Introducing experimental design into real-world practice settings: Protocol for a cross-sectional stepped-wedge effectiveness trial of the New York City HIV Care Coordination Redesign

Mary Irvine,[1] Bruce Levin,[2] McKaylee Robertson,[3] Katherine Penrose,[1] Jennifer Carmona,[1] Graham Harriman,[1] Sarah Kulkarni,[3] Sarah Braunstein,[1] and Denis Nash[3]

1. Bureau of HIV/AIDS Prevention and Control, New York City Department of Health and Mental Hygiene; New York City, New York USA; 2. Department of Biostatistics, Mailman School of Public Health, Columbia University; New York, NY USA; 3. Institute for Implementation Science in Population Health (ISPH), Graduate School of Public Health and Health Policy, City University of New York (CUNY); New York City, New York USA.

Contact:
Mary Irvine, MPH, DrPH
mirvine@health.nyc.gov
1-347-396-7712

Background

- Successful HIV treatment at the individual level requires consistent adherence to antiretroviral therapy (ART) resulting in sustained suppression of HIV-1 viral load (VL) in plasma to levels below the detection limit of HIV RNA tests used by healthcare providers. At the population level, **viral suppression (VS) is key to the dual goals of improving health/survival among people with HIV (PWH) and preventing HIV transmission.**
- In New York City (NYC), a multi-component Ryan White Part A–funded medical case management intervention, **the HIV Care Coordination Program (CCP) was launched in 2009 to meet the needs of PWH with a recent HIV diagnosis or a history of suboptimal HIV care outcomes.**
- In its first 8 years, **the CCP showed significant benefits for care retention and VS**,[1,2,3] particularly for the most vulnerable clients. **Yet room for improvement remained**, and some CCP design features curbed client and provider engagement.[4]
- In response to identified implementation barriers and the evolving intervention literature, **CCP service model revisions were integrated into the Health Department's late-2017 request for proposals (RFP) initiating a competitive selection process for future NYC Care Coordination service delivery contracts.** The RFP outlined plans for agency randomization to an early or delayed start of the revised model, for an experimental evaluation of effectiveness.

Purpose & Hypothesis

- **Our purpose is to inform practice-driven intervention research**, particularly in the context of generating evidence **for the optimization of safety-net service delivery strategies**.
- The overall goal of the study known as Program Refinements to Optimize Model Impact and Scalability based on Evidence (PROMISE) is **to investigate the impact and implementation of empirically driven course corrections to an *already effective intervention model*.**
- **We will test the combined effect of program revisions in a cluster-randomized controlled trial** applying a cross-sectional, stepped-wedge design to the rollout of the revised model in previously funded, re-awarded CCP agencies.
- Drawing upon an implementation science framework and RE-AIM principles, **we posit that the model revisions will minimize logistical and administrative barriers to service delivery and increase program engagement** (among staff and clients), reach, fidelity and effectiveness.
- Specifically, **we hypothesize that a higher proportion of unsuppressed PWH enrolled in the Care Coordination revised (CCR) program will achieve timely VS, as compared with unsuppressed PWH enrolled in the original CCP** during the same period.

Participants, Intervention & Outcomes

Intervention (CCR) & Control (CCP) Conditions

- The control condition is the site-level continuation of CCP delivery; the intervention condition is a site-level change to deliver the CCR. *(See Table 1 for key differences.)*

Table 1: CCR features expected to boost uptake, fidelity, engagement, effectiveness and reach

	Added Components			Changed		Removed
	Self-management assessment	Use of video chat tools (optional)	iART (optional)	Eligibility criteria	Payment structure	Rigid program tracks
Uptake (provider)						X
Fidelity (provider)		X			X	X
Engagement (client)	X	X				X
Intervention effectiveness	X	X	X		X	X
Population reach/impact	X	X	X	X	X	X

Outcome Measurement

- **Timely VS (TVS)** is defined as VL <200 copies/mL on the last VL test reported to the NYC HIV surveillance registry in the four months following CCP/CCR enrollment (TVS=1). Those without any follow-up VL will be considered unsuppressed (TVS=0), given their lack of documented clinical monitoring since their last unsuppressed VL.

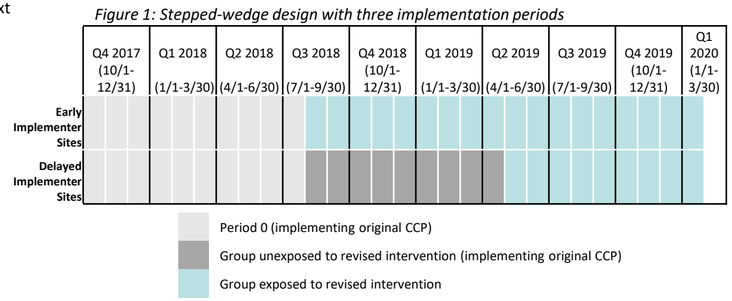
Eligibility Criteria for Trial

- **Clients:** newly enrolled in the CCP/CCR with unsuppressed VL (HIV RNA ≥200 copies/mL) at their latest test in the year prior to enrollment or with no VL test result in that year.
- **Agencies:** 17 previously funded/re-awarded agencies (that could be assigned to continue CCP delivery uninterrupted or begin CCR delivery in the initial implementation phase).

Timeline

- Figure 1 illustrates the three nine-month periods used in the stepped-wedge design: Period 0, with CCP at all 17 agencies and no CCR; Period 1, with CCR implementation only at sites randomized to an early start (and thus encompassing the months of simultaneous CCP and CCR operation); and Period 2, representing CCR implementation at all 17 sites.

Figure 1: Stepped-wedge design with three implementation periods



Assignment of Interventions

Randomization

- The unit of randomization is the Care Coordination provider agency (i.e., cluster).
- The 17 agencies were matched and randomized within pairs (including one case in which two smaller agencies were matched to a larger one). Matching accounted for characteristics plausibly related to the TVS outcome: agency type, primary location/borough and program size (measured via a combination of CCP caseload at time of re-award and award amount).

Assignment of Interventions (continued)

Table 2: Agency characteristics, pairings and study arm assignments (truncated: 4 agencies, 2 matched pairs)

Agency ID	Award increased >20% from prior year?	Typical (prior) Caseload	Borough	Type of Site	Pair	Phase (study arm)
21	Yes	84	Bronx	CBO/no clinical services	1	1
1	No	101	Bronx	CBO/no clinical services	1	2
20	Yes	109	Brooklyn	Public Hospital	2	1
14	No	151	Brooklyn	Public Hospital	2	2
28	Yes	87	Brooklyn	Private Hospital	3	1
24	No	96	Brooklyn	Community Health Center	3	2
25	No	62	Manhattan	Community Health Center	4	1
9	No	78	Manhattan	Community Health Center	4	2
23	No	228	Manhattan	Private Hospital	5	1
18	No	220	Manhattan	Private Hospital	5	2
13	Yes	82	Bronx	Public Hospital	6	1
11	Yes	82	Queens	Public Hospital	6	2
5	No	202	Bronx	Private Hospital	7	1
4	No	181	Manhattan	Private Hospital	7	2
8	Yes	77	Staten Island	CBO/no clinical services	8	1
16	No	63	Brooklyn	Community Health Center	8	1
2	No	184	Manhattan	Community Health Center	8	2

Statistical Analysis for the Matched-Pairs Stepped-Wedge Trial

Analysis Approach & Assumptions

- For each pair of sites, we will produce two 2x3 tables (one table per site in pair), cross-classifying the number of TVS and non-TVS outcomes in Period 0, Period 1 and Period 2. *For identification purposes, we refer to "Site 1" within a matched pair as the site randomized to switch in Period 1 (early start) and "Site 2" as the site randomized to switch in Period 2 (delayed start).*
- We assume **the following logistic regression model for the three binomial outcomes**: the logit of the probability of TVS for a given site, period and intervention equals an intercept representing an arbitrary, pair-specific log odds on TVS for Site 2 in the pair, plus an arbitrary log odds ratio (LOR) for Site 1 vs. Site 2, plus two arbitrary pair-specific LORs for Period 1 and Period 2 effects relative to Period 0, plus one structural LOR of interest, the global intervention effect (present at Site 1 in Period 1 and at both sites in Period 2). The exponent of this parameter is the target of statistical inference: the odds ratio (OR) for TVS vs. non-TVS in CCR vs. CCP.
- The key assumption (that any site effects apply in each period and any period effects apply to each site, independent of the intervention effect) will be tested and the model elaborated if needed.

Estimating the CCR Intervention Effect

- By conditioning on the marginal totals (eligible clients in each period and TVS and non-TVS outcomes) within each site, the joint distribution of the numbers of TVS outcomes for Site 1 by period becomes a non-central multiple hypergeometric distribution with three parameters: the period LORs and intervention LOR; the conditional distribution does not depend on nuisance site parameters.
- By further conditioning on the sum of TVS outcomes across the two sites in each period, the fully conditional joint distribution depends on only one parameter, the intervention effect; i.e., it does not depend on the nuisance site nor nuisance period effects.
- Thus, the sufficient statistic for the intervention LOR in the fully conditional likelihood function is simply the number of TVS outcomes from Site 1 in Period 1. It is then straightforward to calculate the marginal distribution of this outcome as a function of the intervention effect. We will calculate that distribution for each of the 8 matched pairs (including the case of two programs jointly matched to a third) and convolute those distributions to obtain the sampling distribution of the sum of sufficient statistics.
- Once we obtain the fully conditional sampling distribution of the sufficient statistic as described above, we will report the conditional maximum likelihood estimate of the intervention LOR with an exact, test-based 95% confidence interval.
- The test of the null hypothesis at the two-tailed 0.05 significance level will be based on the exact two-tailed P-value, and will form the primary outcome analysis. In sensitivity analyses, we will report the Wald, Score and Likelihood Ratio test results.

Sample Size & Power

- Table 3 provides the detectable effect size and power values given actual, post-randomization numbers of unsuppressed enrollees for Periods 0 (N=169) and 1 (N=389), a conservative estimate of unsuppressed enrollees for Period 2 (N=266), and assumed TVS proportions based on actual proportions for Period 0. Because the randomization of sites within pairs has already been set, the simulations for Table 3 condition on this fact; i.e., early- and late-implementing sites are considered fixed as randomized. The detectable effect size (80% power with exact Type I error rate ≤0.05 two-tailed) is currently an OR of 2.90, corresponding to RRs between 1.49 and 1.74. Power estimates range between ~76% and 83% for true ORs between 2.75 and 3.00, respectively.

Table 3: Power calculations for the CCR effect on TVS

Reference P [TVS]	Detectable P [TVS]	Risk ratio at Detectable P [TVS] for True OR=2.90	True OR	Power (%)
0.35	0.610	1.74	3.00	82.8
0.40	0.659	1.65	2.95	81.1
0.45	0.704	1.56	2.90	80.6
0.50	0.744	1.49	2.85	78.7
			2.80	77.8
			2.75	75.8

Note: Average P[TVS] among all sites in base period = 0.465. Monte Carlo standard error for power values is less than 0.5%.

Discussion

- The PROMISE trial, conducted in real-world service settings, uses secondary analyses of surveillance and programmatic data to assess the effects of revised (CCR) vs. original Care Coordination on VS.
- **To meet stakeholder expectations for rapid CCR rollout, the study applies a stepped-wedge design with a nine-month gap between implementation phases**, prompting use of a short-term (4-month) outcome and a brief (5-month) lead-in time for client enrollments.
- **Randomization at the agency level minimizes crossover** between the intervention conditions, since providers would otherwise struggle logistically and ethically with simultaneously delivering two distinct intervention models based on random client assignment.
- Randomization within matched agency pairs offers advantages akin to those of stratified random assignment: increasing statistical power when the number of units of randomization is small, by maximizing equivalency between intervention and control groups.
- **'Phasing in' an intervention with random assignment to early or delayed implementation offers a means of rigorously evaluating a set of changes to a major public-services program, while ensuring fair, uninterrupted access** to its benefits in the eligible population.
- **Through robust health department-university partnerships that include joint planning of research in advance of key policy or practice initiatives, locally important research questions can be answered without substantially slowing the pace of desired change, and with methods that support knowledge generation**.

Acknowledgements

- This work is supported through the NIH grant R01MH117793 and a Ryan White HIV/AIDS Program Part A services grant (H89HA00015).
- The authors wish to thank Care Coordination service providers and clients; our CCP/CCR Quality Management colleagues Gina Gambone, Scarlett Macias and Tyeirra Seabrook; the entire PROMISE study team and advisory board; evaluators Nadine Alexander and Sarah Kozlowski; and study advisor Dr. Julie Dombrowski.

References

1. 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;21:1572-9.
2. Nash D, Robertson MM, Penrose K, et al. Short-term effectiveness of HIV care coordination among persons with recent HIV diagnosis or history of poor HIV outcomes. PLoS One 2018;13:e0204017.
3. Robertson MM, Penrose K, Irvine MK, et al. Impact of an HIV care coordination program on durable viral suppression. JAIDS 2019;80:46–55.
4. Macias S, Estem K, and Carmona J. Challenges to caseload management in Ryan White Part A funded NYC Care Coordination Programs. National Conference on Social Work and HIV/AIDS. Washington, D.C. 2018.