Supplementary Online Content

Yancy WS Jr, Crowley MJ, Dar MS, et al. Comparison of group medical visits combined with intensive weight management vs group medical visits alone for glycemia in patients with type 2 diabetes: a noninferiority randomized clinical trial. *JAMA Intern Med*. Published online November 4, 2019. doi:10.1001/jamainternmed.2019.4802

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Statistical Methods Technical Appendix

1. Model Selection and Fitting process

The predictors in all models included time effect(s) and associated time-by-arm interaction terms. This model assumes the study arms have equal baseline means, which is appropriate for a randomized control trial and is equivalent in efficiency to an ANCOVA model.(1, 2)

The process for selecting the best model for each outcome was a two-step process.

- In the first step we determined the "best" covariance structure by fitting both "hybrid" models with a random effect for clustering of counseling group over time and different covariance structures that included compound symmetry (CS in SAS commands), autoregressive (AR(1)), Toeplitz (TOEP), spatial exponential (SP(EXP)), spatial power (SP(POW)), and unstructured (UN) for the serial correlation between time points and a set of random coefficient models that included 1) random effect for group and random intercept and linear slope for subjects. These models were fit using REML and AIC model selection criteria were assessed to determine the best fit model.
- In the second step, we used the covariance structure identified in step 1 for each outcome to determine the best mean structure. In this step, we fit separate models using linear time, quadratic time, cubic time and dummy coded time for the fixed effects for each outcome. These models were fit using ML and AIC model selection criteria were assessed to determine the best fit model.

Following this process for each outcome we ran the "best fit" model including stratification variables and estimated arm differences at 48 weeks from these models. All the final models were fit using REML. The final model for hemoglobin A_{1c} as determined by this process was used for all subsequent sensitivity analysis. Final model code is available upon request.

2. Missing data and Multiple Imputation Procedure

Longitudinal models fit in our analysis used all available data, including data from participants who had missing observations and/or were lost to attrition, with the estimation procedure implicitly accommodating missing values when related to prior outcome or to other baseline covariates in the model (i.e., missing at random (MAR)).

We also conducted a sensitivity analysis using a multiple imputation (MI) approach that included additional baseline variables beyond those in our random effects models to strengthen the MAR assumption. As a first step, we used t-tests, Wilcoxon rank sum tests and chi-square tests as appropriate to assess each potential baseline variable's association with missingness at 48 weeks, and any variable with an association p-value of 0.25 or less was included in the imputation model. Baseline variables assessed included gender, race, education level, diabetes medicine regime, marital status, distance from VA hospital, whether taking hypertension medications or cholesterol lowering medications, age, age at diagnosis of diabetes, weight, BMI, waist circumference, MES score, PAID score, daily calories consuming, daily grams of fat, carbohydrate, and protein. Of these, the following were associated with missing status at week 48 and therefore included in the imputation model: age, age at diabetes diagnosis, weight, gender, marital status, waist circumference, education level, whether taking a cholesterol medication or not, daily calories and grams of carbohydrate and protein. The imputation model additionally included randomization arm, stratification variables (baseline hemoglobin A_{1c} 7.5-8.9% vs. ≥9% and use of multiple types of insulin vs. one type or no insulin, group cohort ID, and all collected hemoglobin A_{1c} at the 4 possible time points. Missing hemoglobin A_{1c} measurements at any of the 4 time points were imputed using a Markov chain Monte Carlo (MCMC) algorithm with 10 imputations. The imputation provided results that were very similar to the main analysis. For models run on the imputed datasets, WM/GMV was found to be non-inferior to GMV at 48 weeks (estimated mean difference = -0.1%, 95%CI -0.5%, 0.2%; upper 95% CI <0.5%) but not superior to GMV (p=0.44).

3. Per-protocol analysis

For the per-protocol analysis, we fit the primary model to the subset of patients in each arm who attended at least 75% of group sessions (at least 10 of 13 sessions in WM/GMV arm and at least 7 of 9 sessions in GMV arm; n=77 of 127, 61% in WM/GMV arm; n=75 of 136, 55% in GMV arm). In the subset of patients who attended at least 75% of group sessions, only one subject was missing

the week 48 hemoglobin A_{1c} measurement. For the per-protocol analysis WM/GMV was found to be non-inferior to GMV at 48 weeks (estimated mean difference = -0.1%, 95%CI -0.5%, 0.3%; upper 95% CI <0.5%) but not superior to GMV (p=0.72).

eTable 1. Mean (SD) Calorie and Macronutrient Intake for Weight Management/Group Medical Visit (WM/GMV) and Group Medical Visit (GMV) Arms by Time Point

Measurement	N	WM/GMV	N	GMV
Calories (kcal)				
Baseline	112	2042.7 (875.3)	105	1879.4 (592.0)
16 weeks	100	1848.6 (992.3)	93	1822.4 (839.9)
32 weeks	82	1810.2 (616.0)	87	1780.5 (675.6)
48 weeks	94	1738.8 (685.3)	98	1824.6 (686.1)
Carbohydrates (g)				
Baseline	112	194.4 (85.0)	105	187.7 (85.8)
16 weeks	100	90.3 (53.5)	93	179.0 (99.3)
32 weeks	82	106.0 (68.6)	87	165.5 (71.5)
48 weeks	94	113.9 (63.6)	98	176.2 (87.1)
Fat (g)				
Baseline	112	96.0 (49.7)	105	87.3 (30.0)
16 weeks	100	111.7 (70.0)	93	85.7 (50.0)
32 weeks	82	102.1 (40.5)	87	82.9 (41.1)
48 weeks	94	96.6 (46.7)	98	85.0 (35.0)

Protein (g)				
Baseline	112	103.0 (77.1)	105	89.4 (30.4)
16 weeks	100	121.3 (87.2)	93	86.6 (32.8)
32 weeks	82	115.7 (56.6)	87	89.9 (37.5)
48 weeks	94	103.4 (46.5)	98	90.9 (29.8)

eTable 2. Minutes of Activity Per Week for Weight Management/Group Medical Visit (WM/GMV) and Group Medical Visit (GMV) Arms by Time Point

	WM/GMV			GMV		
	N	**Mean (SD)**	**Median (IQR)**	**N**	**Mean (SD)**	**Median (IQR)**
Minutes Per Week of Vigorous Activity						
Baseline	115	102.8 (266.7)	0 (0)	116	156.4 (413.6)	0 (120.0)
16 weeks	93	101.7 (233.8)	0 (60.0)	94	110.0 (293.0)	0 (90.0)
32 weeks	87	63.9 (172.8)	0 (20.0)	84	75.6 (233.8)	0 (0)
48 weeks	92	61.0 (194.4)	0 (0)	88	118.3 (394.3)	0 (0)
Minutes Per Week of Moderate Activity						
Baseline	115	443.3 (564.9)	180.0 (705.0)	116	625.3 (796.4)	300.0 (865.0)
16 weeks	93	629.5 (757.6)	360.0 (870.0)	94	480.5 (551.1)	315.0 (540.0)
32 weeks	87	480.7 (703.9)	240.0 (630.0)	84	518.7 (704.6)	180.0 (840.0)
48 weeks	93	375.3 (520.6)	240.0 (480.0)	89	412.0 (664.9)	180.0 (495.0)
Minutes Per Week of Walking						
Baseline	115	431.2 (776.6)	120.0 (480.0)	116	489.3 (698.1)	180.0 (707.5)
16 weeks	93	404.5 (614.0)	180.0 (480.0)	94	432.5 (726.5)	210.0 (415.0)

32 weeks	87	322.6 (502.5)	120.0 (410.0)	84	357.4 (500.5)	205.0 (457.5)
48 weeks	93	296.3 (460.5)	120.0 (370.0)	89	468.9 (730.8)	210.0 (570.0)

eTable 3. Estimated Means and Mean Differences (95% CI) of Clinical and Laboratory Outcomes for Weight Management/Group Medical Visit (WM/GMV) and Group Medical Visit (GMV) Arms by Time Point

Measurement	WM/GMV Group	GMV Group	Mean Difference (WM/GMV− GMV) (95% CI)	*P* Value
Systolic blood pressure[+]				
Baseline	129.6	129.6		
16 weeks	130.0	130.3	-0.3 (-3.6, 3.0)	
32 weeks	130.9	130.4	0.5 (-3.3, 4.2)	
48 weeks	132.1	129.9	2.2 (-2.1, 6.5)	0.31
Diastolic blood pressure[+]				
Baseline	79.1	79.1		
16 weeks	78.6	78.3	0.3 (-0.7, 1.3)	
32 weeks	78.1	77.5	0.6 (-1.4, 2.7)	
48 weeks	77.7	76.8	0.9 (-2.1, 4.0)	0.55
Total cholesterol[#]				
Baseline	153.5	153.5		
16 weeks	147.1	152.6	-5.6 (-15.6, 4.4)	
32 weeks	156.1	152.5	3.6 (-6.8, 14.1)	

48 weeks	156.0	154.9	1.1 (-9.8, 11.9)	0.84
HDL cholesterol[#]				
Baseline	40.8	40.8		
16 weeks	42.2	41.1	1.1 (0.3,1.9)	
32 weeks	43.6	41.5	2.2 (0.5, 3.8)	
48 weeks	45.0	41.8	3.2 (0.8, 5.7)	0.01
LDL cholesterol[^]				
Baseline	91.3	91.3		
16 weeks	89.3	89.4	-0.1 (-5.2, 5.0)	
32 weeks	90.3	90.0	0.3 (-5.9, 6.4)	
48 weeks	94.5	93.3	1.2 (-5.9, 8.3)	0.74
Triglycerides[#]				
Baseline	167.7	167.7		
16 weeks	141.6	163.7	-22.1 (-44.3, 0.1)	
32 weeks	163.8	176.0	-12.2(-39.9, 15.6)	
48 weeks	160.8	176.0	-15.2 (-41.3, 10.9)	0.25
Creatinine[~]				
Baseline	1.1	1.1		

16 weeks	1.1	1.1	-0.02 (-0.03, -0.00)	
32 weeks	1.1	1.1	-0.03 (-0.06, -0.00)	
48 weeks	1.1	1.1	-0.05 (-0.09, -0.00)	0.040

[+]Baseline data are missing for 9 participants. Follow-up data at week 16 are missing for 41 participants (19 in WM/GMV and 22 in GMV), at week 32 for 65 participants (31 in WM/GMV and 34 in GMV), and at week 48 for 54 participants (26 in WM/GMV and 28 in GMV).

[#]Follow-up data at week 16 are missing for 36 participants (17 in WM/GMV and 19 in GMV), at week 32 for 50 participants (25 in WM/GMV and 25 in GMV) and at week 48 for 38 participants (18 in WM/GMV and 20 in GMV).

[^]Follow-up data at week 16 are missing for 36 participants (17 in WM/GMV and 19 in GMV), at week 32 for 50 participants (25 in WM/GMV and 25 in GMV) and at week 48 for 39 participants (19 in WM/GMV and 20 in GMV).

[~] Follow-up data at week 16 are missing for 35 participants (17 in WM/GMV and 18 in GMV), at week 32 for 50 participants (25 in WM/GMV and 25 in GMV) and at week 48 for 38 participants (19 in WM/GMV and 19 in GMV).

eReferences

1. Liu GF, Lu K, Mogg R, Mallick M, Mehrotra DV**.** Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? Stat Med. 2009;28(20):2509-30.
2. Fitzmaurice GM, Laird NM, Ware JH**.** Applied longitudinal analysis. Hoboken, NJ: Wiley-Interscience; 2004.