Health Psychology

Exercise Advice by Humans Versus Computers: Maintenance Effects at 18 Months

Abby C. King, Eric B. Hekler, Cynthia M. Castro, Matthew P. Buman, Bess H. Marcus, Robert H. Friedman, and Melissa A. Napolitano

Online First Publication, February 18, 2013. doi: 10.1037/a0030646

CITATION

King, A. C., Hekler, E. B., Castro, C. M., Buman, M. P., Marcus, B. H., Friedman, R. H., & Napolitano, M. A. (2013, February 18). Exercise Advice by Humans Versus Computers: Maintenance Effects at 18 Months. *Health Psychology*. Advance online publication. doi: 10.1037/a0030646

Health Psychology

0278-6133/13/$12.00 DOI: 10.1037/a0030646

BRIEF REPORT

Exercise Advice by Humans Versus Computers: Maintenance Effects at 18 Months

Abby C. King, Eric B. Hekler, Cynthia M. Castro, and Matthew P. Buman
Stanford University School of Medicine

Bess H. Marcus
Brown University Center for Behavioral and Preventive Medicine at the Miriam Hospital

Robert H. Friedman
Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts

Melissa A. Napolitano
Brown University Center for Behavioral and Preventive Medicine at the Miriam Hospital

Objective: An automated telehealth counseling system, aimed at inactive midlife and older adults, was shown previously to achieve 12-month physical activity levels similar to those attained by human advisors. This investigation evaluated the sustained 18-month impacts of the automated advisor compared with human advisors. ***Methods:*** Following the end of the 12-month randomized, controlled trial, participants who had been randomized to either the human advisor ($n = 73$) or automated advisor ($n = 75$) arms were followed for an additional 6 months. During that period, human or automated advisor-initiated telephone contacts ceased and participants were encouraged to initiate contact with their advisor as deemed relevant. The primary outcome was moderate-to-vigorous physical activity (MVPA), measured using the Stanford Physical Activity Recall and validated during the major trial via accelerometry. ***Results:*** The two arms did not differ significantly in 18-month MVPA or the percentage of participants meeting national physical activity guidelines (ps $>.50$). No significant within-arm MVPA differences emerged between 12 and 18 months. Evaluation of the trajectory of physical activity change across the 18-month study period indicated that, for both arms, the greatest physical activity increases occurred during the first 6 months of intervention, followed by a relatively steady amount of physical activity across the remaining months. ***Conclusions:*** The results provide evidence that an automated telehealth advice system can maintain physical activity increases at a level similar to that achieved by human advisors through 18 months. Given the accelerated use of mobile phones in developing countries, as well as industrialized nations, automated telehealth systems merit further evaluation.

Keywords: physical activity, computer, automated, intervention, maintenance

Automated telehealth interventions represent a potentially convenient, low-cost platform for facilitating ongoing engagement in key health-promoting behaviors, such as physical activity, yet rarely have been compared directly with human-delivered interventions (Piette, 2000). The comparative effects of telephone-delivered automated versus human advice on *sustained* health behavior change (i.e., behavioral maintenance) when formal advisor-initiated contact ceases also are unknown. Behavioral maintenance represents a key challenge facing all major health behavior fields, including physical activity (Castro, King, & Brassington, 2001).

The Community Health Advice by Telephone (CHAT) randomized, controlled trial evaluated the 12-month effectiveness of telephone-based, theory-informed physical activity guidance and support proactively delivered via trained health educators or an automated telehealth system (King et al., 2007). The results of that

Abby C. King, Eric B. Hekler, Cynthia M. Castro, and Matthew P. Buman, Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine; Bess H. Marcus and Melissa A. Napolitano, Brown University Center for Behavioral and Preventive Medicine at the Miriam Hospital; Robert H. Friedman, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts.

Boston Medical Center (BMC) Inc. has a copyright for the computer-based technology (Telephone-Linked Care [TLC]) that was used in the automated advisor intervention program described. In 1992, BMC gave InfoMedics Inc. the commercial rights in the TLC technology. Dr. Friedman has stock ownership and a consultancy agreement with InfoMedics, and is a member of its Board of Directors.

Correspondence concerning this article should be addressed to Abby C. King, Stanford University School of Medicine, 259 Campus Drive, HRP Redwood Building, Room T-221, Stanford, CA 94305-5405. E-mail: king@stanford.edu

RCT indicated that automated telephone advice was an effective alternative to trained counselors for delivering physical activity information and support to initially inactive midlife and older adults (King et al., 2007). The aim of the current investigation was to evaluate the sustained maintenance effects of the two interventions across a subsequent 6-month period when the more resource-intensive human or automated advisor-initiated telephone contacts were no longer in effect. Several smaller-scale studies have evaluated initial physical activity adoption rates using an automated telehealth system into which patients were required to initiate calls, with mixed results (Jarvis, Friedman, Heeren, & Cullinane, 1997; Pinto et al., 2002). In contrast, no study has systematically evaluated the effects of such a patient-initiated call-in system, relative to calling into a human advisor, on *maintaining* already established physical activity levels. Given previous evidence from similar telephone-assisted physical activity interventions suggesting that continued telephone contact by human advisors beyond 12 months could in some cases undermine subsequent physical activity participation (Castro et al., 2001), we predicted that physical activity maintenance rates would be as good as, and possibly better, in the automated advisor relative to human advisor arms.

Method

Participants and Experimental Design

The methods and major results for the CHAT trial have been described in detail elsewhere (King et al., 2007) and will be described briefly here. Study eligibility criteria included ages 55 years and older, not initially engaged in more than 60 min per week of health-enhancing moderate-intensity or more vigorous physical activity (MVPA) over the previous 6 months, free of any medical condition that would limit participation in moderate-intensity exercise, and able to speak and understand English sufficiently to provide informed consent and participate in study intervention and assessment procedures. Eligible individuals were randomized to a 12-month home-based moderate-intensity physical activity (primarily walking) adoption program delivered via a trained telephone counselor (human advice arm, $n = 73$); a similar program delivered via an automated, computer-controlled interactive telephone system (automated advice arm, $n = 75$); or a general health education control arm ($n = 70$). Participants randomized to the two physical activity intervention arms received one 30-min in-person, one-on-one instructional session followed by a similar number of advisor-initiated telephone contacts across the 12-month period (human arm mean $\pm$ $SD = 13.1 \pm 2.5$; automated arm $= 11.8 \pm 4.8$), which emphasized similar cognitive and behavioral skills and was tailored to each participant's current stage of motivational readiness to change (King et al., 2007). Telephone contacts in both arms were supplemented with informational mailings and use of a Yamax Digi-walker pedometer to provide individualized activity feedback to the participant (AccuSplit, San Jose, CA).

After the 12-month assessment, human or automated advisor-initiated contacts in the two intervention arms ceased, and participants were encouraged to contact their assigned advisor by telephone for continued advice and support during the subsequent 6-month period (labeled the Maintenance period). Persons randomized to the general health education control arm were offered weekly health education classes during the initial 12-month period that focused on a variety of nonphysical activity topics of interest to this age group (King et al., 2007). After the 12-month assessment, their formal study participation ended and they will not be discussed here.

Measurement

Measurement occurred at a Stanford outpatient research facility by trained study staff blinded to participant study assignment. The Stanford 7-day physical activity recall (PAR) (Blair et al., 1985) served as the primary measure of minutes per week engaged in moderate-intensity or MVPA, the form of physical activity constituting the major national public health guideline in the field (Physical Activity Guidelines Advisory Committee, 2008). Physical activity levels reported on the PAR were validated at study midpoint via accelerometry (Actigraph #7164) that recorded body movement continuously for up to 7 days (King et al., 2007). The PAR was collected at baseline and every 6 months across the 18-month period.

Statistical Analyses

The primary objective of the current investigation was to evaluate what happened to MVPA levels in the two intervention arms during the *6-month maintenance period* when participants were no longer receiving human or automated advisor-initiated telephone contacts. To determine this, intervention participants who had either 12-month PAR data ($n = 119$) or other physical activity data suitable for imputation at 12 months (e.g., from the additionally collected Community Healthy Activities Model Program for Seniors [CHAMPS] questionnaire; $n = 8$; King et al., 2007) were evaluated at 18 months (total, $n = 127$; human advice, $n = 66$; automated advice, $n = 61$; between-arm difference $p > .10$). This number represented 86% of the original sample randomized to the two experimental arms ($n = 148$) and is commensurate with sample sizes per arm found in other studies in the field (Goode, Reeves, & Eakin, 2012). To provide continuity with the major trial results, *intent-to-treat* (ITT) principles were applied in the 18-month follow-up analyses (i.e., if the 18-month PAR value was missing and the other relevant 12-month data were not available, the participant's baseline PAR value—a conservative approach assuming that physical activity returned to preintervention levels—was used as per the original trial). This ITT approach was compared with an imputation approach as part of the growth model analysis described below. We sought to answer three questions. The first question evaluated whether MVPA levels attained in each arm at 12 months were adequately maintained at 18 months. To address this question, 12- to 18-month MVPA levels were first compared in each arm separately by using paired-comparison t tests. In the event that neither t test was significant (suggesting no significant change in physical activity levels from 12 to 18 months in either arm), we combined the two arms to allow for increased power to detect differences between the 12- and 18-month time-points.

The second question explored whether the two interventions differed significantly in 18-month minutes per week of MVPA. It was assessed via analysis of covariance (ANCOVA), with 12-month MVPA serving as the covariate. Between-arm differences in proportions of participants meeting the national physical activity recommendations of at least 150 min per week of MVPA at 18 months (Physical Activity Guidelines Advisory Committee, 2008) were explored using χ^2 analysis. The current model was powered at $\beta = 0.80$ to detect medium or greater effect size differences (i.e., $f = 0.28$) based on the ANCOVA model characteristics (i.e., $\alpha = .05$, $N = 127$, Num., $df = 2$, numbers of arms = 2, covariates = 1).

The third question addressed overall patterns of change in physical activity (minutes per week in MVPA) across the entire 18-month study period for the two interventions using a growth model approach (Singer & Willett, 2003). We tested multiple growth trajectories (linear, quadratic, and cubic) and group effects to identify the best model fit. We relied upon multiple imputation methods as specified for growth model analyses (Singer & Willett, 2003).

Results

Subjects

The mean age of the sample was 60 ± 5.5 years, with education averaging 16 ± 1.9 years. Mean body mass index (BMI) and number of medications were 29.5 ± 5.4 and 2 ± 1.4, respectively. Relative to intervention subjects who had 18-month PAR data ($n = 115$, or 78% of the originally enrolled intervention sample of 148), intervention subjects without 18-month PAR data ($n = 33$; 18 in human arm, 15 in automated arm, between-arm difference $p > .10$) were significantly younger ($M_{w18m} = 61.5 \pm 6.0$ vs. $M_{wo18m} = 58.2 \pm 4.9$ years), more affluent in relation to annual household income categories ($M_{w18m} = 7.4 \pm 2.2$ vs. $M_{wo18m} = 8.1 \pm 1.6$), less active (MVPA $M_{w18m} = 93.51 \pm 136.0$ vs. $M_{wo18m} = 44.7 \pm 63.9$ min/week), and weighed more at baseline (BMI $M_{w18m} = 29.1 \pm 5.0$ vs. $M_{wo18m} = 31.2 \pm 5.3$; $t = 2.28$; ps $<.04$). Those with and without 18-month data did not differ in years of education at baseline ($M_{w18m} = 16.2 \pm 1.8$ vs. $M_{wo18m} = 16.5 \pm 1.8$, $t = 0.57$; $p > .56$). Relative to those with 18-month PAR data, subjects without 18-month PAR data were less active ($M_{w18m} = 181.0 \pm 140.1$ vs. $M_{wo18m} = 77.2 \pm 77.0$, $t = -4.4$, $p < .001$) and weighed more at 12 months ($M_{w18m} = 28.7 \pm 4.9$ vs. $M_{wo18m} = 31.4 \pm 5.5$, $t = 2.76$, $p < .01$) than those with 18-month PAR data. In contrast, there were no significant between-arm differences in PAR or BMI values at 12 months ($t = 0.92$; ps $>.35$). There were no significant differences detected in the above variables for the human versus automated arms in relation to participants with and without 18-month PAR data, but statistical power may not have been sufficient to detect such differences. The 78% retention rate has been identified as adequate in this field (Fjeldsoe, Neuhaus, Winkler, & Eakin, 2011).

Significantly fewer individuals in the human arm never called in during the Maintenance phase (31.9% [$n = 23$] never called in) compared with the automated arm (58.1% [$n = 43$] never called in; independent-sample t test $= -3.4$, $p < .001$).

Were There Intervention Arm Differences in Physical Activity Levels at 18 Months?

The 18-month between-arm ANCOVA controlling for 12-month PAR values was not statistically significant ($F = 0.68$, $df = 126$, $p = .41$, $M_{auto} = 145.2 \pm 134.5$, $M_{hum} = 167.0 \pm 135.6$). χ^2 analysis also indicated no significant differences between the automated and human advice arms in meeting physical activity guidelines at 18 months (based on the 127 participants with 12-month PAR data who constitute the target sample for this maintenance paper: 44.3% [$n = 27/61$] in the Automated arm vs. 50.0% [$n = 33/66$] in the Human arm, $\chi^2 = 0.42$, $p = .52$).

Were There Differences in Physical Activity Levels Between 12 and 18 Months?

There were no significant differences when 12- and 18-month MVPA were compared either within each arm (ps $>.50$) or when the two arms were combined (12-month combined MVPA = 167.9 ± 138.0 min/week vs. 18-month MVPA = 156.6 ± 135.0, $t = 0.91$, $df = 126$, $p = .36$).

What Was the Overall Trajectory of MVPA Change Over the 18-Month Time Course?

The results from growth model testing indicated that including differences between the arms did not improve model fit and were therefore dropped from the final model. Based on improved model fit statistics (i.e., −2 Log Likelihood, Akaike Information Criterion [AIC], and Bayesian Information Criterion [BIC]), we found that a cubic effect for time significantly improved model fit over linear and quadratic time specifications, with 15% of the variance in within-person MVPA explained by the cubic model. Figure 1 is a visual representation of the change plot over time and suggests, as per a cubic effect, marked increases in MVPA from baseline (0 months) to 6 months, followed by a relative "flattening" (i.e., steady state) of MVPA between 6, 12, and 18 months. This pattern was found regardless of imputation method (including ITT). To test for significant differences between time-points, post hoc analyses were conducted where time was treated categorically using multiple comparisons, that is, 0 versus 6 months, 0 versus 12 months, 0 versus 18 months, and so forth. Results indicated that all time points after baseline were significantly different from baseline values (ps $<.0001$) but were not significantly different from each other (ps $>.64$).

Discussion

The results indicate that, once established via a 12-month telephone advisor-initiated counseling program, the increases in regular physical activity were successfully maintained among initially inactive midlife and older adults during a subsequent 6-month maintenance period irrespective of whether the telephone advisor was a trained health educator or an automated telephone counseling system. Participants in both arms (but particularly the automated advice arm) made few if any calls to their physical activity advisor (mode = 0 for both arms) during the 6-month maintenance period. In contrast, a significant proportion of participants in both arms (45%–50%) met the recommended minutes per week of MVPA at 18 months, suggesting that, while automated advice may be somewhat less engaging over time than human advice, no untoward effects on longer-term maintenance were indicated. While no between-groups differences were detected at 18 months, we may have lacked sufficient power to detect smaller effect sizes.

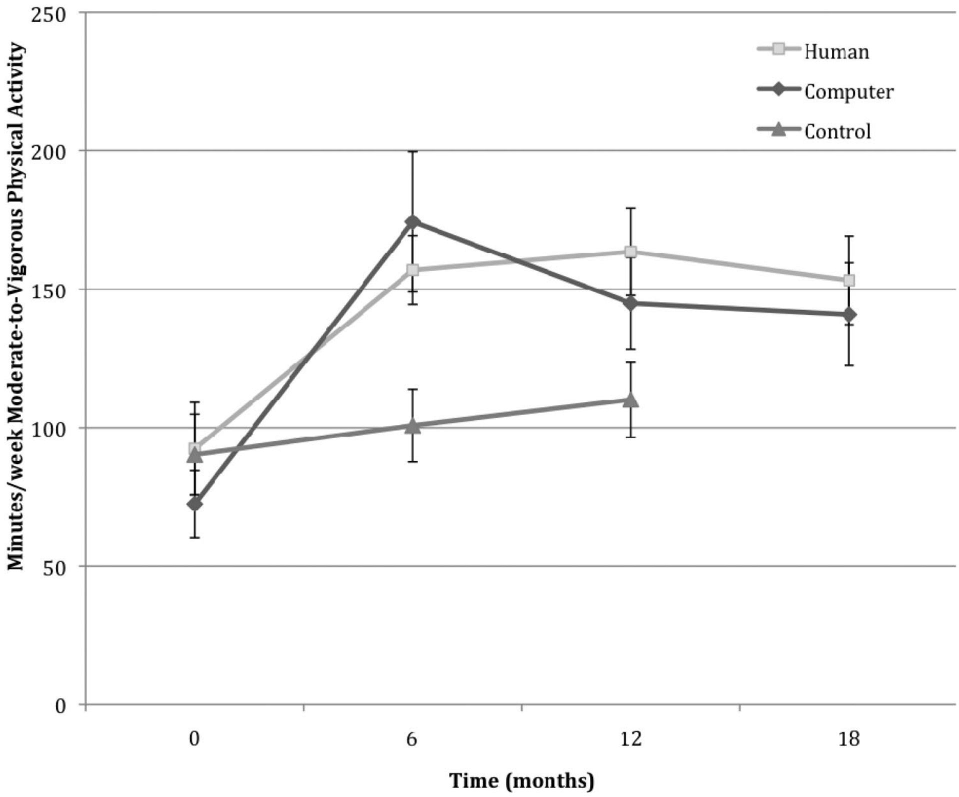

Figure 1. Plot of moderate to vigorous physical activity (MVPA) across 18 months. Control arm ended study participation at end of 12-month randomized, controlled trial period; control arm data are not part of analyses of intervention-specific change over time, but included to provide a reference for the intervention arms' change trajectories. Bars represent *SE*s.

Use of intent-to-treat methods minimized retention bias, although the lack of 18-month data from subgroups that have been frequently recognized as having more difficulty adhering (i.e., those who were initially heavier and less physically active) indicates the importance of continued attention to those subgroups. Because of budgetary constraints, we were unable to continue the control arm past the 12-month major trial period. While we do not know with certainty how the health education control arm would have performed during the 6-month follow-up period when investigator-initiated attention, similar to the two intervention arms, would have been withdrawn, evidence from a recent review that included studies with nonexercise comparison arms indicated minimal physical activity changes in such control arms during the maintenance period (Fjeldsoe et al., 2011). In addition, the physical activity outcome was self-reported. While we were able to validate self-report with objective accelerometry measures during the main 12-month trial, budgetary constraints precluded further accelerometry assessment during the follow-up period.

The results indicate the importance of the first 6 months of physical activity programming in establishing sustained physical activity change across the next year, and add to the small but growing literature supporting the use of interactive communication technologies in promoting ongoing physical activity. The potential population reach, continuing availability, and cost containment value offered by such automated technologies make them attractive intervention tools for managing ongoing changes in behaviors such as physical activity that need to be performed regularly.

References

Blair, S. N., Haskell, W. L., Ho, P., Paffenbarger, Jr., R. S., Vranizan, K. M., Farquhar, J. W., & Wood, P. D. (1985). Assessment of habitual physical activity by seven-day recall in a community survey and controlled experiments. *American Journal of Epidemiology, 122,* 794–804.

Castro, C. M., King, A. C., & Brassington, G. S. (2001). Telephone versus mail interventions for maintenance of physical activity in older adults. *Health Psychology, 20,* 438–444. doi:10.1037/0278-6133.20.6.438

Fjeldsoe, B., Neuhaus, M., Winkler, E., & Eakin, E. (2011). Systematic review of maintenance of behavior change following physical activity and dietary interventions. *Health Psychology, 30,* 99–109. doi:2011-02060-010[pii]10.1037/a0021974

Goode, A. D., Reeves, M. M., & Eakin, E. G. (2012). Telephone-delivered interventions for physical activity and dietary behavior change: An updated systematic review. *American Journal of Preventive Medicine, 42,* 81–88. doi:S0749-3797(11)00742-2[pii]10.1016/j.amepre.2011.08.025

Jarvis, K. L., Friedman, R. H., Heeren, T., & Cullinane, P. M. (1997). Older women and physical activity: Using the telephone to walk. *Women's Health Issues, 7,* 24–29. doi:10.1016/S1049-3867(96)00050-3

King, A. C., Friedman, R. M., Marcus, B. H., Castro, C., Napolitano, M., Ahn, D., & Baker, L. (2007). Ongoing physical activity advice by humans versus computers: The Community Health Advice by Telephone (CHAT) Trial. *Health Psychology, 26,* 718–727. doi:10.1037/0278-6133.26.6.718

Physical Activity Guidelines Advisory Committee. (2008). *Report of the Physical Activity Guidelines Advisory Committee, 2008*. Washington, DC: U. S. Department of Health and Human Services.

Piette, J. D. (2000). Interactive resources for patient education and support. *Diabetes Spectrum, 13,* 110–112.

Pinto, B. M., Friedman, R., Marcus, B. H., Kelley, H., Tennstedt, S., & Gillman, M. W. (2002). Effects of a computer-based, telephone-counseling system on physical activity. *American Journal of Preventive Medicine, 23,* 113–120. doi:10.1016/S0749-3797(02)00441-5

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Received November 18, 2011
Revision received June 8, 2012
Accepted July 3, 2012 ■