A cohort study of factors associated with LBTI treatment initiation and completion in Hamilton, Ontario, Canada

Deborah A. Milinkovic, MA; Stephen Birch, PhD; Fran Scott, PhD; K. Bruce Newbold, PhD; Jessica Hopkins, MD; Beverley M. Essue, PhD

Appendix B

At the individual level, the decision to adhere is a two-stage process. First the individual decides to initiate treatment, and then the individual decides to adhere to treatment given that they had initiated. The same is true for lost to follow-up. The individual first decides to initiate treatment and then is lost to follow-up given that they initiated treatment. Because these two linked decisions may be connected in unobservable ways, there can be correlation in the error terms between the two stages. We tested for a selection effect using a Heckman probit and oprobit model. If there is a selection effect and the model is estimated using a two-part model, the estimates could be bias [1]. Of course, the reverse is true if a two-part model is more appropriate. A feature of the Heckman model is the exclusion restriction for the identification of the first stage: at least one variable that appears in the selection step (initiated treatment) should not be used in the outcome step (completion of treatment), but we could not find any reasonable justification for excluding variables from the second stage and therefore employed the full model for both the selection phase and the outcome phase. We found no evidence of a selection effect, so we proceeded with a two-part model.

We proceeded with the ordinal completion variable under the assumption that higher values were better (1 is better than 2, 2 is better than 3, etc.). Using an ordinal variable in this

way requires that the proportional odds assumption is not violated.[1] Unfortunately, the assumption was violated for some specifications. We therefore relaxed this assumption and ran a generalized ordered probit model as a robustness check (Table 5). We also collapsed the ordinal variable into a binary variable at the 50% cut point as a further robustness check (Table 1). However, the assignment of what constitutes "good" completion is subjective, so the binary model was not our preferred model.

We also conducted several other robustness checks: the ordinal dependent variables were collapsed into four levels (categories 0 and <50% were combined) (Table 1), individuals with unknown origin were recoded as foreign (Table 2), FSA was recoded using last known address (Table 4), and individuals that did not complete due to medical reasons were recoded as complete (Table 3). The results reported in the study were robust to model specification and recoding.

RESULTS

Originating from an endemic country, immigrating less than five years prior to the LTBI episode, and the interaction between these variables are not statistically significant. All model specifications show a strong positive association between having been identified through the Canadian medical screening program and initiation and completion and a strong negative association with lost to follow-up ($p<.01$). Relative to younger adults (18-30 years), middle-aged adults (31-49 years) continue to have a strong negative association with completion ($p > .05$) for all models except for the binary model recoded to FSA of last known address and the binary model with unknown origin recoded as foreign-born. Last known address was taken from

[1] The proportional odds assumption, also called parallel line assumption, assumes that the coefficient remains constant whether measured from 1 to 2, 2 to 3, 3 to 4, etc.

addresses noted on laboratory test results and was inconsistently recorded in the database, so it was not our preferred specification. Some models suggest that relative to younger adults (18-30 years), older adults (50-64 years) are less likely to be adherent and more likely to be lost to follow-up, but these results are not consistent. Females show a strong positive association with completion and a negative association with lost to follow-up, but these results are not robust to model specification. FSA-level Median Individual Income less than or equal to $35,000 was not robust to model specification. Variations on this variable noted above were not statistically significant.

TABLE 1: Ordinal variable with reduced levels (FSA at episode)

Variables	(5) completion5 ordinal collapsed at 50% (cl-FSA)	(2) completion2 ordinal-4 levels (cl-FSA)	(9) completion4 ordinal-4 levels (cl-FSA)
Age under 18	0.486	0.345	0.240
	(0.408)	(0.366)	(0.362)
Age 31 to 49	-0.672**	-0.577***	-0.407**
	(0.291)	(0.220)	(0.207)
Age 50 to 64	-0.0231	0.132	-0.721**
	(0.359)	(0.268)	(0.290)
Age over 64	0.106	0.357	0.0487
	(0.617)	(0.449)	(0.581)
Female	0.123	-0.0180	0.451***
	(0.221)	(0.205)	(0.166)
High TB birth country	0.0373	-0.196	0.238
	(0.220)	(0.198)	(0.184)
Immigrated < 6 years	-0.397	-0.500	-0.0767
	(0.375)	(0.321)	(0.296)
FSA income =<$32,000	-0.266	-0.0218	-0.356
	(0.272)	(0.357)	(0.258)
Identified by immigration	omitted	0.841***	1.235***
		(0.307)	(0.392)
Observations	179	214	276
Pseudo R-squared	0.0580	0.0653	0.112

F test	0.000590	1.83e-05	1.32e-08

* Immigration screening was omitted in the model 5 due to perfect prediction.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

TABLE 2: Recoded unknown origin to foreign-born (FSA at episode)

VARIABLES	(1) Initiations binary (boot/cl-ro)	(2) completion1 ordinal (boot/cl-ro)	(6) completion3 binary (boot/cl-ro)	(9) Lost binary (cl-FSA)
Age under18	0.701***	0.347	0.311*	-0.129
	(0.125)	(0.246)	(0.182)	(0.176)
Age 31 to 49	-0.130	-0.378**	-0.166	0.108
	(0.0851)	(0.151)	(0.141)	(0.158)
Age 50 to 64	-0.256***	-0.142	-0.372**	0.356**
	(0.0720)	(0.0935)	(0.157)	(0.149)
Age over 64	-0.284*	0.380	0.210	-0.358
	(0.157)	(0.234)	(0.261)	(0.376)
Female	0.140*	0.106	0.246**	-0.300***
	(0.0734)	(0.114)	(0.0954)	(0.113)
FSA income =<$35,000	0.0379	0.0629	-0.0470	-0.240
	(0.0901)	(0.246)	(0.161)	(0.157)
Identified by immigration	0.659***	0.790***	0.782***	-0.925***
	(0.119)	(0.219)	(0.214)	(0.275)
Observations	2,895	479	658	658
Pseudo R-squared	0.0277	0.0379	0.0409	0.0485
F test	0	7.42e-07	3.16e-07	1.17e-09

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

TABLE 3: Non-adherence due to medical reasons recoded completed (FSA at episode)

Variables	(9) completion4 binary (cl-FSA)
Age under 18	0.240
	(0.362)
Age 31 to 49	-0.407**
	(0.207)
Age 50 to 64	-0.721**
	(0.290)
Age over 64	0.0487
	(0.581)
Female	0.451***
	(0.166)
High TB birth country	0.238
	(0.184)
Immigrated < 6 years	-0.0767
	(0.296)
FSA income =<$32,000	-0.356
	(0.258)
Identified by immigration	1.235***
	(0.392)
Observations	276
Pseudo R-squared	0.112
F test	1.32e-08

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

TABLE 4: FSA recoded to last known address

Variables	(1) Initiation	(2) completion1 ordinal 5-levels	(3) completion2 ordinal 4-levels	(7) completion3 binary	(11) Lost binary
	(boot/ro-cl)	(boot/ro-cl)	(boot/ro-cl)	(cl-fsa)	(cl-fsa)
Age under 18	0.588***	0.413	0.342	0.430	-0.118
	(0.194)	(0.323)	(0.341)	(0.277)	(0.332)
Age 31 to 49	-0.135	-0.543**	-0.545**	-0.373*	0.526*
	(0.120)	(0.223)	(0.234)	(0.207)	(0.293)
Age 50 to 64	-0.0561	0.102	0.107	-0.664**	0.572
	(0.133)	(0.239)	(0.253)	(0.262)	(0.355)
Age over 64	-0.0152	0.351	0.348	0.314	0.164
	(0.237)	(0.367)	(0.396)	(0.559)	(0.655)
Female	0.172*	-0.0176	-0.0312	0.326**	-0.437**
	(0.101)	(0.194)	(0.206)	(0.150)	(0.171)
High TB birth country	0.158	-0.227	-0.193	0.242	-0.159
	(0.115)	(0.166)	(0.176)	(0.205)	(0.211)
Immigrated < 6 years	0.161	-0.626**	-0.549*	-0.371	0.0359
	(0.124)	(0.305)	(0.297)	(0.229)	(0.233)
FSA income =<$32,000	0.114	-0.0720	0.0140	-0.401	0.130
	(0.105)	(0.299)	(0.328)	(0.250)	(0.276)
Identified by immigration	0.709***	0.891***	0.862***	0.956***	-1.364***
	(0.161)	(0.296)	(0.295)	(0.235)	(0.400)
Observations	1,359	214	214	276	276
Pseudo R-squared	0.0365	0.0644	0.0666	0.106	0.106
F test	6.83e-06	0.000112	0.000730	8.76e-10	1.26e-09

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

TABLE 5: Generalized probit (FSA at episode) cluster at FSA

Variables	(1) Completion (<50%)	(2) Completion (50-79%)	(3) Completion (80-99%)	(4) Completion (100%)
Age under 18	5.024***	0.598	0.553	-3.831***
	(0.426)	(0.441)	(0.379)	(0.352)
Age 31 to 49	-1.110***	-1.506***	-0.0827	-0.310
	(0.339)	(0.523)	(0.249)	(0.419)
Age 50 to 64	-0.0676	0.639	-0.394	4.165***
	(0.663)	(0.634)	(0.253)	(0.260)
Age over 64	3.834***	-5.384***	4.392***	4.539***
	(0.592)	(0.449)	(0.329)	(0.589)
Female	0.616	0.221	0.246	-0.489
	(0.535)	(0.273)	(0.238)	(0.323)
High TB birth country	0.613	0.857**	-0.275	-0.213
	(0.465)	(0.403)	(0.180)	(0.433)
Immigrated < 6 years	-0.227	-0.0392	-0.425**	0.0338
	(0.438)	(0.355)	(0.204)	(0.420)
FSA income =<$32,000	-4.729***	-0.584	-0.109	4.087***
	(0.537)	(0.513)	(0.354)	(0.325)
Identified by immigration	0	6.343***	1.495***	-0.267
	(0)	(0.517)	(0.396)	(0.548)
Observations	213	213	213	213

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

REFERENCES

1. Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics, 27*(2), 300-307.