# Identifying a Sample of HIV-Positive Beneficiaries From Medicaid Claims Data and Estimating Their Treatment Costs

Arleen A. Leibowitz, PhD, and Katherine Desmond, MS

*Objectives.* We sought to identify people living with HIV/AIDS from Medicare and Medicaid claims data to estimate Medicaid costs for treating HIV/AIDS in California. We also examined how alternate methods of identifying the relevant sample affect estimates of per capita costs.

*Methods.* We analyzed data on Californians enrolled in Medicaid with an HIV/AIDS diagnosis reported in 2007 Medicare or Medicaid claims data. We compared alternative selection criteria by examining use of antiretroviral drugs, HIV-specific monitoring tests, and medical costs. We compared the final sample and average costs with other estimates of the size of California's HIV/AIDS population covered by Medicaid in 2007 and their average treatment costs.

*Results.* Eighty-seven percent (18 290) of potentially identifiable HIV-positive individuals satisfied at least 1 confirmation criterion. Nearly 80% of confirmed observations had claims for HIV-specific tests, compared with only 3% of excluded cases. Female Medicaid recipients were particularly likely to be miscoded as having HIV. Medicaid treatment spending for Californians with HIV averaged $33 720 in 2007.

*Conclusions.* The proposed algorithm displays good internal and external validity. Accurately identifying HIV cases in claims data is important to avoid drawing biased conclusions and is necessary in setting appropriate HIV managed-care capitation rates. (*Am J Public Health.* 2015;105:567–574. doi:10.2105/AJPH.2014.302263)

In 2010, the White House Office of National AIDS Policy outlined an ambitious National HIV/AIDS Strategy for the United States that called for evaluation strategies that would "obtain data (core indicators) that capture the care experiences of people living with HIV without substantial new investments."[1] Surveillance systems already in place in each state provide the Centers for Disease Control and Prevention with comprehensive data on incident HIV and AIDS cases.[2] However, much less is known about the medical treatments received by people living with HIV/AIDS and the cost of those treatments.

Much of the cost of HIV/AIDS treatment is borne by public insurance programs, principally Medicaid and Medicare. These 2 programs provide health insurance for more than half of people living with HIV/AIDS who are receiving care.[3,4] The importance of Medicaid as a source of funding for HIV/AIDS treatment of low-income persons will grow substantially after full implementation of the Affordable Care Act, which eliminates the additional disability requirement for Medicaid eligibility in states accepting the Medicaid expansion, thereby extending coverage to nondisabled, low-income people living with HIV/AIDS in those states.

Because of its prominent role in insuring low-income people living with HIV/AIDS, Medicaid can provide a rich source of data on the types and costs of treatments delivered to some of the most vulnerable individuals with HIV/AIDS. Insurance claims data can potentially allow us to monitor HIV/AIDS treatment without substantial new investments because most claims data are stored as computerized records. Claims data provide a comprehensive picture of medical care received from a variety of providers in multiple settings (outpatient, inpatient, laboratory, pharmacy), contain procedure codes that detail the services provided, and include cost of the treatment. By contrast, medical records tend to have smaller scope, in terms of both numbers of patients and services covered. Furthermore, medical records most often lack payment information.

Insurance claims data can provide information on a large number of individuals, even among those with relatively low-prevalence conditions, which is valuable in reducing the variability of estimates of per capita expenditures. However, the greater precision afforded by large administrative data sets is of little value if estimates are based on an inappropriate sample. Claims data are primarily designed for billing purposes; thus, they generally lack clinical detail important for selecting cases with a particular disease.[3,5] For example, claims data will document whether a laboratory test was performed, but not the test result. Therefore, analysts must rely on the diagnosis information on insurance claims.[6] Professional medical records specialists code diagnoses on inpatient claims, leading to greater accuracy and reliability of diagnosis information coming from inpatient stays. However, diagnosis coding is more error-prone in the outpatient sector, which has accounted for an increasing percentage of HIV/AIDS care since 1996 when antiretroviral medication (ARV) began to dramatically reduce hospitalization for HIV/AIDS.[7] This has increased the challenges of identifying people living with HIV/AIDS from insurance claims data.

We applied a practical algorithm for identifying people living with HIV/AIDS in insurance claims data to estimate Medicaid costs for treating HIV/AIDS in California. We also examined how alternate methods of identifying the relevant sample affect estimates of per capita costs.

## METHODS

At our request, the Centers for Medicare and Medicaid Services (CMS) prepared a file containing medical claims for all Californians enrolled in Medicare, Medi-Cal (California's Medicaid program), or both, whose records indicated an HIV/AIDS infection in 2007. Individuals were included in the CMS file if (1) any Medicare or Medi-Cal claims during calendar year 2007 contained a possible *International Classification of Diseases, Ninth Revision* (*ICD-9*)[8] diagnosis code for HIV; or (2) Medicare or Medi-Cal records indicated a drug prescription for an ARV. Medicare and Medi-Cal claims were linked, and individuals were retained in the analysis if they were enrolled either in Medi-Cal only, or were dually enrolled in both Medi-Cal and Medicare.

The process of identifying people living with HIV/AIDS is illustrated in Figure 1. Our analysis included individuals aged 16 years or older who lived in California. The initial file request cast a wide net so as not to exclude any Medi-Cal enrollees who were possibly living with HIV. We dropped from the initial file those who received HIV counseling services only, who were missing eligibility data, or who were eligible only for partial benefits, such as emergency services, or for pregnancy, family planning, or breast or cervical cancer treatment. We also dropped individuals enrolled in managed care. Nearly half (46%) of the full-benefit Medi-Cal enrollees were dually enrolled in both Medicare and Medi-Cal. For these individuals, Medicare is the primary payer, and the bulk of their care would be expected in the Medicare claims data. However, there are no claims or encounter data available for care received under Medicare managed care, and thus no opportunity to identify HIV-positive persons through diagnosis codes recorded in Medicare claims. Therefore we limited our analysis to these enrolled in fee-for-service (FFS) care only.

### Development of Classifications for Confirmed Diagnoses

We refer to the resulting file as the "Any DX/ARV" file, as FFS enrollees are included if they had at least 1 HIV/AIDS diagnosis code included anywhere in the claims data or any drug claims for ARV medications. Diagnosis coding is notably subject to error—both random error and miscoding of "rule-out diagnoses." Peabody et al. found that only 57% of administrative records had the correct primary diagnosis coded.[5] To address the errors typically found in diagnosis coding,[3,5-7,9–12] we developed 2 criteria for adding a record to a confirmed HIV/AIDS file.

The inpatient–2 visit criterion required an HIV diagnosis code coded in an inpatient claim, or in at least 2 outpatient claims, spaced at least 1 month apart. The strategy of requiring that a diagnosis appear twice in the claims has been endorsed by many analysts. Rector et al.[10] evaluated 38 different algorithms for identifying Medicare plus Choice enrollees with 6 common chronic diseases (hypertension, heart failure, chronic lung disease, arthritis, glaucoma, diabetes). They concluded that algorithms requiring 2 or more claims with diagnoses increased the specificity (the proportion of cases identified who actually have the condition) but reduced sensitivity (the proportion of individuals with the condition who are identified) of their algorithm. The National Cancer Institute also proposes assigning a diagnosis only if a patient's diagnosis appears on 2 different claims spaced more than 30 days apart.[13] An algorithm to identify HIV-positive enrollees from Medicaid files developed by Walkup et al.[14] required at least 1 inpatient claim or 2 outpatient claims containing an HIV diagnosis code, spaced at least 1 month apart. Chesnut et al.[15] also proposed an algorithm based on HIV-specific diagnosis codes, procedure codes, and ARV use.

The nonscreening criterion required a diagnosis coded in an outpatient setting on a day when HIV screening tests were not performed. There is evidence that rule-out diagnoses sometimes remain in the claims even after the condition has been ruled out,[16] so we did not consider a sole diagnosis occurring on the same day as an HIV screening examination, on its own, sufficient evidence of HIV infection. Therefore, the nonscreening criterion required an HIV diagnosis recorded in an outpatient claim on at least 1 day when there was not also an HIV antibody screening test.

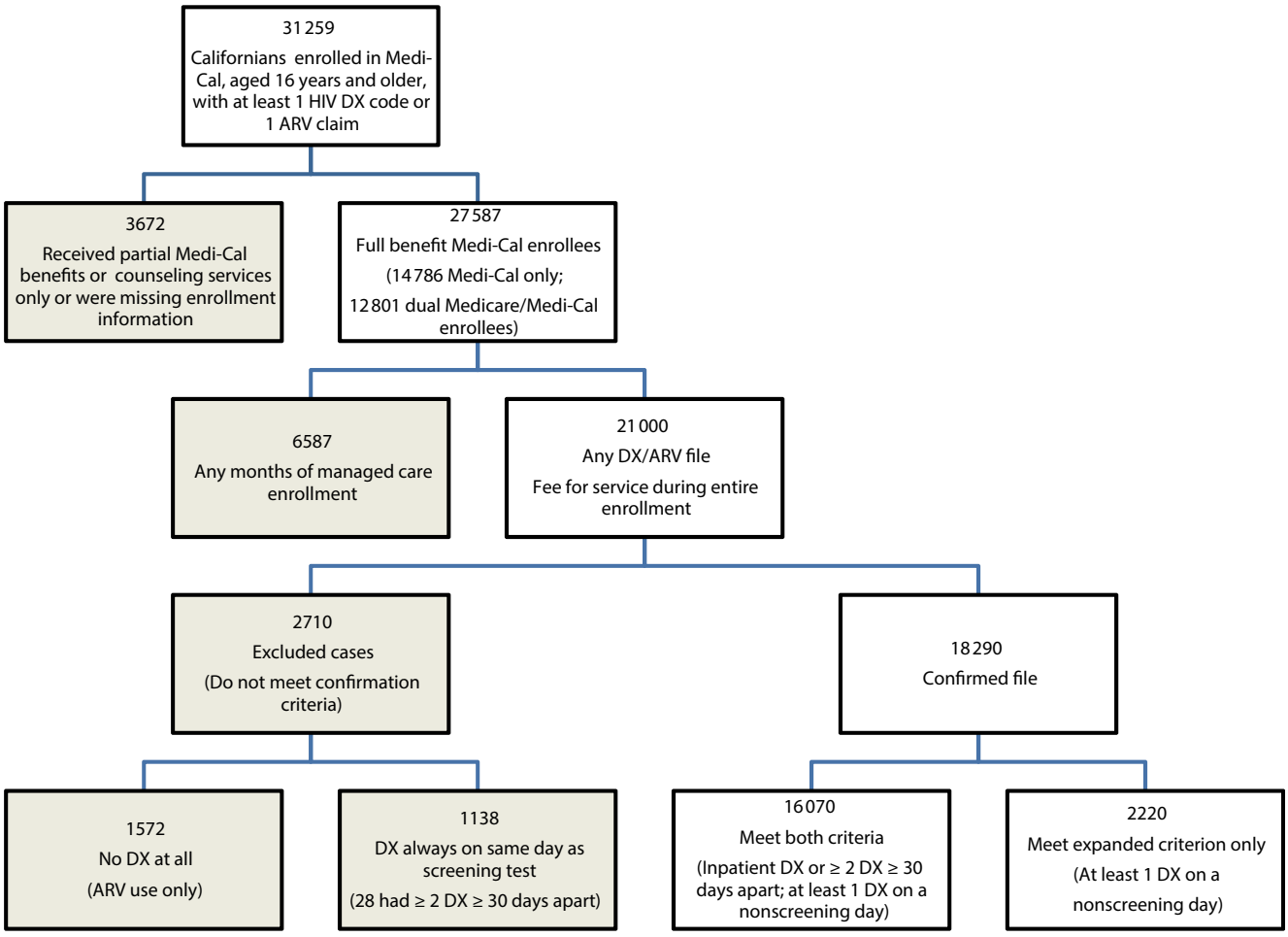
Although the initial file requested from CMS included individuals on the basis of ARV use, we did not continue to use that as an inclusion criterion. Some ARV regimens can also be used in the treatment of hepatitis B virus (HBV). Although it is not uncommon for HIV-positive persons to have a concurrent HBV infection, it is possible that the use of ARVs to identify persons as HIV-positive may lead to erroneously designating HBV-only persons as being people living with HIV/AIDS. We therefore do not use ARV claims alone to identify HIV.

From the pool of persons potentially identifiable as people living with HIV/AIDS in the Any DX/ARV file, we constructed 2 levels of a confirmed HIV/AIDS file, as shown in Figure 1, by adding individuals to the confirmed file in order of the strength of evidence, as follows:

1. Include persons with diagnosis confirmed by both the inpatient–2 visit and the nonscreening criteria. We refer to this as the strict definition of HIV-positive status.
2. Add persons with diagnosis confirmed only by the nonscreening criterion (as shown subsequently in this article, very few individuals had multiple HIV diagnoses 30 or more days apart, but only on screening days). We refer to this as the expanded definition of being HIV-positive.

### Evaluation of HIV-Positive Classification

We examined whether selected aspects of treatment history for individuals in the confirmed files were consistent with HIV infection and whether the treatment history for persons excluded from the confirmed files was inconsistent with HIV infection. Monitoring of viral loads (counts of HIV RNA in the blood) and levels of CD4 lymphocyte cells is an important part of HIV care. The presence of claims for these tests in the data is consistent with HIV-positive status. The 2007 Centers for Disease Control and Prevention guidelines for ARV medications did not call for their use by all HIV-positive persons; nevertheless, their presence in the claims data was also consistent with HIV. Finally, we looked for evidence of HBV infection, through diagnosis codes or use of ARV regimens used for HBV and not HIV. Indications of greater HBV infection among persons excluded from the confirmed file would indicate the likelihood that their initial inclusion was because of treatment of HBV.




*Note.* ARV = antiretroviral; DX = diagnosis.

**FIGURE 1—Identification of Medi-Cal persons living with HIV/AIDS: California, 2007.**

We present frequencies of treatment of each classification group by program enrollment (Medi-Cal–only enrollees, dual Medicare–Medi-Cal enrollees, and total). We performed the $\chi^2$ test to determine whether the rates of treatment differ significantly between the confirmed and excluded cases.

For each group, we also calculated annual per capita expenditures by summing all paid medical and pharmaceutical claims for each individual. We annualized expenditures of part-year enrollees who did not die during the year. We did not change expenditures of those who died during the year. We performed the *t* test to determine whether mean spending differed significantly between the confirmed and excluded cases. Per capita expenditures are reported in 2007 dollars to facilitate comparison with other sources, except where noted that they have been adjusted to 2013 dollars, using the medical component of the consumer price index for urban consumers, which rose by 29% between July 2007 and July 2013.[17]

## RESULTS

A flowchart of the case confirmation process can be seen in Figure 1. The initial file from CMS included 31 259 Californians aged 16 years and older with any Medi-Cal coverage. Of these, 27 587 remained after we excluded those receiving counseling services only or entitled to only partial Medicaid benefits. Among those, 21 000 had FFS coverage for their entire enrollment, constituting the "Any DX/ARV" file (Table 1). Fifty-two percent (10 941) were enrolled in Medi-Cal only and 48% (10 059) were covered by both the Medi-Cal and Medicare programs, the "dual eligibles" (Table 1).

Among the 21 000 FFS enrollees in the CMS file with any HIV diagnosis or ARV use, 18 290 (87%) satisfied at least 1 of our confirmation criteria. Of the 2710 (13%) excluded, 1572 had no diagnoses recorded anywhere, but had been included originally solely on the basis of ARV use. An additional 1138 had diagnoses recorded, but only on the same day as screening tests. Among these were 28 persons who had multiple diagnoses 30 or more days apart, but they were always on the same day as screening. These cases were not included in the confirmed file. Seventy-seven

**TABLE 1—Numbers of Persons Living With HIV/AIDS With Fee-for-Service Medi-Cal Coverage Under Different Sample Inclusion Criteria: California, 2007**

| Variable | Source: Any DX/ARV[a] | Add in Step 1: Strict Definition[b] | Add in Step 2: Expanded Definition[c] | Confirmed File (Strict Plus Expanded)[b,c] | Excluded Cases[d] |
|---|---|---|---|---|---|
| All enrollees (No., % of Row) | | | | | |
| In Medi-Cal only | 10 941 (100) | 7 996 (73) | 1 512 (14) | 9 508 (87) | 1 433 (13) |
| Dual Medicare/Medi-Cal | 10 059 (100) | 8 074 (80) | 708 (7) | 8 782 (87) | 1 277 (13) |
| All Medi-Cal | 21 000 (100) | 16 070 (77) | 2 220 (11) | 18 290 (87) | 2 710 (13) |
| Women (No., % of all enrollees) | | | | | |
| In Medi-Cal only | 3 122 (29) | 2 110 (26) | 433 (29) | 2 543 (27) | 579 (40) |
| Dual Medicare/Medi-Cal | 1 613 (16) | 905 (11) | 188 (27) | 1 093 (12) | 520 (41) |
| All Medi-Cal | 4 735 (23) | 3 015 (19) | 621 (28) | 3 636 (20) | 1 099 (41) |

*Notes.* ARV = antiretroviral medication; DX = diagnosis. *International Classification of Disease, Ninth Revision*,[9] HIV-related diagnosis codes used: 042, 079.53, 795.71, V08. National Drug Codes used to identify antiretroviral medications available from the authors.

[a]The Any DX/ARV file includes beneficiaries who are Californians aged ≥ 16 years, enrolled in Medi-Cal with full benefits, and had at least 1 claim with an HIV diagnosis code or 1 prescription fill for an ARV medication.

[b]Strict definition requires that beneficiaries have at least 1 inpatient claim with an HIV diagnosis recorded, or 2 or more outpatient claims, at least 30 days apart, with at least 1 of them occurring on a day that was not coincident with an HIV screening test.

[c]Expanded definition requires that there be at least 1 HIV diagnosis recorded in an outpatient claim, occurring on a day that was not coincident with an HIV screening test.

[d]Excluded cases are those beneficiaries in the Any DX/ARV file who were not included in the confirmed file.

percent (16 070) of the individuals in the Any DX/ARV file met the strict definition. There were an additional 2220 individuals (11% of the Any DX/ARV file) who met the expanded definition only—they had at least 1 outpatient diagnosis on a day without screening, but did not have an inpatient diagnosis or multiple outpatient diagnoses recorded 30 days apart.

The selection criteria resulted in differentially excluding women. In the strictly defined HIV/AIDS group, the percentage female was 19%, and the percentage female among excluded cases was 41%. The cases added under the expanded definition included a larger proportion of females (28%) compared with cases included under stricter criteria.

### Treatment History of Confirmed and Excluded Cases

Individuals in the confirmed file had medical care use that was more consistent with treatment of HIV infection than medical care use among the excluded cases. Whereas 78% of individuals in the confirmed file received viral load or CD4 tests, only 3% of the excluded individuals had these tests (Table 2). The difference was most pronounced among those meeting both inclusion criteria, with 84% of the strict HIV/AIDS file receiving one of these tests, and was smaller among those added under the expanded definition (33%).

Eighty percent of individuals in the confirmed file filled prescriptions for ARV drugs, compared with 59% of the excluded cases. This 59% includes those who had an ARV claim but no HIV/AIDS diagnosis code. There was less-frequent ARV use among those included by the expanded definition (39% using ARVs vs 85% among those meeting both inclusion criteria).

Evidence of HBV infection, through diagnosis or HBV-consistent use of ARVs, was greatest among the excluded group (37% vs 7% in the confirmed file; Table 2).

All differences in treatment were statistically significant, with $\chi^2$ $P$ values less than .001.

### Expenditures

Annual expenditures were greatest within the Medi-Cal–only group among persons meeting the strict definition for inclusion ($36 469; Table 2). People living with HIV/AIDS added via the expanded definition had lower average costs ($19 184). Expenditures averaged $33 720 for the confirmed HIV/AIDS file without dual coverage. Spending for the excluded cases was 27% lower ($24 514 per year).

Persons in the confirmed HIV/AIDS file with dual coverage had annual Medi-Cal expenditures of $7003 and Medicare expenditures of $39 046. Thus, although dual enrollees had greater total medical expenditures ($46 050) than persons with only Medi-Cal, Medi-Cal costs were much lower for duals because Medicare paid for the largest share of the duals' medical spending. Part-year enrollment is more common among Medi-Cal–only beneficiaries, who averaged nearly a month less coverage during the year than enrollees with dual coverage. Without annualizing, the spending estimates for Medi-Cal–only beneficiaries would appear 11% lower.

All differences in mean spending were statistically significant, with $t$ test $P$ values less than .001.

## DISCUSSION

Insurance claims data have many advantages for examining the quality and cost of medical treatment of relatively low-prevalence diseases such as HIV. Claims data include the full range of treatment services—outpatient and inpatient care as well as long-term care, medication, laboratory, and support services. The validity of the claims data are high as they report what was actually paid by the insurer and are not subject to erroneous recall of either use or diagnosis by patients, as survey data

**TABLE 2—Use of Medical Care and Expenditures of Fee-for-Service Beneficiaries Included Under Different Sample Inclusion Criteria: California, 2007**

| Variable | Source: Any DX/ARV,[a] No. (%) or $ | Add in Step 1: Strict Definition,[b] No. (%) or $ | Add in Step 2: Expanded Definition,[c] No. (%) or $ | Confirmed File (Strict Plus Expanded),[b,c] No. (%) or $ | Excluded Cases,[d] No. (%) or $ |
|---|---|---|---|---|---|
| Received any viral load or CD4 tests | | | | | |
| Enrolled in Medi-Cal only | 6 566 (60) | 6 034 (75) | 484 (32) | 6 518 (69) | 48 (3) |
| Dual Medicare/Medi-Cal enrollees | 7 701 (77) | 7 416 (92) | 254 (36) | 7 670 (87) | 31 (2) |
| All Medi-Cal enrollees | 14 267 (68) | 13 450 (84) | 738 (33) | 14 188 (78) | 79 (3) |
| Claims for any ARV medications | | | | | |
| Enrolled in Medi-Cal only | 7 836 (72) | 6 346 (79) | 564 (37) | 6 910 (73) | 926 (65) |
| Dual Medicare/Medi-Cal enrollees | 8 294 (82) | 7 336 (91) | 297 (42) | 7 633 (87) | 661 (52) |
| All Medi-Cal enrollees | 16 130 (77) | 13 682 (85) | 861 (39) | 14 543 (80) | 1 587 (59) |
| Any hepatitis B diagnoses or drug claims | | | | | |
| Enrolled in Medi-Cal only | 832 (8) | 357 (4) | 40 (3) | 397 (4) | 437 (31) |
| Dual Medicare/Medi-Cal enrollees | 1 410 (14) | 792 (10) | 45 (6) | 837 (10) | 573 (45) |
| All Medi-Cal enrollees | 2 242 (11) | 1 149 (7) | 85 (4) | 1 234 (7) | 1 010 (37) |
| Expenditures, average annual spending | | | | | |
| Medi-Cal only | 32 515 | 36 469 | 19 184 | 33 720 | 24 514 |
| Dual enrollees–Medi-Cal portion | 7 381 | 6 913 | 8 033 | 7 003 | 9 980 |
| Dual enrollees–Medicare portion | 39 010 | 39 781 | 30 662 | 39 046 | 38 767 |

*Notes.* ARV = antiretroviral medication; DX = diagnosis. *International Classification of Diseases, Ninth Revision,*[9] hepatitis B–related codes used: 070.2, 070.3, 5714x, V02.61. Antiretroviral drug use considered consistent with hepatitis B but not HIV included lamivudine only, emtricitabine only, tenofovir alone or in combination with only emtricitabine, or only lamivudine. Expenditures of part-year enrollees who did not die during the year were annualized by multiplying by a factor of 12 divided by months enrolled. The t test was used to compare confirmed versus excluded cases with regard to mean expenditures were performed for each enrollment group and payer. The *P* value for all such tests was <.001. The percentage reported is percentage of enrollee group.

[a]The Any DX/ARV file includes beneficiaries who are Californians aged ≥ 16 years, enrolled in FFS Medi-Cal with full benefits, and had at least 1 claim with an HIV diagnosis code or 1 prescription fill for an ARV medication.

[b]Strict definition requires that beneficiaries have at least 1 inpatient claim with an HIV diagnosis recorded, or 2 or more outpatient claims, at least 30 days apart, with at least 1 of them occurring on a day that was not coincident with an HIV screening test.

[c]Expanded definition requires that there be at least 1 HIV diagnosis recorded in an outpatient claim, occurring on a day that was not coincident with an HIV screening test.

[d]Excluded cases are those beneficiaries in the Any DX/ARV file who were not included in the confirmed file.

might be. It would also be prohibitively expensive to gather information on such a large number of patients with a relatively low-prevalence disease such as HIV/AIDS by using population-based surveys[6] or by combining data from electronic medical records across providers, each with only a small number of relevant patients. Furthermore, clinical electronic medical records most often lack billing and reimbursement data, so analysts must estimate costs by applying price schedules to utilization data.[18]

Despite these advantages, identifying the relevant population from claims can be challenging. Reports from early in the HIV epidemic also noted the promise of studying HIV by using Medicaid claims data, and noted the difficulties in using a diagnosis-based algorithm for identifying cases. In 1991, Keyes et al. identified the problem as undercounting of HIV cases because of changing definitions and nonreporting of diagnoses resulting from attempts to protect patients from stigma.[12] Our analysis indicates that currently the problem is more likely to be overcounting: a substantial proportion of the initial CMS sample had received HIV screening, but did not have other indications of HIV disease. Including these "HIV screeners" in the population used to assess quality of care, disparities, or costs of HIV care could lead to misleading results.[10]

In an illustration of the importance of properly identifying people living with HIV/AIDS, a recent publication reported that only 21% of a large sample of Medicaid recipients newly diagnosed with HIV followed up with HIV treatment within a year.[19] However, the analysis sample appeared to include people with a single HIV diagnosis that could have been reported in conjunction with their HIV screening tests. As argued previously, this strategy can include many individuals without HIV who did not seek (and did not need) HIV treatment.[11] Similarly, Zhang et al.[20] reported that 37.3% of HIV-positive Medicaid-enrolled pregnant women received no ARV treatment in pregnancy. As screening pregnant women for HIV is the standard of care,[21] this method may include many who were only screened for HIV, but did not have the disease. The algorithm we propose addresses the HIV-specific challenges in identifying people living with HIV/AIDS from claims data files, including coding errors such as miscoding of rule-out diagnoses.

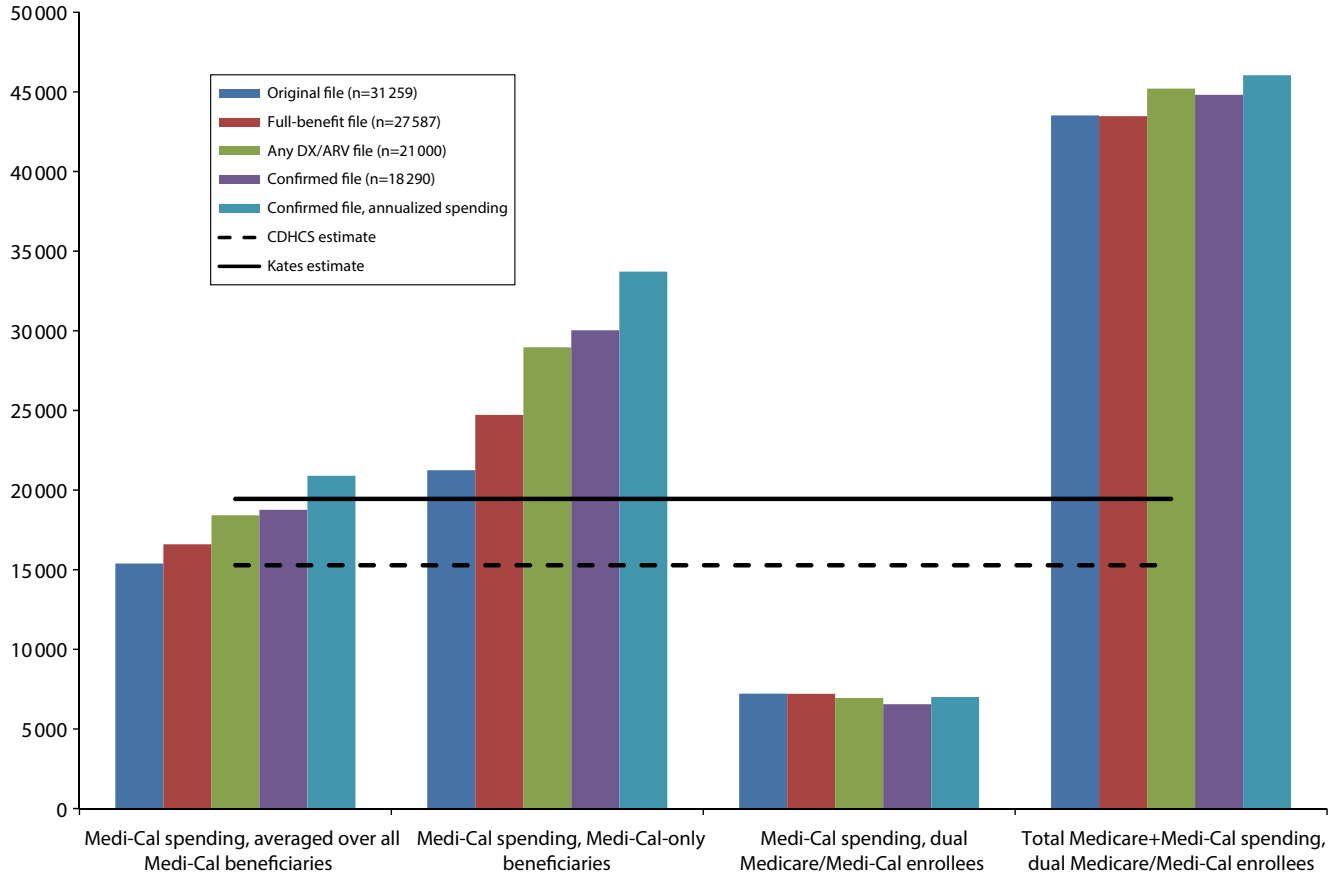
## Using Claims Data to Estimate HIV Treatment Costs

Accurately identifying actual HIV cases in the claims data has potentially large policy

implications because managed care capitation rates are frequently based on calculated levels of FFS spending[22] and the true costs of treating HIV will be underestimated if the estimation sample includes individuals without HIV, especially if HIV screening is more likely to be performed among younger and healthier groups. Figure 2, which relates to Medi-Cal spending for both Medi-Cal and dual beneficiaries, illustrates that each of the adjustments that we have made (limiting the sample to those with full Medi-Cal benefits, removing managed care enrollees, including only confirmed cases, and annualizing expenditures of part-year enrollees) increases the estimate of per capita Medi-Cal spending on HIV. Furthermore, per capita spending levels are significantly higher in each group for persons with Medi-Cal only, because the vast majority of the treatment costs for those dually eligible for Medicaid and Medicare are borne by Medicare, as shown in Figure 2. Thus, although duals had significantly higher total treatment costs than Medi-Cal–only beneficiaries, the costs paid by Medi-Cal were lower.

We based our calculation of annual Medi-Cal treatment costs for persons with confirmed HIV/AIDS on our estimate that there were 18 290 Californians with HIV who had only FFS Medi-Cal coverage. Medi-Cal costs for these individuals averaged $33 720 in 2007 ($43 500 in 2013 dollars). By contrast, the California Department of Health Care Services (CDHCS)[23] and Kates[24] used less targeted definitions to identify people living with HIV/AIDS who were treated by Medicaid. The overestimation of the population of interest resulted in underestimating the average annual Medicaid costs for treating individuals with confirmed HIV/AIDS who were insured by Medicaid only.

The CDHCS reported that in state fiscal year 2008, Medi-Cal paid claims for 33 100 adults living with AIDS, which is close to our estimate of 31 259 Medi-Cal recipients aged 16 years and older with HIV, before deleting from the sample those with partial benefits or managed care coverage, persons not confirmed to be HIV-positive, dual enrollees, and part-year enrollees. The CDHCS calculation is higher because it is based on all fiscal year 2008 Medi-Cal recipients with a diagnosis of HIV/AIDS at any time since July 1994. Averaging Medi-Cal costs over 33 100 recipients results




*Note*. ARV = antiretroviral; CDHCS = California Department of Health Care Services; DX = diagnosis. CDHCS estimates from Leibowitz, Mendes and, Desmond.[23] Kates estimates from *Medicaid and HIV: A National Analysis*.[24] The "Any DX/ARV file" includes beneficiaries who are Californians aged ≥ 16 years, enrolled in fee-for-service Medi-Cal with full benefits, and had at least 1 claim with a qualifying HIV diagnosis code or 1 prescription fill for an ARV medication.

**FIGURE 2—Estimated per-capita annual Medi-Cal spending on persons living with HIV/AIDS, by sample inclusion criteria: California, 2007.**

in an estimate of per capita Medi-Cal spending on people living with HIV/AIDS in fiscal year 2008 of nearly $15 000.[23]

Kates[24] estimated that there were 23 350 Medicaid enrollees with HIV in California in 2007, including those enrolled full and part year, with either full or limited benefits, and both with and without supplemental Medicare coverage. Kates estimated that per capita annual Medicaid spending for this sample of Californians was $19 447. Our estimate of $33 720 addresses a different question: What was the annualized, per capita cost of HIV care for FFS Medi-Cal enrollees without additional Medicare coverage? The CDHCS and Kates estimates seem to suggest that Medicaid treatment costs were well below the $29 808 average calculated by Gebo et al.[18] for individuals with disability that would qualify them for Medicaid (CD4 counts < 200 cells per microliter). The Gebo et al. estimates were imputed from 2006 medical records at 3 HIV Research Network clinics and correspond well to our estimate of per capita costs of $33 720 in 2007. Thus, Medicaid support for HIV treatment appears to be at a similar level to the cost of care in clinics with a mix of private and public payers. Without excluding duals from the average cost calculation, one might erroneously conclude that Medicaid enrollees with HIV received significantly less care than persons with private insurance.

## Limitations

Our analysis is subject to some limitations. Although the algorithm we propose deletes from the confirmed HIV sample persons most likely to have been assigned an HIV diagnosis code erroneously, it is possible that some of the individuals with only a rule-out diagnosis or other data errors may remain in the sample. However, available evidence strongly supports the distinctions between the HIV sample included under the strict definition (80% of whom had evidence of viral load or CD4 testing) and the excluded group (in which only 3% had viral load or CD4 testing). The high prevalence of HBV diagnoses or HBV-consistent use of ARVs among the excluded group (37% vs 7% in the confirmed file; Table 2) confirms that this exclusion was likely effective.

In addition to this strict definition group with the strongest evidence of HIV infections, there is a group of people that did not have diagnoses recorded on multiple visits, but did receive an HIV diagnosis on a nonscreening day. One third of people in this group received viral load or CD4 tests, and 39% had claims for ARVs. The cases added under the expanded definition included a larger proportion of female patients, suggesting that we may have added some cases in which a rule-out diagnosis was recorded in conjunction with a screening test, but did not appear in the claims on the same day.

Although we are less confident that everyone in this latter group is actually HIV-positive, there is enough evidence of some infection that it is difficult to ignore them. Using criteria for inclusion that are too strict will result in missing many HIV cases, particularly individuals with lower levels of health care utilization, who have fewer visits on which to detect an HIV/AIDS diagnosis. The FFS beneficiaries in the expanded definition sample, who had a diagnosis on a nonscreening day, but did not have 2 spaced outpatient visits or 1 inpatient stay with a diagnosis, are likely to include some people who are HIV-positive, but have lower medical use than the group meeting the strict definition. To exclude them from the file of HIV-positive persons could result in an upward bias in estimates of annual expenditures and would also bias quality-of-care measures that require evidence of regular visits.[25]

## Conclusions

As outpatient care for people living with HIV/AIDS continues to increase and HIV screening tests become a routine part of medical care, as recommended by the Centers for Disease Control and Prevention,[26] it becomes increasingly important to be able to identify people living with HIV/AIDS in health insurance claims data to assess the cost and quality of HIV care rendered by different providers. This is particularly true for public insurance programs. Medicare and Medicaid account for more than three quarters of domestic federal spending on HIV/AIDS treatment[27] and in California these 2 programs covered almost half of people living with HIV/AIDS in 2007.[23]

As publicly funded insurance, these programs are of particular importance to policymakers as they seek to monitor the cost and quality of treatments received by publicly insured people living with HIV/AIDS and to set managed care capitation rates for Medi-Cal enrollees living with HIV. California is moving all Medi-Cal beneficiaries, including the disabled, who previously had been allowed to remain in FFS, into managed care. Our method should help policymakers avoid underestimating the cost of treatment of HIV/AIDS (if dual-eligibles, part-year enrollees, or people with rule-out diagnoses are included in average Medicaid costs) and also avoid overestimation (if persons with HIV who receive irregular care are not included in the sample). By setting managed care capitation rates for people living with HIV/AIDS appropriately, policymakers can enhance the sustainability of managed care providers of HIV/AIDS treatment. ■


## About the Authors

*Both authors are with the Department of Public Policy, Luskin School of Public Affairs, University of California Los Angeles.*

*Correspondence should be sent to Arleen A. Leibowitz, Research Professor, Department of Public Policy, UCLA Luskin School of Public Affairs, 6345 Public Affairs Building, Box 951656, Los Angeles, CA 90095-1656 (e-mail: leibowia@luskin.ucla.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.*



*This article was accepted July 31, 2014.*


## Contributors

Both authors jointly designed the analysis plan, interpreted the results, wrote the article, and critically reviewed and approved the report. K. Desmond carried out the empirical analysis.


## Acknowledgments

This research was supported by grants from the California HIV/AIDS Research Program of the University of California (grant RP11-LA-020). This work was also supported by the Center for HIV Identification, Prevention, and Treatment, National Institute of Mental Health (grant MH58107); the University of California Los Angeles Center for AIDS Research (grant 5P30AI028697); and the National Center for Advancing Translational Sciences through University of California Los Angeles Clinical and Translational Science Institute (grant UL1TR000124).

We are grateful to Jennifer Gildner for programming assistance and to Thomas Rice, PhD, and to Raphael Landovitz, MD, MSc, for comments on a previous draft.

**Note:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the California HIV/AIDS Research Program or the National Institutes of Health.


## Human Participant Protection

The University of California Los Angeles institutional review board approved this research by expedited review (UCLA IRB G09-10-016-01).

## References


1. White House Office of National AIDS Policy. National HIV/AIDS Strategy for the United States. 2010. Available at: http://www.whitehouse.gov/sites/default/files/uploads/NHAS.pdf. Accessed May 15, 2012.

2. Centers for Disease Control and Prevention. HIV Surveillance Report, 2010. Vol 22; 2012. Available at: http://www.cdc.gov/hiv/topics/surveillance/resources/reports. Accessed May 15, 2012.

3. Fleishman JA, Gebo KA. HIV Research Network Report to the IOM committee. In: Ford MA, Spicer CM, eds. *Monitoring HIV Care in the United States: Indicators and Data Systems*. Washington, DC: The National Academies Press; 2012.

4. Blair JM, Fagan JL, Frazier EL, et al. Behavioral and clinical characteristics of persons receiving medical care for HIV infection—Medical Monitoring Project, United States, 2009. *MMWR Surveill Summ*. 2014;63(suppl 5):1–22.

5. Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P. Assessing the accuracy of administrative data in health information systems. *Med Care*. 2004;42(11):1066–1072.

6. Crystal S, Akincigil A, Bilder S, et al. Studying prescription drug use and outcomes with Medicaid claims data: strengths, limitations, and strategies. *Med Care*. 2007;45(10 suppl 2):S58–S65.

7. Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention. Epidemiology of HIV infection through 2011. Available at: http://www.cdc.gov/hiv/library/slideSets/index.html. Accessed September 16, 2013.

8. *International Classification of Diseases, Ninth Revision*. Geneva, Switzerland: World Health Organization; 1980.

9. Ferver K, Burton B, Jesilow P. The use of claims data in healthcare research. *Open Public Health J*. 2009;2:11–24.

10. Rector TS, Wickstrom SL, Shah M, et al. Specificity and sensitivity of claims-based algorithms for identifying members of Medicare + Choice health plans that have chronic medical conditions. *Health Serv Res*. 2004;39(6 pt 1):1839–1857.

11. Leibowitz AA, Desmond K. Do only 21% of HIV-positive Medicaid enrollees link to treatment? Challenges in interpreting Medicaid claims data. Letter to the editor. *Sex Transm Dis*. 2013;40(7):582.

12. Keyes M, Andrews R, Mason ML. A methodology for building an AIDS research file using Medicaid claims and administrative data bases. *J Acquir Immune Defic Syndr*. 1991;4(10):1015–1024.

13. National Cancer Institute, Cancer Control and Population Sciences, Health Services and Economics. SEER-Medicare: calculation of comorbidity weights. Available at: http://healthservices.cancer.gov/seermedicare/program/comorbidity.html. Accessed May 16, 2013.

14. Walkup JT, Sambamoorthi U, Crystal S. Use of newer antiretroviral treatments among HIV-infected Medicaid beneficiaries with serious mental illness. *J Clin Psychiatry*. 2004;65(9):1180–1189.

15. Chesnut TJ, Laufer FN, Carrascal AF, Feldman IS. An expenditure analysis of high-cost Medicaid recipients with HIV disease in New York State. *J Health Care Poor Underserved*. 2011;22(1):330–345.

16. Flagg E, Weinstock H. Incidence of neonatal herpes simplex virus infection in the United States, 2006. *Pediatrics*. 2011;127(1):e1–e8.

17. US Department of Labor, Bureau of Labor Statistics. Consumer Price Index–July 2013. Available at: http://www.bls.gov/schedule/archives/cpi_nr.htm#2007. Accessed September 16, 2013.

18. Gebo KA, Fleishman JA, Conviser R, et al. Contemporary costs of HIV healthcare in the HAART era. *AIDS*. 2010;24(17):2705–2715.

19. Johnston SS, Juday T, Seekins D, et al. Patterns and correlates of linkage to appropriate HIV care after HIV diagnosis in the US Medicaid population. *Sex Transm Dis*. 2013;40(1):18–25.

20. Zhang S, Senteio C, Felizzola J, Rust G. Racial/ethnic disparities in antiretroviral treatment among HIV-infected pregnant Medicaid enrollees, 2005–2007. *Am J Public Health*. 2013;103(12):e46–e53.

21. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Morb Mortal Wkly Rep*. 2006;55(RR14);1–17.

22. Medicaid Managed Care and HIV/AIDS. Washington, DC: Office of the Inspector General, US Department of Health and Human Services; 1998. OEI-05-97-00210.

23. Leibowitz AA, Mendes AC, Desmond K. Public funding of HIV/AIDS prevention, treatment, and support in California. *J Acquir Immune Defic Syndr*. 2011;58(1):e11–e16.

24. Kates J. *Medicaid and HIV: A National Analysis*. Washington, DC: The Henry J. Kaiser Family Foundation; 2011.

25. National Committee for Quality Assurance. Recommendations for health insurance exchange quality measure requirements. Available at: http://www.ncqa.org. Accessed September 16, 2013.

26. Centers for Disease Control and Prevention. Advancing HIV prevention: new strategies for a changing epidemic. *MMWR Morbid Mortal Wkly Rep*. 2003;52(15);329–332.

27. Kaiser Family Foundation. U.S. federal funding for HIV/AIDS: the president's FY2013 budget request. HIV/AIDS policy fact sheet. 2012. Available at: http://kff.org/hivaids/fact-sheet/u-s-federal-funding-for-hivaids-the-presidents-fy-2014-budget-request. Accessed June 17, 2013.