Table of Contents

S1. Sample Size

The sample size was derived from an a priori power analysis. Power estimates were calculated by simulation of a two-arm balanced trial with simple binomial randomization and one pre- and one post-intervention measurement of the A1c outcome, analyzed with a random-intercepts linear regression model. In our simulations we targeted a minimum clinically important intervention effect (between arm difference in pre-post difference) of 0.3 percentage points A1c as the alternate hypothesis, and zero difference as the null. The standard deviation of A1c among people with HbA1c >7.0% was assumed to be 1.6 percentage point based on an analysis of data from the 2004 New York City Health and Nutrition Examination Survey data. The intra-class correlation for repeated measures of A1c at one-year intervals was assumed to be 0.59, as obtained in the Framingham Heart Study. Based on these assumptions, we simulated the power we could obtain with sample sizes ranging between 250 and 750 patients per arm (in multiples of 50.) A sample of 400 analyzable subjects per arm (total N=800 after attrition) was found to provide 83% power to detect the 0.3 percentage point at the 0.05 significance level. To allow for 15% of accrued participants failing to complete the study (based on our team's prior experience), we arrived at a planned enrolled sample size of 941 patients.

S2. Randomization

Method. The randomization was achieved through the following steps:

1. The project epidemiologist, who played no role in study recruitment, sampled 1,200 pseudo-random numbers from the uniform distribution on the unit interval using the random number generator in Stata version 10. The random number generator seed was set to the serial number (stripped of non-numeric characters) of a dollar bill from the front of the epidemiologist's wallet. See Stata code below.
2. For those sequence numbers where the corresponding random number was less than or equal to 0.5, assignment to the Print Only group was generated; where the corresponding random number was greater than 0.5, assignment to the Telephone/Print group was generated. The 1,200 assignments were exported to a tab-delimited text file.
3. The tab-delimited text file was input to a Microsoft Word 2003 mail merge onto 3x5 index cards. Each card bore a sequence number between 1 and 1,200 and the corresponding study arm assignment.
4. 1,200 opaque envelopes were sequentially numbered and each index card was placed in the same-numbered envelope by a research assistant not involved with study recruitment.
5. Whenever a verified eligible person consented to participation, the next unopened envelope was opened and the person was assigned to the arm of the study specified on the enclosed index card.

Comments. We generated 1,200 index cards, anticipating that it might be necessary to replace participants who initially consented to participate and then withdrew early in the study or were determined to have not met eligibility requirements after enrollment. As it turns out, study withdrawal was uncommon, so recruitment ended when the original sample size of 941, planned on the basis of a power analysis, was achieved.

The distribution of treatment assignments among the 1,200 index cards was 634 Print Only and 566 Telephone/Print. Among the first 941 index cards, the ones actually used in the study, the distribution was 498 (52.9%) to Print Only and 443 (47.1%) to Telephone/Print. The two-tailed binomial probability, with 0.5 chance per draw, of imbalance greater than or equal to this magnitude is 0.072.

The approach used is a simple random assignment without blocking. Although program staff engaged in recruitment were not aware at the outset of the imbalance in study arm sizes, the project epidemiologist was. Consideration was given, at that time, to re-doing the randomization with a different seed, or using a blocked design that would assure more equal size of study arms. He decided against these alternatives because:

1. The discrepancy in study arm sizes is not large enough to materially impair inference, and there was no reason to believe that the intended simple Bernoulli trial randomization procedure for generating random assignments had been incorrectly implemented.

2. While the study design's power analysis assumed equal numbers in both arms, the degree of difference in this randomization was not enough to materially degrade the power when tested in a secondary power simulation.

3. Blocked designs provide information about upcoming treatment assignments. For example, if a block size of 8 is used, the study assignment of the 8th participant is known with certainty before he or she is actually "randomized." If 4 of the first 6 in a block have already been assigned to one arm, the assignments of the remaining two are known with certainty. Large block sizes reduce the number of forthcoming arm assignments that can be predicted, but do not eliminate it entirely. Because the recruitment methods of the study involved active persuasion of potential participants by study personnel who ultimately delivered the study interventions, as well as a degree of judgment about some eligibility requirements (e.g., signs of cognitive impairment), it was deemed critical that those personnel be at all times completely uninformed about the study arm to which a potential recruit would be assigned, to avoid any possible bias in recruitment into the study arms. The use of a simple randomization achieves this goal.

Code for Randomization Process. The following is a transcript of the output log of the randomization process :

```
      log:  C:\Documents and Settings\Clyde Schechter\My Documents\NYC DOH DM
Initiative\cr_randomization_scheme.smcl
  log type:  smcl
 opened on:  11 Jun 2008, 20:09:07

.
. version 10

.
. // CREATE A SCHEME OF RANDOM ASSIGNMENTS FOR 1200 PEOPLE.
. // 50% PROBABILITY EACH TREATMENT AND CONTROL
. // NO STRATIFIED BLOCKING
.
. clear

. set obs 1200
obs was 0, now 1200

. gen int seq = _n

. set seed 22984043

. gen byte assignment = (uniform() <= 0.5)

. label define assignment 0       "NYCHAR (Control)"      ///
>                                 1       "Phone + NYCHAR (Intervention)"

. label values assignment assignment

. tab assignment

                  assignment |      Freq.     Percent        Cum.
-----------------------------+-----------------------------------
             NYCHAR (Control) |        634       52.83       52.83
Phone + NYCHAR (Intervention) |        566       47.17      100.00
-----------------------------+-----------------------------------
                       Total |      1,200      100.00

.
. isid seq, sort
(data now sorted by seq)
```

```
. save randomization_scheme, replace
file randomization_scheme.dta saved

.
. // SAVE IN FORMAT FOR PRINTING ON 3X5 CARDS
. outsheet using randomization_scheme.txt, replace noquote

.
. log close
      log:  C:\Documents and Settings\Clyde Schechter\My Documents\NYC DOH DM
Initiative\cr_randomization_scheme.smcl
  log type:  smcl
 closed on:  11 Jun 2008, 20:09:07
```

Note that during the planning stages of the project we referred to the Print Only study arm as "NYCHAR (Control)," and the Telephone/Print arm as "Phone + NYCHAR (Intervention)." [NYCHAR is an acronym for New York City Hemoglobin A1c Registry.]

S3: Missing Outcomes

Expanded Time Windows for Outcome Ascertainment

Not all participants obtained an A1c within the primary study window of 6 weeks before and four months after the anniversary of enrollment. For additional analyses, we also sought A1cs within broader windows. If more than one A1c was found within a window, we sought the one closest in time to the randomization anniversary, and if there were two such, we selected the one occurring after the anniversary. The additional windows were defined as in S3 – Appendix Table 1 below. Note that an outcome that falls in any given window is necessarily also within the next window.

Appendix Table 1. Definitions of Windows for End-of-Study A1c, Relative to One-Year Anniversary of Randomization

Window	Starts	Ends
1	*6 weeks before*	*4 months after*
2	3 months before	6 months after
3	6 months before	9 months after
4	6 months before	1 year after
5	6 *weeks* before	1 year after

A primary outcome was found for 695 participants. Outcomes in windows 2, 3, 4, and 5 were found for 791, 863, 872, and 828 participants, respectively.

Appendix Table 2. Estimated Effect of Telephone Intervention Including Outcomes Obtained in Four Time Windows

Outcome Window	Estimated Effect ($\Delta A1c_{Tele/Pr}$-$\Delta A1c_{PrOnly}$)	95% CI	*p*-value (H0: $\Delta A1c_{Tele/Pr} = \Delta A1c_{PrOnly}$)
2	-0.43%	-0.82 to -0.04%	0.029
3	-0.43%	-0.81 to -0.05%	0.028
4	-0.44%	-0.82 to -0.06%	0.024
5	-0.45%	-0.84 to -0.07%	0.020

Single-Imputation Robustness Analyses

In addition to these window outcomes, additional outcomes for robustness analyses were obtained. These additional outcomes are: the last outcome obtained, even if not in any window, provided it was no later than 4 months beyond the randomization anniversary (last outcome), the lowest A1c on record, with no time restrictions, (best outcome), the highest A1c on record, again, with no time restrictions, (worst outcome), and the extreme worst case outcome—defined as the best outcome for those in the Print Only group and the worst outcome for those in the Telephone/Print group. These analyses do not produce unbiased estimates of treatment effectiveness, but they indicate the extent to which adverse patterns of missing outcome have the ability to refute the conclusions of the primary analysis. The results of these analyses are shown below:

Appendix Table 3. Estimated Study Effect Using Robust Proxy Outcomes

Observation Carried Forward	**Estimated Effect ($\Delta A1c_{Tele/Pr}$-$\Delta A1c_{PrOnly}$)**	**95% CI**	***p*-value (H0: $\Delta A1c_{Tele/Pr} = \Delta A1c_{PrOnly}$)**
Last	-0.31%	-0.58 to -0.03%	0.028
Best (Lowest)	-0.28%	-0.56 to 0.00%	0.047
Worst (Highest)	-0.37%	-0.65 to -0.10%	0.008
Extreme Worst (Worst Tele/Pr and Best PrOnly)	+0.35%	0.07 to 0.63%	0.014

As can be seen, only the highly implausible extreme worst case analysis leads to conclusions in an opposite direction. Combined with the expanded time window analyses, these bolster our confidence in the conclusions of the primary analysis.

Multiple Imputation Analysis

Although we do not consider it credible that outcomes in this study are missing at random, because the technique has become popular, we also carried out multiple imputation with missing A1c outcomes specified conditional on age, baseline A1c, and BMI, with a linear regression model and normal error distribution. Based on twenty imputed replicate data sets, with imputation of all 247 missing primary window A1c outcomes, we obtained an effect estimate ($\Delta A1c_{Tele/Pr}$-$\Delta A1c_{PrOnly}$) of -0.37 (95% CI -0.71 to -0.03, p=0.033).

S4: Diabetes Medications Self-Reported by Participants

Appendix Table 4. Medications Reported Taken at Randomization and at End of Study

		Medications at end of study					
Medications at randomization	**Pills only**	**Insulin only**	**Pills and insulin**	**None**	**Pills + other injection**	**Unknown***	**Total**
Pills only	299	7	25	20	7	96	454
	65.86	1.54	5.51	4.41	1.54	21.15	100.00
Insulin only	4	98	21	0	0	31	154
	2.60	63.64	13.64	0.00	0.00	20.13	100.00
Pills and insulin	26	32	117	3	4	43	225
	11.56	14.22	52.00	1.33	1.78	19.11	100.00
None	4	0	0	3	0	1	8
	50.00	0.00	0.00	37.50	0.00	12.50	100.00
Pills + other injection	9	19	45	0	5	19	97
	9.28	19.59	46.39	0.00	5.15	19.59	100.00
Unknown**	0	0	1	1	0	1	3
	0.00	0.00	33.33	33.33	0.00	33.33	100.00
Total	342	156	209	27	16	191	941
	36.34	16.58	22.21	2.87	1.70	20.30	100.00

* Includes non-response to end of program survey and item non-response.
** All participants responded to baseline survey, unknown here denotes only item non-response.
3

Appendix Table 5. Path Analysis for Mediating Effect of Medication Changes

Path	**Coefficient**	**SE**	**z**	**P>z**	**[95% CI]**	
ΔA1c						
Study arm	-0.443	0.178	-2.49	0.013	-0.793	-0.094
Med change	-0.117	0.195	-0.60	0.548	-0.498	0.264
constant	-0.319	0.138	-2.32	0.020	-0.589	-0.050
Med change						
Study arm	-0.091	0.160	-0.57	0.570	-0.404	0.222
constant	-0.794	0.109	-7.28	0.000	-1.008	-0.581
var(e.Δa1c)	4.614	0.271			4.113	5.177

Notes: 1. N=748 complete observations.
2. log likelihood = -1726.7311

From these path coefficients we calculate the results in Appendix Table 6.

Appendix Table 6. Total and Indirect Effects of Study Arm on Change in A1c.

Effect	Coefficient	SE	z	P>z	[95% CI]	
Indirect effect	0.011	0.026	0.41	0.680	-0.040	0.061
Total effect	-0.433	0.180	-2.41	0.016	-0.786	-0.080

Notes: 1. SEs calculated by delta method.
2. Similar results are found (not shown) when the sample for this analysis is restricted to the tier with baseline A1c >9, although statistical significance of all effects is lost due to smaller sample size.
3. Similar results are also found (not shown) when a linear link is used in the Study Arm to Medication Change path.