STATISTICAL ANALYSIS AND FINAL RESULTS OVERVIEW INCLUDING HETEROGENEITY OF TREATMENT EFFECTS

I.   STATISTICAL ANALYSES

The primary outcome of interest is the sustained virological response (SVR) rate in the two study arms (i.e., facilitated telemedicine compared with usual care) as well as the comparison of the SVR rates between the two arms. Other outcomes of interest include: 1) treatment initiation rates as measured by the proportion of individuals who take an initial medication dose; 2) treatment completion rates; 3) patient satisfaction with the telemedicine-based hepatitis C virus (HCV) care scores using the adapted Patient Satisfaction Questionnaire [1,2]; and 4) reinfection rates.

For an explanation of the study scope, aims, setting, description of the population, and implementation strategies, please see [3]. We analyzed treatment initiation and completion rates graphically. Details on the analysis of patient satisfaction scores are reported below as are the analysis of reinfections. We have followed the CONSORT statement for reporting of stepped wedge trials [4].

Ia.   Primary Aim Analysis Plan

Ia1.   Data Collection: Per guidance of the New York State Office of Addiction Services and Supports (OASAS), HCV antibody levels are assessed on admission to the opioid treatment program (OTP) and on an annual basis. Thus, we obtained a list of potentially eligible participants who were then approached regarding study entry. If HCV seropositive individuals consented to study participation, they were scheduled for a screening visit where the laboratory parameters mentioned in Table 1 were obtained. Those individuals who had detectable serum HCV RNA were enrolled in the study and had their levels of fibrosis and inflammation subsequently assessed.

Table I: Laboratory parameters collected at Baseline

Lab test	Reason
HCV antibody	Assess exposure to HCV
HCV RNA	Assess active HCV infection
HCV genotype	Assess type of hepatitis C
CBC	Assess hemoglobin concentration
CMP	Assess metabolic, renal, hepatic parameters
INR	Assess coagulation parameters
Toxicology screen	Assess use of illicit substances
HIV antibody	Assess exposure to HIV
HBV surface antigen	Assess exposure to HBV
HBV surface antibody	Assess immunity to HBV
FibroSure®	Assess liver fibrosis and inflammation levels
HIV RNA	Assess active HIV infection
CD4 cell count	Assess immune system status in HIV infection

Abbreviations: CBC, complete blood count; CMP, complete metabolic panel; HCV, hepatitis C virus; INR, international normalized ratio; HBV, hepatitis B virus; HIV, human immunodeficiency virus; RNA, ribonucleic acid

Ia2.   Missing data: Missing data can occur in any study design and with any data source. In clinical trials, missing data often arise when participants drop out of the study before its conclusion [5]. Li et al (2014) indicate that the single best

approach to deal with missing values is to prospectively prevent their occurrence [6]. To this end, the study personnel attempted to minimize the amounts of missing values by following the study participants closely. The case managers were provided with a list of strategies to avoid missing values, and issues surrounding missing data were reviewed weekly during conference calls.

Specifically, the following strategies were implemented to limit the amount of missing data. Adherence to the treatment schedule was monitored very closely by the case managers at each site. At the level of the overall study, the Biostatistics and Data Management Team monitored the trial weekly. Study participants were informed at the consent stage of the importance of completing all surveys, and case managers facilitated their completion by administering the surveys onsite within the OTP. During the time of Coronavirus Disease 2019 (COVID-19) pandemic restrictions and because recruitment had been completed, the impact of COVID-19 on the study was minimal. Additional strategies we implemented included frequent reminders to study participants for study visits, education of participants of the importance of continuous engagement, keeping participant contact information current, providing monetary incentives to participants, utilizing OTP staff support to promote participant engagement, and collecting information on participants at risk of dropout.

To adequately address statistical challenges around missing data, we collected detailed information on the reasons for missingness when missing data occurred. Whenever a participant discontinued all or some type of study participation, we documented the following: 1) the reason for discontinuation, 2) type(s) of participation that the discontinuation involved, and 3) who decided to discontinue the participant from the study. For example, in the case of ongoing non-adherence with OTP attendance requirements or prolonged absence from the OTP, such as in the case of incarceration, the OTP may "administratively discontinue" an individual. If a study participant was discontinued from the OTP, their study participation ceased. Therefore, opportunities for follow up blood draws and obtaining study-related information were not possible.

When we encountered missing data, we took the following steps. 1) We filed a report documenting the occurrence of missing data, its potential resolution, and the expected resolution date. 2) The case manager at each site investigated whether data retrieval was possible. If not retrievable, we supplemented the report with the reasons for missing data. We requested that each site report the percentage of missing data for each participant on a weekly basis.

The above strategies severely limited the quantity of missing data. When we encountered missing data, the explanations that accompanied the missing data guided the adjudication of the missing data mechanism. We used valid statistical methods, such as multiple imputation, to impute missing data. In our study, the maximum percentage of missing data is 5.814%, and the missing mechanism is missing at random. To perform multiple imputation, we used Multivariate Imputation by Chained Equations (MICE) since the estimated intra-cluster correlation is 0.099, sufficiently low to allow the use of MICE. We generated 20 imputed data sets that were used in the analysis [7,8].

A recent publication from the Center for Disease Control and Prevention illustrated that the prevalence of initiation of HCV treatment with a direct acting antiviral within 360 days of the first positive HCV RNA among those with Medicaid, Medicare, and private insurance was 23%, 28% and 35%, respectively [9]. Anticipating that we would achieve similar results to those mentioned above, we had pre-specified in the protocol that participants who did not initiate treatment within five months of screening would be considered as treatment failures. Thus, these participants were designated as SVR = 0, indicating that an SVR was not achieved if they had not uptake HCV treatment within the first five months of receiving a referral. Furthermore, in chronic HCV infection, spontaneous resolution occurs at 0.36% per person-years of follow up [10]. This percentage is so low that it indicates that it is highly unlikely that spontaneous resolution would occur in our study participants with chronic HCV infection who did not initiate treatment during the allotted 5-month period. All remaining participants initiated treatment. Once participants initiated treatment, only 5.8% of participants discontinued treatment prematurely. In these instances, we used multiple imputation to provide an SVR result.

The missing mechanism is assumed to be missing at random (MAR), and Table 2 presents the variables that are affected by missingness as well as the percentage of missing values.

84
85
86
87

Table 2: Variables with Missing Values and Percentage of Missingness.

Variable	# of missing observations (percentage)
DAST-10	20 (3.32%)
APRI	12 (1.99%)
Prescription drug use for nonmedical reasons	21 (3.49%)
Illegal drug use	21 (3.49%)
SVR	35 (5.81%)

88 Abbreviations: DAST-10-Drug Abuse Screening Test; APRI-Aspartate aminotransferase to platelet ratio index; SVR-
89 sustained virological response.

90
91 We imputed missing data using MICE (R Software, Version 3.15.0), with a publication date of 11/19/2022 for the MICE
92 software version we used. For additional information on the methods the software uses for data imputation, please see
93 [7,8]. To impute the missing SVR values, we built a logistic regression model that included variables thought to affect
94 missingness and potentially satisfy the MAR assumption. Furthermore, we took account of the clustering by
95 incorporating the site as a fixed effect. Details on the selection of these variables are provided in the draft final research
96 report. We developed the imputation models according to the recommendations of Azur et al (2011) [11]. Diagnostic
97 plots, such as boxplots, are also used to evaluate the agreement between imputed and observed data [12].

98
99   We performed analysis using the 20 imputation data sets and summarized the relevant results using Rubin's rules for
100 combining estimates and standard errors. This method also computes the relevant confidence intervals. For details,
101 please see Schafer (1997) [13]. To estimate the effect of facilitated telemedicine, we used two methods of analysis.
102 Generalized linear mixed effects models, adjusted for the time effect, are used in an intention-to-treat analysis. An
103 alternative robust, nonparametric, within period, cluster-level method is also used [14]. Furthermore, we provide graphical
104 analysis that corroborates our discussion on selection bias in cluster randomized trials. If data configurations in which
105 the outcome is determined by specific combinations of the covariates occur, these configurations generate
106 nonconvergence; in this case, we used exact inference methods.
107 I2   Secondary Aim Analyses

108
109 I2a   Heterogeneity of Treatment Effects: Analysis of heterogeneity of the intervention effect, that is, estimation of
110 the intervention effects in subgroups, can be challenging (see [15]). We defined "a priori subgroups of interest", at the
111 request of the sponsor, by the following variables: 1) comorbid medical conditions (specifically anxiety and depression),
112 2) liver fibrosis stage (binary with one level including F3, F3-F4, and F4 participants versus all other stages), 3) location of
113 residence (specifically urban/rural classification). Our goal is the identification of whether or not these factors influence
114 the decision to initiate and adhere to HCV treatment in the telemedicine and usual care arms.

115
116 The percentage of people with opioid use disorder with comorbid medical conditions ranges from 20% to 85% based
117 upon findings reported by others [15,16]. We included depression and anxiety disorders as a specific diagnosis among
118 those with co-morbid conditions since the prevalence varies from 24.5% to greater than 50% [15-17]. The sample size of our
119 study was sufficient to permit stratification by the comorbid conditions variable to allow exploratory analyses.

120
121 Despite the tremendous therapeutic advances in HCV treatment over the past several years, fibrosis stage remains an
122 important factor to explore for its importance as a determinant of successful HCV cure. We based our estimates of
123 advanced fibrosis and cirrhosis from our previous study [18] in which the percentage of participants with advanced fibrosis
124 (F3 and F3-F4) was 16% and 12% had cirrhosis (F4). Correspondingly, we estimated that 28% of the participants in each
125 study arm would have advanced fibrosis or cirrhosis, thereby permitting evaluation of the variable's influence on
126 virologic outcomes. We also evaluated the percentage of study participants residing in urban or rural areas.
127

128 Univariate subgroup analyses were conducted by incorporating a treatment by factor interaction in all relevant models
129 adjusted for time. Forrest plots provide a visualization of the results. Details are provided in the final report. As our
130 secondary analyses are largely exploratory, we followed Wang et al. (2007) in interpreting and reporting the results [19].
131
132 Figure 1 presents the time-adjusted log-odds ratios for the subgroups of interest and their associated 95% confidence
133 intervals. Note that all results are in favor of the intervention (i.e., facilitated telemedicine). The 95% confidence interval
134 that corresponds to rural classification is based on the use of methods that can accommodate data configurations that
135 result in infinite estimates of the intervention effect (nonconvergence). These methods generally generate wide
136 confidence intervals, so they are exact inference methods. In our case, these data configurations were created because
137 arm membership (i.e., whether participants were in the facilitated telemedicine or referral arms) for rural participants
138 determined the clinical outcome. Additionally, all interaction effects illustrated in Figure 1 are nonsignificant at the α =
139 0.05 level.

140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162

Subgroup	SVR Rate = (# participants with SVR=1) / (# of participants)		Time-adjusted log(OR) (95% CI)	Favors Referral / Favors Telemedicine	Interaction P value
	Telemedicine	Referral			
Comorbid Conditions					
Other/No Comorbid Condition	176/200 (88.00%)	80/233 (34.33%)	2.92 (2.08-3.76)		0.0977
Anxiety/Depression	87/90 (96.67%)	30/79 (37.97%)	4.34 (2.15-6.54)		
Ethnicity					
Non Hispanic/Latino	185/201 (92.04%)	86/216 (39.81%)	2.93 (2.07-3.79)		0.6456
Hispanic/Latino	78/89 (87.64%)	24/96 (25.00%)	4.25 (2.51-5.98)		
Fibrosis Score					
F3-F4	83/91 (91.21%)	34/107 (31.78%)	3.75 (1.95-5.55)		0.4624
F0-F2	180/199 (90.45%)	76/205 (37.07%)	2.98 (2.12-3.83)		
Geographic Location					
Urban	219/245 (89.39%)	97/267 (36.33%)	2.88 (2.09-3.67)		0.0774
Rural	44/45 (97.78%)	13/45 (28.89%)	3.89 (0.36-7.43)*		
Race					
White	143/155 (92.26%)	53/151 (35.10%)	3.46 (2.56-4.35)		0.5099
Other	120/135 (88.89%)	57/161 (35.40%)	3.12 (1.96-4.29)		
Sex					
Male	160/175 (91.43%)	77/194 (39.69%)	3.16 (2.18-4.13)		0.4407
Female	103/115 (89.57%)	33/118 (27.97%)	2.98 (1.66-4.29)		

log(OR) (95% CI) axis: -2 0 2 4 6
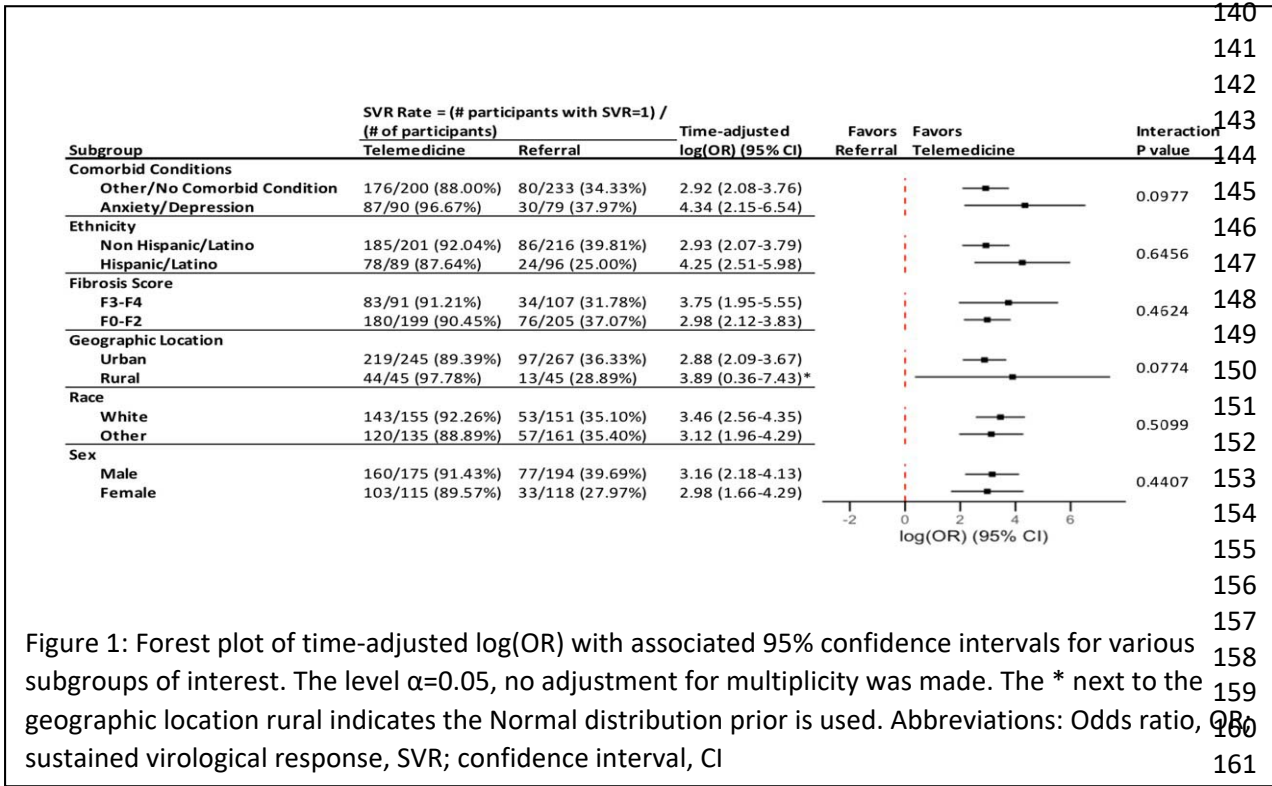
Figure 1: Forest plot of time-adjusted log(OR) with associated 95% confidence intervals for various subgroups of interest. The level α=0.05, no adjustment for multiplicity was made. The * next to the geographic location rural indicates the Normal distribution prior is used. Abbreviations: Odds ratio, OR; sustained virological response, SVR; confidence interval, CI

163 I2b   Patient Satisfaction Analysis: Our mixed-methods analysis of patient satisfaction with HCV treatment delivery
164 comparing in-person and telemedicine has recently been published [20]. Generalized linear mixed models were used to
165 account for the time effect on the Patient Satisfaction Questionnaire that was administered at two timepoints,
166 incorporating adjustments for the clustering effect and the study arm. We modeled the survey response rate as a
167 function of demographic and other relevant covariates.
168
169 I2c   Analysis of Reinfection Study: An exploratory aim of the trial was the follow up of study participants for up to 24
170 months after achieving an SVR to evaluate for HCV reinfection. We compared the reinfection rates between the usual
171 care and telemedicine arms by calculating the cumulative incidence density for both arms.
172
173 Study participants were followed for up to two years after achieving an SVR to assess for reinfection. There were a total
174 of 3 reinfections in the referral arm with a total follow up period of 93.2 person-years. In the telemedicine arm, there
175 were a total of 10 reinfections with a total of 361.0 person-years of follow up. The overall incidence density rate was
176 2.86 per 100 person years of follow up with a rate of 2.77 per 100 person years of follow-up in telemedicine and a rate
177 of 3.29 in referral.

178 A.2.3   Exploratory Analysis
179

180 Distribution of all variables was examined using graphical methods and descriptive statistics. Continuous variables are
181 presented by their means and standard deviations or medians and associated interquartile ranges. Categorical variables
182 are presented by the associated counts and percentages. Distributional assumptions were checked via boxplots, qq-plots
183 or other appropriate graphical techniques, and outliers were identified via the aforementioned graphical analysis. All
184 statistical analyses were conducted using SAS (SAS Institute, Cary, NC) and/or R. The estimated intraclass correlation
185 coefficient (ICC) was .099.
186

187 Demographic description of the population was provided in terms of characteristics such as age, gender, race, and
188 ethnicity. Other variables that contributed to the descriptions included residence type (i.e., urban versus rural), Drug
189 Abuse Screening Test (DAST)-10 scores [21], and variables incorporated within the NIDA Quick Screen questionnaire [22]. We
190 also incorporated variables collected on methadone and HCV adherence.
191

192 **References**

193 1.   Marshall GN, Hays RD. The patient satisfaction questionnaire short-form (PSQ-18). RAND, Santa Monica, CA;
194   1994.
195 2.   Yip M, Chang AM, Chan J, MacKenzie AE. Development of the telemedicine satisfaction questionnaire to
196   evaluate patient satisfaction with telemedicine: a preliminary study. *J Telemed Telecare*. 2003;9(1):46-50.
197   doi:10.1258/135763303321159693
198 3.   Talal AH, Markatou M, Sofikitou EM, et al. Patient-centered HCV care via telemedicine for individuals on
199   medication for opioid use disorder: Telemedicine for Evaluation, Adherence and Medication for Hepatitis C
200   (TEAM-C). *Contemp clin trials*. 2022;112:106632. doi:10.1016/j.cct.2021.106632
201 4.   Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of
202   the CONSORT 2010 statement with explanation and elaboration. *BMJ*. 2018;363:k1614. doi:10.1136/bmj.k1614
203 5.   Little R, Cohen M, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. *Stat Med*.
204   2012;31(28):3433-3443.
205 6.   Li T, Hutfless S, Scharfstein DO, et al. Standards should be applied in the prevention and handling of missing data
206   for patient-centered outcomes research: a systematic review and expert consensus. *J Clin Epidemiol*.
207   2014;67(1):15-32. doi:10.1016/j.jclinepi.2013.08.013
208 7.   Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate imputation by chained equations in R. *J Stat Softw*.
209   2011;45:1-67. doi:https://www.jstatsoft.org/article/view/v045i03
210 8.   Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: A data analyst's perspective.
211   *Multivariate Behav Res*. 1998;33(4):545-71. doi:10.1207/s15327906mbr3304_5
212 9.   Thompson WW, Symum H, Sandul A, et al. Vital signs: hepatitis C treatment among insured adults - United
213   States, 2019-2020. *MMWR Morb Mortal Wkly Rep*. 2022;71(32):1011-1017. doi:10.15585/mmwr.mm7132e1
214 10.   Bulteel N, Partha Sarathy P, Forrest E, et al. Factors associated with spontaneous clearance of chronic hepatitis C
215   virus infection. *J Hepatol*. 2016;65(2):266-72. doi:10.1016/j.jhep.2016.04.030
216 11.   Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it
217   work? *Int J Methods Psychiatr Res*. 2011;20(1):40-9. doi:10.1002/mpr.329
218 12.   Zhao Y. Diagnostic checking of multiple imputation models. *AStA Adv Stat Anal*. 2022;106(2):271-286.
219 13.   Schafer JL. *Analysis of incomplete mulitvariate Data*. Chapman & Hall/CRC; 1997.
220 14.   Thompson JA, Davey C, Fielding K, Hargreaves JR, Hayes RJ. Robust analysis of stepped wedge trials using
221   cluster-level summaries within periods. *Stat Med*. 2018;37(16):2487-2500. doi:10.1002/sim.7668
222 15.   Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO. A framework for the analysis of heterogeneity of treatment
223   effect in patient-centered outcomes research. *J Clin Epidemiol*. 2013;66(8):818-25.
224   doi:10.1016/j.jclinepi.2013.02.009
225 16.   El-Serag HB, Kunik M, Richardson P, Rabeneck L. Psychiatric disorders among veterans with hepatitis C infection.
226   *Gastroenterology*. 2002;123(2):476-82. doi:10.1053/gast.2002.34750

227 17.   Ho SB, Brau N, Cheung R, et al. Integrated care increases treatment and improves outcomes of patients with
228   chronic hepatitis C virus infection and psychiatric illness or substance abuse. *Clin Gastroenterol Hepatol*.
229   2015;13(11):2005-14 e1-3. doi:10.1016/j.cgh.2015.02.022
230 18.   Talal AH, Andrews P, Mcleod A, et al. Integrated, co-located, telemedicine-based treatment approaches for
231   hepatitis C virus management in opioid use disorder patients on methadone. *Clin Infect Dis*. 2019;69(2):323-331.
232   doi:10.1093/cid/ciy899
233 19.   Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine--reporting of subgroup analyses in
234   clinical trials. *N Engl J Med*. 2007;357(21):2189-94. doi:10.1056/NEJMsr077003
235 20.   Talal AH, Sofikitou EM, Wang K, Dickerson S, Jaanimagi U, Markatou M. High satisfaction with patient-centered
236   telemedicine for hepatitis C virus delivered to substance users: a mixed-methods study. *Telemed J E Health*.
237   2023;29(3):395-407. doi:DOI:10.1089/tmj.2022.0189
238 21.   Skinner HA. The drug abuse screening test. *Addict behav*. 1982;7(4):363-71. doi:10.1016/0306-4603(82)90005-3
239 22.   National Institute on Drug Abuse. NIDA Drug Screening Tool. *Available at:*
240   *https://archivesdrugabusegov/nmassist/* Accessed July 3, 2021.
241