**Appendix 3: Detailed analysis plans and results**

# Independence in the pre-post analyses

A total of 740 and 804 patients were marked OOC by the pseudo-RPT in the pre-implementation and implementation phases, respectively. Just under half of patients included in this analysis (n = 350) were marked OOC by the pseudo-RPT in both phases (44% of patients in the implementation phase and 47% of those in the pre-implementation phase). In other words, they were marked OOC by the pseudo-RPT in the pre-implementation phase, reengaged, and marked again in the implementation phase. However, the robust Poisson model assumes that OOC events are independent. Other studies demonstrate that this is a false assumption – past OOC events are predictive of future ones1. To our knowledge, there were no major changes in clinical care practice guidelines or how patients were followed in the years immediately prior to the implementation of Lost & Found. As such, we assume that HIV care histories for these patients are similar to others in each phase. This would mean that the effects of previous OOC events are similarly distributed across the implementation and pre-implementation phases, which would result in non-differential misclassification and further bias to no effect.

# Pre-post analyses: Survival analysis

## Cox proportional hazards model results

The results for the full model Cox proportional hazards model (controlled for Age, Canada, and Sex) are presented in Table A1, overall and stratified by risk category. Adjusted cumulative incidence curves are presented in Figure A1.

| **Table A1:** Results from the Cox proportional hazards model, overall and by risk category |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **All risk categories** (Implementation phase: n = 804; Pre-implementation phase: n = 740) |
| **Imp†** | 1.324 (1.147, 1.530) | < 0.001 |
| **Sex** | 0.947 (0.807, 1.112) | 0.506 |
| **Age** | 1.003 (0.998, 1.009) | 0.273 |
| **Canada** | 0.661 (0.567, 0.772) | < 0.001 |
| **High risk category** (Implementation phase: n = 199; Pre-implementation phase: n = 128) |
| **Imp†** | 1.301 (0.952, 1.777) | 0.099 |
| **Sex** | 0.738 (0.540, 1.008) | 0.056 |
| **Age** | 1.007 (0.997, 1.018) | 0.182 |
| **Canada** | 0.507 (0.354, 0.726) | < 0.001 |
| **Intermediate risk category** (Implementation phase: n = 481; Pre-implementation phase: n = 493) |
| **Imp†** | 1.196 (1.002, 1.428) | 0.048 |
| **Sex** | 0.992 (0.810, 1.214) | 0.936 |
| **Age** | 1.003 (0.996, 1.010) | 0.433 |
| **Canada** | 0.800 (0.664, 0.964) | 0.019 |
| **Low risk category** (Implementation phase: n = 124; Pre-implementation phase: n = 119) |
| **Imp†** | 2.384 (1.465, 3.881) | < 0.001 |
| **Sex** | 0.971 (0.574, 1.642) | 0.912 |
| **Age** | 0.992 (0.972, 1.012) | 0.441 |
| **Canada** | 0.404 (0.244, 0.669) | < 0.001 |
| \*Hazard ratio†‘Imp’ is a binary variable for being marked OOC in the implementation phase, compared to the pre-implementation phase. |

**Figure A1:** Cumulative incidence curves of reengagement over time among OOC patients by risk category, adjusted for age, sex, and being born in Canada



## Testing model assumptions

The tests of the proportional hazards assumption using scaled Schoenfeld residuals are summarised in Table A2 and Figure A2, and graphs of the deviance and dfbeta residuals in Figures A3 and A4.

**Figure A2:** Directed acyclic graph, only ‘Can’ as a confounder



| **Table A2:** Tests of the proportional hazards assumption for Cox regression model fit, using scaled Schoenfeld residuals |
| --- |
| **Variable** | **chisq** | **df** | **p\*** |
| **Imp** | 1.86017523 | 1 | 0.173 |
| **Sex** | 0.01914676 | 1 | 0.890 |
| **Age** | 1.13715691 | 1 | 0.286 |
| **Canada** | 0.35236210 | 1 | 0.553 |
| **GLOBAL** | 4.03610758 | 4 | 0.401 |
| **\***p value under 0.05 indicative of non-proportional hazards  |

**Figure A4:** Graph of dfbeta residuals

**Figure A3:** Graph of the deviance residuals


# Pre-post analysis: Propensity score analyses

## Directed acyclic graphs (DAGs)

Among variables in the Lost & Found dataset, four were possible confounders in the relationship between being marked OOC in the implementation phase (*Imp*) and being reengaged (*Re*):

* Sex;
* Age;
* Being born in Canada (*Can*); and
* History of HCV (*HCV*), defined as any positive HCV RNA test.

*Sex*, *Age*, and *Can* were included in multivariable model outlined in the protocol. *HCV* was added here, since it is likely associated with reengagement (people with HCV include a large number of people who inject drugs, who are likely more difficult to reengage than the general clinical population) and *may be* associated with being OOC in the implementation phase (i.e. there may be more or fewer people with HCV infection histories in the pre-implementation phase compared to implementation.)

Two directed acyclic graphs (DAGs) are considered here. In the first, all covariates are considered confounders (Figure A5). In the second, *Sex*, *Age*, and *HCV* are treated only as ancestors of the outcome, while *Can* is considered a confounder and an ancestor of the other three variables (Figure A6). There was a change in the number and proportion of asylum seekers (thus, people born in Canada) in the implementation phase compared to the pre-implementation phase, many of whom likely became OOC. These asylum seekers were younger, more likely to be female, and less likely to have contracted HCV than the clinical population. It is unclear whether *Sex*, *Age*, or *HCV* vary by implementation phase, other than when mediated by *Can*. Thus, *Sex*, *Age*, and *HCV* may not be true confounders. Nonetheless, if they are not confounders, when controlling for *Can*, they become pure predictors of the outcome, and are thus worth including in the model2.

Unmeasured confounding is included as the covariate $U$ in both DAGs. This unmeasured confounding might include some clinic-level events that would be associated with the exposure and outcome. For example, some people may have been more likely to become OOC and less likely to reengage due to certain doctors’ limited availabilities.

**Figure A5:** Directed acyclic graph, all covariates as confounders

**Figure A6:** Directed acyclic graph, only ‘Can’ as a confounder


## Propensity score model

A logistic regression model (Eq. 1), based on the DAGs in figures A5 and A6, were used to determine propensity scores for each participant:

Eq. 1: $logit(Imp)=β\_{0}+β\_{1}Age+β\_{2}Can+β\_{3}HCV+β\_{4}Sex$

where $Imp$ is a dummy variable, in which $Imp=1$ indicates that the patient was OOC during the implementation phase and 0 for the pre-implementation phase, $Sex$ accounts for the possible differences in the proportion of men and women in the pre-implementation or implementation phases, $Age$ accounts for differences in the ages of patients between the two periods, $HCV$ accounts for differences in proportion of patients with histories of HCV between the two periods, and $Canada$ accounts for differences between the two periods for the proportion of patients born outside of Canada.

The results for this model are presented in Table A3, and the distribution of propensity scores in Figure A7.

**Figure A7:** Propensity score distributions, by study phase

| **Table A3:** Propensity score model |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 1.224 (1.032, 1.453) | 0.020 |
| **Sex** | 0.881 (0.702, 1.105) | 0.272 |
| **Age** | 1.000 (0.992, 1.008) | 0.972 |
| **Canada** | 0.809 (0.650, 1.007) | 0.058 |
| **HCV** | 1.146 (0.838, 1.570) | 0.394 |
| \*Odds for '(Intercept)', odds ratio otherwise |

Of note, the distributions of propensity scores between phases are relatively concordant, indicating that the participants in each phase are very similar.

## Outcome models

The results of the unadjusted, multivariable adjusted, and propensity score adjusted models are presented in tables A4 to A6. Each model is a Poisson regression model with robust variance estimation.

| **Table A4:** Unadjusted model |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.508 (0.457, 0.565) | < 0.001 |
| **Imp** | 1.191 (1.033, 1.372) | 0.016 |
| \*Risk for '(Intercept)', risk ratio otherwise |
| **Table A5:** Multivariable adjusted model |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.576 (0.500, 0.664) | < 0.001 |
| **Imp** | 1.179 (1.022, 1.359) | 0.023 |
| **Sex** | 0.918 (0.783, 1.076) | 0.290 |
| **Age** | 0.999 (0.993, 1.005) | 0.795 |
| **Canada** | 0.799 (0.686, 0.931) | 0.004 |
| \*Risk for '(Intercept)', risk ratio otherwise |
| **Table A6:** Propensity score adjusted model |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.053 (0.013, 0.220) | < 0.001 |
| **Imp** | 1.176 (1.020, 1.356) | 0.025 |
| **Propensity score** | 76.702 (5.057, 1163.353) | 0.002 |
| Propensity scores included as covariates in the model.\*Risk for '(Intercept)', risk ratio otherwise |

The point estimates for the causal effect in each model are about even (ranging from 1.18 in the PS-adjusted model to 1.19 in the unadjusted model). These similar effect sizes can likely be attributed to the balance already achieved prior to adjustment. The unadjusted model had the largest standard errors, while the propensity score weighted model had the lowest (and thus, the thinnest CIs).

# Pre-Post analyses: Robust poisson regression

## Univariate models and contingency tables

### I) Univariate models and contingency tables for reengagement by study phase

| **Table A7:** Contingency table, reengagement by study phase, counts |
| --- |
|  | **ReEng** |  |
|  |  | **0** | **1** |  |
| **Imp** | 0 | 413 | 327 | 740 |
| 1 | 374 | 430 | 804 |
|  |   | 787 | 757 | 1544 |

| **Table A8:** Contingency table, reengagement by study phase, proportions |
| --- |
|  | **ReEng** |  |
|  |  | **0** | **1** |  |
| **Imp** | 0 | 0.267 | 0.212 | 0.479 |
| 1 | 0.242 | 0.278 | 0.521 |
|  |   | 0.510 | 0.490 | 1.000 |

| **Table A9:** Univariate model, reengagement as a function of study phase |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.508 (0.457, 0.565) | < 0.001 |
| **Imp** | 1.191 (1.033, 1.372) | 0.016 |
| \*Risk for 'Intercept', risk ratio otherwise |

### II) Univariate models and contingency tables for reengagement by age

| **Table A10:** Mean age by study phase |
| --- |
| **Imp** | **Mean age** | **Standard deviation** |
| **0** | 50.33539 | 12.39439 |
| **1** | 50.27421 | 12.52399 |

| **Table A11:** Univariate model, reengagement as a function of age |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.560 (0.522, 0.600) | < 0.001 |
| **Age** | 0.998 (0.993, 1.004) | 0.528 |
| \*Risk for 'Intercept', risk ratio otherwise |

### III) Univariate models and contingency tables for reengagement by sex

| **Table A12:** Contingency table, study phase by sex, counts |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Sex** | 0 | 488 | 252 | 740 |
| 1 | 541 | 263 | 804 |
|  |   | 1029 | 515 | 1544 |

| **Table A13:** Contingency table, study phase by sex, proportions |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Sex** | 0 | 0.316 | 0.163 | 0.479 |
| 1 | 0.350 | 0.170 | 0.521 |
|  |   | 0.666 | 0.334 | 1.000 |

| **Table A14:** Univariate model, study phase as a function of sex |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.562 (0.516, 0.612) | < 0.001 |
| **Sex** | 0.986 (0.850, 1.145) | 0.858 |
| \*Risk for 'Intercept', risk ratio otherwise |

### IV) Univariate models and contingency tables for reengagement by Canadian birth

| **Table A15:** Contingency table, study phase by Canadian birth, counts |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Canada** | 0 | 398 | 342 | 740 |
| 1 | 463 | 341 | 804 |
|  |   | 861 | 683 | 1544 |

| **Table A16:** Contingency table, study phase by Canadian birth, proportions |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Canada** | 0 | 0.258 | 0.222 | 0.479 |
| 1 | 0.300 | 0.221 | 0.521 |
|  |   | 0.558 | 0.442 | 1.000 |

| **Table A17:** Univariate model, study phase as a function of Canadian birth |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.610 (0.557, 0.667) | < 0.001 |
| **Canada** | 0.811 (0.702, 0.937) | 0.004 |
| \*Risk for 'Intercept', risk ratio otherwise |

### V) Univariate models and contingency tables for risk category by study phase

#### High risk category

| **Table A18:** Contingency table, study phase by risk category (high risk or not), counts |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **High risk cat.** | 0 | 612 | 128 | 740 |
| 1 | 605 | 199 | 804 |
|  |   | 1217 | 327 | 1544 |

| **Table A19:** Contingency table, study phase by risk category (high risk or not), proportions |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **High risk cat.** | 0 | 0.396 | 0.083 | 0.479 |
| 1 | 0.392 | 0.129 | 0.521 |
|  |   | 0.788 | 0.212 | 1.000 |

| **Table A20:** Univariate model, study phase as a function of risk category (high risk or not) |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.540 (0.499, 0.585) | < 0.001 |
| **High risk category** | 1.167 (0.990, 1.375) | 0.066 |
| \*Risk for 'Intercept', risk ratio otherwise |

| **Table A22:** Contingency table, study phase by risk category (intermediate risk or not), proportions |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Int. risk cat.** | 0 | 0.160 | 0.319 | 0.479 |
| 1 | 0.209 | 0.312 | 0.521 |
|  |   | 0.369 | 0.631 | 1.000 |

#### Intermediate risk category

| **Table A21:** Contingency table, study phase by risk category (intermediate risk or not), counts |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Int. risk cat.** | 0 | 247 | 493 | 740 |
| 1 | 323 | 481 | 804 |
|  |   | 570 | 974 | 1544 |

| **Table A23:** Univariate model, study phase as a function of risk category (intermediate risk or not) |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.518 (0.459, 0.584) | < 0.001 |
| **Intermediate risk category** | 1.127 (0.971, 1.307) | 0.115 |
| \*Risk for 'Intercept', risk ratio otherwise |

| **Table A25:** Contingency table, study phase by risk category (low risk or not), proportions |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Low risk cat.** | 0 | 0.402 | 0.077 | 0.479 |
| 1 | 0.440 | 0.080 | 0.521 |
|  |   | 0.843 | 0.157 | 1.000 |

####  Low risk category

| **Table A24:** Contingency table, study phase by risk category (low risk or not), counts |
| --- |
|  | **Imp** |  |
|  |  | **0** | **1** |  |
| **Low risk cat.** | 0 | 621 | 119 | 740 |
| 1 | 680 | 124 | 804 |
|  |   | 1301 | 243 | 1544 |

| **Table A26:** Univariate model, study phase as a function of risk category (low risk or not) |
| --- |
| **Variable** | **Estimate\* (95% CI)** | **p-value** |
| **(Intercept)** | 0.595 (0.553, 0.641) | < 0.001 |
| **Low risk category** | 0.592 (0.466, 0.752) | < 0.001 |
| \*Risk for 'Intercept', risk ratio otherwise |

# References

1. Woodward B, Person A, Rebeiro P, Kheshti A, Raffanti S, Pettit A. Risk prediction tool for medical appointment attendance among HIV-infected persons with unsuppressed viremia. *AIDS patient care and STDs.* 2015;29(5):240-247.

2. Alam S, Moodie EE, Stephens DA. Should a propensity score model be super? The utility of ensemble procedures for causal adjustment. *Statistics in medicine.* 2019;38(9):1690-1702.