Published in final edited form as:
Eval Program Plann. 2008 May ; 31(2): 125–135. doi:10.1016/j.evalprogplan.2008.02.003.

Cross-system Data Linkage for Treatment Outcome Evaluation: Lessons learned from the California Treatment Outcome Project

Yih-Ing Hser and **Elizabeth Evans**
UCLA Integrated Substance Abuse Programs, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles

Abstract

Using administrative data to evaluate health care outcomes has become increasingly common, but the reliability and validity of outcome measures based on cross-system data linkage have been little scrutinized. Applying a deterministic data matching methodology, we linked 6,545 Californians admitted to 43 substance abuse treatment programs between 2000 and 2001 to administrative data acquired from three state agency databases. We compared self-reported treatment outcome measures to equivalent measures derived from official records on motor vehicle driving incidents, criminal history, and mental health services utilization. Administrative data provided analogous results in some measures (e.g., percent of people using mental health services, percent ever arrested) and more accurate results in others (e.g., frequency of service utilization, and frequency of arrests). Similar to findings based on the interview data, the administrative data also revealed improvements in several domains one-year post-treatment compared to one-year pre-treatment. Experiences with data linkage procedures and strategies for enhancing record linkage accuracy are discussed. Findings illustrate the value of using administrative records for substance abuse treatment outcome evaluation, while highlighting areas for improvement for future cross-system data linkage efforts.

Introduction

Over the past few years, interest in analyzing agency administrative data from health care delivery systems to evaluate health care outcomes has been increasing (See, for example, Lasson 2000; Evans et al., 2007). Although applications of this method are limited in the evaluation of alcohol and drug (AOD) treatment, several studies have used administrative data on arrests, employment, welfare, and/or health care utilization (For a recent review, see Alterman et al., 2001; Evans et al., 2007). Several advantages of using administrative data as AOD treatment outcome measures have been identified (McCarty et al., 1998). For example, health and social services databases typically contain information on large numbers of individuals spanning many years. Administrative data also provide information on a wide range of outcomes, including employment and crime, allowing for an assessment of whether dollars spent for AOD treatment are offset by savings to society in other areas (e.g., Ettner et al., 2006; Gerstein et al., 1994; Wickizer et al., 2006). Hence, state agency administrative data may prove to be a valuable resource for the evaluation of AOD treatment outcomes.

However, cross-system data linkage, like all research and evaluation methods, is limited in certain significant ways. Perhaps most importantly, administrative databases provide no direct information concerning a person's drug use status (e.g., abstinent, using occasionally, or using regularly). Further, administrative databases typically contain few details about treatment

Correspondence: Yih-Ing Hser, Ph.D., UCLA Integrated Substance Abuse Programs, 1640 S. Sepulveda Blvd., Suite 200, Los Angeles, CA 90025. Tel: (310) 267-5388, Fax: 310-473-7885, email: yhser@ucla.edu.

processes, rendering linked data of limited use in identifying clinical "best practices" or developing clinical guidelines. Additionally, because the methodology for linking administrative data is relatively new in the evaluation research field, many issues, such as data errors, mismatch of records, and missing data, remain (Campbell, Deck, & Krupski, 2008; Saunders & Heflinger, 2004). The reliability and validity of outcome measures based on cross-system data linkage have been little scrutinized. Thus, the utility of using administrative data for outcome evaluation requires further assessment before being routinely utilized.

As part of the California Treatment Outcome Project (CalTOP), an effort to pilot test an outcome monitoring system for California's alcohol and other drug (AOD) system of care, procedures were developed for tracking clients within other California health, social service, and criminal justice management information systems. Administrative data were collected as a complementary data source to personal interview data, and major findings of the study have been reported elsewhere (Hser et al., 2003). The purpose of this article is to describe the process and procedures used in CalTOP's cross-system data linkage efforts, focusing on findings regarding the accuracy and results of CalTOP cross-system data linkage, including some limited comparisons between information derived from administrative data and self-report interview data. Following a brief review of data linkage procedures, we describe methods and findings resulting from record matching, and we conclude by offering lessons learned from this study.

Data Linkage

A link occurs when two or more records in a set of databases are determined to represent the same individual and are joined as one case in a new database. True matches describe linked records that actually represent one person. False positives describe linked records of different individuals and, conversely, false negatives describe records on the same individual that are not matched. An effective record linkage procedure is a process that maximizes true matches and minimizes false matches, however handling false matches to increase the true match rate can be problematic. When matching criteria are eased, a greater proportion of false matches could result, but when stricter criteria are imposed, the rate of true matches may decrease.

A trade-off is often established between acceptance of incorrectly matched records and maximizing the match rate. In some situations a researcher may select a lower match probability threshold in order to maximize the linkage results, thereby accepting a few incorrectly matched records. In other situations, the goal may be to maintain a higher probability threshold so that incorrectly matched records are rejected. In reality, data errors or insufficient information frequently cause data linkage difficulties. Matching records from two independent data sources is not always an unambiguous choice between a perfectly-matched and a definitely-unmatched case. The two record matching methods that are widely used, i.e., the probabilistic method and the deterministic method (for further discussion of these two methods, see Leeper et al., 2006; Saunders & Heflinger, 2004; Whalen et al., 2001), heavily rely on computerized matching routines, however human judgment is still usually required to decide either to accept or reject some linkages as true matches. Choosing an appropriate data linkage method is a key decision that largely depends on the integrity of the data and the severity of the consequences associated with false matches.

Especially in a large-scale data linkage effort, many cases may be ambiguously linked and manually classifying each case as a match or a non-match is impracticable. To resolve this problem, researchers set up a mathematical criterion, or a probability, for high-speed computers to classify the cases, separating the matches from the non-matches. This probabilistic matching method is flexible enough to allow researchers to set their tolerance level of accepting incorrect matches in order to maximize the total number of matched cases. However, setting the

mathematical criterion requires *a priori* knowledge of the distributional characteristics of the identifier values, information gathered from research experience as well as from prior analysis of the data. Each variable used in the probabilistic matching procedure is assigned a specific weight based on the prior analysis and is analogous to the weights assigned in surveys with non-equal selection probabilities. In cases without any prior knowledge, the starting value of the mathematical criterion is arbitrarily set and the linkage becomes a trial-and-error process to search for an acceptable result, a process that can significantly increase the energy, time, and cost expended.

The deterministic method can be viewed as an extreme version of the probabilistic method. This method accepts only those cases that fully fulfill the matching criteria. The deterministic method is particularly appropriate when the scrutinized outcomes have serious consequences, such as injuries, fatalities, legal issues, and treatment outcomes. The adverse trade-off is that any data errors in the identifiers (i.e., the linking variables) will result in a non-match. In effect, the deterministic method inherently has the potential to fail to link records that represent the same individual. The magnitude of under-linkage depends upon the accuracy of the identifiers in the databases to be linked in addition to the lag time between the date an event actually occurs and the date that same event is registered into a data system. The deterministic data linkage method provides a high level of linkage specificity, and cases can only be matched if they satisfy specific criteria. At the same time, this method inevitably sacrifices the sensitivity of the linkage, i.e., some cases that should be matched will not be matched due to data error.

The Present Study

In the present study, the deterministic method was used as the most conservative method to link CalTOP clients to records maintained by three state agencies. Although the data linkage process varied by database, as will be described in a later section, a combination of personal identifiers, including client's name, Social Security Number (SSN), and date of birth served as the primary linking variables. The deterministic method was expected to result in under-linkage. However, this limitation was outweighed by the high certainty of linkage associated with deterministic linkage and the corresponding level of confidence in resulting findings. In this article, first we describe the procedures taken to link CalTOP clients to each data system, and then, where equivalent measures are available, we compare information derived from data linkage to comparable information provided by client self-report.

Methods

Study Design

CalTOP was a multisite, multicounty, prospective treatment outcome study which was part of the national Treatment Outcomes and Performance Pilot Studies [see Hser et al. (2003, 2004) and Evans & Hser (2004) for a detailed description of the CalTOP study and overall findings]. Data collection began in April 2000 from all adult patients consecutively admitted to 43 substance abuse treatment programs (24 outpatient drug-free, 11 residential, 4 methadone maintenance, and 4 mixed-modality) in counties throughout California: Alameda, El Dorado, Kern, Lassen, Orange, Riverside, Sacramento, San Benito, San Diego, San Francisco, San Joaquin, San Luis Obispo, and San Mateo. These counties cover a wide geographic area (e.g., the northern, central, and southern regions of California) and include both urban and rural locales.

Subject recruitment began in April 2000 and lasted for approximately two years. All patients were assessed at intake, and a sub-sample of individuals who entered CalTOP treatment between April 1, 2000 and May 31, 2001 was assessed at 3-months and 9-months following treatment admission. Self-reported measures collected at intake and the 9-month follow-up

used the Addiction Severity Index (ASI-Lite/CF), which is a semi-structured interview that assesses problem severity in seven areas: alcohol use, drug use, employment, family and social relationships, legal status, mental health status, and medical status (McLellan et al., 1980; 1992). Relevant items in the mental health domain include number of times treated for psychological/emotional problems in inpatient or outpatient setting over lifetime, and days stayed in a hospital in past 30 days and 6 months. Items in the legal domain include number of times in the lifetime arrested for various charges and in the past 30 days and 6 months. Administrative records on all clients were obtained to provide 12-months of follow-up data. The Institutional Review Board at UCLA and at the California Health and Human Services Agency approved all study protocols.

Study Sample

CalTOP client data collected at treatment admission were used as the primary source for cross-system data linkage. Personal identifiers used for linkage included SSN, name, date of birth, and other supporting information (e.g., sex, race/ethnicity). Linkage efforts involved a sample of 6,545 clients who were consecutively enrolled in CalTOP treatment programs during the first year of the project (April 2000 to May 2001). The basic demographic and socioeconomic statistics of the sample are as follows: 45% female, 53% white, 23% Hispanic, 17% African American, and 7% of other race. Approximately 43% of the clients were reportedly not in the labor force (i.e., retired, disabled, etc.) at the time of admission; about one–third were employed; 45% were single or never married; and nearly 80% had attained 12 or less years of formal education.

Administrative Data Sources

CalTOP client data were linked to the following three databases: the California Department of Motor Vehicle (DMV) master file, the Automated Criminal History System (ACHS) of the California Department of Justice (DOJ), and the Client and Service Information (CSI) database of the California Department of Mental Health (DMH). As reported in this article, the resulting linked file provided detailed information on five events: (1) arrests for driving under the influence; (2) motor vehicle crashes, injuries, and fatalities; (3) arrest history; (4) incarcerations and sentences; and (5) mental health services utilization.

Personal client identifiers collected by CalTOP treatment providers at intake made it possible to track individuals through the CalTOP data system as well as through other administrative databases. Identifying elements included: first and last name, gender, date of birth, SSN, place of birth, and mother's first name.[1] Only personal identifiers that both the CalTOP database and an administrative database had in common were used for cross-system data linkage, most commonly SSN, date of birth, and name. All administrative records are event-based (e.g., a record is created for each occurrence of the event). Data linkage was conducted to search for information on events that occurred any time before clients' CalTOP admission and 12 months post-admission.

Each agency reported using routine data management and validation processes to verify the accuracy of their data however, due to space limitations, we provide only minimal information on those procedures in this manuscript. Detailed information on the procedures used to link data is provided in the results section for each data source.

[1]These elements were based on the Family Health Outcomes Project from UCSF as presented to the California Substance Abuse Research Consortium (10/97). Gender, date of birth, mother's first name, client's first name and patient's last name are considered primary elements. The other elements such as SSN or place of birth were to be confirmatory elements. Middle initial was not considered.

Results

We report, separately for each of the three databases, the accuracy of record linkage as well as the prevalence of major events abstracted from the process. We conducted prevalence analysis at 12-month intervals before and after the CalTOP treatment admission date. The event date (e.g., the arrest date in the ACHS data, or the motor vehicle accident date in the DMV data) was used to determine whether or not an event should be included in the study periods. We also report results of comparisons between equivalent measures available from both administrative data and self-report data. The presence (yes/no) and number of events of interest were counted by the number of relevant records for that individual during the specified time period for each relevant analysis. Comparisons between self-report and administrative records were conducted by analyzing the percent of agreement for dichotomized variables (number of presence or absence of an event determined by both sources over the total sample) and by t-test for continuous variables (number of times of the event during the specified time period).

Motor Vehicle Driving Records

The DMV data include the driving records of all licensed California residents. Data are retained per legal requirements related to licensure. The longest retention period is ten years for information on drunk driving incidents while information on accidents and minor violations is retained for 39-months. A substantial lag exists between the date of a driving violation citation and the date that event is reflected in the DMV data. Additionally, the actual lag time between a violation and a record of conviction varies from court to court. The average time to adjudicate a DUI violation to conviction is three months, and another three months is needed for a conviction to appear in the DMV data. To address these issues, during the data linkage process we allowed for a three-month lag, which may have undercounted recent convictions.

Procedures used in data linkage—The personal identifier information maintained in DMV records is relatively up-to-date and accurate due to the regular renewal processes for both drivers' licenses and motor vehicle registration. The personal identifiers common to both CalTOP and the DMV that were used for linking the two databases included SSN, name, and date of birth. (Driver's license numbers, which might have resulted in the most effective and efficient link, were not collected from CalTOP clients.) Because 1.5% of SSNs were missing from the CalTOP database and also due to the fact that at the time data linkage was conducted not all drivers' SSNs were available in the DMV data, DMV staff, in consultation with UCLA/ISAP staff, linked data in stages. First, records were linked by name and date of birth. CalTOP records not linked in stage one were returned to UCLA/ISAP staff who checked for data accuracy against various information sources. In some cases, several possible versions of the name (e.g., Robert, Bob, Bert, Robbie) were generated to maximize data linkage outcomes. The improved data file was re-linked by name and date of birth. At both stages, multiple linkages or duplicates were generated due to data errors or insufficient information and these duplicates were examined and removed upon further manual examination by UCLA staff.

Overall Matching Rate and Accuracy—Of 6,545 CalTOP clients included in data linkage, 5,725 (or about 87.5%) matched to a DMV driving history record, a satisfactory record linkage rate, especially considering that only about 83.0% of Californians of driving age had driver's licenses in the year 2000 (U.S. DOT, 2001). Quality control procedures examined the accuracy of the linkage by cross-validating personal information obtained from the DMV against the same information (e.g., sex) from the CalTOP admission records. A computerized process followed by a manual examination revealed 649 questionable links and confirmed that 41 of the 5,725 linked records may have been matched incorrectly. However, there was insufficient evidence to determine that these cases were mis-linked, so we included them in

analyses as linked data and we report this uncertainty as a potential source of error. The error rate for this data linkage process was estimated at ± 0.7% (41/5,725).

Outcomes Based on Motor Vehicle Driving Records—Linkage between CalTOP and DMV data files provided information in multiple dimensions, two of which have serious public safety consequences: driving under influence (DUI) and car accidents. The linked records indicated that in the 12 months prior to treatment admission, the DUI rate among CalTOP clients was 5.3%, five times the rate for all California drivers in 1997. Compared to pre-treatment admission, the rates of DUI arrest were consistently lower during post-admission, and across all four time periods examined (2.0% vs. 0.8% for the 3-month pre- and post-admission periods, 3.5% vs. 1.5% for the 6-month period, 4.5% vs. 2.3% for the 9-month period, and 5.3% vs. 2.7% for the 12-month period). In terms of motor vehicle accidents, injuries, and fatalities, compared to the year before treatment admission, there appeared to be a slight decrease in motor vehicle accidents (from 5.6% to 4.4%) and also in the number of injuries (from 173 to 108) during the year after treatment admission.

To summarize, the DMV data comprise an up-to-date and accurate database. Mainly based on name and date of birth, the data linkage process produced a satisfactory link rate, matching about 88% of CalTOP cases with their DMV driving history records. For those linked cases, the error rate was estimated to be very low. During the 12 months before the CalTOP admission, about 5% of the clients were arrested for a DUI, 6% had accidents, 173 accidents resulted in injuries, and 3 ended in death. During the 12 months after the CalTOP admission, about 3% were arrested for a DUI, 4% had accidents, 108 accidents resulted in injuries, and 4 ended in death. Overall, treatment participation appeared to be associated with reduced motor vehicle accidents and injuries among CalTOP clients.

Criminal Identification and Information (CII)

CalTOP clients were linked to databases maintained by the California Department of Justice (DOJ) to obtain information on criminal justice events that occurred any time prior to the date of data linkage. The search covered all Criminal Identification and Information (CII) data within their Automated Criminal History System (ACHS), a database that contains all misdemeanor and felony arrests that occurred from 1977 forward. Compared to the other administrative data systems examined, a relatively shorter lag period exists from the time a criminal justice event occurs to the time the information is registered in the database. We assumed a three-month lag period in the data linkage process.

Procedures Used in Data Linkage—One of the important features of ACHS is that it maintains all of the different known names, aliases and SSNs for individual offenders. Analysis revealed that many offenders had multiple names and SSNs. This multiplicity of information complicates data linkage efforts, but, if handled properly, this information also provides a unique opportunity to search records thoroughly to maximize match rates. Based on name, date of birth and an "unknown" sex, DOJ staff searched the criminal history database for matches. A hit was considered a match if the exact last name (including aliases), first initial, and date of birth was found. If only one record was found, that record was considered matched and the subject’s criminal history transcript was generated. If more than one record was found, a search by SSN (if one was provided in the CalTOP file) was conducted against all records found for that individual. If the SSN was found in only one record, that record was selected as a match. If the SSN was found in more than one record, the case was categorized as "too many to identify" and treated as an unmatched case. This linkage procedure employed an exhaustive search, meaning all possible matches were tested, making the possibility of under-linkage very low.

Overall Matching Rate and Accuracy—Of the 6,545 CalTOP clients included in data linkage, 4,919 (about 75%) matched to at least one record indicating an arrest, detention, citation, or custody in their lifetime. Detailed examination of the matched cases showed that SSN alone could have linked about 85% of individuals to their criminal records. Of those not linked by SSN, about 14% could have been linked by name. The remaining 1% could have been linked by date of birth and gender.

Both name and SSN alone are unreliable personal identifiers for linking to criminal information. Of the linked cases, almost all reported more than one name for an individual. Twelve CalTOP clients had registered more than 32 names, with an extreme case having 73 different names. Furthermore, about 7% of the linked cases had no SSN in the criminal history records. Table 1 presents information on how often different SSNs were recorded in the CII data for CalTOP clients. While SSNs are considered to be a highly reliable piece of personal identification, nearly 32% of the criminal cases recorded had more than one SSN. These results indicate that linkage to criminal databases is not reliable if it is based solely on either name or SSN. Successful linkage to criminal histories depends on the use of multiple personal identification variables.

Quality control procedures were conducted to examine the accuracy of linkage by cross-validating personal information obtained from the criminal history record against the corresponding personal information (e.g., sex, race/ethnicity) included in the CalTOP intake data. A computerized process followed by a manual examination revealed 380 questionable links and confirmed that 64 out of the 4,956 linked records may have been linked incorrectly. However, we did not have definite evidence to determine that these cases were really mis-linked, therefore we included them as linked cases for analyses and we report this uncertainty as a potential source of error. The error rate for this data linkage process was estimated at ± 1.3% (64/4,956).

Outcomes Based on Criminal Justice Records—Record linkage between CalTOP and DOJ criminal history records provided a wide range of information in four major areas of criminal activity: arrest, detention, and citation; incarceration and supervision; probation; and detailed court actions. Arrest and incarceration are the two most frequent criminal events registered in the ACHS database. Almost every linked case had multiple records of arrest. As an example, in the most extreme case, one individual was arrested 90 times (for a total of 192 offenses) between July 1972 and December 1999. During the 12 months before the CalTOP admission, 44% of CalTOP clients were arrested, which is 11 times greater than the arrest rate observed in the general adult population (4% in 1997). Additionally, the linked criminal history data indicated that during the year before treatment admission, about 5%, or 312 clients, were sentenced to prison 384 times. During the 12-month post-treatment admission period, the arrest rate decreased to 30% and the incarceration rate increased slightly to 6%.

In summary, the DOJ criminal history database is a rich but complex research resource. The most important feature of the database is that it maintains all of the different names, aliases and SSNs that have ever been associated with an offender, which constitutes a unique and valuable source for enhancing matching processes. About 75% of CalTOP clients were linked to at least one record indicating arrest, detention, citation, or custody during their lifetime. The record-matching rate may have been slightly diminished due to various factors. During the 12 months before the CalTOP admission, about 44% of clients were arrested and 5% were incarcerated. During the 12 months after the CalTOP admission, about 30% of clients were arrested and 6% were incarcerated. Overall, treatment participation appeared to be associated with reduced criminal justice system involvement among CalTOP clients.

Mental Health Services Data

CalTOP data were linked to the Department of Mental Health's (DMH) Client and Service Information (CSI) database for information on mental health services received in the three and half years starting from July 1998. The database includes all Medi-Cal and state-funded mental health services, including information on 24-hour day services and outpatient programs, but excluding state hospitals. Over half a million consumers of mental health services and about 45 million service records are included in this database. Of California's 58 counties, three had not submitted data to the database by the linkage date. However, this omission did not affect our data linkage outcomes because these three counties did not participate in the CalTOP study.

Procedures Used in Data Linkage—We used personal identifiers common to both CalTOP and CSI to link records, specifically SSN, date of birth, and last and first name. Cognizant of erroneous and missing information in the personal identifiers in both databases, DMH staff, in consultation with UCLA/ISAP staff, employed a two-step strategy to improve the specificity of the data linkage. Records were first linked by SSN and date of birth, and then unlinked records were fed into a program for linkage by first and last name and date of birth.

Overall Matching Rate and Accuracy—The CSI system is an event-based database, i.e., each service received is registered as a record. Therefore, if a client receives more than one service, that client can have multiple records. The data linkage process matched 101,988 records to 6,545 of the subjects included in the search. About 30% of clients received mental health services at least once from hospitals or other mental health providers during the period. All of the linked cases were included in a quality control procedure that involved a computerized check followed by manual examination. It was found that 2.6% of the clients sent for record linkage were over-linked, mainly due to missing SSN, as well as by a duplication of records in an early version of the input data file. When SSN was not available, name and birth-date were used for linkage, which increased the possibility of over-linking one CalTOP client to more than one DMH client with the same name and birth date.

Outcomes Based on Mental Health Services Utilization—A total of 1,960 received mental health services from county hospitals or clinics during the 3.5 year time period when mental health service records were searched for CalTOP clients. As an average, 58% of these clients received at least one mental health service every six months, 21% of them sought services every month, 6% accessed services every week, and about 1% reportedly received services every other day. Within the 12-month period before CalTOP treatment admission, about 3.9% of 6,545 CalTOP clients received in-patient mental health services at county hospitals and 19.7% received services through outpatient clinics. Within the 12-month period after CalTOP treatment admission, about 2.7% of 6,545 CalTOP clients received in-patient mental health services at county hospitals and 21.3% received services through outpatient clinics. Compared to the pretreatment period, there appeared to be a slight decrease in in-patient mental health services and an increase in outpatient services after CalTOP admission among these clients.

In summary, the CSI system includes all Medi-Cal and state-funded mental health services in county hospitals or clinics. Data linkage was based on SSN, date of birth, and name. About 30% of CalTOP clients received mental health services at least once from county hospitals or clinics during the 3.5-year period observed via data linkage. Within the 12-month period prior to CalTOP admission, about 3.9% of the CalTOP clients received inpatient mental health services and 19.7% received services through outpatient clinics. As an average, less than one percent of these clients sought mental health services every other day, 6% of them sought services at least every week, and 21% at least every month. Participation in treatment at

providers in the CalTOP system appeared to be associated with a slight decrease in in-patient services and a slight increase in outpatient services.

Comparisons of Administrative and Interview (i.e., Self-Report) Data

CalTOP collected outcome measures using self-reported data through interviews and also data from official records through cross-system data linkage. We compared measures taken at admission that were available from both sources and the results are provided in Table 2. Self-reported measures relied on ASI items and variables from administrative data were analyzed to provide an equivalent measure of the event (e.g., inpatient, outpatient treatment), its frequency (e.g., times, days), and observation period (e.g., 30 days, 6 months, lifetime). Comparisons in mental health services utilization generally showed greater percentages of people reporting utilization but in lower frequencies by self-report than was recorded in administrative records. In addition to recall bias, several factors could explain the observed discrepancies. The mental health services data only covered a 3.5 year time-period and it included only those facilities that received public funding. Thus, DMH data on lifetime measures of psychiatric treatment were incomplete and thus resulted in an undercount. The higher levels of services utilization in the self-report data could be due to services received from private facilities that do not report their data to DMH. However, self-reports appeared to considerably undercount the number of times ever in psychiatric treatment, while this discrepancy was smaller in both the past 6 months and the past 30 days.

Regarding criminal justice measures, compared to the DOJ arrest records, the self-reported rates of arrest were higher over the lifetime, lower in the past 6 months, and almost equivalent in the past 30 days. This pattern was also reflected in the measure of agreement, showing increasing concordance over time periods (from 79% for the lifetime measure, 82% for the 6 months, to 92% for the past 30 days). Again, the frequency of arrest, however, was consistently underreported by self-report across the three time periods. We could not compare administrative and self-reported incarceration data because the two datasets used different definitions of incarceration: the self-reported data included any detention or incarceration, including same-day arrests and releases, while DOJ records included any incarceration in jail or prison.

Discussion

Administrative data represent an alternative, or at least complementary, method for substance abuse treatment outcome evaluation. CalTOP's cross-systems database linkages provided reasonable information on outcomes related to motor vehicle driving incidents, criminal history, and mental health services utilization. The overall pattern of improvements in criminal justice involvement during the post-treatment period relative to the pre-treatment period is consistent with those observed based on interview or self-report data. One significant finding from CalTOP cross-system data linkages is the change in the arrest rate from 44% during the year prior to treatment admission to 30% during the year after. Rates of DUI arrests and motor vehicle accidents also showed some reductions. Other reductions, such as inpatient mental health services, are relatively minor. These findings are quite consistent with what has been reported by other studies (Claus, Orwin, Kissin, Krupski, Campbell, & Stark, 2007; Green, Rockhill, & Furrer, 2007; Luchansky, Nordlund, Estee, Lund, Krupski, & Stark, 2006; TOPPS II Interstate Cooperative Study Group, 2003). Future efforts to include other databases (e.g., healthcare, employment) can provide a more comprehensive assessment of treatment outcomes.

The experiences of matching CalTOP clients to administrative records demonstrate that such data are valuable resources for the evaluation of AOD treatment outcomes. Of the three administrative databases from which CalTOP acquired data, the DMV database provided the

most accurate linking results, in part because the personal identifier information was up-to-date and accurate as a result of the regular renewal processes for drivers' licenses. The Criminal Identification and Information (CII) data system, maintained by DOJ, was the most sophisticated and complete data source for identifying a particular person, resulting in a low under-linkage probability, which is a critical feature for a data-linkage process based on personal identifiers. It is difficult to assess uncertainties regarding linkage with the mental health service database because information needed to make such an assessment was unavailable.

Comparisons between administrative and self-reported data, although limited, provide some basis for a better understanding of the strengths and weaknesses of different data sources and data collection methods. Sources of biases associated with self-report (e.g., recall error, intentional under-reporting or over-reporting) are better known (Anglin et al., 1993; Cherpitel et al., 2007; Langenbucher & Merrill, 2001; Vitale et al., 2006), while little is known about the reliability and validity of measures resulting from record-matching. Our analyses have suggested that, taking into consideration coverage in length of time and reporting facilities (e.g., publicly funded programs), outcome measures obtained from cross-system data linkage provide equivalent results in some measures (e.g., percent of people using mental health services, or percent ever arrested) and perhaps more accurate results in other measures (e.g., frequency of service utilization, and frequency of arrests), compared to self-report. Future analyses should also examine if the degree of biases from different data sources vary according to population or different characteristics. Further examination of individuals with "incorrect matches" might also yield useful information on whether linkage rates are associated with particular characteristics or are missing at random.

There are benefits to collecting self-reported information, particularly when studying a wide variety of stigmatized behaviors (e.g., illicit drug use, criminal activity, HIV risk behavior) that can often only be assessed by talking directly with research participants. Administrative data are instrumental for observing system-wide impacts of large policy changes (i.e., changes in managed care, or Medicaid enrollment policies) that result in unexpected or unintended consequences (Deck et al., 2002; McFarland et al., 2005; Bray et al., 2007). All data presents some kind of measurement error. One way in which self-reported data and administrative data are complementary is that they speak to somewhat different aspects of the same phenomenon and thus they may be used in combination to establish and verify the frequency and timing of events, lengthen observation periods, and, strengthen study designs by providing information on events, individuals, and system impacts that might not otherwise be revealed when using just one data source. Using different sources of complementary data allows researchers to triangulate the issue under study thereby providing options for understanding and minimizing potential biases due to measurement errors.

The accuracy of CalTOP linkage efforts may have been limited by several factors (see Table 3). An individual might have moved out of the geographic area covered by a particular reporting agency. Thus, failure to match, for example, a criminal record for a CalTOP client in the post-treatment period, may have resulted in an overestimate of treatment effectiveness. Also, we used the deterministic linkage method, as opposed to the probabilistic method, because we had sufficient personal identification information to allow for decisions regarding the accuracy of matches. However, because stringent matching criteria were applied, the deterministic method likely underestimated the match rate. It must be emphasized that accurate record linkage results rely more on the accuracy of the personal identifiers used as linking variables and less on the linking strategy itself. This was particularly true in criminal history records when offenders reported many inaccurate SSNs and names. Finally, rules governing record reporting and preservation influence the completeness and interpretation of findings. For example, although the DMV data contains detailed and accurate information about a driver's traffic safety history,

the information obtained through data linkage may be artificially truncated by data retention and reporting policies and thus may be incomplete for research purposes. For example, regarding one-point violations, if it is the first violation and the driver has not had another violation in the previous 18 months, then the driver is allowed to attend traffic school and have the violation masked (i.e., the violation never appears) in the DMV record system. Additionally, information retention periods vary for different violations. Studies on the long-term effect of drug treatment need to approach DMV data cautiously because some serious driving violations are not retained for more than seven years. It should be stressed that CalTOP used pre- and post-treatment comparisons. Thus, biases caused by inaccurate record matching apply to both periods, and have minimal impact on the pre- and post-admission differences that are used for outcome evaluation. Finally, the focus of this paper is limited to reporting on the accuracy and results of information extracted from administrative data sources and thus we do not report in detail on changes in self-reported behaviors pre vs. post-treatment, a topic that has been addressed elsewhere (Hser et al., 2004;2005).

In summary, these limitations do not warrant discarding data linkage as an outcome evaluation method, but instead highlight the care that must be employed when examining, accepting, and interpreting results. Successful data linkage depends on the interplay between strategy, methodology, and the quality of data. It is important to thoroughly examine factors contributing to linkage uncertainty in order to interpret and utilize the results appropriately. While the CalTOP record linkage outcomes appeared reasonable, the accuracy of record matching could be improved in the future. Below, we summarize our experience regarding data matching accuracy to inform and improve future efforts at cross-system data linkage.

Evaluating the precision of data linkage outcomes is complicated by several factors and the most influential factor affecting matching accuracy is the quality of personal identification information. In linking CalTOP data with other administrative databases, we determined that the most common personal identifier elements are SSN, name, and date of birth, and using these elements in combination with one another increased matching accuracy. Although CalTOP collected several unique personal identifiers, some items were not collected or recorded consistently. For example, 1.5% of CalTOP records were missing SSN and another 1% had more than one record in the CalTOP database and some of these duplicates remained undetected and were accidentally counted as different individuals. This missing and duplicative data reduced the likelihood of correct record matching both within and across databases.

Furthermore, some personal identifiers may not be sufficiently unique to distinguish one individual from another within a system while other identifiers may not be shared by any other system (e.g., mother's first name). Personal identifiers containing missing information or errors caused by human factors, such as data entry errors and deliberate falsification of personal information, also create matching problems. Added to this concern, in computer-based record linkage processes, software program will indiscriminately treat such data errors as correct. Application of multi-staged iterative linking procedures allowed us to accommodate the unique characteristics of specific administrative databases and to maximize record-matching with those databases. This strategy, in combination with good working relationships with agency staff, enhanced record matching rates. Finally, to minimize opportunities for under-linkage, sufficient time needs to be allocated to allow for data reporting to be completed. Information regarding other problems and biases incurred when linking two databases could not be assessed as all linkages were performed by agency staff.

The magnitude of data linkage uncertainty varies from database to database. But by understanding relevant issues, applying acceptable reasoning and analytic methods, and checking the consistency of findings with empirical data, the accuracy of record linkage can be evaluated. Considering the CalTOP linkage strategy and the nature of data involved, and

assuming minimum data errors, we believe that the CalTOP cross-system data linkage efforts produced reasonable estimates, but may, to some extent, be biased towards underestimation. However, because CalTOP used pre- and post-treatment comparisons, biases caused by inaccurate record matching apply to both periods, and therefore have minimal impact for outcome evaluation.

The process for accessing and obtaining administrative data is time consuming and labor intensive. Each agency has its own application process, regulations and rules for data sharing and security procedures. By working with individual agencies to satisfy their specific requirements and preferences, we successfully received data from several agencies. Data were shared at no financial cost, however, on average, it took one to two years from initial data request to receipt of data. In some cases data sharing was facilitated by our prior experiences with particular agencies, the ability of our research findings to address the source agency's aims and mission, and also by our partnership with the California Department of Alcohol and Drug Programs. Concerns about confidentiality and workload were obstacles to agency cooperation and caused delays in data delivery (for a summary of other considerations related to data sharing, see Duran et al., 2005). A considerable amount of time and resources were devoted to navigating regulatory and procedural challenges posed by acquiring administrative data. Given the current environment placing greater and stringent constraints on releasing personal information (Brady, Grand, Powell, & Schnik, 2001; Hotz et al., 1998), it may be increasingly difficult to obtain cooperation from agencies. Before making administrative data a key component of future studies, researchers are advised to carefully weigh the benefits to be gained from analyzing administrative data against the resources required to obtain and derive adequate information from it.

The methodology for using cross-system linkage to support outcome evaluation is still in an early stage. Although, as the first study to make direct comparisons between the accuracy of information derived from self-reported versus administrative data sources, the present study provides promising supportive evidence, further research is required to improve the methodology in order to ensure more accurate findings. Additionally, some measures (e.g., drug use) that are important for clinical and policy decisions (e.g., best clinical practice) cannot be obtained by record linkage. Thus, an outcome monitoring system that combines data from both self-report and administrative records is still necessary.

Acknowledgments

This research was supported in part by the California Alcohol and Drug Programs (contract No. 98-00245) and NIDA (grant P30DA016383). Dr. Hser is also supported by a Senior Scientist Award (K05DA017648). The content of this publication does not necessarily reflect the views or policies of the funding agencies. Special thanks are due to Dr. Haikang Shen and other staff at the UCLA Integrated Substance Abuse Programs for data collection, data analysis, and manuscript preparation.

References

Alterman AI, Langenbucher J, Morrison R. State-level treatment outcomes studies using administrative databases. Eval Rev 2001;25(2):162–183. [PubMed: 11317715]

Anglin MD, Hser YI, Chou C. Reliability and validity of retrospective behavioral self-report by narcotics addicts. Eval Rev 1993;17(1):91–108.

Brady, H.; Powell, A.; Grand, S.; Schink, W. Access and confidentiality issues with administrative data. National Research Council Studies of Welfare Populations: Data collection and research issues. Washington: National Academics Press; 2001.

Bray J, Vandivort R, Dilonardo J, Dunlap L, Schroeder D, Forhan C, Miller K. Healthcare utilization of individuals with opiate use disorders: An analysis of integrated medicaid and state mental health/substance abuse agency data. J Behav Health Serv Res. 2007

Campbell, Deck, Krupski. Record linkage software in the public domain: A comparison of link plus, the link king, and a “basic” deterministic algorithm. Health Informatics Journal 2008;14(1)

Cherpitel CJ, Ye Y, Bond J, Borges G, Macdonald S, Stockwell T, et al. Validity of self-reported drinking before injury compared with a physiological measure: Cross-national analysis of emergency-department data from 16 countries. J Stud Alcohol Drugs 2007;68(2):296–302. [PubMed: 17286349]

Claus RE, Orwin RG, Kissin W, Krupski A, Campbell K, Stark K. Does gender-specific substance abuse treatment for women promote continuity of care? J Subst Abuse Treat 2007;32(1):27–39. [PubMed: 17175396]

Deck DD, McFarland BH. Medicaid managed care and substance abuse treatment. Psychiatr Serv 2002;53 (7):802. [PubMed: 12096160]

Duran, F.; Wilson, S.; Carroll, D. Putting administrative data to work: A toolkit for state agencies on advancing data integration and data sharing efforts to support sound policy and program development. Farmington, CT: Child Health and Development Institute of Connecticut; 2005.

Ettner SL, Huang D, Evans E, Ash DR, Hardy M, Jourabchi M, et al. Benefit-cost in the California Treatment Outcome Project: Does substance abuse treatment "pay for itself?”. Health Serv Res 2006;41(1):192–213. [PubMed: 16430607]

Evans E, Hser YI. Pilot-testing a statewide outcome monitoring system: Overview of the California Treatment Outcome Project (CalTOP). J Psychoactive Drugs, SARC 2004:109–114.

Evans E, Grella CE, Murphy AB, Hser YI. Using administrative data for longitudinal substance abuse research. Under review. 2007

Gerstein, DR.; Johnson, RA.; Harwood, H.; Fountain, D.; Suter, N.; Malloy, K. Evaluating recovery services. The California Drug and Alcohol Treatment Assessment (CALDATA.). Sacramento, CA: State of California Department of Drug and Alcohol Programs; 1994.

Green BL, Rockhill A, Furrer C. Does substance abuse treatment make a difference for child welfare case outcomes? A statewide longitudinal analysis. Child Youth Serv Rev 2007;29(4):460–473.

Hotz, VJ.; Goerge, R.; Balzekas, J.; Margolin, F. Administrative data for policy-relevant research: Assessment of current utility and recommendations for development. A report of the advisory panel on research uses of administrative data. Northwestern University/University of Chicago: Joint Center for Poverty Research; 1998.

Hser Y-I, Evans E, Huang D. Treatment outcomes among women and men methamphetamine abusers in California. J Subst Abuse Treat 2005;28:77–85. [PubMed: 15723735]

Hser YI, Evans E, Huang D, Anglin MD. Relationship between drug treatment services, retention, and outcomes. Psychiatr Serv 2004;55(7):767–774. [PubMed: 15232015]

Hser YI, Evans E, Teruya C, Hardy M, Ettner S, Urada D, et al. The California Treatment Outcome Project (CalTOP) Final Report. Submitted to the California Department of Alcohol and Drug Programs. 2003

Langenbucher J, Merrill J. The validity of self-reported cost events by substance abusers. Limits, liabilities, and future directions. Eval Rev 2001;25(2):184–210. [PubMed: 11317716]

Lasson, MJ. Planning for data linking to enhance your study. Presented at the treatment outcomes and performance pilot studies enhancement TOPPS II grantees meeting. Center for Substance Abuse Treatment; Orlando: Florida; 2000 Mar 13–14.

Leeper, TE.; Heil, SKR.; Nalty, D. Using State administrative records to manage substance abuse treatment system performance. Rockville, MD: U.S. Department of Health and Human Services: Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment; 2006.

Luchansky B, Nordlund D, Estee S, Lund P, Krupski A, Stark K. Substance abuse treatment and criminal justice involvement for SSI recipients: Results from Washington state. Am J Addict 2006;15(5):370–379. [PubMed: 16966193]

McCarty D, McGuire TG, Harwood HJ, Field T. Using state information systems for drug abuse services research. Am Behav Sci 1998;41(8):1090–1106.

McFarland BH, Deck DD, McCamant LE, Gabriel RM, Bigelow DA. Outcomes for medicaid clients with substance abuse problems before and after managed care. J Behav Health Serv Res 2005;32(4): 351–367. [PubMed: 16215446]

McLellan AT, Luborsky L, Woody GE, O'Brien CP. An improved diagnostic evaluation instrument for substance abuse patients: the Addiction Severity Index. J Nerv Ment Dis 1980;168:26–33. [PubMed: 7351540]

McLellan AT, Cacciola J, Kushner H, Peters F, Smith I, Pettinati H. The fifth edition of the Addiction Severity Index: cautions, additions and normative data. J Subst Abuse Treat 1992;5:312–316.

Saunders RC, Heflinger CA. Integrating data from multiple public sources: Opportunities and challenges for evaluators. Evaluation: Int J Theory, Res, Prac 2004;10(3):349–365.

TOPPS-II Interstate Cooperative Study. Drug treatment completion and post-discharge employment in the TOPPS-II Interstate Cooperative Study. J Subst Abuse Treat 2003;25(1):9–18. [PubMed: 14512103]

U.S. Department of Transportation. 2000 national statistics. Washington, DC: U.S. Department of Transportation; 2001.

Vitale SG, van de Mheen H, van de Wiel A, Garretsen HFL. Substance use among emergency room patients: Is self-report preferable to biochemical markers? Addict Behav 2006;31(9):1661–1669. [PubMed: 16446045]

Whalen, D.; Pepitone, A.; Graver, L.; Busch, JD. Linking client records from substance abuse, mental health and Medicaid state agencies. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration; 2001.

Wickizer TM, Krupski A, Stark KD, Mancuso D, Campbell K. The effect of substance abuse treatment on Medicaid expenditures among general assistance welfare clients in Washington state. Milbank Q 2006;84(3):555–576. [PubMed: 16953810]

Table 1

Frequency distribution of social security numbers in the criminal justice records

Different SSNs	Number of Frequency	Percent
1	3140	68%
2	931	20%
3	329	7%
4	132	3%
5–15	94	2%

Table 2

Comparisons of Administrative Data and Self-reported Data

	Administrative Data % or Mean (SD)	Self-reported Data % or Mean (SD)	Agreement %
Department of Mental Health Data*			
Ever in psychiatric treatment, %	24.0	38.8	72.1
Mean number of times	7.8 (40.5)	2.3 (8.6)	
Ever in inpatient treatment, %	5.1	22.2	81.4
Mean number of times	0.1 (0.8)	0.8 (3.8)	
Ever in outpatient treatment, %	23.9	30.3	72.7
Mean number of times	7.7 (40.1)	1.5 (7.2)	
Treated in hospital in past 6 months, %	2.5	5.4	95.9
Mean number of days	0.4 (5.8)	0.5 (5.7)	
Treated in hospital in past 30 days, %	1.3	2.7	97.7
Mean number of days	0.1 (1.4)	0.2 (1.4)	
Department of Justice Data			
Ever arrested, %	75.7	82.6	79.0
Mean number of arrest	10.7 (15.0)	5.7 (10.3)	
Arrested in past 6 months, %	40.7	37.1	81.5
Mean number of arrest	0.90 (1.8)	0.49 (1.9)	
Arrested in past 30 days, %	8.2	8.4	92.2
Mean number of arrest	0.13 (0.6)	0.09 (0.6)	

* Note that administrative data on mental health services only covered a 3.5 year time-period, thus restricting the observation period for analysis of events that “ever” occurred.

Table 3

Factors affecting linking accuracy

Factors that could cause under-linkage of records:
- The deterministic method, which adheres to stringent matching criteria
- Incompleteness of the targeted agency databases
- Missing information in both the CalTOP file and the targeted databases
- Lag time between events and reports

Factors that could cause over-linkage of records:
- The probabilistic method, which tolerates a percentage of false positives
- Insufficient personal identification or duplicate records counted as different cases

Factors that could cause either under- or over-estimation:
- Data error