsweise kann ein Drogentoter im Vorjahr, d.h. zum Zeitpunkt, auf den sich die Schätzung bezieht, noch nicht im Untersuchungsgebiet gewohnt haben und erst später zugezogen sein. Im log-linearen Modell haben wir deshalb auf die Datenquelle „Drogentod“ verzichtet.

Zusammenfassend ist davon auszugehen, dass sich die Prävalenz der i.v. Opioidabhängigkeit seit den 1980er Jahren leicht erhöht hat. Obwohl keine Vergleichszahlen vorliegen, weisen die aktuellen Schätzungen zum problematischen Kokain- und Stimulanzienkonsum auf eine Diversifizierung des Drogenmarktes in Berlin vergleichbar mit anderen europäischen Großstädten hin. Damit ist trotz einer hohen Überschneidung mit dem Opioidkonsum (polyvalenter Konsum) von einer insgesamt höheren Prävalenz drogenbezogener Störungen in Berlin auszugehen. Hinweise zur Stagnation der Prävalenz bzw. des Rückgangs der Inzidenz der Opioidabhängigkeit finden sich in verschiedenen europäischen Städten (de Angelis, Hickman & Yang, 2004; Nordt & Stohler, 2006; Sanchez-Niubo et al., 2009). Während die Inzidenz zurückgeht, nimmt die Prävalenz aufgrund des flächendeckenden Ausbaus der Substitution mit der Folge einer Überlebenssicherung nur langfristig ab.

6 Danksagung

Die Studie wurde von der Senatsverwaltung für Gesundheit und Soziales finanziell gefördert. Für die Unterstützung bei der Datenerhebung möchten sich die Autoren recht herzlich bedanken bei Herrn Prof. Dr. Peter Bräunig, Frau Prof. Dr. Katrin Rothkopf und Herrn Dr. Sven-Olaf Görlich (Vivantes - Netzwerk für Gesundheit GmbH); Herrn Michael Kallin, Herrn Stephan de Reese, Herrn Ragnar Schlage und Frau Annette George (Landeskriminalamt Berlin, Dezernat Rauschgiftdelikte); Herrn Dieter Oremus (Sozialpädagogisches Institut Berlin (SPI)) und Herrn Michael Strobl (Gesellschaft für Standarddokumentation und Datenauswertung). Für die Unterstützung bei der Auswertung der Daten (log-lineare Regression) mit GLIM gilt unser Dank Herrn Dr. Gordon Hay und Frau Jane Oyston (The Centre for Public Health, Liverpool John Moores University).

7 Literatur

Arendt, M., Munk-Jorgensen, P., Sher, L. & Jensen, S. O. (2011). Mortality among individuals with cannabis, cocaine, amphetamine, MDMA, and opioid use disorders: a nationwide follow-up study of Danish substance users in treatment. *Drug and Alcohol Dependence, 114* (2-3), 134-139.

Babor, T. F., Caulkins, J. P., Edwards, G., Fischer, B., Foxcroft, D. R., Humphreys, D., Obot, I. S., Rehm, J., Reuter, P., Room, R., Rossow, I. & Strang, J. (2010). *Drug policy and the public good*. Oxford: Oxford University Press.

Bernard, C. (2013). Leben und Alltag in der Frankfurter Straßen-Drogenszene. In *Frauen in Drogenszenen* (S. 139-265). Wiesbaden: Springer Fachmedien.

Coffin, P. O., Galea, S., Ahern, J., Leon, A. C., Vlahov, D. & Tardiff, K. (2003). Opiates, cocaine and alcohol combinations in accidental drug overdose deaths in New York City, 1990-98. *Addiction, 98* (6), 739-747.

Darke, S. & Hall, W. (1995). Levels and correlates of polydrug use among heroin users and regular amphetamine users. *Drug and Alcohol Dependence, 39* (3), 231-235.

Darke, S. & Ross, J. (2000). Fatal heroin overdoses resulting from non-injecting routes of administration, NSW, Australia, 1992-1996. *Addiction, 95* (4), 569-573.

de Angelis, D., Hickman, M. & Yang, S. (2004). Estimating long-term trends in the incidence and prevalence of opiate use/injecting drug use and the number of former users: Back-calculation methods and opiate overdose deaths. *American Journal of Epidemiology, 160* (10), 994-1004.

Degenhardt, L. & Hall, W. (2012). Extent of illicit drug use and dependence, and their contribution to the global burden of disease. *The Lancet, 379* (9810), 55-70.

Fienberg, S. E. (1972). Multiple-recapture census for closed populations and incomplete contingency tables. *Biometrika, 59*, 591-603.

Groenemeyer, A. (2012). Drogen, Drogenkonsum und Drogenabhängigkeit. In G. Albrecht & A. Groenemeyer (Hrsg.), *Handbuch soziale Probleme* (S. 433-493). Wiesbaden: Springer VS.

Jones, H. E., Hickman, M., Welton, N. J., De, A. D., Harris, R. J. & Ades, A. E. (2014). Recapture or precapture? Fallibility of standard capture-recapture methods in the presence of referrals between sources. *Am J Epidemiol, 179* (11), 1383-1393.

Kirschner, W., Kunert, M., Grünbeck, P., Markert, St. & Tiemann, F. (1994). *Umfang und Struktur der Population i. v. Drogenabhängiger in Berlin*. Berlin: AIDS-Zentrum im Bundesgesundheitsamt.

Korf, D. J. (2004). Die Stadt im Blickpunkt der Drogenforschung. *Sucht, 50* (1), 8-11.

Laux, G. & Möller, H. J. (2011). *Memorix Psychiatrie und Psychotherapie* (2. Auflage). Stuttgart: Georg Thieme Verlag.

Leri, F., Bruneau, J. & Stewart, J. (2003). Understanding polydrug use: review of heroin and cocaine co-use. *Addiction, 98* (1), 7-22.

Nordt, C. & Stohler, R. (2006). Incidence of heroin use in Zurich, Switzerland: a treatment case register analysis. *Lancet, 367* (9525), 1830-1834.

Nutt, D. J., King, L. A. & Phillips, L. D. (2010). Drug harms in the UK: A multicriteria decision analysis. *Lancet, 376* (9752), 1558-1565.

Pabst, A., Kraus, L., Gomes de Matos, E. & Piontek, D. (2013). Substanzkonsum und substanzbezogene Störungen in Deutschland im Jahr 2012. *Sucht, 59* (6), 321-331; Open Access: http://www.psycontent.com/content/946m87w0g4812vu8/fulltext.pdf.

Pfeiffer-Gerschel, T., Steppan, M. & Brand, H. (2013). *Deutsche Suchthilfestatistik 2012. Alle Bundesländer. Tabellenband für ambulante Beratungs- und/oder Behandlungsstellen, Fachambulanzen und Institutsambulanzen (Typ 3 und 4). Bezugsgruppe: 10 ZB - irgendeine substanzbezogene HD Abhängigkeit* (Verfügbar unter: www.suchthilfe statistik.de [letzter Zugriff: 05.08.2014]). München: IFT Institut für Therapieforschung.

Sanchez-Niubo, A., Fortiana, J., Barrio, G., Suelves, J. M., Correa, J. F. & Domingo-Salvany, A. (2009). Problematic heroin use incidence trends in Spain. *Addiction, 104* (2), 248-255.

Schäfer, M. (2011). Opioide. In P. H. Tonner & L. Hein (Hrsg.), *Pharmakotherapie in der Anästhesie und Intensivmedizin* (S. 109-130). Berlin: Springer.

Skarabis, H. & Patzak, M. (1981). *Die Berliner Heroinszene*. Weinheim: Beltz.

Statistisches Bundesamt (2014). *Fortschreibung des Bevölkerungsstandes*. Online verfügbar unter: https://www-genesis.destatis.de/genesis/online/data;jsessionid=8693CF192476E208F6687C27B48E1092.tomcat_GO_2_1?operation=abruftabelleAbrufen&selectionname=12411-0012&levelindex=1&levelid=1407248817024&index=12 [letzter Zugriff: 05.08. 2014].

Steppan, M. (2013). *Multiple Recapture Formula. Capture-Recapture Calculations for more than two samples*. München: IFT Institut für Therapieforschung.

Strathdee, S. A., Hallett, T. B., Bobrova, N., Rhodes, T., Booth, R., Abdool, R. & Hankins, C. A. (2010). HIV and risk environment for injecting drug users: the past, present, and future. *The Lancet, 376* (9737), 268-284.

Tretter, F. & Kraus, L. (2004). Stadtspezifische Prävalenz des Drogenkonsums und ihre Ursachen [Editorial]. *Sucht, 50* (1), 5-7.

Wittes, J. T. (1974). Applications of a multinomial capture-recapture model to epidemiological data. *Journal of the American Statistical Association, 69* (345), 93-97.

Anhang I

8 Schätzverfahren 45

8.1 Capture-recapture Methode 45

8.2 Probabilistisches Modell 47

9 Tabellen und Abbildungen der Schätzungen für das Jahr 2011 51

Anhang I

8 Schätzverfahren

8.1 Capture-recapture Methode

Eine detailierte Beschreibung der Capture-recapture Methode und der log-linearen Regression findet sich bei Hay und McKegany (European Monitoring Centre for Drugs and Drug Addiction, 1999).

Die Fälle, die in den zur Analyse verwendeten Datenquellen registriert bzw. nicht registriert wurden (Fall-Überschneidung), lassen sich in einer Kontingenztabelle darstellen. Tabelle 1 zeigt die Überschneidungsbereiche zwischen den Datenquellen am Beispiel von drei Datenquellen:

Tabelle A: Kontingenztabelle zur Bestimmung der Überschneidungen zwischen den Datenquellen S1, S2 und S3: registriert (ja), nicht registriert (nein).

		S1			
		ja		nein	
		S2			
		ja	nein	ja	nein
S3	ja	*a*	*b*	*e*	*f*
	nein	*c*	*d*	*g*	*x*

Diejenigen Fälle, die in allen drei Datenquellen registriert wurden, sind mit *a* gekennzeichnet, *c* bezeichnet alle Fälle, die in den Datenquellen 1 und 2, jedoch nicht in Datenquelle 3 vorkommen. Fälle, die nur in Datenquelle 3 registriert wurden, sind mir *f* gekennzeichnet. Die unbekannte Population, d.h. diejenige Fälle, die das Ziel der Capture-recapture Schätzung sind, sind mit *x* gekennzeichnet.

Die Anzahl von Fällen in jeder Datenquelle (N1, N2, N3) lassen sich berechnen mit

$$N1 = a + b + c + d$$

$$N2 = a + c + e + g$$

$$N3 = a + b + e + f$$

Die Gesamtzahl der beobachteten Population entspricht:

$$N = a + b + c + d + e + f + g$$

N entspricht der Gesamtzahl von Fällen, die in einem Capture-recapture Verfahren identifiziert werden.

Die Daten können alternativ in einem Venn Diagramm dargestellt werden (Abbildung 1):

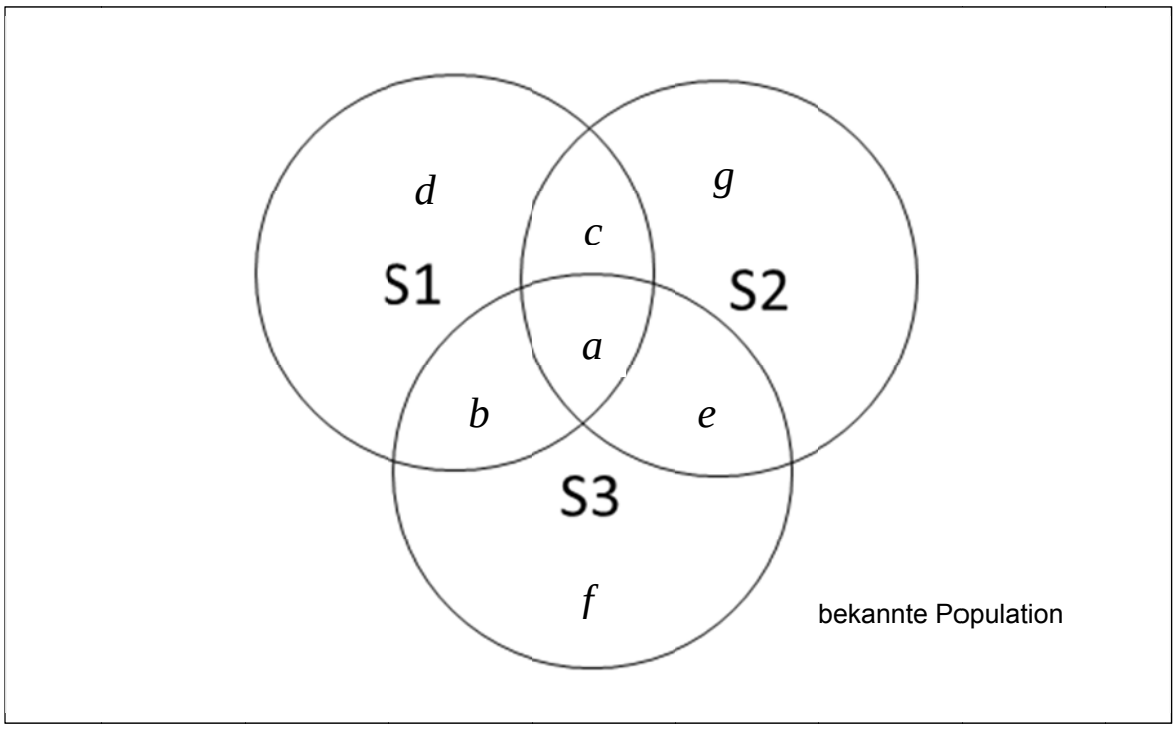


Abbildung A: Venn Diagramm zur Darstellung der Datenstruktur der Capture-recapture Methode am Beispiel von drei Datenquellen (S1, S2 und S3).

Log-lineares Modell

Die in Tabelle A und Abbildung A dargestellten Überschneidungen einer multiplen Datenregistrierung (am Beispiel von drei Datenquellen) können in einem log-linearen Modell zur Schätzung der unbekannten Population x verwendet werden. Zur Einführung in log-lineare Modelle sei auf Bishop, Fienberg und Holland (1975) verwiesen. Unter Einbeziehung verschiedener Interaktionen zwischen den Datenquellen werden mehrere log-lineare Modelle an die Daten angepasst. Ein Beispiel für eine Interaktion zwischen den Datenquellen ist eine höhere Vermittlungswahrscheinlichkeit in eine stationäre Rehabilitation von Personen, die in einer ambulanten Beratung um Hilfe aufsuchen. Entscheidungen über das Vorhandensein von Interaktionen lassen sich auf der Basis des Devianz Maßes G^2 vornehmen. Mit G^2 wird die Güte der Modellanpassung vergleichbar mit einem χ^2 Test zur Messung des Zusammenhangs in Kontingenztabellen bezeichnet. Da im vorliegenden Fall nur sieben statt acht Zellen vorhanden sind, lassen sich nur zwei-fach Interaktion schätzen.

Die dreifach Interaktion, die im Modell nicht geschätzt werden kann, wird Null gesetzt. Die Validität dieser Annahme spiegelt sich im Devianz Maß zur Auswahl der besten Modellanpassung. Eine große Abweichung des Modells mit der besten Modellanpassung bedeutet, dass die Schätzung mit einem Bias versehen ist. Nach Modellanpassung können Konfidenzintervalle geschätzt werden.

Modellauswahl

Eine Diskussion der Methoden zur Modellauswahl und zur Beurteilung der Modellunsicherheit findet sich bei Hook und Regal (1997). Neben dem Akaike Informations-Kriterium (AIC) (Akaike, 1985) zur Beurteilung ob und welche Interaktionen ins Modell aufgenommen werden sollen, wird häufig auch das Bayes Informations- Kriterium (BIC) verwendet.

Die Auswahl des Modells, das die Datenstruktur am besten beschreibt ist inhaltlich verknüpft mit der Wahl der Datenquellen und der Falldefinition. Zudem findet das Kriterium der Modellsparsamkeit Anwendung im dem Sinne, dass das sparsamere Modelle gegenüber komplexeren Modellen bevorzugt werden. Aufgrund leerer Zellen ist es nicht immer möglich, bestimmte Modelle an stratifizierte Daten (Alter und Geschlecht, Region etc.) anzupassen. Neben diesem strukturellen Problem „keiner Überschneidung" zwischen Datenquellen, sind auch artifizielle Überschneidungen zu berücksichtigen, d.h. die Registrierung in einer Datenquelle führt notgedrungen , also „automatisch", zur Registrierung in einer anderen Datenquelle.

Der Vorteil der multiplen Capture-recapture Methode besteht auch darin, dass bestimmte Interaktionen ins Modell aufgenommen werden können, auch wenn dies auf der Grundlage des Akaike Informations-Kriterium (AIC) nicht empfohlen wird. Interaktionen, die auf Expertenwissen beruhen, wie z.B. eine erwartete Abhängigkeit zwischen Einrichtungen der Suchthilfe und eine erwartete Unabhängigkeit zwischen „Suchthilfe" und „Strafverfolgung" können aus rationalen Überlegungen ins Modell aufgenommen werden.

Konfidenzintervalle

Zur Schätzung der Konfidenzintervalle werden Methoden bevorzugt, die berücksichtigen, dass die „unbekannt Population" auf der Basis von Daten mit einer asymmetrischen Verteilung geschätzt wurde (Cormack, 1992; Regal & Hook, 1984). Damit wird vermieden, dass das untere Intervall der Gesamtschätzung niedriger ist als die Anzahl der bekannten Fälle. Die geschätzten Obergrenzen können jedoch unplausible Werte annehmen, insbesondere wenn mehrere Interaktionen ins Modell aufgenommen werden.

Referenzen

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716-723.

Bishop, Y., Fienberg, S. & Holland, P. (1975) *Discrete multivariate analysis*. MIT Press, Cambridge MA.

Cormack, R. M. (1992) Interval estimation for mark-recapture studies of closed populations. *Biometrics, 48*, 567-576.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (1999). *Methodological guidelines to estimate the prevalence of Problem Drug Use on the local level*. Lisbon: EMCDDA.

Hook, E. B. & Regal, R. R. (1997). Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *American Journal of Epidemiology, 145*, 1138-1144.

Regal, R. R. & Hook, E. B. (1984) Goodness-of-fit based confidence intervals for estimates of the size of a closed population. *Statistics in Medicine, 3*, 287-291.

8.2 Probabilistische Methode

Eine detaillierte Beschreibung der Probabilistischen Methode (Probabilistic Multiple Recapture Method) findet sich bei Steppan (2013). Ausgehend von der bekannten Formel der Capture-recapture Methode zur Schätzung einer unbekannten Population mit zwei Datenquellen (Formel 1). Die Gesamtzahl der Population $\widehat{N}$ kann aus der Anzahl der Fälle in der ersten Datenquelle n_1 (capture), der Anzahl der Fälle in der zweiten Datenquelle n_2 (recapture) und der Anzahl der Fälle *x* (overlap), die in beiden Datenquellen vorkommen, geschätzt werde. Im Fall von mehr als zwei Datenquellen (k>2) werden in der Regel log-lineare Modelle angewendet (Tilling & Sterne, 1999).

$$\widehat{N} = \frac{n_1 \cdot n_2}{X} \qquad \text{[Formel 1]}$$

Für die Lösung von Capture-recapture Situationen mit mehr als 2 Datenquellen werden normalerweise aufwändige log-lineare Modelle angewendet, die das Problem statistisch lösen. Formel 2 stellt eine Verallgemeinerung von Formel 1 und damit eine mathematische Lösung für den Fall k>2 dar. Grundlage ist die bekannte Anzahl von Fällen der zu schätzenden Population in mehr als zwei Datenquellen, wobei $\prod_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_i}$ das Produkt der über alle Stichproben und *k* die Anzahl der Stichproben repräsentieren. Die Gesamtpopulation $\widehat{N}$ wird durch Division des Produkts durch die Anzahl von Fällen, die in allen *k* Stichproben vertreten sind (Overlap), und anschließender Ziehung der (k-1)-ten Wurzel des Produkts geschätzt.

$$\widehat{N} = \sqrt[k-1]{\frac{\prod_{i=1}^{k} n_i}{X}} \qquad \text{[Formel 2]}$$

Formel 3 stellt eine Approximation der Standardabweichung der Schätzung dar. Die Formel basiert auf der allgemeinen Formel für $\sigma = \sqrt{n \cdot p \cdot (1-p)}$, wobei p mit X/n ersetzt wird.

$$\sigma_{\widehat{N}} \approx \sqrt[k-1]{\frac{\prod_{i=1}^{k} n_i}{X - \sqrt{x - \frac{x^2}{\widehat{N}}}}} - \widehat{N} \qquad \text{[Formula 3]}$$

Mit Hilfe von Formel 3 lassen sich die 95%-Konfidenzintervalle approximativ berechnen (Formel 4).

$$KI_{95\%} = \left[\sqrt[k-1]{\frac{\prod_{i=1}^{k} n_i}{X + 1{,}96 \cdot \sqrt{x - \frac{x^2}{\widehat{N}}}}} ; \sqrt[k-1]{\frac{\prod_{i=1}^{k} n_i}{X - 1{,}96 \cdot \sqrt{x - \frac{x^2}{\widehat{N}}}}} \right] \qquad \text{[Formel 4]}$$

Referenzen

Tilling, K., & Sterne, J. A. (1999). Capture-recapture models including covariate effects. *American Journal of Epidemiology, 149* (4), 392-400.

Steppan, M. (2013). *Multiple Recapture Formula. Capture-recapture calculations for more than two samples*. München: IFT Institut für Therapieforschung.

9 Tabellen und Abbildungen der Schätzungen für das Jahr 2011

Tabelle 1a. Stichprobengrößen nach Datenquelle, Datenjahr, Geschlecht und Substanz (Datenjahr 2011)

		Substanz			Geschlecht			
	Datenjahr	Opioide	Kokain	Stimu-lanzien	männlich	weiblich	Gesamt	i.v.**
Entgiftung	2011	667	275	221	680	232	912	99
Landeskriminalamt (LKA)	2011	747	585	684	1.415	258	1.673	183
Ambulante Einrichtungen	2011	5.497	3.185	2.179	5.829	1.789	7.618	2.330
Stationäre Einrichtungen	2011	1.012	935	923	1.389	329	1.718	355
Drogentodesfälle	2012*	80	32	21	80	24	104	28

Anmerkung. Gesamtzahl als Fallzahlen (N).
* Drogentodesfälle wurden jeweils ein Jahr später erfasst, da dieses Ereignis die Wahrscheinlichkeit für die Registrierung in einer anderen Datenquelle beeinflusst.
** i.v.=intravenöser Konsum (Daten liegen nur für die Einrichtungen der Suchthilfe vor).

Tabelle 2a. Stichprobengrößen nach Altersgruppe, Substanz, Geschlecht und i.v. Konsum (Datenjahr 2011)

	Substanz			**Geschlecht**			
	Opioide	Kokain	Stimu-lanzien	männlich	weiblich	Gesamt	i.v.*
15-24 Jahre	898	678	1.003	1.387	478	1.865	182
25-34 Jahre	2.449	1.760	1.526	3.089	912	4.001	893
35-64 Jahre	3.388	1.647	758	3.486	936	4.422	1.393

Tabelle 3a. Populationsschätzung anhand log-linearer Regression nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2011)

OPIOIDE

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	872	508	1.820	S2xS3, S2xS4	176.483	4,9
weiblich	25 bis 34	1.681	1.217	2.495	S2xS4	265.858	6,3
weiblich	35 bis 64	1.461	1.223	1.793		691.164	2,1
weiblich	15 bis 64	4.864	3.962	6.124	S1xS2, S2xS3	1.133.505	4,3
männlich	15 bis 24	1.698	1.282	2.382	S1xS2, S2xS3, S1xS3, S2xS4	169.984	10,0
männlich	25 bis 34	4.602	3.904	5.529	S1xS2, S2xS3	261.966	17,6
männlich	35 bis 64	8.551	7.132	10.447	S1xS2, S2xS3	697.456	12,3
männlich	15 bis 64	14.459	12.880	16.362	S1xS2, S2xS3, S1xS3, S1xS4, S2xS4	1.129.406	12,8
beide	15 bis 24	2526	1.951	3.420	S1xS2, S2xS3, S1xS3, S2xS4	346.467	7,3
beide	25 bis 34	6.244	5.379	7.357	S1xS2, S2xS3, S1xS4	527.824	11,8
beide	35 bis 64	10.830	9.278	12.815	S1xS2, S2xS3	1.388.620	7,8
GESAMT	15 bis 64	19.317	17.448	21.518	S1xS2, S2xS3, S1xS3, S1xS4, S2xS4	2.262.911	8,5

KOKAIN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	288	211	435		176.483	1,6
weiblich	25 bis 34	912	659	1.350	S1xS2	265.858	3,4
weiblich	35 bis 64	570	445	765		691.164	0,8
weiblich	15 bis 64	1.759	1.466	2158	S1xS2	1.133.505	1,6
männlich	15 bis 24	1.145	929	1.460	S1xS2	169.984	6,7
männlich	25 bis 34	2.954	2.620	3.369	S1xS2, S1xS3	261.966	11,3
männlich	35 bis 64	3.987	3.232	5.057	S1xS2	697.456	5,7
männlich	15 bis 64	8.115	7.302	9.084	S1xS2, S2xS3, S2xS4	1.129.406	7,2
beide	15 bis 24	1.446	1.201	1.790	S1xS2	346.467	4,2
beide	25 bis 34	3.788	3.375	4.295	S1xS2, S2xS3, S2xS4	527.824	7,2
beide	35 bis 64	4.728	3.923	5.821	S2xS3, S2xS4	1.388.620	3,4
GESAMT	15 bis 64	10.000	9.076	11.085	S1xS2, S2xS3, S2xS4	2.262.911	4,4

Tabelle 3a (Fortsetzung). Populationsschätzung anhand log-linearer Regression nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2011)

STIMULANZIEN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	1.019	684	1.658		176.483	5,8
weiblich	25 bis 34	1.184	774	2.006	S1xS2	265.858	4,5
weiblich	35 bis 64	213	169	284		691.164	0,3
weiblich	15 bis 64	2.793	2.018	4.100	S1xS2	1.133.505	2,5
männlich	15 bis 24	2.201	1.751	2.866	S1xS2	169.984	12,9
männlich	25 bis 34	2.744	2.404	3.174	S1xS2, S1xS3	261.966	10,5
männlich	35 bis 64	1.744	1.418	2.203	S1xS2	697.456	2,5
männlich	15 bis 64	6.642	5.976	7.436	S1xS2, S2xS3, S2xS4	1.129.406	5,9
beide	15 bis 24	3.817	3.067	4.872	S1xS2	346.467	11,0
beide	25 bis 34	3.654	3.209	4.212	S1xS2, S1xS3	527.824	6,9
beide	35 bis 64	2.641	2.028	3.580	S2xS3, S2xS4	1.388.620	1,9
GESAMT	15 bis 64	8.830	7.993	9.816	S1xS2, S2xS3, S1xS3, S2xS4	2.262.911	3,9

ALLE SUBSTANZEN

Geschlecht	Alter	Schätzung	KI 2.5%	KI 97.5%	Modell	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	2.248	1.388	4.171	S1xS2	176.483	12,7
weiblich	25 bis 34	3.564	2.548	5.319	S1xS2	265.858	13,4
weiblich	35 bis 64	2.675	2.183	3.361	S1xS2	691.164	3,9
weiblich	15 bis 64	8.474	6.950	10.558	S1xS2, S2xS3, S2xS4	1.133.505	7,5
männlich	15 bis 24	4.347	3.699	5.184	S1xS2, S2xS3, S2xS4	169.984	25,6
männlich	25 bis 34	7.947	7.244	8.765	S1xS2, S2xS3, S2xS4	261.966	30,3
männlich	35 bis 64	12.957	10.899	15.653	S1xS2	697.456	18,6
männlich	15 bis 64	27.026	24.272	30.286	S1xS2, S2xS3, S2xS4	1.129.406	23,9
beide	15 bis 24	6.742	5.370	8.730	S1xS2, S2xS3, S2xS4	346.467	19,5
beide	25 bis 34	13.340	11.552	15.590	S1xS2, S2xS3, S1xS4, S2xS4	527.824	25,3
beide	35 bis 64	15.770	13.612	18.485	S1xS2, S2xS3, S2xS4	1.388.620	11,4
GESAMT	15 bis 64	35.508	32.278	39.254	S1xS2, S2xS3, S2xS4	2.262.911	15,7

Anmerkung. Schätzungen gerundet auf ganze Zahlen. Konfidenzintervall entspricht den mittleren 95% und liegt asymmetrisch um die zentrale Schätzung (links näher als rechts). Modell enthält die dargestellten Interaktionen zwischen Datenquellen (S1=ambulant, S2=ambulant, S3=Polizei, S4=Entgiftung). Bevölkerungszahlen der jeweiligen Altersgruppe gemäß des Bevölkerungsstandes Berlin zum Stichtag 31.12.2011 (Statistisches Bundesamt, 2014). Prävalenz p.T. („pro tausend Einwohner“) der jeweiligen Altersgruppe.

Tabelle 4a. Populationsschätzung anhand der probabilistischen Methode nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2011)

OPIOIDE

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prä-valenz p.T.
weiblich	15 bis 24	373	373	>239	3.312	176.483	2,1
weiblich	25 bis 34	1.054	1.054	>607	7.060	265.858	4,0
weiblich	35 bis 64	1.323	1.323	>766	7.764	691.164	1,9
weiblich	15 bis 64	2.866	2.866	>1.612	4.288	1.133.505	2,5
männlich	15 bis 24	1.053	1.053	>654	1.968	169.984	6,2
männlich	25 bis 34	2.847	2.847	>1.832	8.355	261.966	10,9
männlich	35 bis 64	5.055	5.055	>2.594	12.307	697.456	7,2
männlich	15 bis 64	9.273	9.273	>5.080	16.777	1.129.406	8,2
beide	15 bis 24	1.453	1.453	2.060	5.490	346.467	4,2
beide	25 bis 34	4.057	4.057	>2.439	6.880	527.824	7,7
beide	35 bis 64	6.521	6.521	4.854	16.034	1.388.620	4,7
GESAMT	15 bis 64	11.919	11.919	>6.692	18.771	2.262.911	5,3

KOKAIN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prä-valenz p.T.
weiblich	15 bis 24	174	174	>155	1.322	176.483	1,0
weiblich	25 bis 34	600	600	>318	7.404	265.858	2,3
weiblich	35 bis 64	448	448	>270	1.201	691.164	0,6
weiblich	15 bis 64	1.374	1.374	>742	2.363	1.133.505	1,2
männlich	15 bis 24	912	912	>521	1.807	169.984	5,4
männlich	25 bis 34	2.305	2.305	>1.438	6.221	261.966	8,8
männlich	35 bis 64	2.565	2.565	>1.366	9.852	697.456	3,7
männlich	15 bis 64	5.847	5.847	>3.325	9.732	1.129.406	5,2
beide	15 bis 24	1.040	1.040	>676	3.905	346.467	3,0
beide	25 bis 34	2.821	2.821	>1.756	7.329	527.824	5,3
beide	35 bis 64	3.356	3.356	>1.635	6.856	1.388.620	2,4
GESAMT	15 bis 64	7.120	7.120	>4.067	11.289	2.262.911	3,1

Tabelle 4a (Fortsetzung). Populationsschätzung anhand der probabilistischen Methode nach Substanz, Geschlecht und Altersgruppen (Datenjahr 2011)

STIMULANZIEN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	582	582	>262	8.039	176.483	3,3
weiblich	25 bis 34	857	857	>300	7.502	265.858	3,2
weiblich	35 bis 64	235	235	>127	565	691.164	0,3
weiblich	15 bis 64	2.100	2.100	>689	5.716	1.133.505	1,9
männlich	15 bis 24	1.859	1.859	>739	3.090	169.984	10,9
männlich	25 bis 34	2.874	2.874	1256	4.493	261.966	11,0
männlich	35 bis 64	1.288	1.288	>625	12.923	697.456	1,8
männlich	15 bis 64	6.183	6.183	>2.583	10.318	1.129.406	5,5
beide	15 bis 24	2.801	2.801	>1.001	11.306	346.467	8,1
beide	25 bis 34	3.673	3.673	1.837	5.508	527.824	7,0
beide	35 bis 64	1.836	1.836	>752	4.463	1.388.620	1,3
GESAMT	15 bis 64	8.299	8.299	>3.272	14.312	2.262.911	3,7

ALLE SUBSTANZEN

Geschlecht	Alter	Schätzung	Schätzung (korr.)*	KI 2.5%	KI 97.5%	Bevölkerung	Prävalenz p.T.
weiblich	15 bis 24	1.297	1.297	>476	9.621	176.483	18,4
weiblich	25 bis 34	2.415	2.415	>903	7.860	265.858	25,4
weiblich	35 bis 64	1.967	1.967	>932	3.235	691.164	12,2
weiblich	15 bis 64	6.357	6.357	3.450	9.264	1.133.505	16,5
männlich	15 bis 24	3.256	3.256	>1.381	5.853	169.984	7,6
männlich	25 bis 34	6.763	6.763	>3.079	12.528	261.966	9,2
männlich	35 bis 64	8.423	8.423	3.752	13.094	697.456	2,8
männlich	15 bis 64	18.706	18.706	>7.911	30.572	1.129.406	5,6
beide	15 bis 24	4.663	4.663	>1.857	8.064	346.467	13,5
beide	25 bis 34	9.154	9.154	>3.982	15.823	527.824	17,3
beide	35 bis 64	10.705	10.705	>4.383	18.164	1.388.620	7,7
GESAMT	15 bis 64	24.511	24.511	13.166	35.855	2.262.911	10,8

Anmerkung. Schätzungen gerundet auf ganze Zahlen. Konfidenzintervall entspricht den mittleren 95% und liegt teilweise asymmetrisch um die zentrale Schätzung. Bevölkerungszahlen der jeweiligen Altersgruppe gemäß des Bevölkerungsstandes Berlin zum Stichtag 31.12. 2011 (Statistisches Bundesamt, 2014). Prävalenz p.T. („pro tausend Einwohner“) der jeweiligen Altersgruppe.
* Sofern die Schätzung unterhalb der registrierten Anzahl von Personen („Hellfeld“) liegt, wurde diese Zahl als das korrigierte Schätzminimum und als unteres Minimum des Konfidenzintervalls angenommen.

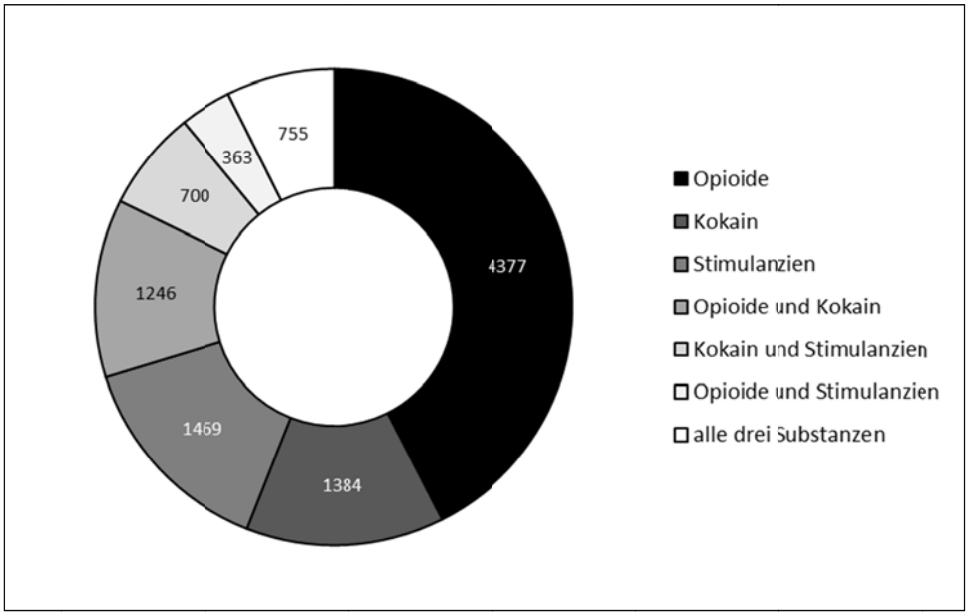


Abbildung 1a. Substanzkonsum der registrierten Drogenkonsumenten (Datenjahr 2011).

Anhang II

A Antrag auf datenschutzrechtliche Beurteilung des Forschungsvorhabens

B Berliner Beauftragter für Datenschutz und Informationsfreiheit: Datenschutzrechtliche Beurteilung

C Der Diözesandatenschutzbeauftragte

D Technische Beschreibung Dateneingabe und Verschlüsselung

Anhang A

Antrag auf datenschutzrechtliche Beurteilung des Forschungsvorhabens

ANTRAG AUF DURCHFÜHRUNG DES FORSCHUNGSVORHABENS

Capture-Recapture-Studie zur Schätzung des Umfangs und der Struktur der Drogenszene in Berlin

1. Angaben zu Rahmenbedingungen des Vorhabens

Das Projekt *Capture-Recapture-Studie zur Schätzung des Umfangs und der Struktur der Drogenszene in Berlin* wird vom Land Berlin vertreten durch die Senatsverwaltung für Gesundheit und Soziales mit einer Laufzeit vom 15. November 2012 bis 31. Juli 2013 finanziell gefördert.

Zentrales Ziel des Forschungsvorhabens ist es, mit Hilfe eines Capture-Recapture-Verfahrens die Anzahl der Opiat-, Kokain- und Amphetaminkonsumenten bzw. der problematischen Drogenkonsumenten in Berlin zu schätzen und Aussagen zu deren Struktur und Problematik zu treffen. Zu diesem Zweck wird auf Grundlage der Erfassung und Wiedererfassung von Personen (Capture-Recapture) in verschiedenen Einrichtungen der Suchthilfe oder Behörden (z.B. ambulante Beratung, Substitution, Polizei, Staatsanwaltschaft etc.) die Anzahl der Personen statistisch geschätzt, die in keiner der verwendeten Quellen („hidden population“) erfasst wurden.

2. Angaben zum Gegenstand und zum Verfahren des Vorhabens

<u>Vorhaben</u>

Konkrete Ziele der Studie sind (1) die Schätzung des Umfangs von Drogenkonsumenten in Berlin, (2) eine Differenzierung nach verschiedenen Substanzen (Opiate, Kokain, Amphetamine) und Konsumformen (Missbrauch, Abhängigkeit, problematischer Konsum, intravenöser Konsum), (3) die Stratifizierung nach Altersgruppen und Geschlecht sowie (4) für Opiatkonsumenten ein Vergleich der Ergebnisse mit früheren Schätzungen (1979 und 1992).

<u>Methodik und Durchführung</u>

Zentraler Aspekt der Datenerhebung ist die Zusammenstellung der zu berücksichtigenden Einrichtungen bzw. Institutionen, die über geeignete Daten verfügen. Notwendig ist die Berücksichtigung der folgenden Informationen:

(1) Ambulante Suchtberatungsstellen
(2) Stationäre Entwöhnungskliniken
(3) Substitutionseinrichtungen bzw. Daten der Substituierten
(4) Entgiftungsstationen
(5) Daten zu Rauschgiftkriminalität und Drogentoten

Die Anwesenheit einer bestimmten Person in einer oder mehreren Quelle(n) wird über personenbezogene Codes identifiziert (Pseudonymisierung), die die Anonymität der erfassten Personen gewährleistet. Unter Verwendung von Informationen zu Geschlecht und Alter werden stratifizierte Schätzungen vorgenommen, so dass geschlechts- und altersgruppenspezifische Aussagen zum Umfang der Drogenabhängigkeit möglich sind. Als Kriterien zur Charakteristik der Zielgruppe sollen verschiedene Definitionen zur Anwendung kommen. Unter Verwendung der ICD-10-Diagnostik lassen sich klinisch relevante Fälle von Personen mit einem Drogenmissbrauch bzw. einer Drogenabhängigkeit (Opiate, Kokain, Amphetamine) schätzen. Zudem lassen die angestrebten Quellen die Identifikation von Fällen mit problematischem Drogenkonsum zu, der mit hoher Wahrscheinlichkeit zu einem Missbrauchs-, bzw. Abhängigkeitsverhalten führt, sowie die Identifikation von Personen, die Opiate intravenös konsumieren.

Einrichtungen bzw. Institutionen, die die entsprechenden Daten verwalten, werden kontaktiert und bei der Extraktion und Kodierung von Daten unterstützt. Darüber hinaus sind die erhobenen Daten der unterschiedlichen Quellen auf ihre Verwendbarkeit zu überprüfen und ggf. weitere Datenquellen zu erschließen. Nach Kodierung der Personen in den verschiedenen Datenquellen erfolgen die Auswertung und die Dokumentation der Ergebnisse.

Datenquellen

Für die einzuschließenden Daten wurden die folgenden Quellen und Datenhalter identifiziert. Name und Anschrift der Datenhalter finden sich in Anhang 1.

Nr.	**Daten**	**Quelle**	**Datenhalter**
1	Ambulante Angebote	Deutsche Suchthilfestatistik Berlin (ambulant)	Ambulante Suchtberatungsstellen und ambulante Entwöhnungseinrichtungen (siehe Anlage)
2	Stationäre Entwöhnung	Deutsche Suchthilfestatistik Berlin (stationär)	Stationäre Entwöhnungseinrichtungen (siehe Anlage)
3	Substitution	Substitutionsregister	Kassenärztliche Vereinigung Berlin
4	Entgiftung	Krankenhausdiagnosestatistik, Fallpauschalensystem	Entgiftungsstationen (siehe Anlage)
5	Rauschgiftkriminalität	Kriminalstatistik	Landeskriminalamt
6	Drogentote	Drogentodesstatistik	Landeskriminalamt

Einschlusskriterien

- Daten aus den Kalenderjahren 2010 und 2011 für Datenquellen 1 bis 5 und 2011 und 2012 für Datenquelle 6
- ICD-10-Diagnosen F11.- (Opiate), F14.- (Kokain), F15.- (Stimulanzien) für Datenquellen 1 bis 4; Opiat-, Kokain-, Amphetaminkonsum für Datenquellen 5 und 6
- Wohnort Berlin (ausschließlich für Datenquelle 2)

Durchführung der Studie

- Kontaktaufnahme mit den Datenhaltern
- Extraktion der relevanten Daten durch den Datenhalter
- Kodierung der einzelnen Datenquellen bei den Datenhaltern (Personen-Code, siehe Punkt 3)
- Datenprüfung
- Zusammenführung der einzelnen Datenquellen
- Statistische Analyse
- Dokumentation der Ergebnisse

Verwertung der Daten

- Projektbericht für das Land Berlin
- Ggf. Publikation der Ergebnisse

3. Angaben zu den erhobenen Daten und zum Datenschutz

Erhobene Daten

In jeder Datenquelle werden folgende Daten erhoben: Name, Vorname, Geschlecht, Alter, Diagnose, Jahr der Erfassung sowie ggf. zusätzlich verfügbare relevante Informationen wie z.B. intravenöser Konsum. Name und Vorname sind zur Bildung eines Pseudonyms notwendig; diese Informationen werden nicht an den Antragsteller weitergegeben.

Kodierung der Daten und Erstellung von Personen-Codes

Name, Vorname und andere Identifikationsmerkmale werden durch ein Kennzeichen pseudonymisiert zu dem Zweck, dass die Bestimmung einzelner Personen ausgeschlossen bzw. wesentlich erschwert wird. Das Pseudonymisierungsverfahren ist begrenzt durch die in den Datenquellen enthaltenen Informationen und die Pseudonyme müssen so viele Informationen enthalten, dass das wiederholte Auftreten eines Kennzeichens unwahrscheinlich ist. In der vorliegenden Studie wird ein HIV-Code verwendet, wie er beispielsweise auch im Rahmen der Basisdokumentation im Suchtbereich (BADO) im Hamburg verwendet wird. Dieser Code basiert auf Normalisierungen des Vor- und Nachnamens (z.B. Umwandlung von Umlauten und ß, Entfernen von Zusatzzeichen und Leerzeichen), der Länge des Vor- und Nachnamens, dem Geschlecht und dem Geburtsjahr. Fehlende Informationen werden dabei jeweils als „*“ eingetragen. Die genauen Bestandteile des HIV-Codes sind folgende:

(1) Dritter Buchstabe des „normalisierten Vornamens“
(2) Länge des „normalisierten Vornamens“
(3) Dritter Buchstabe des „normalisierten Nachnamens“

(4) Länge des „normalisierten Nachnamens“

(5) Geschlecht („M“ für männlich, „W“ für weiblich)

(6) Geburtsjahr (letzte zwei Ziffern des vierstelligen Geburtsjahres)

Um auszuschließen, dass für eine gegebene Person durch Rückgriff auf die im Verfahren gespeicherter Pseudonyme Aussagen über ihren Drogenkonsum abgeleitet werden können, um weiterhin zu verhindern, dass durch den Ausschluss der Existenz eines Datensatzes im Datenbestand auf eine Person rückgeschlossen werden kann, und um ferner auszuschließen, dass aus der Existenz eines "passenden" Pseudonyms die Möglichkeit für eine Person besteht, dass die Aussage, sie drogenabhängig, mit höherer Wahrscheinlichkeit zutrifft als für die allgemeine Population, werden weitere Maßnahmen ergriffen. Zu diesem Zweck wird sichergestellt, dass die ober beschriebenen Pseudonyme vor Übermittlung an den Projektträger kryptografisch transformiert werden und die Transformation von keiner anderen Stelle nachvollzogen werden kann.

Datenschutz

Die Kodierung der individuellen Datensätze und die Bildung der HIV-Codes geschieht bei den jeweiligen Datenhaltern, ggf. mit Unterstützung des Antragstellers. Die Unterstützung erfolgt in Form von Anweisungen für die Normalisierung von Klarnamen und die Bildung des HIV-Codes. Auf Wunsch des Datenhalters kann durch den Antragsteller auch ein (ausschließlich lokal operierendes) Programm zur Verfügung gestellt werden, dass diese beiden Schritte automatisiert vornimmt, vorausgesetzt es liegen die erforderlichen Datenfelder in einem standardisierten Dateiformat und in der korrekten Reihenfolge vor (CSV-Format). Die für die Pseudonymisierung erforderlichen Datenfelder sowie die an das IFT übermittelten Datenfelder sind in Anlage II dargestellt.

Die Kryptographierung der Daten erfolgt auf Basis eines durch den Antragsteller zur Verfügung gestellten lokal operierenden Programms (Visual Basic). Die Chiffrierung erfolgt mit Hilfe sogenannter mathematischer Einwegfunktionen. Zusätzlich ist die Chiffrierung durch einen willkürlichen 10-stelligen Code der Datenhalter und durch einen ebensolchen Code von Seiten des Antragstellers abgesichert, sodass selbst unter Kenntnis des Algorithmus immer beide Informationen benötigt werden, um die Information zu entschlüsseln. Nach Zusammenführung der Datensätze aller Datenhalter mit Hilfe der so erstellten ein-eindeutigen Pseudonyme, erfolgt eine weitere zufallsgenerierte Chiffrierung, sodass selbst unter Kenntnis aller Codes, und des Chiffrierungsalgorithmus eine Rückführung auf die ursprünglichen HIV-Codes nicht einmal mehr theoretisch möglich ist.

Die an den Antragsteller weitergegebenen Daten enthalten somit den Namen, Vornamen und Geburtsdatum der Person nicht mehr. Die getrennten Datensätze werden zum Zweck der Identifizierung der

Registrierung einer Personen in einer oder mehrerer Datenquellen zusammengeführt. Der kryptografisch transformierte pseudonymisierte vollständige Datensatz der Studie wird auf dem Server des IFT Institut für Therapieforschung gesichert und ist nicht öffentlich zugänglich.

Aufbewahrung

Die pseudonymisierten Daten werden am IFT mit einem Code gesichert aufbewahrt, der nur den am Projekt beteiligten Mitarbeitern bekannt ist. Sie werden nach Abschluss des Projektes, d.h. nach Abschluss des Projektberichts bzw. der vorgesehenen Publikationen, spätestens aber nach zwei Jahren vollständig gelöscht.

Anlage I

1 Ambulante Angebote

Ambulante Suchtberatungsstellen

Träger: Notdienst für Suchtmittelgefährdete und –abhängige Berlin e.V. (Ansbacher Straße 11, 10787 Berlin) mit den Beratungsstellen

- Drogennotdienst, Ansbacher Str. 11, 10787 Berlin
- LogIn, Jugend- und Suchtberatung, Kaiser-Friedrich-Str. 82, 10585 Berlin

Träger: Caritasverband für das Erzbistum Berlin e.V. (Residenzstraße 90, 13409 Berlin) mit den Beratungsstellen

- Integrative Suchtberatung „Königsberger 11“, Königsberger Straße 11, 12207 Berlin
- Jugend- und Suchtberatung Spandau, Hasenmark 3, 13585 Berlin
- Integrative Suchtberatung „Große Hamburger 18“, Große Hamburger Str. 18, 10115 Berlin

Träger: vista gGmbH (Alte Jakobstraße 85/86, 10179 Berlin) mit den Beratungsstellen

- Ambulante Suchtberatung Pankow, Erich-Weinert-Str. 145, 10409 Berlin
- Drogen- und Suchtberatung Marzahn-Hellersdorf, Allee der Kosmonauten 47, 12681 Berlin
- Suchtberatung Confamilia, Lahnstr. 84, 12055 Berlin
- Alkohol- und Drogenberatung Mitte, Stromstr. 47, 10551 Berlin
- Drogenberatung Misfit, Cuvrystr. 1, 10997 Berlin

Träger: Stiftung SPI Berlin (Belforter Straße 20, 10405 Berlin) mit den Beratungsstellen

- Drogenberatung Nord, Alt-Reinickendorf 23/24, 13407 Berlin
- Integrative Suchtberatung Lichtenberg, Einbecker Str. 32, 10317 Berlin

Träger: La Vida gGmbH (Nazarethkirchstr. 42, 13347 Berlin) mit der Beratungsstelle

- Frauenladen, Nazarethkirchstr. 42, 13347 Berlin

Ambulante Entwöhnungseinrichtungen

Träger: KOKON Verein für ambulante Drogentherapie (Galvanistraße 14, 10587 Berlin) mit den Einrichtungen

- KIBO – Zentrum für ambulante Drogentherapie, Obentrautstraße 57, 10963 Berlin
- KOKON – Zentrum für ambulante Drogentherapie, Galvanistraße 14, 10587 Berlin

Träger: Therapieladen e.V. (Potsdamer Str. 85/86, 10179 Berlin) mit der Einrichtung

- Therapieladen, Potsdamer Straße 85/86, 10179 Berlin

Träger: Tannenhof Berlin-Brandenburg (Meierottostraße 8-9, 10719 Berlin) mit der Einrichtung

- Ambulantes Zentrum des THBB – AZ, Blissestraße 2-6, 10713 Berlin

Träger: Drogentherapie-Zentrum Berlin e.V.(Frankfurter Allee 40, 10247 Berlin) mit der Einrichtung

- Ambulante Therapie DTZ, Boddinstraße 16, 12053 Berlin

2 Stationäre Entwöhnung

Stationäre Entwöhnungseinrichtungen

Träger: Tannenhof Berlin-Brandenburg e.V. (Meierottostraße 8-9, 10719 Berlin) mit den Einrichtungen

- Der Tannenhof, Mozartstraße 32-36, 12307 Berlin
- Die Pfalzburger, Pfalzburger Straße 35-38, 10717 Berlin
- Adaptionshaus Berlin, Dröpkeweg 10/12, 12353 Berlin

Träger: Drogentherapie-Zentrum Berlin e.V. (Frankfurter Allee 40, 10247 Berlin) mit den Einrichtungen

- Lago – Kurzzeittherapie, Am Großen Wannsee 29-31, 14109 Berlin
- Lago-Langzeittherapie und Adaption, Am Großen Wannsee 29-31, 14109 Berlin
- Entgiftungshaus Count-Down, Frankfurter Allee 40, 10247 Berlin

Träger: Anti-Drogen-Verein (ADV gGmbH Perleberger Straße 27, 10559 Berlin) mit den Einrichtungen

- Fachklinik F 42, Flughafenstraße 42, 12053 Berlin
- Briese, Briese 1-2, 16547 Birkenwerder
- ADV gGmbH Nokta (eig. gGmbH), Perleberger Straße 27, 10559 Berlin

Träger: Lenné gGmbH (Perleberger Staße 27, 10559 Berlin) mit der Einrichtung

- Haus Lenné, Breitehornweg 51-54, 14089 Berlin

3 Substitution

Kassenärztliche Vereinigung Berlin
Masurenallee 6A
14057 Berlin
Andreas von Blanc
Tel: 030-31003439

4 Entgiftung

Entgiftungsstationen

Vivantes Netzwerk für Gesundheit GmbH, Berlin
Am Nordgraben 2
13509 Berlin
Klinik für Psychiatrie, Psychotherapie und
Psychosomatik
Zentrum für Affektive Störungen (Depressionszentrum)

Vivantes deckt etwa 60% der drogenbedingten Entgiftungen in Berlin ab. Zusätzlich werden die folgenden Kliniken in die Studie einbezogen:

St. Hedwig-Krankenhaus
Große Hamburger Str. 5–11
10115 Berlin
Internet: http://www.alexianer-berlin-hedwigkliniken.de/st_hedwig_krankenhaus/wir_ueber_uns/st_hedwig_krankenhaus/

Jüdisches Krankenhaus Berlin
Heinz-Galinski-Straße 1
13347 Berlin-Mitte
Internet: http://www.juedisches-krankenhaus.de/home.html

Klinik für Psychiatrie und Psychotherapie
Charité–Universitätsmedizin Berlin
Campus Charité Mitte
Charitéplatz 1
10117 Berlin
Internet: http://psy-ccm.charite.de

5 Rauschgiftkriminalität

Landeskriminalamt
Tempelhofer Damm 12
12101 Berlin

6 Drogentote

Landeskriminalamt
Tempelhofer Damm 12
12101 Berlin

ANHANG II

Tabelle: Felder der Original-Datenquellen [1)] und Auswertungsdaten, die an das IFT gehen**.**

	1	2	3	4	5	6	Auswertungs-daten IFT
Vorname	x	x	x	x	x	x	-
Nachname	x	x	x	x	x	x	-
Geschlecht	x	x	x	x	x	x	x
Geburtsdatum (T-M-J; xx-xx-xxxx)	x	x	x	x	x	x	-
HIV Code [2)]	-	-	-	-	-	-	-
Kryptographischer Code	-	-	-	-	-	-	x
Alter [3)]	-	-	-	-	-	-	x
ICD-10 F11.-	x	x	x	x			x
ICD-10 F14.-	x	x	-	x			x
ICD-10 F15.-	x	x	-	x			x
Opiate					x	x	x
Kokain					x	x	x
Amphetamin/Stimulanzien					x	x	x

1) 1=Ambulante Suchtberatung, 2=Stationäre Entwöhnungskliniken, 3=Substitutionseinrichtungen, 4=Entgiftungsstationen, 5=Rauschgiftkriminalität, 6=Drogentote.

2) Der HIV Code wird kryptographisch transformiert

3) Zum Zeitpunkte der Aufnahme/Erfassung

Anhang B

Berliner Beauftragter für Datenschutz und Informationsfreiheit: Datenschutzrechtliche Beurteilung

Berliner Beauftragter für Datenschutz und Informationsfreiheit
Bereich Recht II
Forschung

Berliner Beauftragter für Datenschutz und Informationsfreiheit
An der Urania 4 - 10, 10787 Berlin

IFT Institut für Therapieforschung
Herrn Prof. Dr. phil. Ludwig Kraus
Parzivalstraße 25
80804 München

Nur per E-Mail

nachrichtlich:
Senatsverwaltung für Gesundheit und Soziales
– I B 31a –
Monika Wojak

GeschZ. (bitte angeben)	Bearbeiter(in)	Tel.: (030) 13 889-0 Durchwahl 13 889 App.:	Datum
54.3506.9	Frau Kamp	205	21. Juni 2013

Forschungsvorhaben „Capture-Recapture-Studie zur Schätzung des Umfangs und der Struktur der Drogenszene in Berlin"

Datenschutzrechtliche Einschätzung

Sehr geehrter Herr Prof. Dr. Kraus,

auf der Grundlage unseres Gesprächs am 13. März 2013, dem Vorgang des Hauptausschusses Nr. 0645 und der mit E-Mail vom 19. April 2013 übersandten Projektbeschreibung kommen wir zu der folgenden datenschutzrechtlichen Einschätzung des o. g. Projekts:

I. Allgemeines

Als Grundlage für die datenschutzrechtliche Zulässigkeit der Änderung, Übermittlung, Erhebung und Nutzung der in der Projektbeschreibung genannten personenbezogenen Daten scheidet die Einwilligungserklärung der Betroffenen vorliegend aus. Eine Anonymisierung der Datensätze ist nicht vorgesehen. Soweit die Daten also pseudonymisiert und in dieser Form an das IFT übermittelt werden, müssen gesetzliche Regelungen die Verarbeitung, Erhebung und Nutzung der Daten zu Forschungszwecken erlauben.

II. Datenschutzrechtliche Übermittlungsbefugnisse für die einzelnen Datenhalter

1. Suchtberatungsstellen / Entwöhnungseinrichtungen

Soweit ersichtlich handelt es sich bei den ambulanten Suchtberatungsstellen sowie den ambulanten und stationären Entwöhnungseinrichtungen um nicht-öffentliche Stellen, so dass die Datenübermittlungen gem. § 28 Abs. 2 Nr. 3 sowie § 28 Abs. 6 Nr. 4 Bundesdatenschutzgesetz (BDSG) für besondere Arten personenbezogener Daten zulässig sein muss. Da alle Beratungsstellen und Entwöhnungseinrichtungen Diagnosen (Angaben zu Diagnosen nach ICD-10 F11.-, F14.-, F15.-) übermitteln sollen, sind immer auch Angaben über die Gesundheit gem. § 3 Abs. 9 BDSG betroffen. In allen Fällen müssen daher die Anforderungen des § 28 Abs. 6 Nr. 4 BDSG erfüllt sein. Im Gegensatz zu §

28 Abs. 2 Nr. 3 BDSG wird in § 28 Abs. 6 Nr. 4 BDSG nicht auf eine „Forschungseinrichtung" Bezug genommen, so dass vielfach davon ausgegangen wird, dass nur die Eigenforschung durch § 28 Abs. 6 Nr. 4 ermöglicht werden soll (vgl. Simitis, BDSG-Kommentar, § 28 Rn. 309). Gleichwohl ist nach § 28 Abs. 8 Satz 1 BDSG eine zweckändernde Übermittlung von besonderen Arten personenbezogener Daten zulässig, wenn die Anforderungen des § 28 Abs. 6 Nr. 1-4 oder Abs. 7 Satz 1 BDSG erfüllt sind. Danach muss eine Übermittlung und Nutzung von besonderen Arten personenbezogener Daten an Forschungseinrichtungen grundsätzlich möglich sein. Zwar wird im Allgemeinen darauf hingewiesen, dass die Vorgaben des § 28 Abs. 8 widersprüchlich sind, da sich der Gesetzgeber bei sensitiven Daten grundsätzlich für einen sehr restriktiven Umgang entschieden hat (Simitis, BDSG-Kommentar, § 28 Rn. 320) und Zweckbeschränkungen dabei wesentlich sind. Für den Forschungsbereich kann allerdings nichts anderes gelten, als für die Datenübermittlung durch öffentliche Stellen, die in §§ 15, 16 BDSG geregelt ist. Sowohl § 15 Abs. 1 Nr. 2 (Übermittlung an öffentliche Stellen) als auch § 16 Abs. 1 Nr. 2 BDSG (Übermittlung an nicht-öffentliche Stellen) verweisen auf § 14 bzw. auf § 14 Abs. 5 u. 6 BDSG (bei besonderen Arten personenbezogener Daten). § 14 Abs. 5 Nr. 2 BDSG enthält eine wortgleiche Regelung zu § 28 Abs. 6 Nr. 4 BDSG. Im Zusammenhang mit § 14 Abs. 5 Nr. 2 BDSG wird festgehalten, dass die Vorschrift unabhängig davon gelten soll, ob es sich um eine öffentlich-rechtlich oder privatrechtliche Forschungseinrichtung handelt (Simitis, BDSG-Kommentar, § 14 Rn. 89). Auch der Richtlinie 95/46/EG sind keine gegenteiligen Vorgaben zu entnehmen. Zwar enthält der Ausnahmekatalog des Art. 8 Abs. 2 selbst keine Regelung zu Forschungsdatenverarbeitung. In Art. 8 Abs. 4 werden die Mitgliedstaaten allerdings ermächtigt, weitere Ausnahmeregelungen i. S. d. Artikel 8 Abs. 2 zu treffen. In Erwägungsgrund 34 wird dabei explizit auf die wissenschaftliche Forschung als ein wichtiges öffentliches einen Ausnahmetatbestand rechtfertigendes Interesse hingewiesen. Da Art. 8 Abs. 2 generell die Verarbeitung in den Blick nimmt und keine Beschränkungen im Hinblick auf die Art des Datenverarbeitungsschrittes enthält, ist auch die Übermittlung zu Forschungszwecken als Gegenstand eines Ausnahmetatbestandes nach der RL95/46/EG nicht ausgeschlossen.

Nach alledem dürfen besondere Arten personenbezogener Daten zu Forschungszwecken, d. h. auch zum Zwecke der Fremdforschung, übermittelt werden, wenn die Vorgaben gem. §§ 28 Abs. 8 i. V. m. § 28 Abs. 6 Nr. 4 BDSG eingehalten werden.

Gem. § 28 Abs. 6 Nr. 4 BDSG muss die Verarbeitung der besonderen Arten personenbezogener Daten zur Durchführung wissenschaftlicher Forschung erforderlich sein, das wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens muss das Interesse des Betroffenen an dem Ausschluss der Verarbeitung erheblich überwiegen und der Zweck der Forschung darf auf andere Weise nicht oder nur mit unverhältnismäßigem Aufwand erreicht werden.

Nach den Darlegungen der Senatsverwaltung für Gesundheit und Soziales im Vorgang des Hauptausschusses Nr. 0645 ist eine Schätzung der Drogenszene für eine verlässliche Planung und Steuerung der Berliner Drogenhilfeangebote erforderlich. Nur aufgrund der Ergebnisse einer wissenschaftlich fundierten Untersuchung könne beurteilt werden, ob das zur Verfügung stehende Hilfeangebot dem Bedarf entsprechend angepasst wurde und nach wie vor ausreichend sei bzw. ob evtl. weitere Anpassungen in den unterschiedlichen Segmenten der Hilfen notwendig würden. Diese Interessen an der Durchführung des Forschungsprojekts sind grundsätzlich als sehr gewichtig zu bewerten. Diesen stehen die schutzwürdigen Belange der Betroffenen gegenüber. Vorliegend sollen Angaben zur Drogenabhängigkeit, d. h. sehr sensitive Daten, verarbeitet werden, an denen die Betroffenen ein äußerst nachvollziehbares Interesse der Geheimhaltung haben. Auf der anderen Seite sollen die Ergebnisse, Hilfsangebote verbessern. Solche Steuerungsmaßnahmen müssen auf der Grundlage von Informationen über die tatsächlichen Gegebenheiten erfolgen, da ansonsten insbesondere die Effektivität der Hilfsangebote nicht gewährleistet werden kann. Nach den Beschreibungen im Vor-

gang Nr. 0645 sowie nach den Ausführungen in der Projektbeschreibung bestehen keine Anhaltspunkte dafür, dass das Studienkonzept der IFT die entsprechenden Informationen und Ergebnisse nicht liefern kann. Es ist plausibel, dass das Forschungsziel erreicht werden kann und die aufgezeigten Datenverarbeitungen dafür erforderlich sind.

Zudem ist zu berücksichtigen, dass die Daten nur pseudonymisiert verarbeitet werden sollen. Schon der erste Schritt der Pseudonymisierung durch Bildung des HIV Codes macht es ohne weiteres Wissen über den betroffenen Personenkreis unmöglich, aus den Pseudonymen die klaren Identitätsdaten abzuleiten. Der zweite, kryptografische Schritt der Pseudonymisierung (s. Ziff. IV) verhindert nach Abschluss der Erhebung darüber hinaus, dass umgekehrt die Identifikationsdaten einer gegebenen Person mit den erhobenen Daten verknüpft werden können: Sind die Identifikationsdaten einer Person bekannt, kann weder ausgeschlossen werden, dass sie dem Betroffenenkreis angehört, noch kann aus der Existenz eines „passenden" Pseudonyms abgeleitet werden, dass für die Person die Aussage, dass sie drogenabhängig sei, mit höherer Wahrscheinlichkeit zutrifft als für die die allgemeine Population. Selbst für eine bekannte betroffene Person ist es nicht möglich, die sie betreffenden Daten zu selektieren.

Zusammengefasst bedeutet dies, dass die vorgesehene Datenerhebung und -verarbeitung zur Durchführung wissenschaftlicher Forschung erforderlich ist und das wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens das Interesse des Betroffenen an dem Ausschluss der Verarbeitung erheblich überwiegt. Darüber hinaus ist nicht ersichtlich, wie der Forschungszweck auf andere Weise, etwa mit anonymisierten Datensätzen, erreicht werden könnte. Die Suchtberatungsstellen und Entwöhnungskliniken dürfen daher gem. § 28 Abs. 6 Nr. 4 BDSG pseudonymisierte und verschlüsselte Daten an das IFT übermitteln, soweit auch die weiteren Vorgaben beachtet werden (siehe insbesondere Ziff. IV.).

Für die Suchtberatungsstellen, die vom Caritasverband für das Erzbistum Berlin e. V. getragen werden, ist der Berliner Beauftragte für Datenschutz und Informationsfreiheit als Aufsichtsbehörde nicht zuständig. In diesem Fall ist der kirchliche Datenschutz anzusprechen (Diözesandatenschutzbeauftragter Herr Lutz Gramman, Engelbosteler Damm 72, 30167 Hannover, +49(511) 81 93 15, info@datenschutz-kirche.de).

2. Kassenärztliche Vereinigung (KV)

Für die Kassenärztliche Vereinigung existiert auf gesetzlicher Grundlage keine Befugnis zur Übermittlung personenbezogener Daten aus dem Bereich der Leistungsabrechnung (Sozialdaten) zu Forschungszwecken. Zwar lässt § 287 Abs. 1 SGB V die Auswertung von leistungserbringer- bzw. fallbezogenen Daten zu. Dies gilt jedoch nur für KV oder Krankenkassen interne Forschungsvorhaben. Eine Übermittlung an Dritte, wenn auch in pseudonymisierter Form, ist davon nicht erfasst. Für eine solche können auch die Vorschriften des §§ 75, 76 SGB X nicht herangezogen werden, da diese auf die KV nicht anwendbar sind (BT-Drs. 11/2237, S. 239).

Für die KV sehen wir daher keine Möglichkeit, personenbezogene Daten ohne Einwilligung der Betroffenen an das IFT zu übermitteln.

3. Entgiftungsstationen

Bei den Entgiftungsstationen sollen Krankenhäuser im Land Berlin angefragt werden, für die das Landeskrankenhausgesetz (LKG) gilt. Nach § 25 Abs. 3 LKG dürfen dabei Daten in pseudonymisierter Form auch an einrichtungsübergreifende Forschungsvorhaben übermittelt wer-

den. Dies darf gem. § 25 Abs. 1 Satz 2 Nr. 2 LKG auch ohne Einwilligung der Betroffenen erfolgen, wenn es nicht zumutbar ist, die Einwilligung einzuholen und schutzwürdige Belange der Patientin oder des Patienten nicht beeinträchtigt werden. Die Datenerhebung auf Einwilligungsbasis scheidet vorliegend aus. Es handelt sich um Daten aus den Jahren 2010 und 2011, so dass die Betroffenen nicht mehr vor Ort um ihre Zustimmung gebeten werden können. Da das IFT nur pseudonymisierte Daten erhalten soll, müssten die Krankenhäuser, Kontaktmöglichkeiten der ehemaligen Patienten recherchieren, diese kontaktieren und um eine Einwilligung bitten. Dies stellt einen hohen Aufwand dar, der von den Krankenhäusern nicht zusätzlich geleistet werden kann, so dass die Einholung der Einwilligung unzumutbar ist. Durch die Pseudonymisierung und zusätzliche Verschlüsselung ist den schutzwürdigen Belangen der Patientinnen und Patienten ausreichend Rechnung getragen (vgl. Ziff. I. 1.).

Die Entgiftungsstationen dürfen daher gem. § 25 Abs. 3 i. V. m. § 25 Abs. 1 Satz 2 Nr. 2 LKG pseudonymisierte und verschlüsselte Daten an das IFT übermitteln, soweit auch die weiteren Vorgaben beachtet werden (siehe insbesondere Ziff. IV.).

4. Landeskriminalamt

Soweit wir dem Gespräch vom 13. März 2013 entnehmen konnten, sollen Daten einer „zentralen Rauschgiftdatei" beim LKA ausgewertet werden. Eine solche zentralisierte Datei, die Identifikations- und Konsumdaten von Personen zum spezifischen Thema „Rauschgift" enthält, ist in unserem Hause nicht bekannt. Da wir keine Informationen über die Rechtsgrundlage für die Speicherung der Daten in einer solchen Datei beim LKA haben, können wir auch keine Einschätzung dazu abgeben, ob und wie diese Daten zulässigerweise zu Forschungszwecken genutzt werden könnten. Wir empfehlen daher, beim LKA und dem entsprechenden behördlichen Datenschutz anzufragen, auf welcher Rechtsgrundlage die Daten aus einer solchen Datei zu Forschungszwecken genutzt werden dürften und uns ein entsprechendes Votum vorzulegen.

Nach unseren Informationen existiert eine sog. „Falldatei Rauschgift" (siehe z. B. „Bundeslagebericht Rauschgift 2011" des BKA). Diese Verbunddatei wird beim BKA vorgehalten und von den LKAs mit Daten aus den Ländern bestückt. Die LKAs können selbst wiederum zur Aufgabenerfüllung Daten aus der Datei abrufen. Die Datei enthält u. a. Informationen zur konsumierten Drogenart. Die Übermittlung von Daten aus dieser Datei zu Forschungszwecken richtet sich nach den Vorgaben des Gesetzes über das Bundeskriminalamt und die Zusammenarbeit des Bundes und der Länder in kriminalpolizeilichen Angelegenheiten (BKAG), insbesondere § 29 BKAG. Für das BKA als übermittelnde Stelle ist der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit zuständig.

Wie bereits im Gespräch am 13. März 2013 erörtert, bestände ansonsten die Möglichkeit, Daten aus dem Strafverfolgungsbereich anzufragen. Auch hier gibt es gesetzliche Möglichkeiten der Datenübermittlung zu Forschungszwecken nach § 476 StPO. Im Gespräch wurde allerdings von Seiten des IFT bereits darauf hingewiesen, dass dies kein erfolgsversprechender Weg sei, da die Staatsanwaltschaften solche Anfragen aus Kapazitätsmangel in der Regel ablehnen würden.

Auch für die Zulässigkeit der Speicherung von Daten, die Grundlage für die Erstellung der Drogentotenstatistik beim LKA sein sollen, fehlen uns die erforderlichen Informationen. Hier ist ebenfalls nicht klar, auf welcher Rechtsgrundlage die Daten durch das LKA erhoben werden, so dass für uns nicht prüfbar ist, ob für die Nutzung dieser Daten zu Forschungszwecken eine gesetzliche Grundlage besteht. Auch in diesem Fall sollte ein entsprechendes Votum des behördlichen Datenschutzes des LKA eingeholt und uns vorgelegt werden.

III. Schweigepflicht

Eine Offenbarung im Sinne des § 203 Strafgesetzbuch ist dann nicht gegeben, wenn weder aus dem Pseudonym selbst, noch aus den mitgelieferten Daten oder aus sonstigen Zusammenhängen heraus eine Bestimmung des Berechtigten für den Dritten möglich ist. Für die Weitergabe pseudonymisierter Daten muss ausgeschlossen sein, dass der Datenempfänger die Pseudonyme (z. B. durch Zusatzwissen) aufheben kann. Davon ist vorliegend auszugehen, so dass die Schweigepflicht gewahrt bleibt.

IV. Pseudonymisierung

Die Pseudonymbildung muss kryptografisch abgesichert werden, um auszuschließen, dass für eine gegebene Person durch Rückgriff auf den Gesamtbestand der in den erhobenen Daten enthaltenen Pseudonyme mit größerer Sicherheit Aussagen über ihren Drogenkonsum abgeleitet werden können, als ohne diesen Zugriff. Die kryptografische Komponente muss daher zwingend Einwegeigenschaften ausweisen und einen geheimen Schlüssel involvieren, welcher von dem pseudonymen Datenbestand getrennt gespeichert und nach Abschluss der Erhebung gelöscht wird. Die Algorithmen aus Abschnitt 5.1 der Technischen Richtlinie TR-02102 des Bundesamtes für Sicherheit in der Informationstechnik sind geeignet. Für eine Reihe der dort aufgeführten Algorithmen (z. B. für HMAC-SHA1) existieren kostenfrei zu nutzende Implementierungen in Java, C und anderen Programmiersprachen.

Soweit das IFT die Datenhalter bei der Pseudonymisierung unterstützt, muss sichergestellt sein, dass Mitarbeiter des IFT nicht auf Klardaten zugreifen können. Wenn in diesem Zusammenhang Hilfestellungen zu Normalisierung der Namen gegeben werden, so dürfen dabei die Namen selbst nicht genannt werden, sondern es müssen die Probleme abstrakt dargestellt und gelöst werden

Für Rückfragen stehe ich gerne auch telefonisch zur Verfügung.

Mit freundlichen Grüßen

Meike Kamp

Anhang C

Der Diözesandatenschutzbeauftragte

Der Diözesandatenschutzbeauftragte

der Erzbistümer Berlin und Hamburg,
der Bistümer Hildesheim, Magdeburg, Osnabrück und
des Bischöflich Münsterschen Offizialats in Vechta i.O.

DATENSCHUTZ
IN DER KATHOLISCHEN KIRCHE

Datenschutzbeauftragter der norddeutschen Bistümer • Engelbosteler Damm 72 • 30167 Hannover

Engelbosteler Damm 72
30167 Hannover
Fon: 0511 / 81 93 15
Fax: 0511 / 81 21 35
E-Mail:
info@datenschutz-kirche.de
Internet:
www.datenchutz-kirche.de

IFT Institut für Therapieforschung
- Herrn Prof. Dr. Ludwig Kraus -
Parzivalstraße 25

80804 München

EINGEGANGEN
22. Juli 2013

Hannover, den 16. Jul. 2013

Forschungsvorhaben "Capture-Recapture-Studie"
Ihr Zeichen: 50.614/lu

Sehr geehrter Herr Prof. Dr. Kraus,

Ihre Anfrage mit Schreiben vom 11.07.2013, nebst der Kopie der Stellungnahme des Berliner Beauftragten für Datenschutz und Informationsfreiheit vom 21.06.2013 und des Antrags auf Durchführung des Forschungsvorhabens vom 19.04.2013 ist hier am 15.07.2013 eingegangen.

Die katholische Kirche hat die Vorschriften für privatwirtschaftliche Datenverarbeiter des Bundesdatenschutzgesetzes nicht in das kirchliche Recht übernommen. Die Anordnung über den kirchlichen Datenschutz - KDO - enthält keine dem § 28 BDSG entsprechende Vorschrift. Die Ausführungen zur Anwendbarkeit dieser Vorschrift des Berliner Datenschutzbeauftragten können daher nicht auf kirchliche Einrichtungen angewandt werden.

Eine Datenübermittlung an Empfänger außerhalb der Kirche ist nur unter den Voraussetzungen von § 12 Abs. 1 Nr. 2 KDO zulässig. Die Tatsache, dass Sie ein berechtigtes Interesse an der Kenntnis der Daten glaubhaft dargelegt haben, ist schon vom Berliner Datenschutzbeauftragten ausführlich geprüft und bejaht worden. Dabei kommt es aber wesentlich auch auf die Übermittlung von besonderen Arten personenbezogener Daten, besonders aus dem Bereich der Gesundheit (§ 2 Abs. 10 KDO) an. Solche Daten dürfen aber nach § 12 Abs. 1 Nr. 2 Satz 2 KDO nur weitergegeben werden, wenn die Voraussetzungen des § 10 Abs. 5 und 6 KDO gegeben sind. Im vorliegenden Falle könnte insoweit eine Änderung des Nutzungszwecks auf Grund von § 10 Abs. 5 Nr. 2 KDO gegeben sein, wenn die Übermittlung einem Forschungsprojekt dient und das kirchliche Interesse, an dem Vorhaben, die zu berücksichtigenden Interessen der Betroffenen erheblich überwiegt. Der Berliner Datenschutzbeauftragte hat hierzu sehr ausführlich ausgeführt, dass das Forschungsvorhaben, der Steuerung, Planung und Steigerung der Effektivität der Hilfsangebote auf dem Gebiet des Drogenmissbrauchs dient. Es liegt sicherlich auch im Interesse unserer, hier tätigen Einrichtungen, insoweit wissenschaftlich fundiert unterrichtet zu werden. Ein kirchliches Interesse würde ich, gemeinsam mit dem Berliner Kollegen, bejahen, wenn sichergestellt werden kann, dass auch unsere Anbieter vom Ergebnis Ihrer Untersu-

chungen durch Übersendung des Forschungsberichts informiert werden. In diesem Fall hätten kirchliche Suchtberatungsstellen ebenfalls die Möglichkeit, ihr Angebot nach wissenschaftlichen Empfehlungen auszurichten.

Bei der Übermittlung der Daten muss nach § 10 Abs. 6 KDO weiterhin die Verschwiegenheitspflicht nach § 203 StGB gewahrt werden. Diese Voraussetzung wird nach Meinung des Berliner Datenschutzbeauftragten erfüllt durch eine Zurverfügungsstellung in pseudonymisierter Form. Dabei hat er der die zwei von Ihnen angegeben Schritte (Pseudonymisierung des HIV-Codes und kryptografische Transformation) als ausreichend angesehen, um eine Identifikation der Betroffenen durch Mitarbeiter Ihres Forschungsvorhabens zu verhindern. Ich schließe mich dieser Auffassung an und gehe davon aus, dass die Verschwiegenheitspflicht durch die Weitergabe von Daten, die keine Rückschlüsse auf die betreffende Person zulassen, beachtet wird.

Gegen eine Beteiligung von Caritas-Einrichtungen an dem Forschungsvorhaben bestehen insoweit keine Bedenken, wenn die Vorgaben aus Ihrem Antrag vom 19.04.2013 eingehalten werden und die angesprochenen Beratungsstellen eine Abschrift Ihres Berichts zur Weiterentwicklung ihrer Angebote erhalten.

Die Vorschriften der KDO möchten Sie bitte der PDF-Datei auf meiner Webseite unter
http://www.datenschutz-kirche.de/sites/default/files/file/download/KDO%2020011-b.pdf
entnehmen.

Der Caritasverband für das Erzbistum Berlin hat eine Abschrift dieses Schreibens erhalten.

Die St. Hedwig Kliniken Berlin GmbH stehen in Trägerschaft eines Ordens päpstlichen Rechts, der Gesellschaft der Alexianerbrüder, Krausnickstraße 12 A, 10115 Berlin. Zum Datenschutzbeauftragten wurde dort Herr Dr. Heiko Haaz, Mitarbeiter des UIMC, Nützenberger Straße 119, 42115 Wuppertal ernannt.

Mit herzlichen Grüßen

Lutz Grammann
Diözesandatenschutzbeauftragter

Anhang D

Technische Beschreibung von Dateneingabe und Verschlüsselung

Capture-Recapture-Studie zur Schätzung des Umfangs und der Struktur der Drogenszene in Berlin

Technische Beschreibung Dateneingabe und Verschlüsselung

1 Allgemein

Für die Abschätzung der Populationsgrößen substanzbezogener Störungsgruppen ist es notwendig, Fälle über verschiedene Datenquellen hinweg zu identifizieren, d.h. zu „matchen". Um dies bewerkstelligen zu können, ist eine *ein-eindeutige* Zuordnung eines Codes zum äquivalenten Code in den weiteren Datensätzen erforderlich, d.h. ein „unique identifier". Der Klarname einer Person ist als „unique identifier" ungeeignet, zum einen aus Gründen des Datenschutzes, zum anderen sind manche Klarnamen nicht eindeutig genug (z.B. Stefan Maier). Es werden aus diesem Grunde folgende Daten von den Datenhaltern benötigt, die in zwei Schritten verschlüsselt (kryptographiert) werden, sodass eine Rückführung auf den ursprünglichen Klarnamen unmöglich ist: Vorname, Nachname, Geschlecht, Geburtsjahr. Zusätzlich benötigen Sie Angaben zum Drogenkonsum (Opiate, Kokain, Stimulanzien) bzw. den Diagnosen (F11, F14, F15), der Art des Opiatkonsums (intravenös) und dem Jahr der Datenerhebung (2010 oder 2011).

2 Technisches Vorgehen

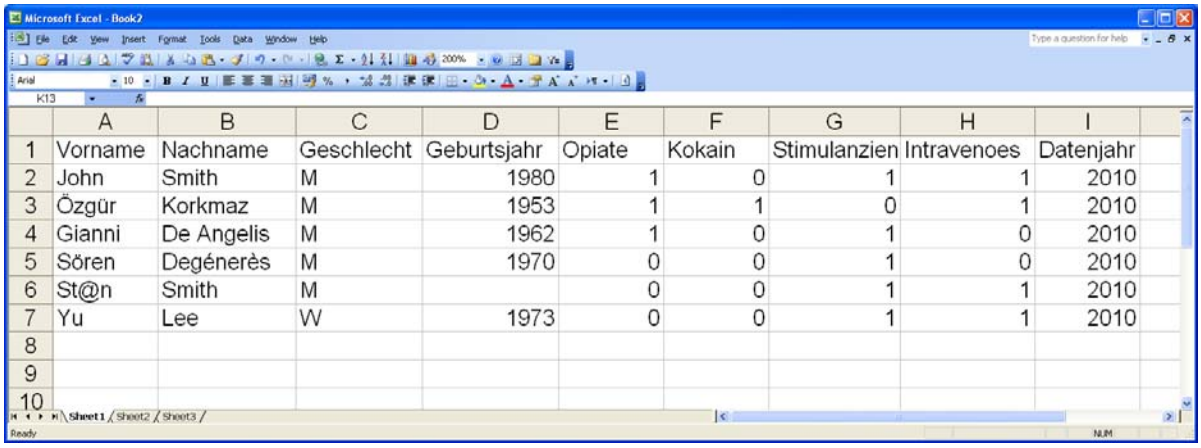
Die Verschlüsselung der ersten vier Variablen (Vorname, Nachname, Geschlecht, Geburtsjahr) erfolgt auf Basis des sogenannten **HIV-Codes** (einer Verschlüsselungsvorschrift, die aus dem dritten Buchstaben des Vornamens, der Länge des Vornamens, dem dritten Buchstaben des Nachnamens, der Länge des Nachnamens, dem Geschlecht [M oder W] und dem Geburtsjahr besteht). Außerdem werden noch weitere HIV-Codes mit ähnlicher Vorschrift gebildet, um die Fehleranfälligkeit und Eindeutigkeit dieses Codes zu überprüfen. Im Anschluss werden diese verschiedenen Codes auf Basis eines **H-MAC Algorithmus** und **eines 12-stelligen Schlüssels** weiter kryptographiert. Zwei der vier für die Verschlüsselung benötigten Variablen (Vorname, Nachname) werden nicht in den finalen Datensatz übernommen. Geschlecht und Geburtsjahr werden für die Gruppenbildung nach Geschlecht und Altersgruppen benötigt. Der finale Datensatz besteht aus den Kryptogrammen, Geschlecht, Geburtsjahr und den Variablen Opiate (Opiate bzw. F11), Kokain (Kokain bzw. F14), Stimulanzien (Stimulanzien bzw. F15), Intravenoes (intravenöser Opiatkonsum) und Datenjahr (Datenerhebung 2010 oder 2011).

3 Anleitung

Die genannten Verschlüsselungsvorgänge müssen von den Datenhaltern nicht selbst durchgeführt werden, sondern liegen als Programm auf einer Website vor. Die Datenhalter müssen lediglich die oben beschriebenen neun Variablen in der genannten Reihenfolge in einem dafür geeigneten Programm vorbereiten, z.B. Excel. Der Datenhalter bereitet dabei diese Variablen in exakt der genannten Reihenfolge als Spalten auf, wobei die Zeilen Fälle darstellen. Eine Korrektur des

Datensatzes hinsichtlich doppelt vorkommender Fälle ist nicht notwendig. Die neun Spaltennamen müssen lauten: Vorname, Nachname, Geschlecht, Jahr, Opiate, Kokain, Stimulanzien, Intravenoes, Datenjahr (siehe Abbildung1). Die Kodierung ist wie folgt:

Vorname	Klarname
Nachname	Klarname
Geschlecht	M=Mann; W=Frau
Geburtsjahr	Geburtsjahr (4-stellig)
Opiate	1=Opiatkonsum vorhanden bzw. Diagnose F11 erfüllt; 0=Nein/Missing
Kokain	1=Kokainkonsum vorhanden bzw. Diagnose F14 erfüllt; 0= Nein/Missing
Stimulanzien	1=Stimulanzienkonsum vorhanden bzw. Diagnose F15 erfüllt; 0= Nein/Missing
Intravenoes	1=i.v. Opiatkonsum; 0= Nein/Missing
Datenjahr	Jahr der Datenerhebung (4-stellig: 2010 oder 2011)



	A	B	C	D	E	F	G	H	I
1	Vorname	Nachname	Geschlecht	Geburtsjahr	Opiate	Kokain	Stimulanzien	Intravenoes	Datenjahr
2	John	Smith	M	1980	1	0	1	1	2010
3	Özgür	Korkmaz	M	1953	1	1	0	1	2010
4	Gianni	De Angelis	M	1962	1	0	1	0	2010
5	Sören	Degénerès	M	1970	0	0	1	0	2010
6	St@n	Smith	M		0	0	1	1	2010
7	Yu	Lee	W	1973	0	0	1	1	2010
8									
9									
10									

Abbildung 1. Datenaufbereitung in Excel anhand der auf der Webseite vorgegebenen Beispieldaten [1]

Anschließend speichert der Datenhalter diese Datei im Formt „csv“ = comma separated variable (siehe Abbildung 2).

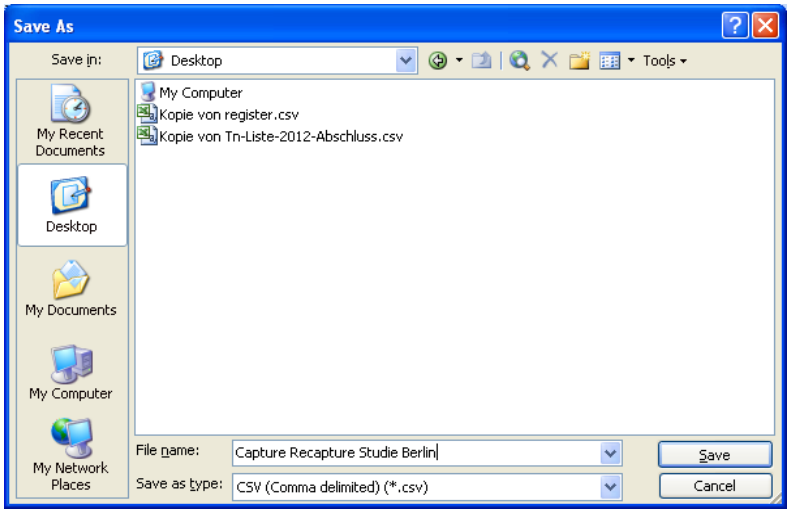


Abbildung 2. Datei > Speichern unter (Englisch: „File > Save as…“) CSV-Dateityp in Excel

[1] Zur Datenaufbereitung vgl. auch Anhang 1

Öffnen Sie dann die gespeicherte Datei mit Hilfe des Programm „Windows Editor“. Dieses Programm findet sich normalerweise unter Start > Programme > Zubehör. Sie können auch ein äquivalentes einfaches Textprogramm verwenden (siehe Abbildung 3).

Wichtig an dieser Stelle ist, dass die Variablen mit „Semikolon“ getrennt sind. Sofern Excel oder ein ähnliches Programm Komma anstatt Semikolon verwendet (siehe Abbildung 3), müssen alle Kommata ersetzt werden. Sie können alternativ auch die Tastenkombination STRG+H verwenden, um alle Kommas durch Strichpunkte zu ersetzen. Achten Sie bereits bei der Datenbearbeitung in Excel darauf, dass in Ihren Daten in den Felder optimalerweise keine zusätzlichen Strichpunke vorhanden sind. Diese gefährden die Spaltenreihenfolge.

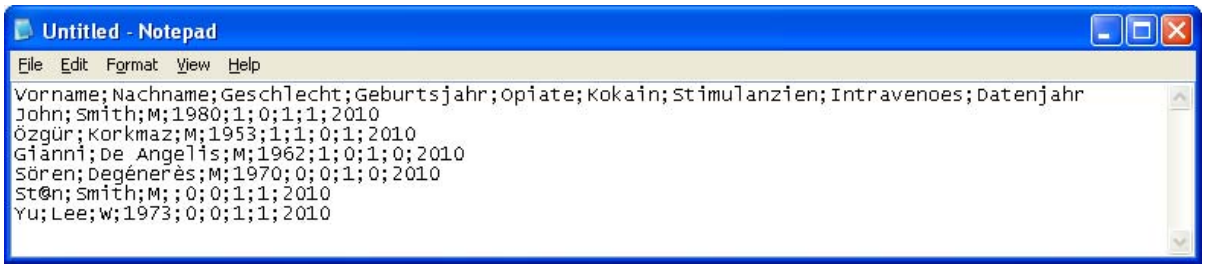


Abbildung 3. Beispiel für ein csv-Dokument nach Speichern in Excel und Öffnen mit Windows Editor

Abbildung 3 zeigt ein Beispiel einer csv-Datei, nachdem sie im Windows Editor geöffnet wurde, in dem fünf Fälle enthalten sind. Sobald der Datenhalter diese csv-Datei erstellt hat, kann der Datenhalter das Internetprogramm unter der Website http://www.iscpt.org/Epitoolbox/BerlinCRC.aspx öffnen.

Abbildung 4. Startseite der Internetverschlüsselungsseite (http://www.iscpt.org/Epitoolbox/BerlinCRC.aspx)

Abbildung 4 oben zeigt die Startseite des Internetprogramms, das für den Zweck der Verschlüsselung genutzt wird. Die csv-Datei wird nun in vier Schritten, die im Folgenden beschrieben werden, in eine finale Datei mit den oben beschrieben Verschlüsselungen umgewandelt:

Schritt 1: Trimmen der Klarnamen, d.h. Normalisierung von Buchstaben wie z.B. ö, ü, oder ß in oe, ue und ss

Schritt 2: Bildung der HIV-Codes; d.h. Umwandlung von Klarnamen (Vor- und Nachname), Geschlecht und Geburtsjahr in einen Code bestehend aus dem dritten Buchstaben des Vornamens, der Länge des Vornamens, dem dritten Buchstaben des Nachnamens, der Länge des Nachnamens, dem Geschlecht [M oder W] und dem Geburtsjahr.

Schritt 3: Eingabe notwendiger Angaben für die Verschlüsselung: Datenquellenname und 12-digit Schlüssel.

Schritt 4: Lokale Speicherung des verschlüsselten Datensatzes und Übergabe an das IFT Institut für Therapieforschung.

Durch Drücken der Schaltfläche „Kryptographierung starten“ kommt man zu dem Fenster, in das die Daten eingefügt werden sollen. Als Standardeinstellung sind hier bereits Beispieldaten eingetragen, die durch die Daten des Datenhalters überschrieben werden. Diese Beispieldaten decken sich der Einfachheit halber mit den hier verwendeten Beispieldaten (siehe Abbildung 5).

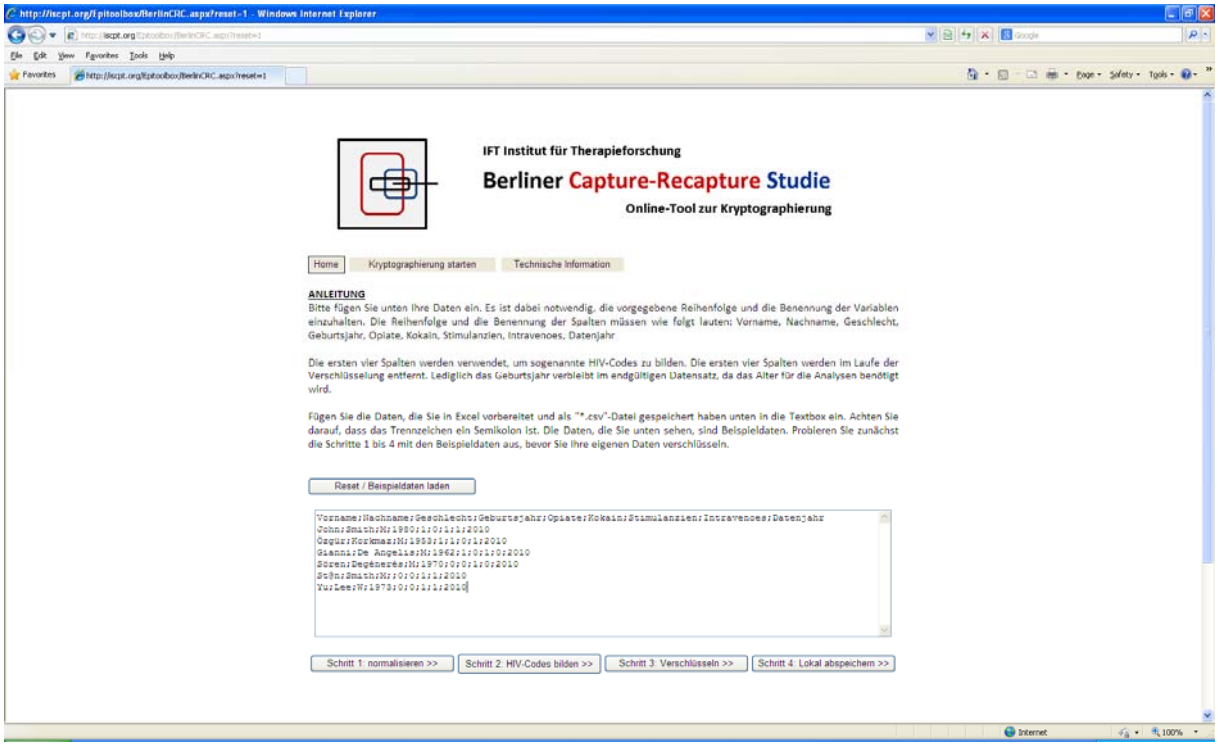


Abbildung 5. Unterseite „Kryptographierung Starten“ der Internetverschlüsselungsseite (http://www.iscpt.org/Epitoolbox/BerlinCRC.aspx)

Abbildung 5 zeigt das Eingabefenster mit den Beispieldaten, die immer automatisch geladen werden. Im Anschluss geht man wieder auf das Editorfenster (Abbildung 2) und betätigt „Bearbeiten – Alles markieren“ (auch möglich als Tastenkombination: STRG+A), dann „Bearbeiten – Kopieren“ (auch möglich als STRG+C). Anschließend fügt man diese Daten aus dem Editor in das Eingabefenster auf

der Internetseite (Abbildung 5) ein, wobei man zuvor die Beispieldaten löscht. Es muss dabei der Cursor in der Textbox sein, in der zuvor die Beispieldaten waren. Einfügen kann man entweder durch rechten Mausklick und Einfügen oder durch die Tastenkombination STRG+V.

Im Anschluss werden die **vier horizontalen Schaltflächen Schritt 1 bis 4 nacheinander betätigt**. Im ersten Schritt werden die Klarnamen automatisch auf Sonderzeichen, Umlaute etc. bereinigt. Im zweiten Schritt werden HIV-Codes automatisch gebildet (siehe Abbildung 6). Nach erfolgreicher Durchführung jedes Schritts erscheint ein grünes Häkchen im Kreis.

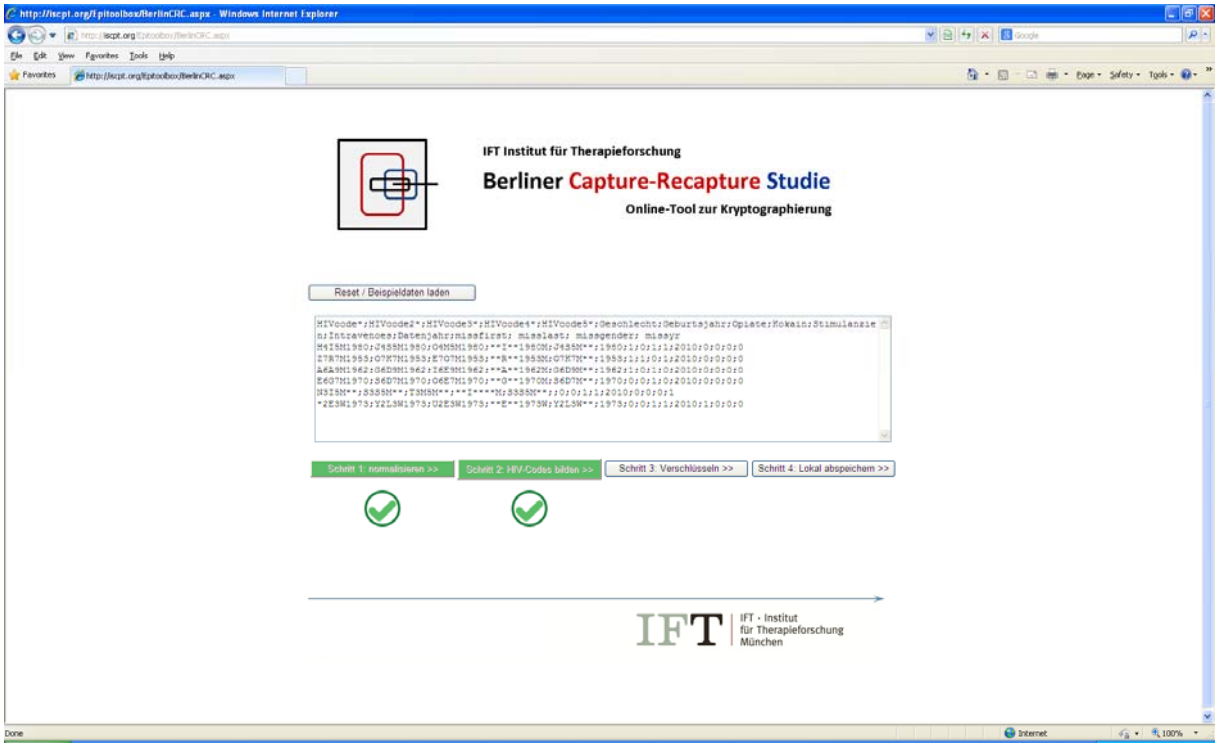


Abbildung 6. Trimmen der Klarnamen (Schritt 1) und Bildung der HIV-Codes (Schritt 2)

Ab dem dritten Schritt (Abbildung 7) ist ein 12-stelliger Schlüssel notwendig. Diesen Schlüssel erhalten die Datenhalter von der Senatsverwaltung für Gesundheit und Soziales, die den Schlüssel auswählt. Dieser Schlüssel ist nur der Senatsverwaltung und den beteiligten Datenhaltern bekannt. Zudem ist im dritten Schritt die Angabe des Namens der Datenquelle erforderlich, z.B. „Drogentotenstatistik“. Außerdem muss der 12-stellige Schlüssel zweimal angegeben werden, auf Basis dessen die Verschlüsselung läuft (siehe Abbildung 7).

Wichtig: Beachten Sie hier, dass Sie den 12-stelligen Schlüssel mit größtmöglicher Sorgfalt eingeben. Sofern Sie zweimal den falschen Code eingeben, sind die Daten für eine weitere Verwendung nicht mehr verwendbar, da die Verschlüsselung zu keinen ein-eindeutigen Kryptogrammen führt.

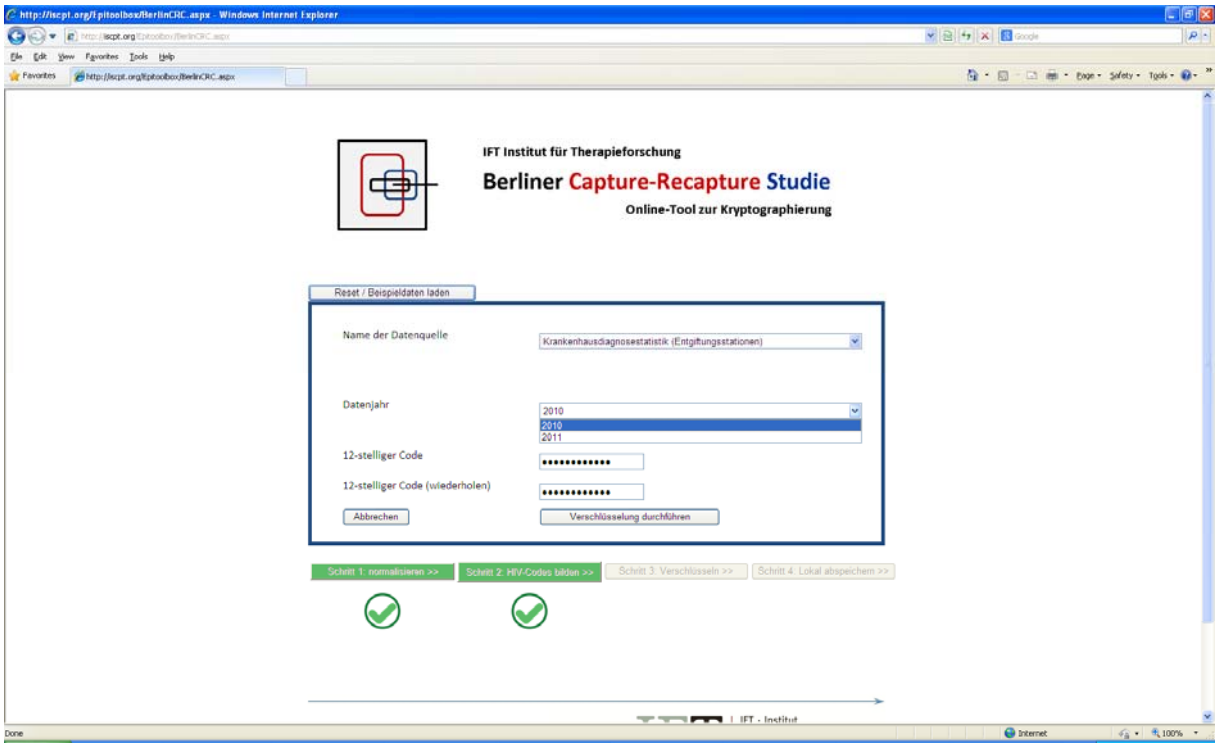


Abbildung 7. Eingabe des Datenquellennamens, des Datenjahres und des 12-stelligen Schlüssels (Schritt 3)

Sobald alle Felder ausgefüllt sind, kann der dritte Schritt „Verschlüsselung durchführen“ ausgeführt werden. Das Ergebnis des dritten Schrittes ist in Abbildung 8 dargestellt.

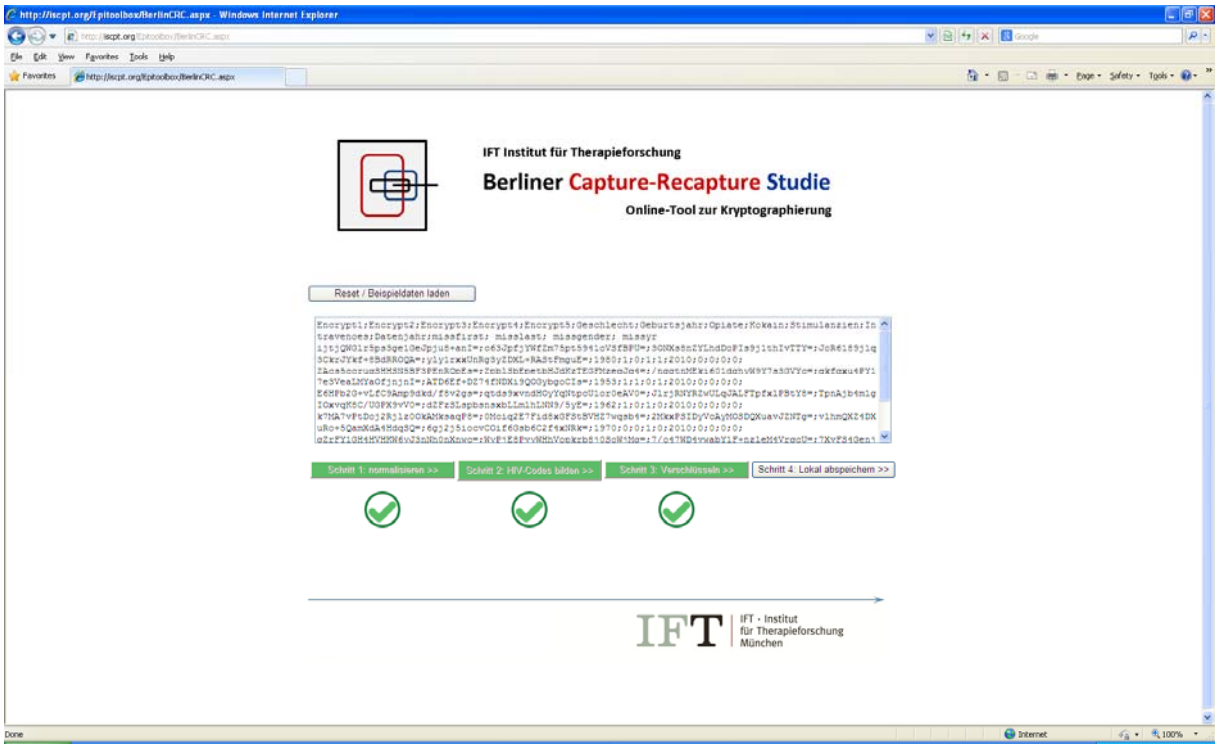


Abbildung 8. Ergebnis der Verschlüsselung durch den H-Mac Algorithmus (Schritt 3)

Abbildung 8 zeigt das Ergebnis der Verschlüsselung durch den H-Mac Algorithmus. Verschlüsselt wurden dabei lediglich die verschiedenen HIV-Codes. Zusätzlich enthält nun das Datenfeld auch Informationen, ob eine der vier Schlüssel-Informationen gefehlt hat (1) oder nicht (0) (missfirst, misslast, missgender, missyear).

Mit der Durchführung des vierten Schritts (siehe Abbildung 8), können die Daten lokal abgespeichert werden. Betätigen Sie dazu die Schaltfläche „Datei lokal speichern“. Während der Verwendung des Online-Programms sind die Daten ausschließlich im Arbeitsspeicher des lokalen Computers abgelegt. Bitte speichern Sie die Datei exakt unter dem Namen ab, den das Programm vorgibt.

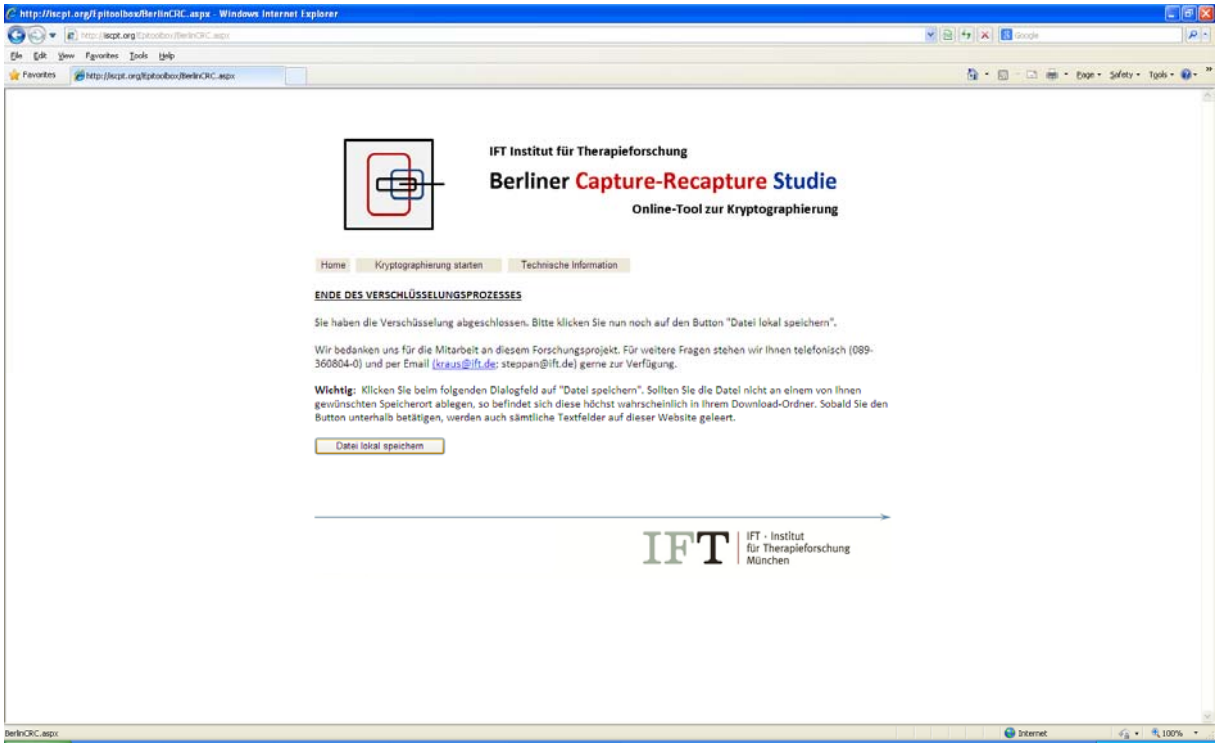


Abbildung 9. Lokale Speicherung der verschlüsselten Daten (Schritt 4)

Schlussbemerkung: Die Verschlüsselung der Daten mit Hilfe der dargestellten Internetseite sollten für einen geübten Computeranwender mit Hilfe dieser Anleitung keine große Schwierigkeit darstellen. Die Verschlüsselung mit Hilfe der Seite kann eigenständig durchgeführt werden, aber auch mit Hilfe von Assistenz seitens des IFT Instituts für Therapieforschung, München.

Für die Extraktion der Daten [z.B. polizeiauffälliger Konsumenten von Opiaten, Kokain oder Stimulanzien oder Entgiftung (qualifizierter Entzug von Opiaten, Kokain oder Stimulanzien) im Krankenhaus] aus den Datensystemen der jeweiligen Datenhalter bedarf es bestimmter Filter, die mit dem Projektleiter am IFT abgestimmt werden sollten. Es sollten getrennte Datenläufe für die drei Substanzgruppen erfolgen. Die Resultate dieser Läufe sollen dann in eine Excel-Datei übertragen werden. Beispielsweise weisen alle Opiatfälle in der Spalte „Opiate“ den Wert 1 auf, für „intravenös“ den Wert 0 oder 1 (je nach Verfügbarkeit der Information). Im Anschluss werden die Kokainfälle mit den entsprechenden Werten in der Spalte Kokain = 1 übertragen. Entsprechend wird bei Stimulanzien verfahren. Die Datensätze sind für die Jahre 2010 und 2011 getrennt zu behandeln.