



Prescription Drug Monitoring Program Training and Technical Assistance Center

# Technical Assistance Guide

## No. 03-13

# Patient Linking Software Options

Brandeis University The Heller School  
For Social Policy and Management

---

Telephone: (781) 609-7741 | Fax: (888) 705-8053 | Email: [info@pdmpassist.org](mailto:info@pdmpassist.org) | Website: [www.pdmpassist.org](http://www.pdmpassist.org)

August 16, 2013

## Introduction

Data from Prescription Drug Monitoring Programs (PDMPs) is increasingly being used by prescribers, dispensers and other authorized stakeholders. Furthermore, the proliferation of interstate data sharing and the increased interest for integration of PDMPs with Health Information Exchanges (HIEs) has placed an expectation and demand on PDMPs to ensure that complete and accurate data is being provided to authorized users. While PDMPs are addressing the quality of their data through the implementation of validation standards, error resolution procedures and technology, the ability to accurately and confidently match patient information continues to be a focus of PDMPs. To satisfy this demand for completeness and accuracy, patient linking programs are increasingly becoming a critical component of many PDMPs.

Patient linking matches the same patient information within a single dataset or across several datasets where variations in data may be present. These variations include, in part: use of nicknames, differences in spelling of patient names or addresses, use of abbreviations, varied identifiers for patients, lack of matching fields, differences in field lengths, typographical errors, and incompatible date formats or data coding.

Examples of data variations:

- Patient name – William, Will, Willy, Bill, Billy
- Patient address – S., S, South; Highway, Hwy; 12<sup>th</sup>, Twelfth; Drive, Dr
- Typographical errors – mistyped dates of birth, patient names/addresses, or transposed characters
- Date of Birth – mmddyyyy, yyyyymmdd
- Data Formats – text, binary
- Data coding - M for male, 1 for male

Patient linking programs enable PDMPs to match all the different variations together and identify the person as being the same.

Some of the benefits of patient record linking are efficient and accurate compilation of a patient's entire prescription history within a PDMP, ability of connecting patient prescription information to other PDMPs, and accurate de-identification of patient health information data for researchers to analyze.

The PDMP Training and Technical Assistance Center (TTAC) has produced this guide for PDMP administrators to facilitate their efforts to identify and implement a patient linking program that links patient records with a high degree of confidence and little manual intervention. This guide provides a sampling of available commercial and public domain programs. The listed programs were obtained from two (2) comprehensive evaluations on patient linking (see Software Evaluations page 4). The TTAC does not endorse nor recommend any specific package or methodology. PDMP administrators are encouraged to research all options to determine which will be the most effective for their state and consult with their IT staff before adopting any patient linking program. The TTAC will continually update the guide as more information on patient linking becomes available. The TTAC encourages PDMP administrators to provide the TTAC with other linking programs they are utilizing, so they may be included in this guide.

### **Software Options** *(in alphabetical order)*

- *Link King* is public domain software combining probabilistic record linkage and deterministic record linkage protocols. The probabilistic protocols were adapted from an algorithm that was developed by MEDSTAT for the Substance Abuse and Mental Health Services Administration's (SAMHSA) Integrated Database Project. The deterministic protocols were developed at Washington State's Division of Alcohol and Substance Abuse. This software requires a SAS license, but no SAS programming experience; employing a graphical user interface (GUI) to guide the user. View the [website](#).
- *Link Plus* is public domain software using probabilistic record linkage protocols developed at the Division of Cancer Prevention and Control in support of the National Program of Cancer Registries at the Centers for Disease Control. Even though Link Plus was originally developed to be used by cancer registries, it can be used on any data in fixed width or delimited format. The program is written in Microsoft Visual Basic and .NET Framework. View the [website](#).

- *LinkageWiz* is commercially available software with features for record linkage, de-duplication, and data cleaning. It uses probabilistic algorithms on common identifiers from data sets in a variety of formats. It does not require the use of external software programs, such as SAS. View the [website](#).
- *DQ Global Match* is commercially available software with features for record linkage, de-duplication, and data cleaning. It is designed to link directly with a dataset without requiring export/import of data. The software contains extensive, advanced standardization libraries to facilitate record linkage. – View the [website](#).

## Software Evaluations

- The *California HealthCare Foundation* sponsored research in 2004 to assess commercially available record linking software packages. View the [evaluation](#).
- The *Washington State Division of Alcohol and Substance Abuse* 2008 comparison of the accuracy of a deterministic record linkage algorithm, the Link King, and Link Plus. View the [evaluation](#).
- *The MITRE Corporation* sponsored research in 2012 on multicultural name matching and identity resolution. The resulting documents cover software packages, analytical tools, and evaluation of name matching systems.
  - [International Multicultural Name Matching Competition: Design, Execution, Results, and Lessons Learned](#)
  - [FEIRI: Extending ISLE's FEMTI for the Evaluation of a Specialized Application in Information Retrieval](#)
  - [An Infrastructure, Tools and Methodology for Evaluation of Multicultural Name Matching Systems](#)
  - [A Ground Truth Dataset for Matching Culturally Diverse Romanized Person Names](#)

## Linking Considerations

- The *Substance Abuse and Mental Health Services Administration (SAMHSA)* compiled a document concerning data integration in 2008 (revised 2011). Appendix I of the document contains useful resources and considerations when undertaking a record linking project. View the [Guide](#) (Appendix I, pages 21-46)

*This project was supported by Grant No. 2010-DG-BX-K088 awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, and the Office for Victims of Crime. Points of view or opinions in this document are those of the author and do not represent the official position or policies of the U.S. Department of Justice.*