



HHS Public Access

Author manuscript

BMJ Support Palliat Care. Author manuscript; available in PMC 2019 November 16.

Published in final edited form as:

BMJ Support Palliat Care. 2017 March ; 7(1): 67–72. doi:10.1136/bmjspcare-2014-000814.

Assessing the internal consistency and temporal stability of advance directives generated by an interactive, online computer program

Jane R. Schubart¹, Fabian Camacho², Michael J. Green³, Kimberly A. Rush⁴, Benjamin H. Levi⁵

¹Departments of Surgery, Public Health Sciences, and Medicine, The Pennsylvania State University, College of Medicine, Hershey, PA USA

²Department of Public Health Sciences, University of Virginia, School of Medicine, Charlottesville, VA USA

³Departments of Humanities and Medicine, The Pennsylvania State University, College of Medicine, Hershey, PA USA

⁴Department of Pediatrics, The Pennsylvania State University, College of Medicine, Hershey, PA USA

⁵Departments of Humanities and Pediatrics, The Pennsylvania State University, College of Medicine, Hershey, PA USA

Abstract

Objective—Evaluate the internal consistency and temporal stability of advance directives (ADs) generated by an interactive, online computer program.

Methods—33 participants completed the program at three visits, two weeks apart. Agreement rates were calculated for the *General Wishes* component of the AD. The test-retest method was used to examine the temporal stability of the *Specific Wish for Treatment* component which contains five clinical scenarios.

Results—*General Wishes* remained stable with 94% selecting the identical response at each visit. For the *Specific Wish for Treatment* scale, significant variations in test-retest correlations existed (i.e., $\rho = 0.32$ to 0.78 between time points 1 vs. 2), however within scenario, correlations did not significantly vary between time points. Temporal stability was lower in the *Specific Wish for*

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ Supportive and Palliative Care and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (<http://group.bmj.com/products/journals/instructions-for-authors/licence-forms>).

Correspondence to: Jane R. Schubart, PhD MS MBA, Department of Surgery, The Pennsylvania State University, College of Medicine, 500 University Drive, Hershey, PA 17033, Tel: 717-531-1262, Fax: 717-531-0480, jschubart@hmc.psu.edu.

Ethics approval

This study was approved by the Penn State College of Medicine's Institutional Review Board before the study began.

Competing Interest

Competing interests MJG and BHL have intellectual property and copyright interests for the decision aid used for this study, Making Your Wishes Known: Planning Your Medical Future (MYWK), which could result in future royalties or income to them.

Treatment scale compared to *General Wishes* (avg $\rho = 0.59$ between time points 1 and 2, and $\rho = 0.75$ between time points 2 and 3).

Conclusion—ADs generated by an online decision aid demonstrate good temporal stability, with highest stability for *General Wishes* and moderate stability for *Specific Wish for Treatment* regarding medical treatments in specific clinical scenarios. Internal consistency for wish for treatment across all time points and scenarios was high (Cronbach alpha > 0.90).

Keywords

advance care planning; advance directive; decision aid; computer

INTRODUCTION

Advance care planning (ACP) is a process of planning for future medical treatment in the event that a person cannot speak for him or herself. This is usually accomplished by completing an advance directive (AD), a document that outlines specific healthcare instructions and/or designates a proxy decision-maker. Up to 75% of adults lack decision-making capacity when life-or-death medical decisions must be made¹, and studies have shown that neither family members nor doctors accurately predict what patients want.²⁻³ The lack of advance planning can lead to unfavorable outcomes including moral distress⁴ and conflict for those who must make the decisions⁵, medical care that is inconsistent with an individual's wishes⁶, and unintended financial burdens to patients, their families, and society.⁷

Although there is general agreement that people ought to plan for their medical futures⁸, there remain significant barriers to implementing ADs.⁹ Key elements within ADs are often poorly understood¹⁰, and there are concerns that: discussing death and dying might diminish hope and raise anxiety¹¹; many patients lack the knowledge to complete informed ADs¹²; ADs often fail to accurately reflect a person's actual values, goals and preferences for healthcare^{5,13}; information contained in ADs is difficult for family members or healthcare providers to interpret¹⁴; and patients change their minds about which medical treatments they want.¹²

Making Your Wishes Known: Planning Your Medical Future (MYWK) is an interactive computer-based decision aid developed to address some of these concerns. The program guides users through the ACP process by providing tailored education, exercises to clarify values, and a decision-making algorithm based on multi-utility attribute theory (MAUT) that generates a personalized AD. In prior work with patients and healthy volunteers, we have demonstrated that users of the program are highly satisfied with *MYWK*, that even patients with advanced illness find *MYWK* easy to use, and that doing so does not raise users' anxiety or decrease their sense of hope.¹⁵⁻¹⁶ Study results also demonstrate that use of *MYWK* improves patients' knowledge about ACP, generates an AD that users report accurately reflects their wishes regarding future medical decisions, and can help healthcare providers make decisions on behalf of patients who cannot speak for themselves.¹⁷

To establish the validity of this decision aid, its internal consistency and temporal stability must also be examined. As such, the present study explores whether (in the absence of major life changes) the AD generated by *MYWK* remains stable over time in articulating an individual's values and preferences. In this study, we asked participants to complete *MYWK* three times, separated by two-week intervals.

METHODS

Recruitment

Study participants were recruited in Summer 2011 from the Penn State Hershey Medical Center using flyers placed in outpatient clinics, on-hold messages for telephone callers, and electronic message screens in public areas. Participants also were recruited from a list of individuals who had previously expressed interest in participating in advance care planning research.

Procedure

Eligible individuals were invited by phone to attend an in-person session at which a member of the research team elicited informed consent and screened for eligibility—8th grade reading level (≥ 26 on WRAT-3)¹⁸, cognitively able to use the program (≥ 25 on Mini-Mental State Examination)¹⁹, and not having “moderate/severe” or “severe” depression (≥ 19 on Beck Depression Inventory-II).²⁰ Depressed individuals were excluded because depression is associated with a diminished will to live and greater desire for death; as such, the presence of depression can distort decisions made during advance care planning.^{21, 22} Study participants completed a demographic questionnaire, a major life events report (recent events that might influence their responses to end-of-life healthcare decisions), and the *MYWK* computer program. During the second and third study visits (each conducted after a two-week interval), participants again completed the major life events questionnaire and *MYWK* program. Each session lasted 1–3 hours, and participants received a \$25 gift certificate after the first and second study visits, and \$50 upon completion of the third visit.

Intervention content and procedure

The computer-based decision aid, *MYWK*, includes six sections.¹⁵ *Getting Started* provides an overview of the program. *Choosing a Spokesperson* reviews surrogate decision-making and then prompts the user to designate primary and alternate spokespersons. *Exploring Your Values* helps the user clarify his or her values and goals regarding medical care, death and dying, and disability. *Your Medical Wishes* explains health conditions that can prevent a patient from communicating preferences for medical treatments, and describes interventions that commonly involve life-or-death decisions. The user is prompted to make a series of decisions involving specific conditions and treatments; these data are used in the program's decision-making algorithm to generate an AD that individuals review in *Putting It All Together*. Finally, *The Next Step* provides practical tips for communicating one's wishes to those who might be involved in medical decision-making.

At each study visit, participants completed the *MYWK* program, starting anew each time. Participants' previous responses were not disclosed on subsequent visits, and participants

did not have access to *MYWK* between visits to practice. In completing the program, participants confirmed, selected an alternative, and/or edited the *General Wishes* statement chosen by the computer program to represent their values and goals (see Appendix 1). They also reviewed and edited the *Specific Wish for Treatment* generated by the decision aid's algorithm regarding desires for eleven life-sustaining medical treatments (mechanical ventilation < 24 hours, up to a month, >1 month; cardio-pulmonary resuscitation; kidney dialysis < 1 month, > 1 month; feeding tube up to one month, > 1 month; surgery; medicines; and blood transfusion) for five clinical scenarios (coma that would improve within a year; coma that would not improve within a year; moderate/severe stroke that would improve within a year; moderate/severe stroke that would not improve within a year; and dementia). This review and confirmation process resulted in a tailored, printable AD for each study visit, whose final contents were then used for data analysis.

Statistical methods

Two components of the AD document generated by the *MYWK* computer program were examined: 1) *General Wishes* score; and 2) *Specific Wish for Treatment* score for five clinical scenarios. The final *General Wishes* score consisted of an ordinal response at each of three time points, with 6 levels ranging from 1 ("want any and all medical treatments") to 6 ("do not want any medical treatments"). The final *Specific Wish for Treatment* score at each time point and scenario consisted of a vote count regarding how many of the 11 life-sustaining treatments a participant wanted. A high score indicated a desire for more extensive life-prolonging treatment, whereas a low score indicated less desire for life-prolonging treatment.

For the multi-item *Specific Wish for Treatment*, a Longitudinal Confirmatory Factor Analysis²³ on binary items was fit to examine the assumption of a latent value driving item responses. Path coefficients between the latent factor and each item were constrained to be equal across occasions. The MPLUS program was used to fit the measurement model for each of 5 scenarios, and the default WLSMV estimator was used.²⁴

The stability of responses was assessed by the test-retest method using Pearson correlations. Agreement was assessed by weighted kappa coefficients.²³ Internal consistency was assessed by Cronbach alpha, or the Kuder-Richardson coefficient for binary responses.²⁶ Weights for the kappa coefficients were based on the squared difference of the levels (Fleiss-Cohen version), shown to be equal to an intraclass correlation coefficient (ICC) in a randomly sampled person by occasion design.²⁷

RESULTS

Thirty-three participants completed the study (79% female; mean age 52 years, range: 31–78), of whom 61% reported being college graduates, 94% being comfortable using a computer, 24% having previously created an AD, and 18% having previously assigned a healthcare spokesperson. To reach the recruitment goal, 63 individuals were telephoned, of whom 26 could not be reached and one declined participation. Of the 36 people who agreed to participate, two did not show up for the study visit and one screen-failed.

At the second study visit, 29/33 (88%) self-reported (by survey questionnaire) no change in their medical wishes for treatment, and at the third visit, 28/33 (85%) self-reported no change from the second visit. At both the second and third study visits, 22/33 (67%) reported sharing their advance directives with others since the prior visit; and 2/33 (6%) had changed their mind about one of their spokespersons. Current health was rated as excellent or very good by 20/33 (61%) and good or fair by 13/33 (39%). Two reported a major life event in the 4–6 weeks prior to visit 2, and two reported a major life event in the 4–6 weeks prior to visit 3.

Specific Wish for Treatment

Confirmatory factor analysis suggested that the unidimensional measurement model fits the item response data. The root mean square error of approximation²⁸ (RMSEA) ranged 0.0–0.03 indicating excellent fit by scenario. CFI (Comparative Fit Index)/TLI (Tucker-Lewis Index of Non-normed Fit Index) indices of fit exceeded 0.99 (where 1.00 indicates perfect fit). Standardized factor loading averages for the 11 items were high, ranging 0.94–0.98, with the minimum loading being 0.86. These results at best confirm and at worst do not contradict our view that sum score of items measures an underlying latent “wish” or desire for treatment.

Test-retest stability between time 1 and time 2 ranged 0.32–0.78 (test of equality²⁹, $p=0.02$) and between time 2 and time 3 ranged 0.58–0.83 (test of equality, $p=0.19$). Weighted kappa coefficients closely track the correlations and ranged 0.32–0.82, with kappa agreement particularly low (0.32) for time 1 to time 2 comparison for the dementia scenario. Cronbach alpha scores were consistently high (>0.90).

Although Pearson correlations are higher when scores are correlated between time 2 and time 3 compared to time 1 and time 2, a pair-wise comparison of correlation coefficients²⁹ did not find a significant difference for any of the five scenarios. Because statistically significant differences were not found across scenarios (except in one instance), there is not sufficient evidence to conclude correlations between time points vary.

General Wishes

For 30/33 (91%) of participants, their final *General Wishes* statement was identical for each of the three study visits. Additionally, 28/33 indicated that “quality-of-life” was a major determinant in their *General Wishes*, and that so long as they would have a good quality-of-life they would want any/all life-sustaining medical treatments. As shown in Table 1, for time 1 vs time 2, the weighted kappa is 0.12 (95% CI: –1.00,1.00); for time 2 vs time 3 the weighted kappa is 0.94 (95% CI: 0.83,1.00). It should be noted that the low kappa for time 1 vs time 2 is driven by a single outlier participant. In fact, only two participants changed their responses from time 1 to time 2, and one changed from time 2 to time 3; only 1 participant changed by more than one unit. Thus, we view this measure as particularly stable across the three time periods despite the low kappa attributable to an abnormal response from time 1 to time 2. Table 1 shows characteristics of *Specific Wish for Treatment* and *General Wishes* scales.

DISCUSSION

Making Your Wishes Known (MYWK) was highly reliable in representing users' *General Wishes* preferences for future medical treatment when administered three times, separated by two-week intervals, but less reliable regarding *Specific Wish for Treatment* preferences. Despite the low kappa due to an outlier from time 1 to time 2, the high agreement rate for the *General Wishes* (91%) across the three time periods illustrates the stability of the measure. This stability may be helped by the lower number of categories a respondent can choose from for the *General Wishes* score.

To better understand whether reliability was influenced by *MYWK* itself, (i.e., impact of the program's content on individuals' preferences), participants completed the program three times rather than twice. However, in evaluating stability across these three visits, no such differences (T1→T2 versus T2→T3) were identified despite consistently higher within scenario stability at T2→T3 compared to T1→T2 for the *Specific Wish for Treatment* scale. We surmise that within scenarios comparison tested nonsignificant due to the small sample size, and that the consistent increase in stability between scenarios is indicative of our hypothesis.

Given the highly controlled study conditions, these findings raise the question concerning the larger than expected variability seen in individuals' *Specific Wish for Treatment* as patient desire for treatment is assumed to be a fundamentally stable trait during the time frame of the study. In this study, this is supported by 29/33 participants reporting at visit 2 and 28/33 reporting at visit 3 that their wishes for medical treatment had not changed. Sources of instability as measured by the test-retest correlation can be separated into two components pertaining to transient error and random response error.³⁰

Random response error is caused by "momentary changes in attention, mental efficiency, distractions"³⁰ during a given occasion which may lead to different item responses in *MYWK* even when overall patient preference remains unchanged. In light of prior research showing that individuals may lack awareness of changes in their preferences and/or have faulty recollections^{31–35}, this type of error is a potential threat to the stability of a measure and to the use of static documents compared to good verbal communication. As random response diminishes with increasing number of items in the scale and higher Cronbach alphas,^{30,36} which are high (> .90) in this study, our assessment is that this type of error does not explain the major reason for instability of the scale.

The other source of instability is then transient error, which reflects temporal variation in the underlying trait which may be attributable to mood, disposition, or other time varying states.³⁰ If transient error plays the major role in affecting the stability, one would question why other markers (such as the *General Wishes* score) suggest that the desire for treatment remains unaffected by transience. Among plausible reasons for this discrepancy, we suspect the scenario dependence of the *Specific Wish for Treatment* scale may play a role; stability could be adversely affected by the specific scale's use of the hypothetical clinical scenarios (e.g., an imagined decision which does not capture the emotions surrounding a real-life decision may be more susceptible to influence to transient errors). On the other hand, use of

hypothetical clinical scenarios makes sense given that ACP is premised on anticipating future events. Consequently, it is important to examine whether *MYWK* can reliably produce ADs that accurately reflect an individual's wishes regarding life-or-death medical decisions.

Limitations of this study include a relatively small sample size which may have affected the power to detect significant differences between correlations between scenarios and time points, a single geographic location, and a predominance of female participants. The small sample size also contributes to kappa coefficients that are variable. A larger sample that includes diverse ethnic and socioeconomic groups is desired to better characterize how the instrument performs in the real world. Also, because we excluded individuals with "moderate/severe" or "severe" depression, there is a potential for bias towards increased stability and a favorable response to *MYWK*; thus, future studies should consider inclusion of those with depression. Such studies are justified on the basis of this study.

CONCLUSION

Using *MYWK* generates an advance directive that demonstrates good temporal stability. In addition to a very high (91%) agreement rate for *General Wishes* statements across time-points, internal consistency of participants' *Specific Wish for Treatment* within scenario and occasion was found to be ideal (> 90%). Within scenario, *Specific Wish for Treatment* had a lower stability across multiple occasions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This study was funded by grants from the National Institute of Nursing Research (R21NR008539), and Penn State University (Social Science Research Institute, Woodward Endowment for Medical Science Education, and Tobacco Settlement Fund Award).

REFERENCES

1. No Authors Listed. Guidelines for the appropriate use of do-not-resuscitate orders. Council on Ethical and Judicial Affairs, American Medical Association. *JAMA* 1991;265:1868–71. [PubMed: 2005737]
2. Shalowitz DI, Garrett-Mayer E, Wendler D. The accuracy of surrogate decision makers: a systematic review. *Arch Intern Med* 2006;166:493–97. [PubMed: 16534034]
3. Fagerlin A, Ditto PH, Danks JH, et al. Projection in surrogate decisions about life-sustaining medical treatments. *Health Psychol* 2001;20:166–75. [PubMed: 11403214]
4. Elpern EH, Covert B, Kleinpell R. Moral distress of staff nurses in a medical intensive care unit. *Am J Crit Care* 2005;14:523–30. [PubMed: 16249589]
5. Hawkins NA, Ditto PH, Danks JH, et al. Micromanaging death: process preferences, values, and goals in end-of-life medical decision making. *Gerontologist* 2005;45:107–17. [PubMed: 15695421]
6. Fagerlin A, Schneider CE. Enough. The failure of the living will. *Hastings Cent Rep* 2004;34:30–42. [PubMed: 15156835]
7. Emanuel EJ, Emanuel LL. The economics of dying: the illusion of cost savings at the end of life. *N Eng J Med* 1994;330:540–44.

8. Emanuel LL, Danis M, Pearlman RA, et al. Advance care planning as a process: structuring the discussions in practice. *J Am Geriatr Soc* 1995;43:440–46. [PubMed: 7706637]
9. Ditto PH, Hawkins NA. Advance directives and cancer decision making near the end of life. *Health Psychol* 2005;24(4 Suppl):S63–70. [PubMed: 16045421]
10. Upadya A, Muralidharan V, Thorevska N, et al. Patient, physician, and family member understanding of living wills. *Am J Respir Crit Care Med* 2002;166:1430–35. [PubMed: 12406822]
11. Curtis JR, Patrick DL, Caldwell ES, et al. Why don't patients and physicians talk about end-of-life care? Barriers to communication for patients with acquired immunodeficiency syndrome and their primary care clinicians. *Arch Intern Med* 2000;160:1690–96. [PubMed: 10847263]
12. Stelter KL, Elliott BA, Bruno CA. Living will completion in older adults. *Arch Intern Med* 1992;152:954–59. [PubMed: 1580721]
13. Jordens C, Little M, Kerridge I, et al. From advance directives to advance care planning: current legal status, ethical rationales and a new research agenda. *Intern Med J* 2005;35:563–66. [PubMed: 16105160]
14. Forrow L. The green eggs and ham phenomena. *Hastings Cent Rep* 1994;24:S29–32. [PubMed: 7860277]
15. Green MJ, Levi BH. Development of an interactive computer program for advance care planning. *Health Expect* 2009;12:60–69. [PubMed: 18823445]
16. Hossler C, Levi BH, Simmons Z, et al. Advance care planning for patients with ALS: feasibility of an interactive computer program. *Amyotrophic Lateral Scler* 2011;12:172–77.
17. Levi BH, Heverley S, Green MJ. Accuracy of a decision aid for advance care planning: simulated end-of-life decision making. *J Clin Ethics* 2011;22:223–38. [PubMed: 22167985]
18. Wilkinson GS. WRAT-3: Wide Range Achievement Test Administration Manual. Lutz, FL: Wide Range, Inc. (Par, Inc.); 1993.
19. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98. [PubMed: 1202204]
20. Beck AT, Steer RA. Internal consistencies of the original and revised Beck Depression Inventory. *J Clin Psychol* 1984;40:1365–67. [PubMed: 6511949]
21. Winter L, Parks S, Diamond J. Ask a different question, get a different answer: why living wills are poor guides to care preferences at the end of life. *J Palliat Med* 2010;13:567–72. [PubMed: 20377499]
22. Navaie-Waliser M, Feldman PH, Gould DA, Levine C, Kuerbis AN, Donelan K. When the caregiver needs care: the plight of vulnerable caregivers. *Am J Public Health* 2002;92:409–13. [PubMed: 11867321]
23. Schubart JR, Levi B,H, Camacho F, et al. Reliability of an interactive computer program for advance care planning. *J Palliat Med* 2012;15:637–42. [PubMed: 22512830]
24. Muthén B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 1984;49:115–32.
25. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–20. [PubMed: 19673146]
26. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons; 2004.
27. Fleiss JL, Cohen J. The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973;33:613–19.
28. Steiger J, Lind J. Statistically-based tests for the number of common factors. Paper presented at: Spring Meeting of the Psychometric Society 1980; Iowa City, IA.
29. Raghunathan T. An approximate test for homogeneity of correlated correlation coefficients. *Qual Quant* 2003;37:99–110.
30. Schmidt FL, Le H, Remus I. Beyond alpha: an empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-difference Constructs. *Psychological Methods* 2003;8:206–24. [PubMed: 12924815]

31. Goethals GR, Reckman RF. The perception of consistency in attitudes. *J Exp Soc Psychol* 1973;9:491–501.
32. Eibach RP, Libby LK, Gilovich TD. When change in the self is mistaken for change in the world. *J Pers Soc Psychol* 2003;84:917–31. [PubMed: 12757138]
33. Loftus EF. Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn Mem* 2005;12:361–66. [PubMed: 16027179]
34. Meyvis T, Ratner RK, Levav J. Why don't we learn to accurately forecast feelings? How misremembering our predictions blinds us to past forecasting errors. *J Exp Psychol Gen* 2010;139:579–89. [PubMed: 20853995]
35. Drivdahl SB, Zaragoza MS, Learned DM. The role of emotional elaboration in the creation of false memories. *Appl Cognitive Psych* 2009;23:13–25.
36. Feldt LS, Brennan RL. Reliability In: Linn RL, ed. *Educational Measurement*. New York City: Macmillan 1989:105–46.

Table 1
General Wish Score and Specific Wish for Treatment scale characteristics at time 1 (T1), T2, T3

Specific wish for treatment		Moderate/severe stroke that would improve within a year	Moderate/severe stroke that would NOT improve	Coma that would resolve within a year	Irreversible coma	Dementia	General wish
Descriptive statistics	Mean (SD) *	T1: 8.79 (3.52) T2: 9.12 (3.32) T3: 9.49 (3.05)	T1: 4.55 (3.89) T2: 3.94 (4.02) T3: 3.39 (3.64)	T1: 8.55 (3.80) T2: 8.49 (3.93) T3: 9.03 (3.39)	T1: 1.94 (2.89) T2: 2.49 (3.78) T3: 2.15 (3.47)	T1: 4.18 (3.75) T2: 3.89 (3.53) T3: 4.33 (3.80)	T1: 3.45 (1.23) T2: 3.34 (1.26) T3: 3.32 (1.27)
Test-retest reliability between T1 and T2	Pearson correlation coefficient (95% CI)	0.78 (0.58,0.88)	0.50 (0.18,0.72)	0.72 (0.49,0.85)	0.62 (0.34,0.79)	0.32 (-0.03,0.59)	0.12 (-0.23,0.45)
Agreement of scale between T1 and T2	Weighted κ (95% CI)	0.77 (0.47,1.00)	0.49 (0.18,0.81)	0.72 (0.45,0.98)	0.59 (0.26,0.92)	0.32 (-0.04,0.68)	0.12 (-1.00,1.00)
Test-retest reliability between T2 and T3	Pearson correlation coefficient (95% CI)	0.83 (0.67,0.91)	0.80 (0.61,0.89)	0.77 (0.57,0.88)	0.76 (0.56,0.87)	0.58 (0.29,0.77)	0.95 (0.89,0.97)
Agreement of scale between T2 and T3	Weighted κ (95% CI)	0.82 (0.62,1.00)	0.78 (0.62,0.95)	0.75 (0.56,0.94)	0.76 (0.49,1.00)	0.57 (0.28,0.87)	0.94 (0.83,1.00)
Internal consistency of scales at each time point	Cronbach α (KR-20)	T1: 0.95 T2: 0.94 T3: 0.95	T1: 0.92 T2: 0.94 T3: 0.92	T1: 0.96 T2: 0.96 T3: 0.94	T1: 0.90 T2: 0.95 T3: 0.95	T1: 0.91 T2: 0.91 T3: 0.92	

* Possible answer choices for Specific Wish Score ranged from 0–11 (0 = No wish for treatment, 11 = Wish for maximum treatment). Possible answer choices for General Wish Score ranged from 1–6 (1 = Wish for all treatments, 6 = Wish for no treatments)