

This chapter discusses the collaboration among the California Community Colleges Chancellor's Office, the California Employment Development Department, and community college practitioners to develop an administrative data-matching system to meet accountability mandates. The chapter delineates the many benefits of collaboration when matching student records with records from other administrative databases and examines how completeness of external databases can affect outcomes.

Collaborative Administrative Record Matching in California

W. Charles Wiseley

California community colleges (CCC), like college systems in every state, have accountability legislation to which they must respond. Prior to the 1990s, most of the accountability reporting requirements were met by the individual colleges in the CCC system. Colleges either sent reports directly to the coordinating agency requiring the report or to the CCC chancellor's office for packaging into a single statewide report. Both the state and the chancellor's office recognized the need for a systematic way of meeting the increasingly burdensome reporting requirements. The chancellor's office and the CCC began implementing a data system in 1989 to consolidate efforts to respond to legislative accountability mandates on the 107 colleges in the system. Under the new data system, colleges submit detailed course enrollment and student demographic data to the chancellor's office. These data are then aggregated to meet the state and federal reporting requirements.

Accountability legislation typically has many aspects of accountability that may include measures of inputs, processes, outputs, and outcomes. Most of the data available for accountability from the new chancellor's office data system address the input, processes, and output measures in legislative reporting requirements of access, successful course completion, retention, degrees conferred, and so forth.

Both state and federal bodies, however, have mandated accountability systems that include outcome measures. Two pieces of federal legislation, in particular, the Carl D. Perkins Vocational and Applied Technology Education Act (1990) and the Student Right to Know Act (1996), have mandated outcome measures that require data not typically available from a college where a student attended. Because the new chancellor's office data system did

include student Social Security numbers (SSNs), however, tracking students in administrative databases outside community colleges where SSN were used as a unique identifier was possible. Linking the chancellor's office database with external administrative databases has provided useful information for specific outcomes contained in each separate external matching source. Each external database, however, usually addresses only one possible outcome. Multiple external sources are required to provide as complete a picture of student flow as possible. Gaps in outcome data for specific groups may produce extremely negative consequences for those groups.

This chapter discusses the collaboration required among state agencies, higher education segments, and college practitioners to meet the Carl D. Perkins Act mandates for accountability. Moreover, it delineates the many benefits of collaboration when matching student records with records from other administrative databases and examines how completeness of external databases can affect outcomes. Past and present collaborative efforts have provided and continue to provide California community colleges with information for accountability purposes and, more important, information useful for instructional improvement and enhancement of student services.

Collaborative Requirements in the Carl D. Perkins Act

The federal Carl D. Perkins Act passed in 1990 not only had specific requirements for an accountability system but included requirements for a collaborative process. Those requirements specified that a committee of practitioners (COP) would be formed to develop outcome measures and performance standards to inform vocational education practitioners about their programs. In California, where both the K–12 and community college systems offer vocational education, the COP membership included California Department of Education (CDE) representatives, K–12 district staff and teachers, and California community college staff and practitioners. The COP had a wide range of interests represented and soon developed into two groups that could focus on the measurement systems available to the two different entities of adult vocational education under CDE and community college vocational education. The community college members included the chancellor's office dean of vocational education, a vocational education specialist from the chancellor's office, and eighteen college vocational deans and faculty members. After developing numerous measures, the California community college section of the COP determined that it would begin with the three measures for which data were available or could be developed in the near future: achievement, retention, and placement.

Data for the achievement and retention measures were available directly from the California Community Colleges Chancellor's Office Management Information System (COMIS) database, which had its first phase fully implemented by the fall of 1991. Although the COMIS database contains a wide variety of information, such as student demographics, enrollments,

grades, student services received, staff and faculty demographics, assignments, salaries, and facilities, the COMIS did not contain any information on student placement after college or any data that would allow follow up via traditional survey methodologies, such as student mailing address and complete name.

Traditional follow-up surveys were being conducted at the individual colleges but few standards were in place for either the data collected on those surveys or the methodologies used to collect the data. California community colleges had been doing some student flow tracking by matching student records with the student record data in the other higher education segments in the region (that is, University of California, California State University, Association of Independent Colleges and Universities, and the University of Nevada). Those data matches were occurring sporadically in California through efforts of researchers at the colleges, the information systems staff at the chancellor's office, and through California's Intersegmental Coordinating Counsel (ICC) data needs task force.

Using education databases to track student flow through and within higher education segments does not inform us about or tell us where the majority of community college students go when they leave the college—into the workforce. A consistent, objective, and low-cost method to follow up that majority of students who did not immediately continue in public higher education in California was needed.

Matching Administrative Student Records with the UI Database

A few states were leading the way toward doing student follow-up in a not so traditional way. Matching student leaver records with administrative databases collected for other purposes such as unemployment insurance (UI) and military and federal payrolls was increasingly providing very high match or return rates for follow-up studies for community colleges in Florida (Pfeiffer, 1990), Texas (Froeschle, 1991), Washington (Seppanen, 1993), and other states.

California Community Colleges Chancellor's Office commissioned a feasibility study in 1991 to determine whether matching student leaver records with the California UI administrative database would provide the same useful low-cost follow-up information experienced in the other states matching with UI systems. The Employment Development Department (EDD), who maintains California's UI database, was contacted by the chancellor's office, and contracts were developed with Dr. Jack Friedlander at Santa Barbara City College to conduct the study and with EDD to do the UI base-wage file matching with community college student records.

Early matches used cohorts of students who had either left the college system or completed degree or certificate programs prior to 1993. The match of COMIS student records to employment records in the UI base-wage file was accomplished electronically and involved the following general procedures:

CCC creates a file of student SSNs from the COMIS where the student either received a degree or certificate (completers) or did not return to any community college in the CCC for one year (leavers) following a cohort year.

CCC submits a file of completer and leaver SSNs to EDD for matching with the UI wage records.

EDD appends quarterly wage record data for the prior year to the SSNs submitted for matching.

Student records with wage data appended are returned to the chancellor's office.

To acquire longitudinal data, each completer/leaver cohort is resubmitted to EDD for matching until a four-year span of wages is accumulated at the chancellor's office.

The California Collaborative Working Group

As the feasibility study progressed and began to show promise, an advisory committee, the Vocational Education Technical Advisory Research and Accountability Committee (RAC), was developed to determine the best methodology for assessing the benefit of community college attendance given the constraints of the available data. The RAC needed to assess how the deficiencies of both the UI database and the COMIS would affect the outcomes being measured when the project was expanded to include more colleges. The RAC, to meet this charge, needed to include members beyond chancellor's office and EDD staff, and brought in college faculty, administrators, and researchers in the development of the reports to address the COP's placement measure. This was the beginning of the expansion of the collaboration beyond the chancellor's office and EDD.

Understanding and documenting details such as where the UI data would not provide employment information (for example, self-employed, military, federal employment) and where the COMIS could not provide accurate information (for example, formal program participation, complete enrollment history) required a broad range of participants in the collaboration process. Although the estimates of those employed but not represented in the UI wage records was only 3 percent nationally (Friedlander, 1996; Stevens, Richmond, Haenn, and Michie, 1992), concerns were raised that questioned the composition of the California labor force as well as the impact on specific college program outcomes for those groups not represented. During the course of the feasibility study and following implementations, the collaboration among EDD, the chancellor's office, and the college practitioners grew ever more important in understanding the data shortcomings as the reports became more refined and results began to be made available for public scrutiny. Without the concerted efforts of the participants, the development of methodologies to provide meaningful data and useful reports would have taken many more months, possibly even years, longer to accomplish. Six years into the study, the collaborative process continues to inform the committees and helps us refine the reports

to develop new understandings of subpopulations, such as traditional degree-seeking students, students returning for skill upgrading or license maintenance, or economically disadvantaged students, in the follow-up reports.

The UI database contained employment records for 80 percent of the student completers and leavers during their last year in college from the two colleges in the feasibility study (Friedlander, 1993). When considering the needs for follow-up in the Carl D. Perkins Act that the feasibility study was testing, however, a number of deficiencies in California's UI base-wage file system were illuminated as well as some deficiencies in the COMIS.

The Working Group's Challenge

To follow up students who left the community colleges and provide useful information to faculty, administrators, and policy makers, some very basic information needed to be defined by the RAC. In concept, the simple question to answer was, Is there an economic benefit for attendance at a California community college?, or more specifically, Is there an economic benefit for attendance in a *vocational program* at a California community college? Two basic pieces of information were required to operationalize this measure indicating possible benefit for community college attendance: (1) the vocational program the student was in college attending, and (2) the impact on earnings that program may have had. The operationalization of these concepts was seen to be problematic, given the available data, for the reasons indicated below.

First, unlike senior universities in California and like community college systems in many other states, no formal acceptance for entry into a program of study is required for most programs offered at California community colleges. Entrance into any California community college is open to anyone who can benefit, whether the student is a high school graduate or not (some health care profession programs such as nursing do require formal acceptance). Moreover, no date of acceptance into the college is collected in the COMIS; therefore, data to determine program entry date could not be determined until a sufficient number of years were available in the COMIS.

Program of study could easily be identified for those students receiving certificates or degrees, since the program is identified as the program in which an award is conferred, which is reported in the COMIS as well as the type of award (for example, AA, AS, two-year certificate). The value of community colleges, however, with missions that address life-long learning, skill upgrading, license maintenance, and so forth, would be understated by tracking only completers. Efforts were started to identify programs using student course-taking behavior that would allow leavers not earning or receiving degrees or certificates to be included in the study. With the additional leaver information, programs could better use the reports to evaluate their impact on different populations of students. Identifying a concentration of enrollments in a program area of sufficient quantity to justify evaluating the program based on those enrollments was identified as a pseudo-indicator of program. For the feasibility study, leavers, completers

who left the institution for at least one year, and those completing but not leaving were followed.

Next, impact on wages using after-college earnings require a baseline for comparison. Without before-college earnings, since no beginning college date for establishing a before-college period was available, last-year-in-college earnings and first-year-out-of-college earnings were determined to be useful comparative baselines for third-year-out-of-college earnings, depending on the population to be examined.

Unlike Florida, Texas, Washington, and other states, California does not collect any data from employers other than gross quarterly earnings that could be used to calculate hourly wage such as hours worked per week, weeks worked in the quarter, or even data to determine full-time/part-time status. No job title or occupational classification indicators or even county of employment were available in the California UI data.

Working closely with EDD staff, chancellor's office staff began to investigate the possibility of expanding the amount of data collected from employers in an employer survey pilot project. From that pilot, CCC found that a majority of employers would not provide that additional data. Those findings were similar to the findings of a study funded by the California Occupational Information Coordinating Committee (COICC) in the late 1980s to examine the possibility of expanding the amount of data collected from employers in the quarterly reports to EDD to include some of those data types. Employers were overwhelmingly opposed to any expansion of the data required on the quarterly statements to EDD or in a supplemental survey. Specifically, smaller businesses saw it as an additional burden that could not be justified. In the pro-business, less-government political atmosphere of California in the 1980s and 1990s, that burden was seen as sufficient cause to abandon any efforts to expand the data collection for the UI system.

Without some indicator of time worked for the earnings reported, no comparable figure such as hourly wage or even classification of full-time/part-time status was available. Comparing gross wages for any amount of work such as part-time or partial quarter seemed invalid, and alternative methods of estimating annual earnings were considered and tested. Decisions were made to use the median of gross wages for those working in all four quarters as well as a proxy measure for full-time employment. The proxy measure for full-time employment was chosen to allow a more reliable figure for comparing earnings from year to year when looking at earning gains. The standard labor market figure for average manufacturing wage in 1991 was used to approximate full-time employment. The figure \$12,875, half the average manufacturing wage (about \$5.50/hour), was considered full-time/full-year employment for the study.

As the feasibility study came to a close and sufficient information was available to support a belief that the project was not only feasible but cost-effective and would produce information currently available in no other forum, the decision to continue the project was made. The RAC, chaired by Jack Friedlander, first had to address the following basic questions in order to

expand the study systemwide: How do we identify the vocational program that will be evaluated based on student leaver employment and earnings? What earnings should be used as the baseline for comparison until we can identify program entry date? Can we identify program entry date without new data collection? When is a student considered a leaver?

The committee worked diligently, pouring over the methodology and conclusions of the feasibility study, and provided invaluable advice for the implementation of the next step—the pilot.

The next study commissioned by the chancellor's office was a pilot conducted by Jack Friedlander in 1993 with eighteen colleges. The linking of the COMIS and UI databases returned the same high match rate and provided a wealth of information to the practitioners at the colleges (Friedlander, 1996).

The Challenge of Implementation Systemwide

The advisory committee, working with the information from the pilot study, began discussions to refine the reports for implementation systemwide. Categories of students, such as *vocational student*, *skill upgrade student*, and *enrollment concentration*, and program assignment methodologies were redefined at the RAC's direction. Additional links were made to the California State University (CSU) system to reclassify those students as continuing their education at CSU and to remove them from the earnings calculations.

Systemwide implementation began in 1996 with the publication of the 1990–1991 leaver cohort draft reports, which included 100 colleges submitting SSNs as unique identifiers to the COMIS. Six months later, in early 1997, after review and further refinement, the first annual report following the 1991–1992 leaver cohort was published.

As of January 1998, California Community Colleges has matched six leaver cohorts with the California EDD's UI base-wage file and the first three cohorts, 1990–1991, 1991–1992, and 1992–1993, with the CSU system. The system set up to provide reports from the matches had to meet federal Privacy Act restrictions and is much more complex than the data matching itself. Even though California Community Colleges contracted with EDD to do the matching as research on our behalf, all of the Family Education Rights and Privacy Act (FERPA) caveats and restrictions had to be observed. Contracting with EDD facilitated the cooperation and support of the staff at EDD to help us understand and meet the constraints put on EDD to meet federal Privacy Act and California Unemployment Insurance Code (sections 1094 and 1095) requirements.

Median annual earnings while in college, one year after college, and three years after college, as well as median earnings gain for those years, were now available on reports with certain privacy-based restrictions and caveats. The reports were designed to inform faculty of the levels of employment their former students were entering into and any increases in earnings for those students over a four-year span. It has become very clear in the past few years that this

information can be invaluable for instructional improvement at the local level and is the most effective use of the data available through the UI base-wage file matches. The objective employment information regarding entry level wages of students and the earning gains over the following years provide insights to faculty about program acceptance in the workplace that are available from no other venue.

Only through collaborative efforts with faculty and faculty's interaction with the employers in the local economies, however, can we understand the complexities of the data coming out of the base-wage file research as we develop standardized reports. For example, movement of major segments of job categories into temporary or independent contractor status will have great impact on certain programs in different economic areas when UI wage data are used to determine employment outcomes. Faculty need to be aware of those conditions to prepare their students for the workplace, and report readers need to be apprised of the labor market conditions for those programs to properly interpret the outcomes in the reports. For this reason, collaborative systems to facilitate communication among employers, faculty, report developers, and the policy makers who read those reports need to be institutionalized.

The Collaborative Process Helps Highlight Problems

Through the efforts of the vocational deans, staff, and faculty participating in the RAC we began to understand that aggregating outcomes derived from the UI base-wage files for programs at a college level may be problematic. Although the aggregation may provide sufficient numbers for valid analysis and inference, the aggregation of unlike programs removed some reliability. For example, nursing programs may include both LVN and RN programs, which have very different outcome expectations (that is, annual income). Aggregating outcomes to a general nursing category therefore would mask those programs, and any comparison between colleges would be directly influenced by the proportion of RN students in the nursing program. Aggregating to higher levels such as nursing statewide would have the same RN ratio influence. Therefore, using aggregations above the college level where local practitioners understand program content can only make sense where programs across institutions have similar core curriculum and outcome goals and objectives for their students.

In California, with few exceptions, the curriculum is not standardized among colleges or universities in the same system or even colleges in the same district. Through cooperative efforts between college practitioners, chancellor's office staff, and EDD staff, measures were constructed that allowed a more valid consolidation of program outcomes across colleges by examining student workload completed and awards conferred. Additional work is being done to differentiate between certificate programs requiring less than two years (less than sixty units) and those lasting two or more years (sixty or more semester units).

Benefits of Including EDD staff on the RAC

The efforts of the EDD staff were invaluable in working with the EDD information security office to develop a trusting relationship. Report development proceeded smoothly with little actual monitoring of the project other than review at the RAC meetings. EDD staff participated in the project's advisory committee to advise CCC research staff of many of the deficiencies of the UI base-wage file, the UI data collection and storage mechanisms, and changes in those systems. Committee participation facilitated EDD's availability for quick additional research to provide information such as employers' reticence to report any additional information. EDD also provided links to other research entities using the UI base-wage files for research, thus helping community colleges develop a deeper understanding of the composition, participant attrition, and other complexities of the UI base-wage file and its contents. EDD staff continue to work on expanding the linkages to other administrative databases, such as the military and federal payroll systems, to increase our knowledge of the 20 percent we don't find in the UI base-wage and CSU enrollment files.

California Community Colleges continues to collaborate with EDD in initiating new studies to complement the UI follow-up reports with relatedness to training measures, and in developing lists of occupations of program leavers and local employers of program leavers. Having EDD staff intimately involved with these projects also meets EDD's need to monitor the project to see that our reporting and handling of the data meet their privacy and UI code requirements.

Policy, Privacy, and Resources

Requirements for unitary data matching have policy, privacy, and resource considerations. The general counsels from all of the data matching sources were involved to assure compliance with FERPA and other privacy legislation and regulations. The basic barriers that the educational institutions face when matching student educational records with records from other data sources are those that protect the privacy of every person in the United States. Additionally, however, educational institutions must meet the requirements of FERPA when disclosing student records to any individual or entity. FERPA does, however, allow disclosure of student records for research under the exception of instructional improvement.

The task of responding to accountability requirements and meeting the requirements of FERPA can only be accomplished with cooperative efforts between institutions that minimize disclosure of student records and maximize the ability to relate the outcomes to instruction. Any cooperative effort between educational institutions and entities outside education must keep the management of privacy breach risks as the highest priority. By developing the appropriate cohorts and reporting levels that meet instructional improvement requirements, executing detailed contracts, and building relationships between

institutional liaisons for research, report approval, and monitoring, Privacy Act considerations can be met while constructing accountability systems.

Adhering to the constraints of privacy legislation does not address policy implications of information provided by the UI base-wage file research. Full cooperation of faculty administering programs at the college level is required to understand student flow into the workplace. Policy makers must be made aware of how the outcomes may be affected by the deficiencies in the UI base-wage file. Local-level participation is necessary in the development of reports and in the reporting of findings for this awareness to be realized. Coordination and cooperation must be ongoing and provide for communication pathways among faculty, researchers, and policy makers.

Once college staff began using the reports to understand the outcomes of students in their programs, researchers began requesting access to the unitary earnings data. College-level research that tests whether specific intervention strategies had an impact on student earnings or placement might more directly inform college staff. Privacy restrictions of FERPA and California UI code restrictions, however, under current uses of available technology, set constraints on local college access to unitary UI wage data.

Without being able to provide college access to unitary outcomes data, human resources at the chancellor's office for implementing and monitoring projects have been strained. New avenues for communication such as e-mail, video conferencing, e-mail discussion lists, teleconferences, video phones, and other electronic media as well as the high-power, low-cost computing resources becoming available will help the chancellor's office meet the expanding information needs of college staff and policy makers with increased timeliness.

We Are Still Learning

The California Community Colleges Chancellor's Office continues to pursue avenues that will focus resources on collecting, analyzing, and disseminating new information to college researchers and faculty as we learn new lessons through the collaborative process. As California government moves closer to performance-based funding for community colleges and noneducation workforce development boards develop outcome measures on which to base that funding, the collaboration and the lessons we have learned through it in the past few years become invaluable. The lessons continue as the chancellor's office and other collaborative working groups sift through comments from over eighty college staff and faculty researching how to better identify which program to assign students who complete varying degrees of course work but do not complete degree or certificate requirements so that instruction can be more directly related to the outcomes. Those collaborative efforts among college faculty, researchers, the CCC chancellor's office, and EDD are helping to develop a new understanding of outcomes by using the UI base-wage and other data and determining their relationship to instruction and educational services provided at the colleges.

References

- Friedlander, J. *Using Wage Record Data to Track the Post-College Employment and Earnings of Community College Students*. Santa Barbara, Calif.: Santa Barbara City College, 1993. (ED 360 007)
- Friedlander, J. *Using Wage Record Data to Track the Post-College Employment Rates and Wages of California Community College Students*. Santa Barbara, Calif.: Santa Barbara City College, 1996. (ED 390 507)
- Froeschle, R. *Pilot Study Findings of the Use of the Texas Employment Commission Unemployment Insurance Wage Records to Document Outcomes of Technical-Vocational Education and Training*. Austin: Texas Higher Education Coordinating Board, 1991.
- Pfeiffer, J. J. *Florida Education and Training Placement Program*. Annual Report. Tallahassee: Florida Department of Education, 1990.
- Seppanen, L. J. *Using Administrative Data Matches for Follow-Up*. Technical Report No. 93-5. Olympia: Washington State Board for Community and Technical Colleges, 1993. (ED 382 250)
- Stevens, D. W., Richmond, P. A., Haenn, J. F., and Michie, J. S. *Measuring Outcomes Using Unemployment Insurance Wage Records*. Washington, D.C.: Research and Evaluation Associates, 1992.

