1

Online-Only Supplemental Table

3

4

eTable 1: Reasons for Withdrawal

Reason	All Participants, n=986			English-speakers, n=541			Spanish-speakers, n=445		
	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value
	n=33 n (%)	n=56 n (%)		n=12 n (%)	n=27 n (%)		n=21 n (%)	n=29 n (%)	
Lost Interest, No. (%)	6 (18.2)	12 (21.4)	0.04*	1 (8.4)	5 (18.5)	0.12*	5 (23.8)	7 (24.1)	0.44*
Too Sick, No. (%)	5 (15.1)	2 (3.6)		3 (25.0)	1 (3.7)		2 (9.5)	1 (3.5)	
Study too long, No. (%)	3 (9.1)	13 (23.2)		0 (0)	4 (14.8)		3 (14.3)	9 (31.0)	
Study upsetting, No. (%)	2 (6.1)	12 (21.4)		1 (8.3)	8 (29.7)		1 (4.8)	4 (13.8)	
Too busy, No. (%)	5 (15.1)	6 (10.7)		3 (25.0)	3 (11.1)		2 (9.5)	3 (10.3)	
Other, No. (%)	12 (36.4)	11 (19.7)		4 (33.3)	6 (22.2)		8 (38.1)	5 (17.3)	

*Fisher's Exact Test. The p-value significance level was set at 0.05. Total withdrawals, n=89 (9.0%), 7.2% for English-speakers and 11.0% for Spanish-speakers. The p-value significance level was set at < 0.05.

eTable 2: Interactions Effects of Patient Characteristics on Advance Care Planning Documentation by Study Arm*

ALL PARTICIPANTS		AD-only n=505 (%)	PREPARE n=481 (%)	p-value
Language	English	38.1	46.2	0.21
	Spanish	25.6	40.1	
Health literacy	Adequate	34.6	45.2	0.70
	Limited	29.6	42.2	
Role in decision making	Makes own decisions	32.7	43.4	0.71
	Doctor decides	31.2	45.8	
Age	< 65 years	28.6	41.5	0.27
	≥65 years	39.9	46.8	
Gender	Women	29.0	40.6	0.66
	Men	38.1	47.8	
Race/ethnicity	White	31.3	43.9	0.42
	Non-white	35.0	42.5	
Health status	Good-to-excellent	29.6	38.5	0.62
	Fair-to-poor	35.0	47.7	
Potential surrogate	Yes	31.9	43.4	0.69
	No	43.9	49.7	
Internet access at home	Yes	35.2	41.0	0.08
	No	29.6	45.8	
SPANISH-SPEAKING PARTICIPANTS ONLY		AD-only n=226 (%)	PREPARE n=219 (%)	p-value*
Doctor-patient language	Concordant	29.8	38.7	0.13
	Discordant	21.1	44.0	

*Percentages represent stratified, adjusted analysis. Interactions were tested by adding interaction terms to the group-by-time variable in adjusted mixed effects models. All dichotomous interaction terms are defined in the protocol.[1] All models were adjusted for health literacy, prior ACP documentation, and clustering by clinician; however, the health literacy interaction model was only adjusted for prior ACP documentation and clustering by clinician. The p-value significance level was set at < 0.05. All variables were assessed for interactions among the total cohort, except doctor-patient language concordance which was assessed only for Spanish-speaking participants.

eTable 3. Effect Size Estimates* of ACP Behavior Change and Actions Over Time and By English and Spanish-speakers

	All Participants, n=986		English-speakers, n=541		Spanish-speakers, n=445	
	AD-only	PREPARE	AD-only	PREPARE	AD-only	PREPARE
ACP Behavior Change						
1 week	0.27	0.58	0.34	0.54	0.21	0.70
3 months	0.41	0.73	0.43	0.63	0.41	0.93
6 months	0.47	0.82	0.48	0.77	0.49	1.00
12 months	0.59	0.93	0.61	0.83	0.63	1.16
ACP Actions						
1 week	0.24	0.56	0.26	0.55	0.24	0.57
3 months	0.37	0.79	0.30	0.74	0.47	0.86
6 months	0.47	0.90	0.43	0.90	0.54	0.90
12 months	0.61	1.09	0.55	1.04	0.74	1.18

*Clinically meaningful effect sizes were calculated from changes from baseline (i.e., mean PREAPRE scores minus mean AD-only scores divided by pooled baseline standard deviations) and are based on standardized criteria (i.e., 0.20-0.49 small, 0.50-0.79 medium, and ≥0.80 large).[2]

29 **eTable 4: Ease-of-use and Satisfaction Between Study Arms and by English- and Spanish-speakers***

	All Participants, n=986			English-speakers, n=541			Spanish-speakers, n=445		
	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value
Ease of use	8.0 (2.2)	8.1 (2.0)	0.46	8.5 (1.8)	8.7 (1.7)	0.31	7.3 (2.4)	7.5 (2.1)	0.72
Satisfaction									
Comfort viewing	4.0 (0.9)	4.1 (0.8)	0.37	4.3 (0.9)	4.4 (0.8)	0.95	3.7 (0.9)	3.8 (0.6)	0.09
Helpfulness	4.2 (0.8)	4.4 (0.7)	<0.001	4.4 (0.8)	4.5 (0.7)	0.11	3.9 (0.7)	4.2 (0.6)	<0.001
Recommend	4.1 (0.8)	4.2 (0.8)	0.36	4.3 (0.9)	4.3 (0.9)	0.72	4.0 (0.7)	4.1 (0.7)	0.26

30

31 *Ease-of-use was measured on a 1 (very hard) to 10 (very easy) point scale. Satisfaction was measured by, "How comfortable were
32 you viewing this guide?"; helpfulness: "How helpful was this guide?"; and recommendations: "How likely are you to recommend this
33 guide to others?" assessed on a 5-point Likert scale (not-at-all to extremely) from our prior work.[1] The p-value significance level was
34 Bonferroni adjusted for this secondary outcome at < 0.025.

35

36

37 **eTable 5: Depression and Anxiety Scores Between Study Arms and by English- and Spanish-speakers**

	All Participants, n=986 mean (95% confidence interval)			English-speakers, n=541 mean (95% confidence interval)			Spanish-speakers, n=445 mean (95% confidence interval)		
	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value	AD-only	PREPARE	p-value
Depression Score (0-24)	4.3 (3.9-4.7)	4.2 (3.8-4.6)	0.69	4.1(3.6-4.6)	4.4 (3.9-5.0)	0.35	4.5(4.0-5.1)	3.9 (3.3-4.4)	0.10
Anxiety Score (0-21)	3.4 (3.1-3.8)	3.2 (2.9-3.5)	0.32	3.2(2.8-3.7)	3.4 (2.9-3.8)	0.70	3.7(3.2-4.2)	3.0 (2.5-3.5)	0.05

38

39 *Adjusted for baseline depression and anxiety scores. Adjusted mean and 95% confidence interval (CI). Depression was measured
40 using the validated Patient Health Questionnaire (PHQ)-8 (scores 0-24) and anxiety was measured with the Generalized Anxiety
41 Disorder (GAD)-7 (scores 0-21).[3] Scores less than 5 represent no clinically meaningful anxiety or depression and scores of 5, 10, 15,
42 and 20 represent mild, moderate, moderately severe and severe depression or anxiety, respectively.[4,5] The p-value significance level
43 was set at < 0.05.

44

45
46
47
48
49
50

Online-Only References:

1. Sudore RL, Barnes DE, Le GM, et al. Improving advance care planning for English-speaking and Spanish-speaking older adults: study protocol for the PREPARE randomised controlled trial. *BMJ open.* 2016;6(7):e011705.

2. Cohen J. *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ:: Lawrence Earlbaum Associates.; 1988.

3. Lowe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord.* 2010;122(1-2):86-95.

4. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord.* 2009;114(1-3):163-173.

5. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097.