Whole-Body Hyperthermia for the Treatment of Major Depressive Disorder
A Randomized Clinical Trial

Clemens W. Janssen, PhD; Christopher A. Lowry, PhD; Matthias R. Mehl, PhD; John J. B. Allen, PhD; Kimberly L. Kelly, MPA; Danielle E. Gartner, BA; Angelica Medrano, BA; Tommy K. Begay, PhD; Kelly Rentscher, MA; Joshua J. White, BS; Andrew Fridman, BS; Levi J. Roberts, BA; Megan L. Robbins, PhD; Kay-u Hanusch, MSc; Steven P. Cole, PhD; Charles L. Raison, MD

IMPORTANCE Limitations of current antidepressants highlight the need to identify novel treatments for major depressive disorder. A prior open trial found that a single session of whole-body hyperthermia (WBH) reduced depressive symptoms; however, the lack of a placebo control raises the possibility that the observed antidepressant effects resulted not from hyperthermia per se, but from nonspecific aspects of the intervention.

OBJECTIVE To test whether WBH has specific antidepressant effects when compared with a sham condition and to evaluate the persistence of the antidepressant effects of a single treatment.

DESIGN, SETTING, AND PARTICIPANTS A 6-week, randomized, double-blind study conducted between February 2013 and May 2015 at a university-based medical center comparing WBH with a sham condition. All research staff conducting screening and outcome procedures were blinded to randomization status. Of 338 individuals screened, 34 were randomized, 30 received a study intervention, and 29 provided at least 1 postintervention assessment and were included in a modified intent-to-treat efficacy analysis. Participants were medically healthy, aged 18 to 65 years, met criteria for major depressive disorder, were free of psychotropic medication use, and had a baseline 17-item Hamilton Depression Rating Scale score of 16 or greater.

INTERVENTIONS A single session of active WBH vs a sham condition matched for length of WBH that mimicked all aspects of WBH except intense heat.

MAIN OUTCOMES AND MEASURES Between-group differences in postintervention Hamilton Depression Rating Scale scores.

RESULTS The mean (SD) age was 36.7 (15.2) years in the WBH group and 41.47 (12.54) years in the sham group. Immediately following the intervention, 10 participants (71.4%) randomized to sham treatment believed they had received WBH compared with 15 (93.8%) randomized to WBH. When compared with the sham group, the active WBH group showed significantly reduced Hamilton Depression Rating Scale scores across the 6-week postintervention study period (WBH vs sham; week 1: −6.53, 95% CI, −9.90 to −3.16, $P < .001$; week 2: −6.35, 95% CI, −9.95 to −2.74, $P = .001$; week 4: −4.50, 95% CI, −8.17 to −0.84, $P = .02$; and week 6: −4.27, 95% CI, −7.94 to −0.61, $P = .02$). These outcomes remained significant after evaluating potential moderating effects of between-group differences in baseline expectancy scores. Adverse events in both groups were generally mild.

CONCLUSIONS AND RELEVANCE Whole-body hyperthermia holds promise as a safe, rapid-acting, antidepressant modality with a prolonged therapeutic benefit.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01625546

Warm-sensitive thermosensory pathways projecting from the skin (and other epithelial linings) to specific subcortical and cortical regions may affect neural activity and behavior in ways relevant to the treatment of major depressive disorder (MDD).[1] For example, in humans, exposure to cutaneous heating (41°C) activates the midorbitofrontal cortex, the pregenual anterior cingulate cortex, and the ventral striatum, with the degree of activation being associated with subjective pleasantness ratings made in response to the warm temperature.[2] Importantly, these and other brain regions most implicated in registering—and reacting to—pleasant thermal signals show decreased activity in patients with MDD.[3] Moreover, patients with MDD display abnormalities in thermoregulation characterized by increased core body temperature and reduced ability to sweat, both of which have been reported to normalize following successful treatment and both of which would be expected consequences of reduced activity in warm-sensitive thermosensory pathways in the periphery, brain, or both.[1]

Based on these considerations, we conducted animal studies demonstrating that whole-body heating activated subdivisions of the dorsal raphe nucleus implicated in mood regulation and antidepressantlike responses, while not activating other dorsal raphe subregions, including those implicated in the facilitation of anxiety states.[4,5] We also found that whole-body heating produced an acute antidepressantlike response (unpublished data; M. W. Hale, J. L. Lukkes, K. F. Dady, K. J. Kelly, E. D. Paul, C.L.R., C.A.L.; October 2010). To conduct a preliminary examination of the relevance of these preclinical findings, we previously conducted a small open trial of whole-body hyperthermia (WBH) in humans.[6] In 16 medically healthy adults with MDD, we found that a single session of WBH was significantly associated with a reduction in depressive symptoms when measured 5 days after treatment.

While intriguing, these findings provided no way of separating direct biological effects of hyperthermia from the many nonspecific and placebo effects that attend any conceptually attractive and invasive procedure such as WBH. In addition, after the day-5 posttreatment assessment, participants in the open trial received other treatments, making it impossible to assess how long the benefits of a single WBH treatment might persist.

To address these issues, the current study used a randomized, double-blind design to compare WBH with a sham procedure that matched all aspects of the active treatment except the intense heat. In addition, the study used a 6-week follow-up period during which participants received no other antidepressant treatment. We hypothesized that a single session of WBH would reduce depressive symptoms 1 week after treatment when compared with sham, and we sought to characterize whether any observed improvements in depressive symptoms would persist across 6 weeks of follow-up. In addition, we sought to better understand the development and time course of adverse events in response to WBH.

Methods

The University of Arizona institutional review board approved the study. Signed informed consent was obtained from all participants after a full description of study procedures and risks and potential benefits was provided and prior to conducting any study procedures. The full study protocol can be found in Supplement 1.

Participants

This study enrolled participants at the Banner University Medical Center in Tucson, Arizona, between February 2013 and May 2015. Participants were recruited via print, radio, posted fliers, email listserv, social media, and television advertising. Eligible participants were men and women, aged 18 to 65 years, who were medically healthy and had MDD for at least 4 weeks prior to signing consent per *DSM-IV-TR* criteria. The study initially required a 17-item Hamilton Depression Rating Scale (HDRS) score of 18 or greater for enrollment, but this cutoff was lowered to 16 or greater as a result of many otherwise eligible individuals presenting with HDRS scores of 16 and 17 at screening. Of the 34 randomized participants, 11 were enrolled with a screening score of 18 or greater and 23 were enrolled with a screening score of 16 or greater. A full listing of study inclusion and exclusion criteria is provided in eTable 1 in Supplement 2. At screening, all participants underwent routine hematological and biochemical laboratory testing, urine toxicology, and pregnancy testing (in premenopausal women only) and received an electrocardiogram.

Randomization and Blinding

Participants were randomized on an equal (ie, 1-to-1) basis in blocks of 6 to a single treatment of WBH or sham based on a computer-generated randomization list that was provided to the study by the Arizona Statistics Consulting Laboratory. This list was kept by a Psychiatry Department administrator who had no contact with any study participants. Participants remained blinded to their randomization status until completion of the last study assessment at posttreatment week 6. A full description of study blinding procedures is provided in the eAppendix in Supplement 2.

Study Design

Participants who signed consent and met eligibility requirements were scheduled to receive an intervention within 25 days

of completing screening. Between screening and baseline assessment, participants completed the Inventory of Depressive Symptomatology–Self-report (IDS-SR) at home (mean [SD], 8.28 [4.17] days after screening). Participants showing a 30% or greater reduction from their IDS-SR score at screening were considered likely placebo responders and were discontinued from the study. On the intervention day, participants arrived at the medical center at 8 AM and completed a baseline assessment comprising questionnaires that assessed all primary and secondary study outcomes. Following this, they rested until commencing the study intervention between noon and 1 PM. On completion of the intervention, participants rested for 1 hour and were released to home. Follow-up assessments were conducted at postintervention days 1, 2, and 3, and weeks 1, 2, 4, and 6. Need for initiation of an antidepressant or psychotherapy during the 6-week follow-up period resulted in study discontinuation.

Study Interventions

For both WBH and sham, the current study used a Heckel HT3000 WBH system (Heckel Medizintechnik GmbH and Hydrosun Medizintechnik GmbH). Sensors continuously monitored core and skin temperatures and heart rate throughout the procedures. See the eFigure in Supplement 2 for a photograph of the Heckel device.

Based on positive results from our prior open trial, we used mild-intensity hyperthermia in the active condition. Participants randomized to active WBH received heating at the level of the chest by infrared lights and at the level of the lower extremities by infrared heating coils until their core body temperature reached 38.5°C, which is the upper limit temperature for mild-intensity WBH.[7] Time to attainment of this core body temperature varied from patient to patient but required a mean (SD) length of 107 (19.4) minutes (range, 81-140 minutes). When core body temperature reached 38.5°C, the infrared lights and heating coils were turned off, and participants remained recumbent in the Heckel device and entered a 60-minute cool-down phase.

All procedures for the sham condition were identical to WBH, except that orange-colored nonheating lights and a false fan were used to produce a similar color and noise as the infrared lights but provide no heat. To increase believability, mild heat was provided within the Heckel device by activating the heating coils situated above participants' lower extremities at the same setting used for active WBH, while keeping the primary infrared lights off. For each participant randomized to the sham WBH condition, time in the Heckel device in the mild-heating phase was matched to the time the prior participant of the same sex undergoing actual WBH spent in the active heating phase. As with WBH, the cool-down phase in the sham condition was 60 minutes. The mean (SD) maximal core body temperature achieved during sham treatment was 37.69°C (0.32), which was significantly lower than the mean (SD) maximal core temperature achieved during WBH (38.85°C [0.45]; $P < .001$). Similarly, mean (SD) core body temperature increased less during sham than during WBH (sham: 0.78°C [0.36] vs WBH: 1.91°C [0.49]; $P < .001$). The mean (SD) maximal skin temperature achieved during sham treatment was

39.79°C (1.32), which was lower than the maximum skin temperature achieved during WBH (40.74 [0.85]; $P = .03$). For a photograph of the Heckel HT3000 delivering hyperthermia, see the eFigure in Supplement 2.
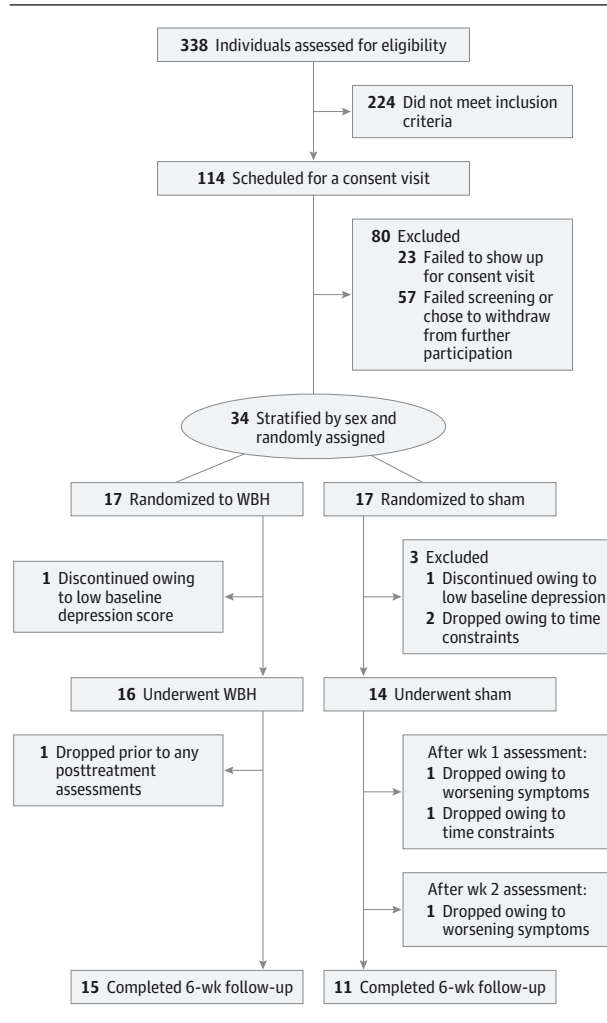
Outcome Measures

The study's a priori primary outcome measure was reduction in depression severity across the 6-week study period as assessed by the 17-item HDRS at 1, 2, 4, and 6 weeks following exposure to either WBH or sham treatment. Trained raters blind to group assignment performed all HDRS assessments. Training to establish interrater reliability was overseen by the principal investigator (C.L.R.) and was conducted according to a standard procedure for the HDRS.[8] Five raters conducted HDRS assessments for the study. Interrater reliability was assessed using a 2-way mixed, consistency, average-measures intraclass correlation coefficient to assess the degree that coders provided consistency in their ratings of HDRS scores.[9] The resulting intraclass correlation coefficient was in the excellent range (intraclass correlation coefficient = 0.985),[10] indicating that coders had a high degree of agreement and suggesting that HDRS scores were rated similarly across coders.

Secondary outcome measures included IDS-SR scores at posttreatment days 1, 2, and 3, and weeks 1, 2, 4, and 6, as well as Sheehan Disability Scale and Quality of Life Enjoyment Satisfaction Scale–short-form scores at posttreatment weeks 1, 2, 4, and 6. Adverse events were assessed immediately after study interventions and at postintervention weeks 1, 2, 4, and 6 with the Sequenced Treatment Alternatives to Relieve Depression Patient Rated Inventory of Side Effects (PRISE) questionnaire. At baseline, the Credibility/Expectation Questionnaire (CEQ) and the Massachusetts General Hospital Antidepressant Treatment History Questionnaire were administered.[11,12] We also assessed length of the current depressive episode and number of past episodes. In addition, to assess the believability of the sham condition, immediately following the study intervention, participants were asked to guess whether they had received the active or sham treatment.

Statistical Analysis

Frequency distributions, means, and SDs were calculated for the primary and secondary outcome measures for all waves of data collection. Distributions were examined for outliers and for significant deviations from normality. The primary study hypothesis was tested with a hierarchical linear model, with an autoregressive covariance structure using 0, 7, 14, 28, and 42 days' measurement of HDRS with a linear model on ln $(t+1)$, where t is time from treatment. The mixed-effect model provides unbiased estimates assuming data are missing at random conditional on information in the model. A similar hierarchical linear model approach was used to evaluate potential between-group differences in adverse events. Cohen d was calculated to assess effect sizes for between-group differences at all posttreatment points based on means and SDs derived from the mixed-effect model. Because baseline expectancy scores differed between groups, possible moderation of

Figure 1. CONSORT Flow Diagram for Whole-Body Hyperthermia (WBH) and Sham Groups

Table 1. Demographics and Baseline Clinical Characteristics

Characteristic	Mean (SD)	
	WBH (n = 17)	Sham (n = 17)
Women, No. (%)	12 (71)	11 (65)
Age, y	36.71 (15.20)	41.47 (12.54)
Range	18-65	24-61
White, non-Hispanic, No. (%)	9 (53)	10 (59)
Clinical characteristics		
HDRS		
Screening	22.18 (4.53)	23.71 (3.80)
Baseline	20.71 (4.87)	22.76 (4.42)
Treatment (CEQ)		
Credibility	0.89 (2.72)	−1.01 (2.50)
Expectancy	1.02 (2.68)[a]	−1.26 (1.91)
Length of current episode, mo	126.07 (163.82)	100.50 (143.67)
Prior depressive episodes	0.87 (1.06)	0.50 (0.76)
Participants with prior antidepressant use, No. (%)	9 (53)	8 (47)
Past antidepressant exposures	0.94 (1.14)	0.76 (1.09)
Hypnotic medication use, No. (%)	3 (18)	0 (0)

Abbreviations: CEQ, Credibility/Expectancy Questionnaire; HDRS, 17-item Hamilton Depression Rating Scale; WBH, whole-body hyperthermia.

[a] Denotes difference between WBH and sham at $P < .05$.

treatment effect by baseline expectations was done by obtaining estimated marginal HDRS means for each participant from the primary hierarchical linear model analysis. A dichotomous expectancy variable was created by categorizing participants with expectancy scores less than the median score of zero as low expectancy and participants with greater than or equal to zero as high expectancy. Treatment condition, expectancy group, and their interaction were entered into an analysis of variance. A statistically significant interaction would indicate that baseline expectancy was a moderator. All tests were 2-tailed, with significance set at $P < .05$. Analyses were conducted with SPSS version 22 for Windows (IBM).

Results

Study Participants

Figure 1 shows the disposition of the 338 individuals screened for study participation. Thirty-four of those screened met inclusion/exclusion criteria and were randomized to receive a study intervention. Thirty received a study intervention

(16 active WBH and 14 sham). One participant in the active WBH group elected to discontinue the study prior to completion of any postintervention assessments, 2 individuals randomized to sham discontinued the study between the postintervention week 1 and week 2 assessments, and 1 discontinued following the week 2 assessment. The treatment groups were well matched at baseline on a range of demographic and clinical measures, as shown in **Table 1**. However, CEQ expectancy scores were significantly higher in the group that subsequently received active WBH than in the group randomized to sham (mean [SD] expectancy score, 1.02 [2.68] vs −1.26 [1.91], respectively; mean [SD] credibility score, 0.89 [2.72] vs −1.01 [2.50], respectively).

Outcome Measures

Supporting the credibility of our sham condition, 10 of 14 participants (71.4%) randomized to sham believed they had received active hyperthermia immediately on completion of the procedure (compared with 15 of 16 [93.8%] receiving active WBH). **Table 2** provides scores for the primary study end point (HDRS scores at postintervention weeks 1, 2, 4, and 6); eTable 2 in Supplement 2 provides scores for relevant secondary outcome measures (IDS-SR, Sheehan Disability Scale, and Quality of Life Enjoyment Satisfaction Scale–short-form). As shown in **Figure 2**, when compared with the sham group, the active WBH group showed significantly reduced HDRS scores across the 6-week postintervention study period (WBH vs sham; week 1: −6.53, 95% CI, −9.90 to −3.16, $P < .001$, $d = 2.23$; week 2: −6.35, 95% CI, −9.95 to −2.74, $P = .001$, $d = 2.11$; week 4: −4.50, 95% CI, −8.17 to −0.84, $P = .02$, $d = 1.66$; and week 6: −4.27, 95% CI, −7.94 to −0.61, $P = .02$, $d = 1.66$).

Table 2. 17-Item Hamilton Depression Rating Scale Scores at Postintervention Assessments in WBH vs Sham Groups

	Participants per Group, No.		Score, Mean (SD)	
Week	WBH	Sham	WBH	Sham
Baseline	16	14	20.71 (4.87)	22.75 (4.42)
1	15	14	14.80 (5.40)	20.86 (3.33)
2	15	12	12.67 (6.78)	18.71 (3.17)
4	15	11	12.93 (4.92)	17.79 (4.06)
6	15	11	12.40 (5.45)	17.21 (4.78)

Abbreviation: WBH, whole-body hyperthermia.

Figure 2. Effect of Whole-Body Hyperthermia (WBH) vs Sham Treatment on 17-Item Hamilton Depression Rating Scale (HDRS) Across the 6-Week Postintervention Period

No. at risk

WBH	16	15	15	15	15	15	15
Sham	14	14	12	11	11	11	11
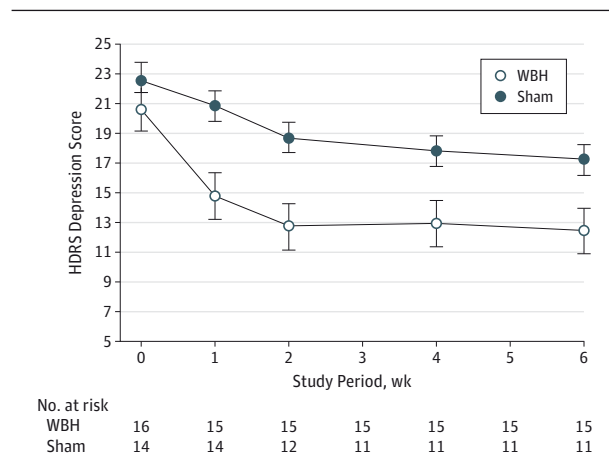
The error bars indicate SEs. Means and SEs were derived from the raw data provided in Table 2.

Cognizant of recent concerns regarding the potential of adjustment for unplanned covariates to produce false findings,[13] covariates were not entered into our primary analysis. However, because baseline expectancy scores differed between groups (mean [95% CI], 1.02 [−0.41 to 2.45] for WBH vs −1.26 [−2.42 to −0.10] for sham; *P* = .02), we conducted a moderator analysis in the 30 participants with CEQ scores. This analysis did not find that CEQ expectancy scores significantly moderated between-group differences in HDRS score (mean [95% CI], WBH/low expectancy, 15.25 [13.31 to 17.19]; WBH/high expectancy, 14.44 [12.94 to 15.95]; sham/low expectancy, 18.25 [16.66 to 19.83]; sham/high expectancy, 20.91 [18.52 to 23.29]; *P* = .07), which remained significant in the moderator analysis (mean [95% CI], WBH, 14.85 [13.62 to 16.08]; sham, 19.58 [18.15 to 21.01]; *P* < .001).

Safety of Study Interventions

A full listing of PRISE-assessed adverse events is provided in eTable 3 in Supplement 2. No significant difference in overall adverse events was observed between treatment groups across the postintervention study period. The most common adverse effects immediately following both study interventions were headache, fatigue, and dry mouth, with no statistical difference between groups. Numerically, participants who received WBH reported more sweating and nausea.

Discussion

To our knowledge, this is the first randomized, double-blind, sham-controlled study of WBH for the treatment of MDD. Consistent with results from a prior small open trial,[6] the current study found that WBH was associated with a substantial reduction in depressive symptoms that was apparent within 1 week of treatment. Moreover, the use of a credible sham condition increases confidence that the effect of WBH on depressive symptoms is not solely the result of placebo factors related to nonspecific aspects of the procedure. Indeed, recognizing the modest effect of sham treatment is important for not "overselling" the therapeutic effects of WBH. Although a single session of WBH produced a clear antidepressant signal, rates of response and remission at each postintervention assessment were lower than are typically observed in antidepressant trials in which the intervention is delivered on a daily basis throughout the study period (eTable 4 in Supplement 2).

That a single treatment of WBH might produce long-term symptomatic improvement is consistent with results from other novel antidepressant interventions, such as ketamine and scopolamine, which have also demonstrated therapeutic effects that outlast their immediate biological actions.[14,15] Based on results from most studies of ketamine for MDD, we anticipated that the magnitude of the antidepressant response to WBH would diminish between postintervention weeks 1 and 6 as participants experienced a relapse in their depressive symptoms, as is common following a single exposure to ketamine.[14] However, 2 points require consideration prior to concluding that WBH may have a longer duration of effect than is typical for ketamine or scopolamine. First, as is apparent from Figure 2 and Table 2, active improvement in mean HDRS scores in the WBH group only occurred during the first 2 weeks after treatment, after which scores were maintained but not further reduced. This suggests a timeframe of biologic effect more in line with the assumed effects of ketamine and scopolamine. Second, the lack of relapse across the 2-week postintervention period was seen in both the WBH and sham groups and may reflect to an important degree the fact that the study sample—although had chronic depression—was not formally treatment resistant. Had a treatment-resistant population been recruited, relapse rates following WBH may have more closely approximated those seen with ketamine in treatment-resistant populations.

In general, the adverse effect profiles of both WBH and the sham comparator were mild and time limited. Adverse effects obviously induced by WBH, such as sweating or thirst, had already resolved when posttreatment adverse effects were assessed approximately 1 hour after treatment. No serious adverse events occurred during the study. Although we did not attempt to measure the patients' subjective response, most participants randomized to WBH found the experience to be pleasant rather than stressful or aversive.

Several limitations warrant discussion. The study sample was of modest size, which constrained the number of tests that could be run on the data without risking type I errors and which limited the ability to test the moderating effects of baseline covariates not balanced by randomization. In addition, although a large proportion of people randomized to the sham (71.4%) guessed incorrectly that they had received active WBH, it does not change the fact that the experience of the sham and WBH treatments was different in terms of the degree of heat experienced. Because this key aspect of the 2 interventions was significantly different, the possibility that functional unblinding contributed to differences between the 2 interventions cannot be dismissed. This is highlighted by the fact that almost all participants who received WBH correctly guessed they had received the active intervention.

Although most participants had experienced continuous depression for an extended period, we did not specifically enroll participants with treatment-resistant depression. Thus, we do not know how effective WBH would be in this specific subpopulation of individuals for whom a new treatment might be of most value. Specifically evaluating the effectiveness of WBH in treatment-resistant depression will be an important next step in determining where the intervention will fit in relation to current treatment algorithms. Nonetheless, we note that with its sustained antidepressant effect and mild effect profile, WBH might be an attractive alternative to antidepressant treatment in the large percentage of individuals with depression who might respond adequately to an antidepressant trial, but who harbor negative beliefs/feelings about antidepressant medications that have been shown to reduce adherence and worsen therapeutic outcomes.[16,17]

Finally, our selection of mild hyperthermia was based on the fact that the same temperature had produced an antidepressant signal in an earlier open trial and the fact that higher temperatures might be more likely to activate sensory pathways that respond to noxious levels of heat and that activate brain areas thought to already be hyperactive in MDD.[6,18] In addition, the risks and adverse effect burden of higher levels of WBH (ie, >38.5°C) are significantly greater,[7] which would reduce the attractiveness of the intervention for prospective patients. However, we do not know whether either higher or lower levels of heat might produce more robust antidepressant responses.

Conclusions

Results from the current study suggest that WBH holds promise as a safe, rapid-acting, antidepressant modality with a prolonged therapeutic benefit. Future studies will be required to identify both the optimal temperature and number and timing of treatments likely to produce the largest and longest-lasting clinical response in most patients.

Author Affiliations: Department of Human Development and Family Studies, School of Human Ecology, University of Wisconsin–Madison (Janssen, Begay, Raison); Norton School of Family and Consumer Sciences, College of Agriculture, University of Arizona, Tucson (Janssen, Medrano, Begay, Raison); Department of Integrative Physiology, Center for Neuroscience, University of Colorado–Boulder (Lowry); Department of Psychology, College of Science, University of Arizona, Tucson (Mehl, Allen, Rentscher); Department of Psychiatry, College of Medicine, University of Arizona, Tucson (Kelly, Gartner, Medrano, Begay, White, Fridman, Roberts, Raison); School of Anthropology, University of Arizona, Tucson (Kelly); Office of the Senior Vice President for Health Sciences, University of Arizona, Tucson (Gartner); Midwestern University, Glendale, Arizona (White); Department of Psychology, University of California, Riverside (Robbins); University of Applied Science Fresenius Idstein Medical School, Idstein, Germany (Hanusch); Research Design

Associates Inc, Yorktown Heights, New York (Cole); Department of Psychiatry, School of Medicine and Public Health, University of Wisconsin–Madison (Raison).

Author Contributions: Dr Raison had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Janssen, Lowry, Mehl, Begay, Hanusch, Raison.
Acquisition, analysis, or interpretation of data: Janssen, Lowry, Mehl, Allen, Kelly, Gartner, Medrano, Begay, Rentscher, White, Fridman, Roberts, Robbins, Cole, Raison.
Drafting of the manuscript: Janssen, Begay, Cole, Raison.
Critical revision of the manuscript for important intellectual content: Lowry, Mehl, Allen, Kelly, Gartner, Medrano, Begay, Rentscher, White, Fridman, Roberts, Robbins, Hanusch, Cole, Raison.
Statistical analysis: Janssen, Mehl, Allen, Cole, Raison.
Obtained funding: Janssen, Gartner, Raison.
Administrative, technical, or material support: Janssen, Mehl, Allen, Kelly, Gartner, Begay, White, Fridman, Roberts, Hanusch, Raison.
Study supervision: Janssen, Kelly, Medrano, Rentscher, White, Robbins, Hanusch, Raison.

Conflict of Interest Disclosures: In the previous 12 months, Dr Raison served on the speakers' bureau for Merck. None of the investigators has a financial

REFERENCES

1. Raison CL, Hale MW, Williams LE, Wager TD, Lowry CA. Somatic influences on subjective well-being and affective disorders: the convergence of thermosensory and central serotonergic systems. Front Psychol. 2015;5:1580.

2. Rolls ET, Grabenhorst F, Parris BA. Warm pleasant feelings in the brain. Neuroimage. 2008;41 (4):1504-1513.

3. Drevets WC, Price JL, Furey ML. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain Struct Funct*. 2008;213(1-2):93-118.

4. Hale MW, Lowry CA. Functional topography of midbrain and pontine serotonergic systems: implications for synaptic regulation of serotonergic circuits. *Psychopharmacology (Berl)*. 2011;213(2-3):243-264.

5. Lowry CA, Hale MW, Evans AK, et al. Serotonergic systems, anxiety, and affective disorder: focus on the dorsomedial part of the dorsal raphe nucleus. *Ann N Y Acad Sci*. 2008;1148:86-94.

6. Hanusch KU, Janssen CH, Billheimer D, et al. Whole-body hyperthermia for the treatment of major depression: associations with thermoregulatory cooling. *Am J Psychiatry*. 2013;170(7):802-804.

7. Wenger CB. Human heat acclimation. In: Pandolf KB, Sawka MN, Gonzalez RR, eds. *Human Performance Physiology and Environmental Medicine at Terrestrial Extremes*. Indianapolis, IN: Benchmark; 1988:153-197.

8. Müller MJ, Dragicevic A. Standardized rater training for the Hamilton Depression Rating Scale (HAMD-17) in psychiatric novices. *J Affect Disord*. 2003;77(1):65-69.

9. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30-46.

10. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6(4):284-290.

11. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry*. 2000;31(2):73-86.

12. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry*. 2003;53(8):649-659.

13. Kraemer HC. A source of false findings in published research studies: adjusting for covariates. *JAMA Psychiatry*. 2015;72(10):961-962.

14. Murrough JW, Iosifescu DV, Chang LC, et al. Antidepressant efficacy of ketamine in

treatment-resistant major depression: a two-site randomized controlled trial. *Am J Psychiatry*. 2013;170(10):1134-1142.

15. Drevets WC, Zarate CA Jr, Furey ML. Antidepressant effects of the muscarinic cholinergic receptor antagonist scopolamine: a review. *Biol Psychiatry*. 2013;73(12):1156-1163.

16. Goldney RD, Fisher LJ, Wilson DH. Mental health literacy: an impediment to the optimum treatment of major depression in the community. *J Affect Disord*. 2001;64(2-3):277-284.

17. Melartin TK, Rytsälä HJ, Leskelä US, Lestelä-Mielonen PS, Sokero TP, Isometsä ET. Continuity is the main challenge in treating major depressive disorder in psychiatric care. *J Clin Psychiatry*. 2005;66(2):220-227.

18. Patapoutian A, Peier AM, Story GM, Viswanath V. ThermoTRP channels and beyond: mechanisms of temperature sensation. *Nat Rev Neurosci*. 2003;4(7):529-539.