Practice of Epidemiology

Using Registry Data to Construct a Comparison Group for Programmatic Effectiveness Evaluation

The New York City HIV Care Coordination Program

McKaylee M. Robertson*, Levi Waldron, Rebekkah S. Robbins, Stephanie Chamberlin, Kate Penrose, Bruce Levin, Sarah Kulkarni, Sarah L. Braunstein, Mary K. Irvine, and Denis Nash

* Correspondence to McKaylee M. Robertson, Institute for Implementation Science in Population Health, Graduate School of Public Health and Health Policy, 55 West 125th Street New York, NY 10027 (e-mail: mckaylee.robertson@sph.cuny.edu).

Many nonrandomized interventions rely upon a pre-post design to evaluate effectiveness. Such designs cannot account for events external to the intervention that may produce the outcome. We describe a method to construct a surveillance registry–based comparison group, which allows for estimating the effectiveness of the intervention while controlling for secular trends in the outcome of interest. Using data from the population-based, human immunodeficiency virus Surveillance Registry in New York City, we created a contemporaneous comparison group for persons enrolled in the New York City human immunodeficiency virus Care Coordination Program (CCP) from December 2009 to March 2013. Inclusion in the Registry-based (non-CCP) comparison group required meeting CCP eligibility criteria. To control for secular trends in the outcome, we randomly assigned persons in the non-CCP, Registry-based comparison group a pseudoenrollment date such that the distribution of pseudoenrollment dates matched the distribution of enrollment dates among CCP enrollees. We then matched CCP to non-CCP persons on propensity for enrollment in the CCP, enrollment dates, and baseline viral load. Registry-based comparison group estimates were attenuated relative to pre-post estimates of program effectiveness. These methods have broad applicability for observational intervention effectiveness studies and programmatic evaluations for conditions with surveillance registries.

comparative effectiveness research; implementation science; population-based surveillance registry; programmatic evaluation; propensity score; retrospective observational studies; viral load suppression

Abbreviations: ART, antiretroviral treatment; CCP, Care Coordination Program; CI, confidence interval; HIV, human immunodeficiency virus; PLWH, persons living with HIV; RR, relative risk; VL, viral load; VLS, viral load suppression.

Often large-scale, programmatic interventions are implemented and scaled without a contemporaneous comparison group for assessment of effectiveness (1). In such situations, a single-sample "pre-post design" is frequently used to evaluate effectiveness (1–11). However, such historical comparisons cannot account for external events (i.e., secular trends) that may influence the outcome independently of the intervention (11–16). For example, persons enrolled in the New York City human immunodeficiency virus (HIV) Care Coordination Program (CCP) between 2009 and 2013 experienced significant improvement in short-term HIV viral load suppression (VLS), based on comparison of the year before to the year after enrollment (3, 17).

However, VLS among all New York City residents in HIV medical care (including those enrolled in the CCP) steadily increased in each year from 2009 to 2013, and improvements coincided with the implementation of multiple population-based HIV treatment strategies (e.g., the recommendation that all persons living with HIV (PLWH) initiate antiretroviral treatment (ART) at diagnosis regardless of CD4+ lymphocyte count) (18, 19).

For many interventions implemented in real-world, programmatic settings, an obvious contemporaneous comparison population may not exist, especially after the program has been rolled out and scaled up. However, short of an advance-planned experimental or quasiexperimental design, a design incorporating a

contemporaneous comparison population, subject to the same external events but not receiving the intervention, can be used to isolate effects of the intervention versus secular trends ([12]). The increasing availability of large administrative, clinical, and surveillance registries, and the ability to link these data sets to one another, are expanding opportunities to use secondary data to create contemporaneous comparison populations and evaluate interventions in real-world settings ([20]). The lack of straightforward methods to select a contemporaneous comparison population and control for secular trends nonetheless limits the usefulness of registries for building and strengthening the evidence base for public health practice.

Here we describe a methodologic approach that builds upon the single-sample pre-post design, to corroborate and/or refine pre-post effect estimates ([3], [17]). The objective of this study was: 1) to select from a surveillance registry a contemporaneous comparison group for evaluation of programmatic effectiveness; and 2) to compare estimates of programmatic effectiveness obtained from a pre-post design with estimates obtained from a registry-based comparison group. To aid in achieving our objective, we present a detailed application that assesses the effect of the CCP on HIV VLS.

METHODS

Intervention description

In December 2009, with Ryan White Part A funding, the New York City Department of Health and Mental Hygiene launched the HIV CCP to support persons at high risk for suboptimal HIV care outcomes ([3], [21]). The intervention has previously been described, and the program materials are available on the New York City health department website ([3], [21]).

Briefly, the CCP combines various evidence-based elements into a package that includes case management, multidisciplinary case conferencing, patient navigation, structured health promotion, and adherence support. These services can be tailored to individual needs. Importantly, the CCP was rolled out as a service program, with no randomization or comparison groups.

Data sources

We constructed an observational cohort of persons enrolled and not enrolled in the CCP by merging provider-reported programmatic data with data from the longitudinal population-based New York City HIV Surveillance Registry ("the Registry"). The Registry contains demographic and laboratory information on all diagnoses of HIV (since 2000) and AIDS (since 1981) reported in New York City and comprehensive HIV-related laboratory reporting (including all CD4 and viral load (VL) test results) starting in 2005. Vital status information is updated through regular matches with local and national death data. All Ryan White Part A service providers are contractually obligated to submit person-level and service-level data through the Electronic System for HIV/AIDS Reporting and Evaluation ([3]).

Using data from the Electronic System for HIV/AIDS Reporting and Evaluation, we identified all persons who enrolled in the CCP from December 1, 2009, to March 31, 2013 (the enrollment period), and excluded clients who died within 12 months of program enrollment ($n = 279$) to ensure adequate

observation time. Using Registry data, we identified persons who were ≥18 years old, were diagnosed with HIV as of March 31, 2013, and had at least one CD4 or VL test reported to the Registry during December 1, 2007, to March 31, 2013. We required a laboratory test in that period to ensure that clinical eligibility could be assessed for the period during which CCP enrollees became eligible for enrollment. The laboratory window started 2 years earlier than the enrollment period, because CCP eligibility was based on clinical status in the 2 years prior to enrollment.

By merging program data with Registry data, we were able to identify, within the Registry, the persons who were enrolled and those who were not enrolled in the CCP; the latter were potentially eligible for inclusion in the Registry-based comparison group of non-CCP PLWH ([3]). To ensure comparability for CCP and non-CCP persons, all data, aside from CCP enrollment status, were taken from the Registry.

This study was approved by the institutional review boards at the City University of New York and the New York City Department of Health and Mental Hygiene. For these secondary analyses of deidentified data, we received a waiver for informed consent under Title 45, Code of Federal Regulations Part 46.116 (d) ([2]).

Constructing the observational comparison groups

Development of the Registry-based comparison group. In 4 steps, we identified a group of PLWH who were similar to CCP enrollees. First, we identified persons who were not enrolled in the CCP but met broad clinical eligibility criteria for enrollment, and we assigned persons eligibility windows— months where the persons appeared eligible for enrollment in the CCP. Second, we assigned eligible non-CCP PLWH pseudoenrollment dates falling within their windows of eligibility. Third, we restricted the data set to PLWH who were residents of New York City. Finally, we matched CCP enrollees to non-CCP PLWH according to: 1) baseline VL status in the year prior to pseudoenrollment/enrollment, 2) pseudoenrollment/enrollment dates, and 3) propensity for enrollment in the CCP.

Assignment of CCP eligibility windows. To identify a subset of persons who were eligible but not enrolled in the CCP, we used Registry-based criteria that aligned closely with CCP eligibility criteria to create CCP eligibility windows: ranges of time (in month-year format) between December 2009 and March 2013 where the person appeared eligible for enrollment in the CCP. The CCP permits enrollment of HIV-infected adults or emancipated minors who are eligible for local Ryan White Part A services (based on residence in the New York City grant area and a household income <435% of federal poverty level) and are: newly diagnosed with HIV, not consistently in medical care, and/or experiencing ART challenges ([3]). For a description of the Registry-based eligibility criteria and enrollment eligibility windows, see Web Table 1 (available at https://academic.oup.com/aje).

Persons could be assigned multiple eligibility windows based on qualifying for the CCP via multiple Registry criteria and/or qualifying for the same criteria at different points in time. For example, if Person X was newly diagnosed as of February 2010 and was not consistently in HIV medical care from February 2011 to December 2011, then Person X was assigned 2 CCP

eligibility windows for being newly diagnosed and for being inconsistently in HIV medical care (Web Table 1). Overlapping eligibility windows were combined. For persons who died, eligibility windows ended ≥12 months prior to the date of death, to ensure adequate time to observe the 12-month outcome; no exclusions were made in later steps based on death.

Assignment of pseudoenrollment dates to the non-CCP group. Persons in the Registry who met CCP eligibility criteria but were not enrolled in the CCP were randomly assigned a pseudoenrollment date: the time point used to start follow-up and outcome assessment that fell within their CCP eligibility window(s). Pseudoenrollment dates were assigned with probabilities so that the distribution of pseudoenrollment dates in the Registry-based comparison group would match the distribution of enrollment dates among CCP enrollees.

We assigned pseudoenrollment dates based on a probability equal to the ratio of the number of persons needed to the number of persons eligible. For example, if 10 persons enrolled in the CCP in March of 2011 and 100 non-CCP persons were eligible in March 2011, then the ratio was estimated as 0.1 (10 needed/100 available). We randomly selected a subset based on the number needed (i.e., 10 persons needed) and assigned them that month as the pseudoenrollment date (i.e., March 2011). To ensure the maximal number of non-CCP persons were assigned a pseudoenrollment date, we started with the month with the greatest need (i.e., highest ratio of persons needed to persons eligible) and proceeded to the month with the second-greatest need and so on, until we reached the month with the least need. For each iteration, we removed from the eligible pool the PLWH who had been assigned a date.

Restriction to New York City residents. After pseudoenrollment dates were assigned, we restricted the eligible pool to CCP and non-CCP persons who had at least 1 CD4 or VL test reported to the Registry in the 24 months after the pseudoenrollment/enrollment date. We required a laboratory test to identify persons who were most likely to be residing in New York City after the pseudoenrollment date; this helped to account for out-migration, which we suspected to have occurred more frequently among non-CCP persons than CCP participants, given that CCP enrollment and services require residence in New York City.

Propensity score model and match. Correctly specified propensity models balance measured confounders across exposure groups (22). We estimated the propensity score by modeling exposure status as a function of the confounders of the relationship between exposure and outcome. To begin, we developed an a priori list of variables considered to be potential confounders of the relationship between enrollment in the CCP and the outcome of VLS (Table 1).

Model development. We hypothesized that outcomes would differ according to baseline VL status in the year before pseudoenrollment/enrollment date and wanted to estimate the effect of the CCP according to baseline status. We created 4 baseline VL status groups, as a proxy for ART and adherence status, for those who were: 1) newly diagnosed (in the 12 months prior to pseudoenrollment/enrollment date and not expected to have a viral suppression pattern given the recency of diagnosis) or previously diagnosed (i.e., not newly diagnosed) and 2) had consistent suppression (≥2 VLs ≥90 days apart and all VLs ≤200 copies/μL in the 12 months prior to the pseudoenrollment/

enrollment date), 3) had no evidence of suppression (i.e., all VLs reported >200 copies/μL or no VL tests reported to the Surveillance Registry in the 12 months prior to pseudoenrollment/enrollment), or 4) had inconsistent suppression (≥1 VL ≤200 copies/μL in the 12 months prior to pseudoenrollment/enrollment, but not all VLs ≤200 copies/μL or not ≥2 VLs ≥90 days apart).

We used logistic regression to estimate the propensity for enrollment in the CCP among 1) persons newly diagnosed, 2) persons with consistent suppression, and 3) persons with inconsistent suppression or with no evidence of suppression (i.e., 3 independent propensity models, including a 2-group pooled model). We started with a model that used all a priori confounders and used backward selection to identify the model with the lowest value of Akaike information criterion. In a sensitivity analysis, we fitted models using all hypothesized confounders. The effect estimates did not differ from the approach described below; however, fewer CCP enrollees were matched.

We created 3 models, as opposed to constructing one pooled model with interaction terms for the baseline VL status, because we hypothesized that these 3 groups had different potential confounders (Web Table 2). For persons with inconsistent suppression and persons with no evidence of suppression, we examined a 2-group pooled model (i.e., number 3 above) and separate models. We used the 2-group pooled model because we were able to match more persons than with separate models. In a sensitivity analysis, the effect estimates from the 2-group pooled model did not differ from the 2 individual models.

Match. Within each of the 4 baseline VL status groups, we matched on propensity scores and pseudoenrollment/enrollment dates. To match on enrollment dates, we subset the population to non-CCP persons who had a pseudoenrollment date within the 3 months before or after a given CCP month of enrollment, and we ordered the match iteration to proceed from months with the highest to the lowest ratio of the number of CCP enrollees to the number of non-CCP persons. We matched on propensity score using a 1:1 greedy match technique, which proceeded sequentially from 8 to 1 decimal places of the propensity score (23, 24). A non-CCP person could be included once in the Registry-based comparison group.

Model and match diagnostics. The degree to which confounding was controlled can be evaluated by examining "balance," or the distribution of potential confounders by exposure status (22). To assess balance, we examined the standardized difference of all confounder variables between CCP and non-CCP PLWH (22, 23, 25). Multilevel categorical confounders were represented using sets of binary indicators (22). We considered a standardized difference of ≥0.1 to indicate an imbalance in the measured confounders between CCP and non-CCP Registry-based groups (22). We evaluated balance of each hypothesized confounder (shown in Table 1) within each baseline VL status group and added interaction terms to the model as needed. The final model and match was chosen based on having no imbalances ≥0.1 and the greatest number of persons matched (22).

Development of the single-sample pre-post comparison group. To ensure comparability of the Registry-based comparison group estimates of CCP effectiveness with the pre-post estimates of effectiveness, we used the propensity-matched CCP enrollees to sample the pre-post comparison group. The propensity-matched CCP enrollees served as their own

Table 1.   Variables and Variable Classification Used in Propensity Models as Potential Confounders of the Relationship Between Enrollment in the Human Immunodeficiency Virus Care Coordination Program and the Outcome of Viral Suppression, New York City, 2018

Variable	Classification	Detail
Sex	Male, female	
Race	Black, Hispanic, white, other	
Age at HIV diagnosis, years	≤24, 25–44, 45–64, ≥65	
Country of birth	United States or United States dependency, foreign-born, unknown	
Transmission risk	Men who have sex with men, heterosexual, injection drug use, other/unknown	
Year of HIV diagnosis	Prior to 1995, 1995–1999, 2000–2004, 2005–2009, 2010–2013	For persons who were newly diagnosed (within 1 year of the pseudoenrollment or enrollment date), year of diagnosis was categorized as 2009–2010, 2011, or 2012–2013.
Baseline viral load	No viral load, ≤200 copies/μL, 201–1,500 copies/μL, >1,500 copies/μL	Based on the last viral load reported before the pseudoenrollment or enrollment date.
Baseline CD4+ lymphocyte count	No CD4, <200, 200–349, 350–499, ≥500	Based on the last CD4 count reported before the pseudoenrollment or enrollment date.
Linkage within 91 days of HIV diagnosis	Yes, no	A viral load or CD4 count or CD4% reported from 0–7 days after diagnosis did not indicate linkage.
Concurrent AIDS diagnosis within 365 days	Yes, no	AIDS diagnosis reported within 365 days of HIV diagnosis.
Number of viral load laboratory events at enrollment	0, 1–3, ≥4	Number of viral load values reported in the 12 months before pseudoenrollment or enrollment date.
Prevalence and poverty at enrollment	High poverty and high prevalence, low poverty and high prevalence, high poverty and low prevalence, low poverty and low prevalence, unknown	Poverty was based on the level of poverty in the zip code at enrollment according to the American Community Survey and was classified as high versus low. Prevalence was based on the HIV prevalence in the zip code at enrollment and was classified as high (prevalence greater than the median HIV prevalence for a given year of enrollment) versus low.
Zip code at enrollment		New York City zip codes
Interaction terms:	Baseline CD4 and viral load, baseline CD4 and race, sex and risk, and year of diagnosis and risk	

Abbreviations: AIDS, acquired immune deficiency syndrome; HIV, human immunodeficiency virus.

controls, comparing the year after with the year prior to enrollment. We restricted the matched group to persons who were previously diagnosed (i.e., excluding the newly diagnosed, who may not have had a measure before enrollment).

Outcome definition and CCP effectiveness estimates

For the pre-post estimate of effectiveness, we compared 12-month postenrollment with pre-enrollment proportions with VLS among CCP enrollees. For the Registry-based comparison group estimate, we compared postenrollment VLS for CCP and non-CCP PLWH. VLS was based on the most recent VL laboratory result reported to the Registry in the 12 months prior to or following the enrollment/pseudoenrollment date and dichotomized as ≤200 copies/μL. Persons with no VL in the Registry for the entire 12-month period were classified as not having VLS, given that a 12-month gap in VL monitoring likely represented loss to care. While rare, 1.3% (87/6,812) of the CCP

and 5.4% (353/6,812) of the non-CCP PLWH were missing a follow-up VL result.

We fitted a log-binomial model to estimate the relative risk of 1) post- versus pre-enrollment and 2) the CCP versus usual care (non-CCP) on VLS. We estimated relative risks using the method of generalized estimating equations in order to account for the repeated measures or the matched pair design, respectively. Relative risks were estimated with the GENMOD procedure in SAS, version 9.3 (SAS Institute, Inc., Cary, North Carolina). The relative effect of the CCP was estimated within baseline VL status groups.

RESULTS

As of March 2013, 90,361 persons ≥18 years of age were diagnosed with HIV and had a laboratory test reported in the Registry from December 2007 to March 2013 (Figure 1). A total of 7,337 persons enrolled in the CCP from December 1,
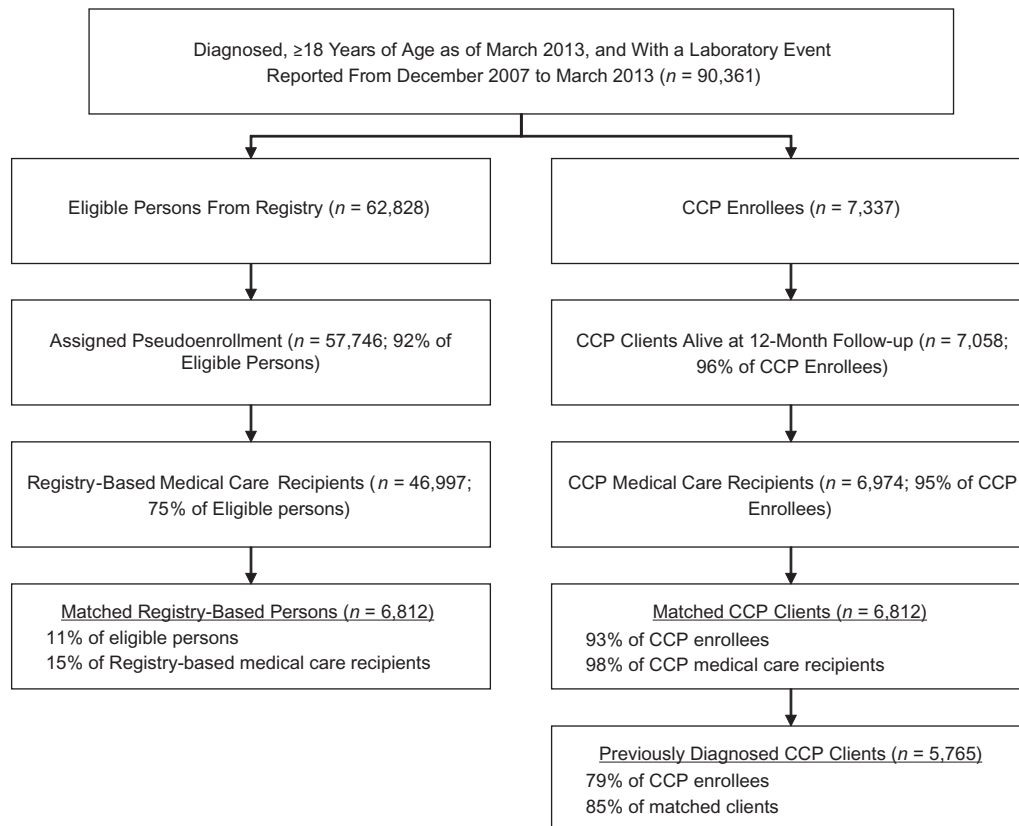
Figure 1. Flow chart for inclusion in the study comparing members of the population-based, human immunodeficiency virus Surveillance Registry with participants in the human immunodeficiency virus Care Coordination Program (CCP) or for inclusion in the study comparing pre-post results for the CCP enrollees, New York City, 2009–2013. Surveillance Registry members who were not enrolled in CCP met CCP eligibility criteria. In both arms, medical care was defined as at least one CD4+ lymphocyte or viral load laboratory test in the 24-month follow-up period.

2009, to March 31, 2013. Of these, 7,058 (96%) were alive at 12-month follow-up, and 6,974 (95%) received HIV care in New York City after their enrollment and were considered to reside in New York City. A total of 62,828 persons were not enrolled in the CCP and met at least 1 of the Registry-based eligibility criteria within the December 2007 to March 2013 period. Of these, 57,746 (92%) were assigned a pseudoenrollment date between December 2009 and March 2013, and 46,997 (75%) received HIV care in New York City after their pseudoenrollment date. A total of 6,812 (93% of all CCP enrollees) were matched to a non-CCP person in New York City based on propensity score, baseline VL status, and enrollment date, and 5,765 CCP enrollees (85% of matched CCP enrollees) were previously diagnosed and included in the pre-post group.

Prior to matching, the group of non-CCP persons included more men and persons aged ≥45 years than the group of CCP persons (75% vs. 64% and 56% vs. 50%, respectively) (Table 2). Pseudoenrollment and enrollment dates differed between CCP and non-CCP PLWH within a baseline VL status group (e.g., the median date among newly diagnosed was November 2011 for non-CCP and August 2011 for CCP persons).

After matching, non-CCP persons were demographically, clinically, and epidemiologically similar to the CCP persons

(Table 2). All 4 matched groups exhibited balance (standardized difference <0.1) across all hypothesized confounders. The post-match pseudoenrollment/enrollment date distribution of the 10 persons mirrored that of the CCP persons, even within baseline VL status groups (Web Figure 1).

The proportion of previously diagnosed CCP enrollees with VLS increased from 34% prior to enrollment to 58% in the 12 months after enrollment (Table 3). In the pre-post analysis, which used individuals enrolling in CCP as their own controls, the previously diagnosed CCP enrollees were 1.7 times as likely to have VLS after enrollment (relative risk (RR) = 1.68 (95% confidence interval (CI): 1.62, 1.74)). Among CCP enrollees with no evidence of suppression, the proportion with VLS rose from 0% (by definition) to 43%, and the relative risk is indeterminate due to the 0% VLS in the denominator. Among the CCP enrollees with consistent suppression, the proportion with VLS dropped from 100% (by definition) to 92% (RR = 0.92, 95% CI: 0.90, 0.94). The group with inconsistent suppression showed significantly higher VLS after versus before CCP enrollment (62% versus 46%, respectively; RR = 1.22, 95% CI: 1.16, 1.28).

In the Registry-based comparison group analysis, 12 months after enrollment/pseudoenrollment CCP enrollees were 1.1 times as likely to have VLS compared with non-CCP enrollees (58% versus 52%, respectively; RR = 1.10, 95% CI: 1.07,

Table 2. Characteristics Before and After Matching of Persons From the Surveillance Registry Who Met Clinical Eligibility Criteria for Enrollment in the Care Coordination Program and Enrollees in the Care Coordination Program, New York City, 2009–2013

Characteristic	Before Matching				After Matching			
	Registry-Based		CCP		Registry-Based		CCP	
	No.	%	No.	%	No.	%	No.	%
Total	46,997	100	6,974	100	6,812	100	6,812	100
Male	33,920	75	4,460	64	4,379	64	4,371	64
Black	21,931	47	3,715	53	3,603	53	3,645	54
≥45 years of age	25,013	56	3,476	50	3,334	49	3,400	50
Baseline CD4 count of <200	6,466	14	2,292	33	2,135	31	2,200	32
Median pseudo-/enrollment date[a]	March 2011 (September 2010, February 2012)		March 2011 (September 2010, February 2012)		March 2011 (September 2010, February 2012)		March 2011 (September 2010, February 2012)	
Newly diagnosed, total[b]	6,590	100	1,092	100	1,044	100	1,044	100
Male	5,256	80	802	73	775	74	778	75
Black	2,484	43	495	45	487	47	481	46
≥45 years of age	1,746	37	262	24	257	25	254	24
Baseline CD4 count of <200	830	13	283	6	238	23	253	24
Median pseudo-/enrollment date[a]	November 2011 (January 2011, August 2012)		August 2011 (October 2010, June 2012)		August 2011 (November 2010, June 2012)		August 2011 (October 2010, June 2012)	
Consistent suppression, total[c]	4,917	100	1,022	100	967	100	967	100
Male	3,714	76	613	60	587	61	587	61
Black	1,982	40	465	46	438	45	445	46
45+	3,113	63	714	70	662	68	664	69
Baseline CD4 count of <200	221	5	76	7	83	9	68	7
Median pseudo-/enrollment date[a]	May 11 (November 10, March 12)		February 11 (September 10, March 12)		February 11 (September 10, December 11)		February 11 (September 10, December 11)	
No evidence of viral suppression, total[d]	18,742	100	2,889	100	2,833	100	2,833	100
Male	13,089	70	1,828	63	1,745	62	1,791	63
Black	9,551	51	1,684	58	1,643	58	1,652	58
≥45 years of age	9,653	52	1,342	46	1,306	46	1,324	47
Baseline CD4 count of <200	3,121	17	1,354	47	1,249	44	1,303	46
Median pseudo-/enrollment date[a]	January 2011 (August 2010, September 2011)		March 2011 (September 2010, January 2012)		February 2011 (September 2010, January 2012)		March 2011 (September 2010, January 2012)	
Inconsistent suppression, total[e]	16,748	100	1,971	100	1,968	100	1,968	100
Male	11,861	71	1,217	62	1,272	65	1,215	62
Black	7,558	45	1,068	54	1,035	53	1,067	54
≥45 years of age	10,501	63	1,158	59	1,109	56	1,158	59
Baseline CD4 count of <200	2,294	14	579	29	565	29	576	29
Median pseudo-/enrollment date[a]	March 11 (September 10, February 12)		February 11 (September 10, January 12)		February 11 (September 10, January 12)		February 11 (September 10, January 12)	

Abbreviations: CCP, Care Coordination Program; IQR, interquartile range.

[a] Values are expressed as median (IQR).

[b] Newly diagnosed within 12 months of pseudoenrollment or enrollment.

[c] Consistent suppression: at least 2 viral load values at least 90 days apart and all viral loads <200 copies/μL in the 12 months prior to pseudoenrollment or enrollment.

[d] No evidence of viral suppression: All viral load values >200 copies/μL in the 12 months prior to pseudoenrollment or enrollment. This included persons missing viral load laboratory results. Persons with 1 viral load value of >200 copies/μL would be in this group.

[e] Inconsistent suppression: previously diagnosed and not in the "consistent suppression" or "no evidence of viral suppression" groups. Persons with 1 viral load value of ≤200 copies/μL would be in this group.

Table 3. Pre-Post Relative Risks for Having Viral Load Suppression Among Care Coordination Program Enrollees With Human Immunodeficiency Virus, New York City, 2009–2013

CCP Baseline Groups	Denominator	CCP Enrollees			
		% VLS Pre[a]	% VLS Post[a]	RR[b]	95% CI[b]
Overall	6,812	NA	59.9	NA	NA
Previously diagnosed[c]	5,768	34.2	57.5	1.68	1.62, 1.74
Baseline viral load status					
Newly diagnosed[d]	1,044	NA	73.3	NA	NA
Consistent suppression[e]	967	100	91.7	0.92	0.90, 0.94
No evidence of viral suppression[f]	2,833	0	42.5	Indeterminate	Indeterminate
Inconsistent suppression[g]	1,968	45.5	62.2	1.22	1.16, 1.28

Abbreviations: CCP, Care Coordination Program; CI, confidence interval; NA, not applicable; RR, relative risk; VLS, viral load suppression.

[a] Proportion with the last viral load result ($\leq$200 copies/μL) in the 12 months prior to enrollment. Persons missing a viral load value are considered not to have suppression (values of >200 copies/μL).

[b] Cannot compute the RR for persons with 0% viral load suppression. RRs were considered statistically significant when the 95% CIs for the relative risk excluded the null value of 1.

[c] Excludes the newly diagnosed baseline viral load status group.

[d] Newly diagnosed within 12 months of pseudoenrollment or enrollment.

[e] Consistent suppression: at least 2 viral load results at least 90 days apart and all viral load values <200 copies/μL in the 12 months prior to pseudoenrollment or enrollment.

[f] No evidence of viral suppression: All viral load values >200 copies/μL in the 12 months prior to pseudoenrollment or enrollment. This included persons missing viral load results. Persons with 1 viral load value of >200 copies/μL would be in this group.

[g] Inconsistent suppression: previously diagnosed and not in the "consistent suppression" or "no evidence of viral suppression" groups. Persons with 1 viral load value of $\leq$200 copies/μL would be in this group.

1.13) (Table 4). The proportion of persons with VLS differed by baseline VL status. Two baseline VL status groups showed significantly higher VLS among CCP versus non-CCP persons in the follow-up year: newly diagnosed PLWH (73% vs 63%, respectively; RR = 1.15, 95% CI: 1.09, 1.23) and PLWH who were previously diagnosed and had no evidence of suppression in the year prior to enrollment/pseudoenrollment (43% versus 32%, respectively; RR = 1.32, 95% CI: 1.23, 1.42). No differences in VLS between CCP and non-CCP persons were observed among those with consistent VLS (92% versus 91%, respectively; RR = 1.01, 95% CI: 0.98, 1.04) or with inconsistent VLS (62% versus 62%, respectively; RR = 0.99, 95% CI: 0.95, 1.05) in the year prior to enrollment/pseudoenrollment.

DISCUSSION

We developed and applied a rigorous method to identify and select from a surveillance registry a contemporaneous comparison population to estimate the effectiveness of a large-scale programmatic initiative, while controlling for both measured confounders and secular trends in the outcome of interest. We combined programmatic data on those receiving an HIV care coordination intervention with data from New York City's longitudinal, population-based HIV Surveillance Registry. The program effect derived from the Registry-based comparison group analysis was weaker or nonsignificant when compared with the program effect from the pre-post analysis that used individuals as their own controls. Our method for developing a contemporaneous comparison population should be applicable to many nonrandomized evaluations, as an alternative or follow-up to a pre-post design (12, 26).

Health outcomes may change over time due to influences (e.g., policy initiatives or scientific advances) outside of a particular intervention exposure (11, 12). To minimize the likelihood that secular trends in the outcome of interest were differentially affecting persons in the intervention and control arms, we applied random assignment of pseudoenrollment dates and further matching on pseudoenrollment dates. The pre-post analysis among those with inconsistent suppression showed a 22% increase in VLS; however, given the null findings from the non-CCP Registry-based comparison group, this pre-post effect may have been driven by secular improvements in VLS, as opposed to a program effect.

Propensity score matching has been proposed as a technique to reduce bias due to observed confounders and their unmeasured correlates when estimating the effect of exposure on outcome (22, 25). If implemented correctly, the propensity-matched group will be balanced on observed confounders; however, subgroups may not be balanced. To examine subgroups we recommend researchers create propensity models and match within groups of interest, as we did for baseline VL status.

For efficient causal inference, the intervention and control groups need to be as similar as possible (23). A randomized experiment uses random treatment assignment to ensure the intervention and control groups are balanced. Propensity score matching attempts to replicate this balance for observed covariates. An advantage of propensity score matching over regression adjustment for causal inference is that matching forces careful consideration of the balance between intervention and control groups (22, 23). This is evidenced by the changes in the demographic, clinical, and epidemiologic characteristics of our Registry-based population following the match. Regression

Table 4. Relative Risks for Having Viral Load Suppression, Among Persons With Human Immunodeficiency Virus Who Were Enrollees in the Care Coordination Program Versus Program-Eligible Persons Drawn From the Surveillance Registry, at 12 Months After Measurement, New York City, 2009–2013

Baseline Groups	Denominator (for CCP or Registry-Based)	CCP % VLS Post[a]	Registry-Based % VLS Post[a]	CCP Versus Registry-Based	
				RR[b]	95% CI[b]
Overall	6,812	59.9	53.9	1.11	1.08, 1.14
Previously diagnosed[c]	5,768	57.5	52.2	1.10	1.07, 1.13
Baseline viral load status					
Newly diagnosed[d]	1,044	73.3	63.3	1.15	1.09, 1.23
Consistent suppression[e]	967	91.7	90.6	1.01	0.98, 1.04
No evidence of viral suppression[f]	2,833	42.5	32.1	1.32	1.23, 1.42
Inconsistent suppression[g]	1,968	62.2	62.3	0.99	0.95, 1.05

Abbreviations: CCP, Care Coordination Program; CI, confidence interval; RR, relative risk; VLS, viral load suppression.

[a] Proportion with the last viral load result ($\leq$200 copies/µL) in the 12 months prior to enrollment. Persons missing a viral load value are considered not to have suppression (values of >200 copies/µL).

[b] RRs were considered statistically significant when the 95% CIs for the relative risk excluded the null value of 1.

[c] Excludes the newly diagnosed baseline viral load status group.

[d] Newly diagnosed within 12 months of pseudoenrollment or enrollment.

[e] Consistent suppression: at least 2 viral load results at least 90 days apart and all viral load values <200 copies/µL in the 12 months prior to pseudoenrollment or enrollment.

[f] No evidence of viral suppression: All viral load values >200 copies/µL in the 12 months prior to pseudoenrollment or enrollment. This included persons missing viral load results. Persons with 1 viral load value of >200 copies/µL would be in this group.

[g] Inconsistent suppression: previously diagnosed and not in the "consistent suppression" or "no evidence of viral suppression" groups. Persons with 1 viral load value of $\leq$200 copies/µL would be in this group.

models have been shown to perform poorly in situations where there is insufficient overlap of covariate distribution between the intervention and control groups; moreover, regression model diagnostics do not involve checking this overlap, and modest overlooked nonlinearity in regression models can increase rather than decrease bias (23, 27). Thus, the match method allows for a more carefully designed nonexperimental study (23, 25).

Our Registry-based comparison group approach had some limitations. First, not all program enrollment eligibility criteria could be translated to Registry-eligibility criteria. For example, persons with comorbidities (e.g., depression) were eligible for enrollment in the CCP, but information on comorbidities is unavailable in the Registry. However, a high proportion (92%) of CCP persons met Registry-based eligibility criteria, which suggests we may have captured most of the non-CCP persons eligible for the CCP. Second, the proposed matching approach may not work well when there is not a high ratio of comparison persons to intervention persons. In our example, nearly all persons in the intervention (98%) were successfully matched, enhancing generalizability of findings. Third, we excluded persons who had died in the first 12 months after pseudoenrollment/enrollment, and exclusion of cohort members with advanced disease and/or rapid disease progression may bias estimates (28). The CCP aims to enroll persons most at risk for poor HIV outcomes. However, persons with advanced HIV disease may be enrolled in the CCP as a last attempt to stop rapid disease progression, which does not happen with non-CCP PLWH because pseudoenrollment date assignment is random. To increase the comparability of the CCP and non-CCP groups, we required that individuals have $\geq$12

months of observation beyond their pseudoenrollment/enrollment. Finally, our approach has the standard limitations of observational research. Notably, confounding may not be fully controlled if measures are inaccurate or omitted from the propensity score. Of note, the pre-post approach has the advantage of controlling for many measured and unmeasured confounders, for which our registry-based comparison group approach may not have been able to account. However, the potential for bias may exist in the pre-post approach, given that we stratified on baseline VL status, which is temporally aligned with the 12 month period before enrollment. We've presented an overall (nonstratified) pre-post estimate that would not be subject to bias, and this bias would not affect the Registry-based comparison group estimates (stratified or nonstratified).

Administrative and clinical databases and surveillance registries are valuable sources for the identification of a contemporaneous comparison population. The conditions for identifying a contemporaneous comparison population were favorable because: 1) enrollment in the program was based largely on clinical criteria that could be emulated in the Registry; and 2) highly complete information on the outcome of interest was available from a single source for the intervention and control groups.

Others interested in adopting these techniques to identify an intervention-eligible yet nonexposed population, when a comparison population is not immediately obvious, need access to a secondary linkable database (i.e., with identifiers), in which they could assess program eligibility criteria and capture outcome information, and must possess the capability to link databases. Public health surveillance registries for conditions

such as tuberculosis, hepatitis C virus, diabetes, cancer, sexually transmitted infections, and immunizations may be utilized for this purpose. Researchers at public health departments and hospitals may be especially well-positioned to incorporate a contemporaneous comparison population into nonrandomized evaluation designs, given that these organizations often merge registries, manage programmatic activities, and maintain disease or clinical registries (29).

In summary, we developed and applied a technique for rigorous assessment of intervention effectiveness that controls for both measured confounders and secular trends in the outcome of interest. Moreover, analyses using a contemporaneous comparison group are useful for triangulation with findings from pre-post analyses. Our approach to develop a registry-based comparison group complements the pre-post design often used to assess the effect of nonrandomized interventions or exposures.

REFERENCES

1. Risher KA, Kapoor S, Daramola AM, et al. Challenges in the evaluation of interventions to improve engagement along the HIV care continuum in the United States: a systematic review. *AIDS Behav*. 2017;21(7):2101–2123.
2. Enriquez M, Farnan R, Cheng AL, et al. Impact of a bilingual/bicultural care team on HIV-related health outcomes. *J Assoc Nurses AIDS Care*. 2008;19(4):295–301.
3. Irvine MK, Chamberlin SA, Robbins RS, et al. Improvements in HIV care engagement and viral load suppression following enrollment in a comprehensive HIV care coordination program. *Clin Infect Dis*. 2015;60(2):298–310.
4. Davila JA, Miertschin N, Sansgiry S, et al. Centralization of HIV services in HIV-positive African-American and Hispanic youth improves retention in care. *AIDS Care*. 2013;25(2):202–206.
5. Mugavero MJ. Improving engagement in HIV care: what can we do? *Top HIV Med*. 2008;16(5):156–161.
6. Gardner LI, Marks G, Craw JA, et al. A low-effort, clinic-wide intervention improves attendance for HIV primary care. *Clin Infect Dis*. 2012;55(8):1124–1134.
7. Hightow-Weidman LB, Smith JC, Valera E, et al. Keeping them in "STYLE": finding, linking, and retaining young HIV-positive black and Latino men who have sex with men in care. *AIDS Patient Care STDS*. 2011;25(1):37–45.
8. Andersen M, Hockman E, Smereck G, et al. Retaining women in HIV medical care. *J Assoc Nurses AIDS Care*. 2007;18(3):33–41.
9. Bradford JB, Coleman S, Cunningham W. HIV System Navigation: an emerging model to improve HIV care access. *AIDS Patient Care STDS*. 2007;21(suppl 1):S49–S58.
10. Higa DH, Marks G, Crepaz N, et al. Interventions to improve retention in HIV primary care: a systematic review of US studies. *Curr HIV/AIDS Rep*. 2012;9(4):313–325.
11. Eliopoulos GM, Harris AD, Bradham DD, et al. The use and interpretation of quasi-experimental studies in infectious diseases. *Clin Infect Dis*. 2004;38(11):1586–1591.
12. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin; 2002.
13. Eliopoulos GM, Harris AD, Lautenbach E, et al. A systematic review of quasi-experimental study designs in the fields of infection control and antibiotic resistance. *Clin Infect Dis*. 2005;41(1):77–82.
14. Mercer SL, DeVinney BJ, Fine LJ, et al. Study designs for effectiveness and translation research: identifying trade-offs. *Am J Prev Med*. 2007;33(2):139–154.e2.
15. Wilson DB, Lipsey MW. The role of method in treatment effectiveness research: evidence from meta-analysis. *Psychol Methods*. 2001;6(4):413–429.
16. Brown CH, Curran G, Palinkas LA, et al. An overview of research and evaluation designs for dissemination and implementation. *Annu Rev Public Health*. 2017;38:1–22.
17. Irvine MK, Chamberlin SA, Robbins RS, et al. Come as you are: improving care engagement and viral load suppression among HIV care coordination clients with lower mental health functioning, unstable housing, and hard drug use. *AIDS Behav*. 2017;21(6):1572–1579.
18. Hanna DB, Felsen UR, Ginsberg MS, et al. Increased antiretroviral therapy use and virologic suppression in the Bronx in the context of multiple HIV prevention strategies. *AIDS Res Hum Retroviruses*. 2016;32(10–11):955–963.
19. New York City Department of Health and Mental Hygiene. HIV Surveillance Annual Report, 2014. 2015; https://www1.nyc.gov/assets/doh/downloads/pdf/dires/2014-hiv-surveillance-annual-report.pdf. Accessed June 20, 2016.
20. Bor J, Geldsetzer P, Venkataramani A, et al. Quasi-experiments to establish causal effects of HIV care and treatment and to improve the cascade of care. *Curr Opin HIV AIDS*. 2015;10(6):495–501.
21. New York City Department of Health and Mental Hygiene. HIV care coordination tools. http://www.nyc.gov/html/doh/html/living/hiv-care-coord.shtml. Accessed April 29, 2016.
22. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399–424.

23. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25(1):1–21.
24. Parsons LS. Performing a 1:N Case-Control Match on Propensity Score. 2003; http://www2.sas.com/proceedings/sugi29/165-29.pdf. Accessed June 20, 2016.
25. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007;26(1):20–36.
26. Wilson DB, Lipsey MW. The role of method in treatment effectiveness research: evidence from meta-analysis. *Psychol Methods*. 2001;6(4), 413–429.
27. Dehejia RH, Wahba S. Propensity score-matching methods for nonexperimental causal studies. *Rev Econ Stat*. 2002;84(1):151–161.
28. Hoover DR, Muñoz A, Carey V, et al. The unseen sample in cohort studies: estimation of its size and effect. Multicenter AIDS Cohort Study. *Stat Med*. 1991;10(12):1993–2003.
29. Pati R, Robbins RS, Braunstein SL. Validation of retention in HIV care status using the New York City HIV surveillance registry and clinical care data from a large HIV care center. *J Public Health Manag Pract*. 2017;23(6):564–570.