Alliance FOR AGING RESEARCH

Research Advocacy Training

Senior Patient and Family Caregiver Network

Chicago Marriott O'Hare
Chicago Ballroom C
8535 W. Higgins Rd.
Chicago, IL 60631

June 11-13, 2018

This training is funded through the Eugene Washington PCORI Engagement Award Program, Award 3445-AAR.

Catalyzing Innovation for Healthy Aging

Alliance
FOR AGING RESEARCH

June 11, 2018

Dear Senior Patient and Family Caregiver Network Advocates, Advisors, and Guests,

Welcome and thank you for joining us!

We are very excited and grateful to have you here for the training of the Senior Patient and Family Caregiver Network.

As you know, the purpose of this program is to develop an older adult patient and family caregiver-led nationwide group of advocates with the following:

- Basic understanding of patient-centered outcomes research (PCOR)
- Ability to develop research questions that are important to older adult patients and their family caregivers and will ultimately help inform research design, encourage broader participation, and produce meaningful health outcomes
- Willingness to provide the patient and family caregiver perspective by participating in PCOR opportunities at the national or local level

You each bring a unique perspective to this training, and that has already helped us to think about this curriculum in ways we would not have considered otherwise.

Some of you have advanced degrees and many years of professional experience. Some of you have less formal education, but extensive personal experience living with a chronic condition or caring for someone who has that condition. Some of you may have variations of one or both.

To set a level from the beginning, let me say this: everyone's experience is valid and important to this process. My request is for all of us to approach this training with an openness and willingness to both teach and learn from each other. If someone you are training with is struggling to understand something, be patient and help them along. If you have a question, be willing to ask—chances are, you are not the only one who has that question.

Please use this opportunity to learn, teach, and connect with each other. The psychologist, Dr. Joanne Cacciatore, states it well: "There is no pharmacy that can fill the need for compassionate interaction with others. There is no panacea. The answer to human suffering is both within us and between us."

Best,

Susan Peschin

Sue Peschin

Alliance
FOR AGING RESEARCH

SENIOR PATIENT AND FAMILY CAREGIVER NETWORK
ADVOCATE TRAINING
JUNE 11-13, 2018

TABLE OF CONTENTS

AGENDA ....TAB 1

SESSION MATERIALS ....TAB 2

SESSION 1: CLINICAL TRIALS...TAB 3
- ALZHEIMER'S ....TAB 4
  - ARTICLES
- ATRIAL FIBRILLATION .... TAB 5
  - PRESENTATION
  - ARTICLES
- CHRONIC PAIN .... TAB 6
  - ARTICLES
- SARCOPENIA...TAB 7
  - PRESENTATION
  - ARTICLES
- ACTIVITY SHEETS ....TAB 8

SESSION 2: PCOR IN ACTION ....TAB 9
- RECENT TRENDS ....TAB 10
- PCORI INVOLVEMENT ....TAB 11
- OPPORTUNITIES FOR ENGAGEMENT...................TAB 12
- FDA ENGAGEMENT ....TAB 13

ACTION PLAN TEMPLATE ....TAB 14

LIST OF ATTENDEES..TAB 15

WEBINARS ....TAB 16

WEBINAR 1: PATIENT-CENTERED OUTCOMES RESEARCH ....TAB 17
WEBINAR 2: CLINICAL TRIALS ....TAB 18

GLOSSARY OF TERMS ....TAB 19

Alliance
FOR AGING RESEARCH

Senior Patient and Family Caregiver Network
Advocate Training

Chicago Marriott O'Hare
8535 W Higgins Rd, Chicago, IL 60631
Chicago Ballroom C

June 11-13, 2018

Agenda

Monday, June 11, 2018

6:00 pm	**Welcome Dinner, Background, and Introductions** Sue Peschin
6:30 pm	**Research Advocacy in Action** Jim and Geri Taylor
7:00 pm	**Understanding the Patient-Centered Outcomes Research Institute** Lia Hotchkiss
8:00 pm	**Announcements and Wrap-Up**

Tuesday, June 12, 2018

8:00 am	**Breakfast**
8:30 am	**Overview of the Day**
8:40 am	**Session One: Clinical Trials** Jack Guralnik, Alan Jacobson, Mary Lyons, George Perry
10:00 am	**Break**

10:15 am	**Research Exercise: ClinicalTrials.gov**
12:00 Noon	**Break and Lunch**
1:30 pm	**Session Two: Patient-Centered Outcomes Research (PCOR) in Practice**
	Recent Trends in Incorporating Patient and Family Caregiver Perspectives Sue Peschin
	Examples of PCOR-related Participation Options from the Experts Penney Cowan, Gail Hunt, Debbe McCall
2:45 pm	**Break**
3:00 pm	**Discussion: What needs to be done by advocates to successfully participate and contribute to the process?**
3:45 pm	**Break**
4:00 pm	**Wrap-Up: Pause and reflect on our day, address questions**
5:00 pm	**Conclude for the day**

Wednesday, June 13, 2018

8:30 am	**Breakfast**
9:00 am	**Session Three: Developing Personal Action Plans** Sue Peschin
10:30 am	**Break**
10:45 am	**Session Three Report-Out**
12:00 Noon	**Wrap Up**
12:30 pm	**Workshop Complete**

doi:10.1038/nature19323

The antibody aducanumab reduces Aβ plaques in Alzheimer's disease

Jeff Sevigny[1]*, Ping Chiao[1]*, Thierry Bussière[1]*, Paul H. Weinreb[1]*, Leslie Williams[1], Marcel Maier[2], Robert Dunstan[1], Stephen Salloway[3], Tianle Chen[1], Yan Ling[1], John O'Gorman[1], Fang Qian[1], Mahin Arastu[1], Mingwei Li[1], Sowmya Chollate[1], Melanie S. Brennan[1], Omar Quintero-Monzon[1], Robert H. Scannevin[1], H. Moore Arnold[1], Thomas Engber[1], Kenneth Rhodes[1], James Ferrero[1], Yaming Hang[1], Alvydas Mikulskis[1], Jan Grimm[2], Christoph Hock[2,4], Roger M. Nitsch[2,4]§ & Alfred Sandrock[1]§

Alzheimer's disease (AD) is characterized by deposition of amyloid-β (Aβ) plaques and neurofibrillary tangles in the brain, accompanied by synaptic dysfunction and neurodegeneration. Antibody-based immunotherapy against Aβ to trigger its clearance or mitigate its neurotoxicity has so far been unsuccessful. Here we report the generation of aducanumab, a human monoclonal antibody that selectively targets aggregated Aβ. In a transgenic mouse model of AD, aducanumab is shown to enter the brain, bind parenchymal Aβ, and reduce soluble and insoluble Aβ in a dose-dependent manner. In patients with prodromal or mild AD, one year of monthly intravenous infusions of aducanumab reduces brain Aβ in a dose- and time-dependent manner. This is accompanied by a slowing of clinical decline measured by Clinical Dementia Rating—Sum of Boxes and Mini Mental State Examination scores. The main safety and tolerability findings are amyloid-related imaging abnormalities. These results justify further development of aducanumab for the treatment of AD. Should the slowing of clinical decline be confirmed in ongoing phase 3 clinical trials, it would provide compelling support for the amyloid hypothesis.

The amyloid hypothesis posits that Aβ-related toxicity is the primary cause of synaptic dysfunction and subsequent neurodegeneration that underlies the progression characteristic of AD[1,2]. Genetic, neuropathological, and cell biological evidence strongly suggest that targeting Aβ could be beneficial for patients with AD[3,4]. So far, attempts at therapeutically targeting Aβ have not been successful[5–7], casting doubt on the validity of the amyloid hypothesis. However, the lack of success may have been due to the inability of the antibodies to adequately engage their target or the proper target in the brain, or selecting the wrong patient population.

We describe the development of an antibody-based immunotherapeutic approach by selecting human B-cell clones triggered by neo-epitopes present in pathological Aβ aggregates. The screening of libraries of human memory B cells for reactivity against aggregated Aβ led to molecular cloning, sequencing, and recombinant expression of aducanumab (BIIB037), a human monoclonal antibody that selectively reacts with Aβ aggregates, including soluble oligomers and insoluble fibrils. In preclinical studies, we show that an analogue of aducanumab is capable of crossing the blood–brain barrier, engaging its target, and clearing Aβ from plaque-bearing transgenic mouse brains. These results prompted the start of clinical trials[8].

We report here interim results from a double-blind, placebo-controlled phase 1b randomized trial (PRIME; ClinicalTrials.gov identifier NCT01677572) designed to investigate the safety, tolerability, pharmacokinetics, and pharmacodynamics of monthly infusions of aducanumab in patients with prodromal or mild AD with brain Aβ pathology confirmed by molecular positron emission tomography (PET) imaging. Together, our data support further development of aducanumab as an Aβ-removing, disease-modifying therapy for AD.

Removal of brain Aβ plaques in patients with AD

In the PRIME study, 165 patients were randomized and treated between October 2012 and January 2014 at 33 sites in the United States. Patients with a clinical diagnosis of prodromal or mild AD and visually positive Aβ PET scan[9] were given monthly intravenous infusions of placebo or aducanumab at doses of 1, 3, 6 or 10 $\mathrm{mg\,kg^{-1}}$ for one year. Of these patients, 125 completed and 40 discontinued treatment, most commonly due to adverse events (20 patients) and withdrawal of consent (14 patients): 25% of the placebo group discontinued compared with 23%, 19%, 17%, and 38% of the 1, 3, 6 and 10 $\mathrm{mg\,kg^{-1}}$ aducanumab dose groups, respectively (Extended Data Fig. 1). Baseline characteristics, including cognitive measures, were generally well-balanced across the groups, although the 1 $\mathrm{mg\,kg^{-1}}$ dose group included a higher proportion of patients with mild AD, and the aducanumab treatment groups tended to have a higher Clinical Dementia Rating—Sum of Boxes (CDR-SB) score (Table 1).

Treatment with aducanumab reduced brain Aβ plaques as measured by florbetapir PET imaging in a dose- and time-dependent fashion (Fig. 1, 2a). The mean PET standard uptake value ratio (SUVR) composite score at baseline was 1.44. After 54 weeks of treatment, this had decreased significantly ($P < 0.001$) in the 3, 6 and 10 $\mathrm{mg\,kg^{-1}}$ dose groups; whereas change for the placebo group was minimal (Fig. 2a, Extended Data Table 1). In the 10 $\mathrm{mg\,kg^{-1}}$ dose group, the SUVR composite score was 1.16 after 54 weeks of treatment, a value near the purported quantitative cut-point of 1.10 that discriminates between positive and negative scans (Fig. 2b)[10]. The adjusted mean changes in SUVR composite scores in the 6 and 10 $\mathrm{mg\,kg^{-1}}$ groups treated for 26 weeks were similar in magnitude to the dose group below (3 and 6 $\mathrm{mg\,kg^{-1}}$, respectively) treated for 54 weeks (Fig. 2a). Reductions in amyloid PET SUVR composite score in aducanumab-treated patients

---

[1]Biogen, Cambridge, Massachusetts 02142, USA. [2]Neurimmune, Schlieren-Zurich 8952, Switzerland. [3]Butler Hospital, Providence, Rhode Island 02906, USA. [4]Institute for Regenerative Medicine, University of Zurich, Zurich 8952, Switzerland.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Table 1 | Baseline characteristics

Characteristic		Placebo ($n=40$)	Aducanumab 1 mg kg^{-1} ($n=31$)	 3 mg kg^{-1} ($n=32$)	 6 mg kg^{-1} ($n=30$)	 10 mg kg^{-1} ($n=32$)	Total ($n=165$)*
Years of age (mean ± s.d.)		72.8 ± 7.2	72.6 ± 7.8	70.5 ± 8.2	73.3 ± 9.3	73.7 ± 8.3	72.6 ± 8.1
Female sex (*n* (%))		23 (58)	13 (42)	17 (53)	15 (50)	15 (47)	83 (50)
ApoE ε4 (*n* (%))	Carriers	26 (65)	19 (61)	21 (66)	21 (70)	20 (63)	107 (65)
	Non-carriers	14 (35)	12 (39)	11 (34)	9 (30)	12 (38)	58 (35)
Clinical stage (*n* (%))	Prodromal	19 (48)	10 (32)	14 (44)	12 (40)	13 (41)	68 (41)
	Mild	21 (53)	21 (68)	18 (56)	18 (60)	19 (59)	97 (59)
MMSE (mean ± s.d.)		24.7 ± 3.6	23.6 ± 3.3	23.2 ± 4.2	24.4 ± 2.9	24.8 ± 3.1	24.2 ± 3.5
Global CDR (*n* (%))	0.5	34 (85)	22 (71)	22 (69)	25 (83)	24 (75)	127 (77)
	1	6 (15)	9 (29)	10 (31)	5 (17)	8 (25)	38 (23)
CDR-SB (mean ± s.d.)		2.66 ± 1.50	3.40 ± 1.76	3.50 ± 2.06	3.32 ± 1.54	3.14 ± 1.71	3.18 ± 1.72
FCSRT sum of free recall score (mean ± s.d.)		15.2 ± 8.5	13.2 ± 9.0	13.8 ± 8.0	14.4 ± 8.3	14.6 ± 8.3	14.3 ± 8.3
PET SUVR composite score (mean ± s.d.)		1.44 ± 0.17	1.44 ± 0.15	1.46 ± 0.15	1.43 ± 0.20	1.44 ± 0.19	1.44 ± 0.17
AD medications use† (*n* (%))		24 (60)	19 (61)	28 (88)	20 (67)	17 (53)	108 (65)

Percentages are rounded to the nearest integer. AD, Alzheimer's disease; ApoE ε4, apolipoprotein E ε4 allele; CDR, Clinical Dementia Rating; CDR-SB, Clinical Dementia Rating—Sum of Boxes; FCSRT, Free and Cued Selective Reminding Test; MMSE, Mini-Mental State Examination; PET, positron emission tomography; SD, standard deviation; SUVR, standard uptake value ratio.
*Number of patients dosed.
†Cholinesterase inhibitors and/or memantine.

were similar in patients with mild and prodromal AD, and apolipoprotein E (ApoE) ε4 carriers and non-carriers (Extended Data Fig. 2a, b). Pre-specified regional analyses of SUVR changes demonstrated statistically significant dose-dependent reductions in all brain regions, except for the pons and sub-cortical white matter, two areas in which Aβ plaques are not expected to accumulate (Extended Data Fig. 3).

Effect on clinical measures

Clinical assessments were exploratory as the study was not powered to detect clinical change. The test of dose response was the pre-specified primary analysis for the clinical assessments. Analysis of change from baseline on the CDR-SB (adjusted for baseline CDR-SB and ApoE ε4 status) demonstrated dose-dependent slowing of clinical progression with aducanumab treatment at one year (dose-response, $P<0.05$), with the greatest slowing for 10 mg kg^{-1} ($P<0.05$ versus placebo) (Fig. 3a, Extended Data Table 1). Sensitivity analysis using a mixed model for repeated measures (MMRM) also showed a trend for slowing of decline on the CDR-SB at one year ($P=0.07$ with 10 mg kg^{-1} aducanumab versus placebo). A dose-dependent slowing of clinical progression on the Mini Mental State Examination (MMSE) with aducanumab treatment was also observed at one year (dose-response, $P<0.05$), with the greatest effects at 3 and 10 mg kg^{-1} aducanumab ($P<0.05$ versus placebo) (Fig. 3b, Extended Data Table 1). On sensitivity analysis using MMRM, the greatest difference was retained for 10 mg kg^{-1} aducanumab ($P<0.05$ versus placebo), with a smaller difference at 3 mg kg^{-1} ($P=0.10$ versus placebo). No changes from baseline after one year were found on the composite neuropsychological test battery (NTB) or the Free and Cued Selective Reminding Test (FCSRT) free recall (Extended Data Table 1), but skewed non-normal (floor) effects at baseline were observed. The floor effects on the NTB were seen in the individual tests; specifically, in the two most clinically relevant components given the stage of the population enrolled: Wechsler Memory Scale-Fourth Edition Verbal Paired Associates II (WMS-IV VPA II) and Rey Auditory Verbal Learning Test (RAVLT) delayed recall of the NTB memory domain.

Safety and tolerability

The most common adverse effects were amyloid-related imaging abnormalities (ARIA), headache, urinary tract infection, and upper respiratory tract infection (Table 2). Using the most specific description of ARIA by magnetic resonance imaging (MRI), ARIA-vasogenic oedema (ARIA-E) abnormalities occurred in no patients receiving placebo compared with 1 (3%), 2 (6%), 11 (37%), and 13 (41%) patients receiving 1, 3, 6 and 10 mg kg^{-1} aducanumab, respectively (Extended Data Table 2). ARIA-E was generally observed early in the course of treatment, MRI findings typically resolved within 4–12 weeks, and of the 27 patients who developed ARIA-E, 15 (56%) continued treatment (Supplementary Information). All cases of symptomatic ARIA were

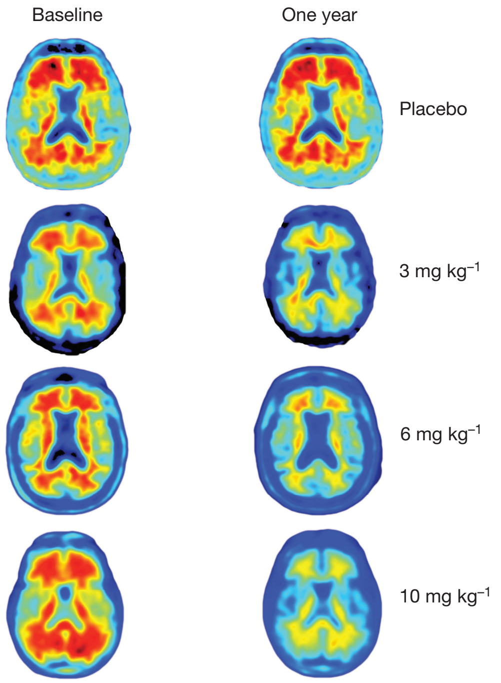
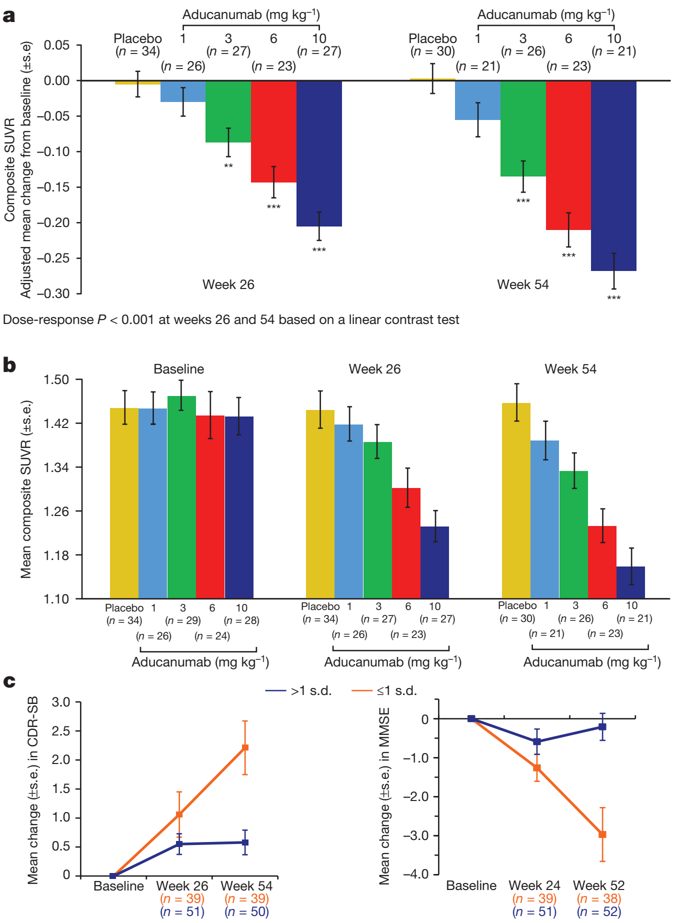


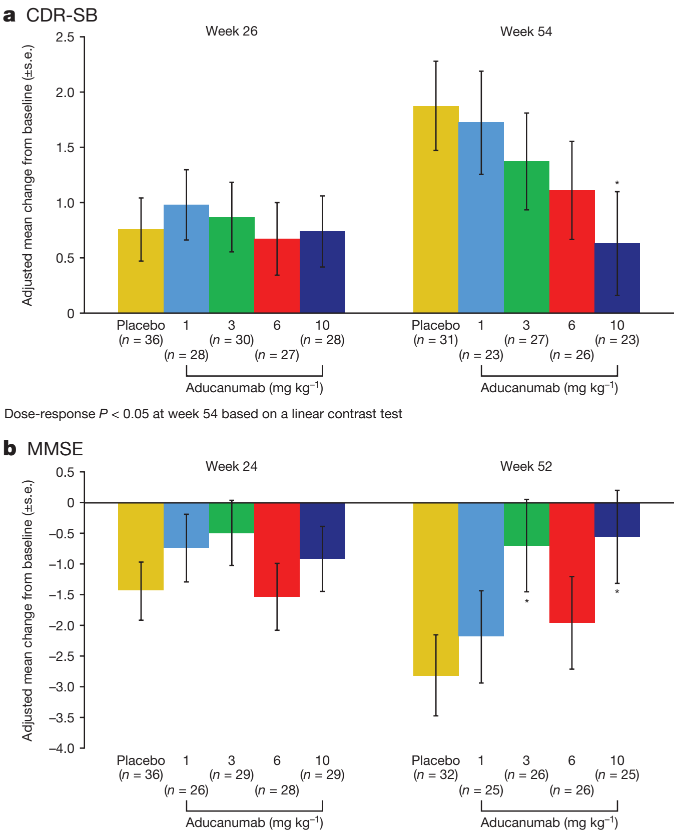
Figure 1 | Amyloid plaque reduction with aducanumab: example amyloid PET images at baseline and week 54. Individuals were chosen based on visual impression and SUVR change relative to average one-year response for each treatment group ($n=40$, 32, 30 and 32, respectively). Axial slice shows anatomical regions in posterior brain putatively related to AD pathology. SUVR, standard uptake value ratio.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



Dose-response $P < 0.001$ at weeks 26 and 54 based on a linear contrast test

Figure 2 | Amyloid plaque reduction with aducanumab. a–c, Change from baseline (**a**, analyses using ANCOVA), SUVR values (**b**), and categorization of change in amyloid PET (**c**) at week 54 and associated change from baseline CDR-SB and MMSE in aducanumab-treated patients (post hoc analysis). Categorization of amyloid PET at week 54 based on s.d. of change from baseline in placebo-treated patients. $**P < 0.01$; $***P < 0.001$ versus placebo; two-sided tests with no adjustments for multiple comparisons. Mean ± s.e. ANCOVA, analysis of covariance; CDR-SB, Clinical Dementia Rating—Sum of Boxes; MMSE, Mini Mental State Examination; SUVR, standard uptake value ratio.



Dose-response $P < 0.05$ at week 54 based on a linear contrast test

Dose-response $P < 0.05$ at week 52 based on a linear contrast test

Figure 3 | Aducanumab effect (change from baseline) on CDR-SB and MMSE. a, b, Aducanumab effect on CDR-SB (**a**) and MMSE (**b**). $*P < 0.05$ versus placebo; two-sided tests with no adjustments for multiple comparisons. CDR-SB and MMSE were exploratory endpoints. Adjusted mean ± s.e. Analyses using ANCOVA. CDR-SB, Clinical Dementia Rating—Sum of Boxes; MMSE, Mini Mental State Examination.

required to be reported as medically important serious adverse effects. No patients were hospitalised for ARIA. The only serious adverse effects (by preferred term) that occurred in more than one patient in any treatment group were ARIA (0, 1 (3%), 1 (3%), 4 (13%), and 5 (16%) of patients receiving placebo, and 1, 3, 6 and 10 mg kg^{-1} aducanumab, respectively) and superficial siderosis of the central nervous system (0, 1 (3%), 0, 2 (7%), and 3 (9%) of patients receiving placebo and 1, 3, 6 and 10 mg kg^{-1} aducanumab, respectively). Owing to the requirement for repeated MRI assessments of those patients who developed ARIA, these individuals were partially unblinded to treatment. Other adverse effects and serious adverse effects were consistent with the patient population. There were no drug-related deaths (Supplementary Information).

Pharmacokinetics

The pharmacokinetics of aducanumab (maximum concentration ($C_{\max}$) and cumulative area under the concentration curve (AUC)) were linear across the dose range in patients who received all 14 planned doses (Extended Data Table 3). The median plasma half-life was 21 days. In total, 3 of 118 evaluable patients (3%) in the combined aducanumab groups developed treatment-emergent anti-aducanumab antibodies within the first year of treatment. Antibody responses were transient, with minimal titres, and had no apparent effect on aducanumab pharmacokinetics or safety.

Brain penetration and binding to Aβ plaques

In the preclinical studies which preceded PRIME, systemically administered aducanumab (single dose, 30 mg kg^{-1} intraperitoneally (i.p.)) bound to diffuse and compact Aβ plaques in the brains of 22-month-old female Tg2576 transgenic mice (‘Target engagement study’; Extended Data Fig. 4a–d). $C_{\max}$ in plasma was 181 μg ml^{-1}, with a terminal elimination half-life ($t_{1/2}$) of 2.5 days. The $C_{\max}$ in brain was 1,062 ng g^{-1} of tissue, and approximately 400–500 ng g^{-1} of drug was measured 3 weeks after dosing, suggesting long-term retention. Consequently, the brain:plasma AUC ratio of 1.3% was higher than the 0.1% frequently reported for systemically administered antibodies[11,12].

Administration of a single dose of aducanumab did not affect plasma (Extended Data Fig. 4b) or brain (data not shown) Aβ concentrations, consistent with the observation that aducanumab does not bind to soluble Aβ monomers. In contrast, the murine bapineuzumab precursor antibody 3D6, which binds to Aβ monomers, triggered a transient plasma Aβ spike (Extended Data Fig. 4b). Similarly, plasma Aβ concentrations were stable after repeated dosing with aducanumab in the PRIME study (data not shown). Within 24 h of dosing, aducanumab bound to parenchymal brain Aβ with a spatial pattern essentially superimposable with *ex vivo* pan-Aβ antibody staining, confirming that aducanumab binds all morphological types of brain Aβ plaques *in vivo*, including diffuse Aβ deposits and compact Aβ plaques (Extended Data Fig. 4c, d). Aducanumab binding to Aβ deposited in cerebral amyloid angiopathy (CAA) lesions within brain blood vessel walls was less

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Table 2 | Summary of adverse events and most common adverse events

		Aducanumab			
Adverse event (*n* (%))	Placebo (*n* = 40)	1 mg kg^{-1} (*n* = 31)	3 mg kg^{-1} (*n* = 32)	6 mg kg^{-1} (*n* = 30)	10 mg kg^{-1} (*n* = 32)
Any adverse event	39 (98)	28 (90)	27 (84)	28 (93)	29 (91)
Serious event	15 (38)	3 (10)	4 (13)	4 (13)	12 (38)
Discontinuing treatment due to an adverse event	4 (10)	3 (10)	2 (6)	3 (10)	10 (31)
Common adverse events					
ARIA	2 (5)	2 (6)	4 (13)	11 (37)	15 (47)
Headache	2 (5)	5 (16)	4 (13)	8 (27)	8 (25)
Urinary tract infection	4 (10)	3 (10)	2 (6)	4 (13)	5 (16)
Upper respiratory tract infection	6 (15)	2 (6)	2 (6)	2 (7)	6 (19)
Diarrhoea	3 (8)	0	6 (19)	1 (3)	3 (9)
Arthralgia	2 (5)	0	6 (19)	2 (7)	1 (3)
Fall	8 (20)	3 (10)	2 (6)	2 (7)	2 (6)
Superficial siderosis of CNS	0	2 (6)	1 (3)	2 (7)	4 (13)
Constipation	0	3 (10)	1 (3)	1 (3)	3 (9)
Nausea	2 (5)	2 (6)	5 (16)	0	1 (3)
Anxiety	4 (10)	4 (13)	1 (3)	1 (3)	1 (3)
Nasopharyngitis	0	1 (3)	5 (16)	0	1 (3)
Cough	2 (5)	3 (10)	1 (3)	0	1 (3)
Alanine aminotransferase increased	0	3 (10)	0	1 (3)	0
Aspartate aminotransferase increased	0	3 (10)	0	0	1 (3)

Common adverse events are those with an incidence of ≥10% in any aducanumab treatment group. Incidence of ARIA based on adverse event reporting. Adverse events of ARIA-E (oedema) and ARIA-H (micro-haemorrhage) are both coded to the MedDRA preferred term of amyloid-related imaging abnormalities, and ARIA-H (superficial siderosis) codes to the MedDRA preferred term of superficial siderosis of the CNS. ARIA, amyloid-related imaging abnormalities; CNS, central nervous system; MedDRA, Medical Dictionary for Regulatory Activities.

prominent than parenchymal Aβ binding, when compared with the total amount of Aβ (Extended Data Fig. 4c, d).

Reduction of brain Aβ in transgenic mice

Exposure in plasma and brain correlated linearly with dose after chronic dosing in plaque-bearing transgenic mice (Extended Data Fig. 5) (Supplementary Information). $^{\mathrm{ch}}$aducanumab, a murine IgG2a/κ chimaeric analogue, dose-dependently reduced Aβ measured in brain homogenates by up to 50% relative to the vehicle control in the diethylamine (DEA) fraction that extracted soluble monomeric and oligomeric forms of Aβ_{40} and Aβ_{42}, and in the guanidine hydrochloride (GuHCl) fraction that extracted insoluble Aβ fibrils (Fig. 4a, b).

Quantitative 6E10 immunohistochemistry showed significant reductions in all forms of Aβ deposits by up to 70% (Fig. 4c, d). Thioflavin S (ThioS) staining of compact Aβ plaques showed dose-dependent and statistically significant reductions in the cortex and hippocampus by up to 63% (Fig. 4c, d). Quantitative histology indicated that $^{\mathrm{ch}}$aducanumab significantly reduced the number of plaques of all sizes, including plaques $>500\,\mu\mathrm{m}^2$ and plaques $<125\,\mu\mathrm{m}^2$ (Extended Data Fig. 6a–c). Quantification of ThioS-positive vascular and parenchymal Aβ plaques separately showed that $^{\mathrm{ch}}$aducanumab did not affect vascular Aβ in either cortex or hippocampus (Fig. 4e–h).

To identify the mechanism of Aβ clearance, we analysed the involvement of microglia which are known to display enhanced phagocytic activities through binding to the Fc region of an antibody[13,14]. $^{\mathrm{ch}}$aducanumab significantly increased recruitment of Iba-1-positive microglia to Aβ plaques, suggesting FcγR-mediated phagocytosis of antibody–Aβ complexes as a possible clearance mechanism (Extended Data Fig. 7a–c and Supplementary Information).

Biochemical characterization

The apparent affinities of aducanumab and $^{\mathrm{ch}}$aducanumab for aggregated Aβ_{42}, with half maximal effective concentration (EC$_{50}$) values of 0.1 nM, were comparable to 3D6 (ref. 13) (Fig. 5a). Neither aducanumab nor $^{\mathrm{ch}}$aducanumab bound monomeric soluble Aβ_{40} at concentrations up to 1 μM, indicating >10,000-fold selectivity for aggregated Aβ over monomer, whereas 3D6 bound soluble Aβ_{40} with an EC$_{50}$ of 1 nM (Fig. 5b). In contrast to 3D6, which immunoprecipitated both monomeric and aggregated Aβ, $^{\mathrm{ch}}$aducanumab bound soluble Aβ_{42} oligomers and insoluble Aβ_{42} fibrils prepared *in vitro*, but not Aβ_{42} monomers (Fig. 5c). Histological staining of autopsy tissue from patients with AD or aged amyloid precursor protein (APP) transgenic mice confirmed binding of aducanumab to bona fide human Aβ fibrils (Fig. 5d, e).

Discussion

The PRIME study shows that aducanumab penetrates the brain and decreases Aβ in patients with AD in a time- and dose-dependent manner. Within 54 weeks of treatment, 3, 6 and 10 mg kg^{-1} doses of aducanumab significantly decreased the amyloid PET SUVR. Patients receiving placebo showed virtually no change in their mean PET SUVR composite scores over one year, indicating that Aβ pathology had already reached an asymptote of accumulation. Considering that it may have taken up to 20 years for Aβ to have accumulated to the levels in these patients at study entry[15], the observed kinetics of Aβ removal within a 12-month time period appears encouraging for a disease-modifying treatment for patients with AD.

The cognitive results for CDR-SB and MMSE provide support for the clinical hypothesis that reduction of brain Aβ confers a clinical benefit. Post hoc analysis showed that those aducanumab-treated patients who had decreased SUVR scores >1 standard deviation unit relative to placebo-treated patients after one year of treatment experienced a stabilization of clinical decline on both CDR-SB and MMSE scores; whereas, those patients with a smaller or no decrease experienced clinical decline similar to placebo patients (Fig. 2c). The apparent clinical benefit observed in PRIME could also be explained by the binding of aducanumab to oligomeric forms of Aβ, which would not be directly detected by PET imaging. The reductions in SUVR scores may be surrogates for reductions in toxic soluble Aβ oligomers which may have had a more functionally relevant impact on cognition. Whereas significant Aβ reduction was detectable by 6 months, clinical effects were not

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

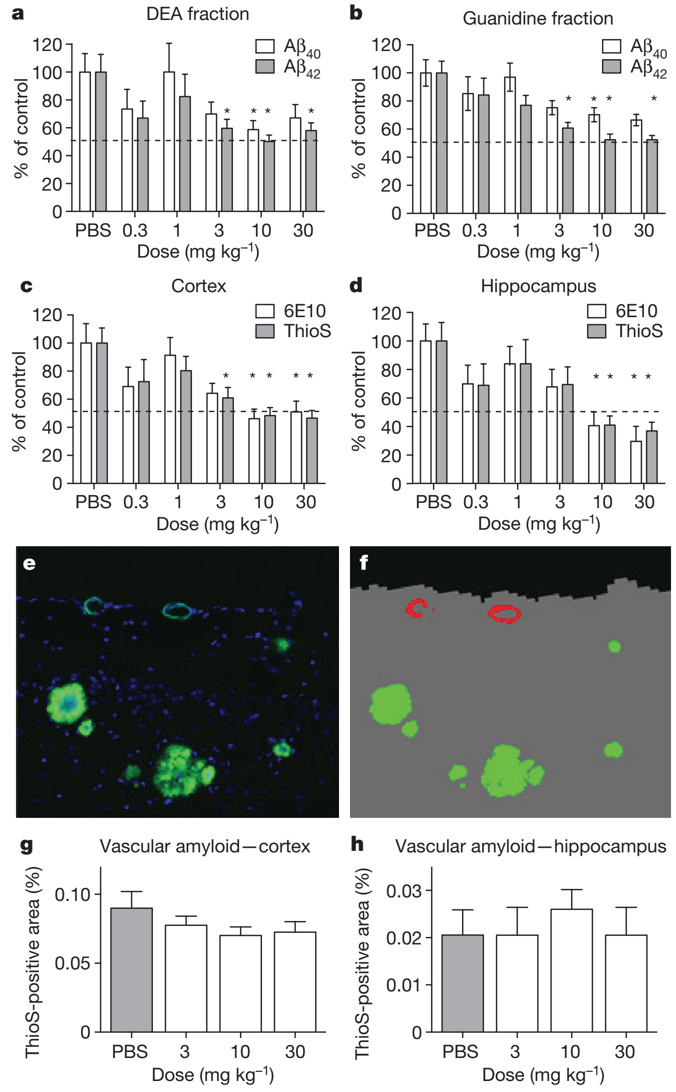


Figure 4 | Reduction of amyloid burden following weekly dosing with ^{ch}aducanumab in 9.5- to 15.5-month-old Tg2576 transgenic mice. **a**, **b**, $A\beta_{40}$ and $A\beta_{42}$ levels in soluble DEA (**a**) and insoluble GuHCl (**b**) brain fractions. **c**, **d**, Total brain Aβ (6E10) and compact amyloid plaques (ThioS) in cortex (**c**) and hippocampus (**d**) (mean ± s.e.; $n = 20$–55; dotted line 50% reduction; $^*P < 0.05$ versus control). **e**–**h**, ThioS staining of amyloid deposits (**e**) and Visiopharm software (**f**) differentiated parenchymal deposits (green) from vascular deposits (red) (representative pictures 10× magnification), and quantified area of vascular amyloid (**g**, **h**; mean ± s.e.; $n = 20$–24).

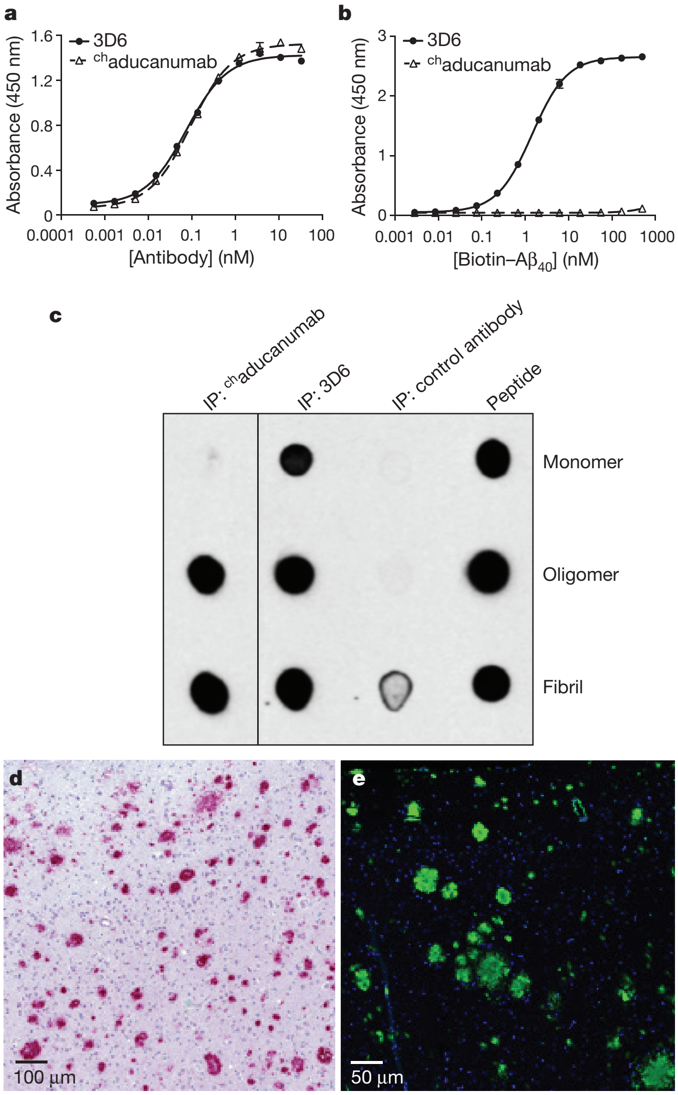


Figure 5 | Aducanumab binds selectively to insoluble fibrillar and soluble oligomeric Aβ aggregates. **a**, Binding of ^{ch}aducanumab or 3D6 to immobilized fibrillar $A\beta_{42}$. Mean ± s.d., in triplicate. **b**, Capture of soluble monomeric $A\beta_{40}$ with immobilized ^{ch}aducanumab or 3D6. Mean ± s.d., in triplicate. **c**, Dot blots of $A\beta_{42}$ monomer, soluble oligomers, or insoluble fibrils immunoprecipitated with ^{ch}aducanumab, 3D6, or irrelevant antibody control. Equivalent concentrations confirmed by direct dot blotting (Peptide). **d**, **e**, Immunostaining of Aβ in autopsy brain tissue from a patient with AD with ^{ch}aducanumab (0.2 μg ml^{-1}) (**d**) and 22-month-old Tg2576 transgenic mouse brain tissue with aducanumab (60 ng ml^{-1}) (**e**).

apparent until one year. Given that clearance of Aβ could be followed by recovery of neuronal function, a lag between reduction of Aβ burden and slowing of disease progression is not altogether surprising.

The main safety finding, ARIA-E, was dose-dependent and more common in ApoE ε4 carriers, consistent with findings with other anti-Aβ monoclonal antibodies[7,16,17]. Although the underlying cause of ARIA is not well understood, it is likely that the MRI signal of ARIA is due to increased extracellular fluid. This may be a result of underlying CAA, changes in perivascular clearance and vascular integrity, or local inflammatory processes associated with Aβ-targeting therapies[17–20] (see Supplementary Information for further discussion).

Study limitations of the PRIME phase 1b study included staggered parallel-group design, small sample sizes, limited region (USA only), and possible partial unblinding due to ARIA-E. Measures were taken to maintain blinding to adverse effects: raters of given tests were not permitted to perform other clinical assessments, and were blinded to other assessments (for example, MMSE and CDR raters were required to be different and neither were permitted to perform other study assessments). Post hoc analyses of change from baseline PET SUVR composite score and cognition by presence/absence of ARIA suggested no apparent difference in treatment effect when comparing patients with and without ARIA-E (Extended Data Table 4). There was overlap in enrolment in Arms 1–3 (aducanumab 1 and 3 mg kg^{-1}, placebo) and Arms 4 and 5 (aducanumab 10 mg kg^{-1}, placebo) but Arms 6 and 7 (aducanumab 6 mg kg^{-1}, placebo) were initiated after enrolment in Arms 1–5 was complete. This was a small study designed for assessment of safety and tolerability, and for detecting a pharmacological effect on brain Aβ levels measured by PET imaging. The trial was not powered for the exploratory clinical endpoints, thus the clinical cognitive results should be interpreted with caution. Primary analyses were based on observed data with no imputation for missing values, nominal P values were presented with no adjustments for multiple comparisons, and they were supported by sensitivity analyses using a MMRM.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

The initiation of the PRIME study and its results are supported by extensive preclinical data. Detection on parenchymal Aβ plaques following a single systemic administration confirmed that aducanumab penetrates the brain to a sufficient extent to allow accumulation on Aβ plaques. This is consistent with earlier findings showing that, in the presence of significant Aβ deposition, plaque-binding antibodies can be detected bound to the target over an extended period[14,21]. The minimal effective dose upon repeated systemic administration of ^{ch}aducanumab in transgenic mice was $3\,\mathrm{mg\,kg^{-1}}$ (corresponding to minimally effective concentrations of $13.8 \pm 1.9\,\mu\mathrm{g\,ml^{-1}}$ in plasma and $99.8 \pm 30.0\,\mathrm{ng\,g^{-1}}$ in brain) with reductions of $A\beta_{42}$ in soluble and insoluble brain fractions of approximately 50%, and reductions in Aβ plaque of approximately 40%. Since exposure at $3\,\mathrm{mg\,kg^{-1}}$ in animals and humans is approximately equivalent, the observed dose-response in the model was consistent with the clinical doses that led to reductions in amyloid PET SUVR. ^{ch}aducanumab cleared plaques of all sizes, suggesting that aducanumab triggered clearance of pre-existing Aβ plaques and prevented formation of new plaques.

In transgenic mice, aducanumab preferentially bound to parenchymal Aβ over vascular Aβ deposits, consistent with the lack of effect on vascular Aβ following chronic dosing. The effect of anti-Aβ antibody therapies on the vascular Aβ compartment could be related to micro-haemorrhages or oedema in transgenic mice, and may relate to ARIA in clinical trials[22]. Nevertheless, the preferential binding of aducanumab to parenchymal versus vascular Aβ may have been critical in allowing the use of relatively high doses in the clinical study so as to achieve robust target engagement in the brains of patients with AD.

Several mechanisms may be involved in aducanumab's Aβ-lowering activity. The clearance of Aβ deposits was accompanied by enhanced recruitment of microglia. Together with the reduced potency of the aglycosylated form of ^{ch}aducanumab (data not shown), and the *ex vivo* phagocytosis data, this suggests that FcγR-mediated microglial recruitment and phagocytosis played an important role in Aβ clearance in these models. Activated microglia appeared to encapsulate the remaining central dense core of plaques in treated animals, possibly isolating them from the surrounding neuropil. It is commonly thought that soluble Aβ oligomers, rather than monomers or plaques, may be the primary toxic species[23,24]. Considering that Aβ plaques might be a source of Aβ oligomers[25–28], this suggests that treatment with aducanumab might slow their release into the neuropil, thereby limiting their toxic effect on neurons[29]. In fact, chronic dosing of 18-month-old Tg2576 transgenic mice with ^{ch}aducanumab led to normalization of neuritic calcium overload in the brain[30]. Other studies have linked calcium dyshomeostasis in neurons and microglia to binding of Aβ oligomers to metabotropic receptors[31–33]. Aducanumab binding to soluble Aβ oligomers may prevent their interaction with those receptors, thereby preventing the detrimental effect of membrane depolarization. Restoration of this functional endpoint suggests that aducanumab treatment may lead to beneficial effects on neuronal network function underlying cognitive deficits.

Together, the clinical and preclinical data support continued development of aducanumab as a disease-modifying treatment for AD. The clinical study results provide robust support to the biological hypothesis that treatment with aducanumab reduces brain Aβ plaques and, more importantly, to the clinical hypothesis that Aβ plaque reduction confers clinical benefit. This concurs with preclinical data demonstrating brain penetration, target engagement, and dose-dependent clearance of Aβ plaques in transgenic mice. The clinical effects of aducanumab need to be confirmed in larger studies. Both the long-term extension (LTE) phase of this study and phase 3 development are ongoing.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 July 2015; accepted 21 July 2016.

Published online 24 August 2016.

1. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297,** 353–356 (2002).
2. Hardy, J. A. & Higgins, G. A. Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256,** 184–185 (1992).
3. Ising, C., Stanley, M. & Holtzman, D. M. Current thinking on the mechanistic basis of Alzheimer's and implications for drug development. *Clin. Pharmacol. Ther.* **98,** 469–471 (2015).
4. Selkoe, D. J. The therapeutics of Alzheimer's disease: where we stand and where we are heading. *Ann. Neurol.* **74,** 328–336 (2013).
5. Cummings, J. L., Morstorf, T. & Zhong, K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res. Ther.* **6,** 37 (2014).
6. Doody, R. S. *et al.* Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* **370,** 311–321 (2014).
7. Salloway, S. *et al.* Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* **370,** 322–333 (2014).
8. Ferrero, J. *et al.* First-in-human, double-blind, placebo-controlled, single-dose escalation study of aducanumab (BIIB037) in mild-to-moderate Alzheimer's disease. *Alzheimers Dement (N Y)* (in the press).
9. Sevigny, J. *et al.* Amyloid PET screening for enrichment of early-stage Alzheimer disease clinical trials: experience in a phase 1b clinical trial. *Alzheimer Dis. Assoc. Disord.* **30,** 1–7 (2016).
10. Landau, S. M. *et al.* Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *J. Nucl. Med.* **54,** 70–77 (2013).
11. Banks, W. A. *et al.* Passage of amyloid beta protein antibody across the blood-brain barrier in a mouse model of Alzheimer's disease. *Peptides* **23,** 2223–2226 (2002).
12. Levites, Y. *et al.* Insights into the mechanisms of action of anti-Abeta antibodies in Alzheimer's disease mouse models. *FASEB J.* **20,** 2576–2578 (2006).
13. Bard, F. *et al.* Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease. *Nat. Med.* **6,** 916–919 (2000).
14. Bohrmann, B.*et al.* Gantenerumab: a novel human anti-Aβ antibody demonstrates sustained cerebral amyloid-β binding and elicits cell-mediated removal of human amyloid-β. *J. Alzheimers Dis.* **28,** 49–69 (2012).
15. Villemagne, V. L. *et al.* Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* **12,** 357–367 (2013).
16. Ostrowitzki, S. *et al.* Mechanism of amyloid removal in patients with Alzheimer disease treated with gantenerumab. *Arch. Neurol.* **69,** 198–207 (2012).
17. Sperling, R. *et al.* Amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with bapineuzumab: a retrospective analysis. *Lancet Neurol.* **11,** 241–249 (2012).
18. Sperling, R. A. *et al.* Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's Association Research Roundtable Workgroup. *Alzheimers Dement.* **7,** 367–385 (2011).
19. Barakos, J. *et al.* MR imaging features of amyloid-related imaging abnormalities. *AJNR Am. J. Neuroradiol.* **34,** 1958–1965 (2013).
20. Zago, W. *et al.* Vascular alterations in PDAPP mice after anti-Aβ immunotherapy: Implications for amyloid-related imaging abnormalities. *Alzheimers Dement.* **9** (Suppl), S105–S115 (2013).
21. Wang, A., Das, P., Switzer, R. C., III, Golde, T. E. & Jankowsky, J. L. Robust amyloid clearance in a mouse model of Alzheimer's disease provides novel insights into the mechanism of amyloid-beta immunotherapy. *J. Neurosci.* **31,** 4124–4136 (2011).
22. Boche, D. *et al.* Consequence of Abeta immunization on the vasculature of human Alzheimer's disease brain. *Brain* **131,** 3299–3310 (2008).
23. Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat. Rev. Mol. Cell Biol.* **8,** 101–112 (2007).
24. Kayed, R. & Lasagna-Reeves, C. A. Molecular mechanisms of amyloid oligomers toxicity. *J. Alzheimers Dis*. **33** (Suppl 1), S67–S78 (2013).
25. Benilova, I., Karran, E. & De Strooper, B. The toxic Aβ oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat. Neurosci.* **15,** 349–357 (2012).
26. Koffie, R. M. *et al.* Oligomeric amyloid beta associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proc. Natl Acad. Sci. USA* **106,** 4012–4017 (2009).
27. Shankar, G. M. *et al.* Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat. Med.* **14,** 837–842 (2008).
28. Condello, C., Yuan, P., Schain, A. & Grutzendler, J. Microglia constitute a barrier that prevents neurotoxic protofibrillar Aβ42 hotspots around plaques. *Nat. Commun.* **6,** 6176 (2015).
29. Jin, M. *et al.* Soluble amyloid beta-protein dimers isolated from Alzheimer cortex directly induce tau hyperphosphorylation and neuritic degeneration. *Proc. Natl Acad. Sci. USA* **108,** 5819–5824 (2011).
30. Kastanenka, K. *et al.* Amelioration of calcium dyshomeostasis by immunotherapy with BIIB037 in Tg2576 mice. *Alzheimers Dement.* **9,** P508 (2013).

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

31. Jarosz-Griffiths, H. H., Noble, E., Rushworth, J. V. & Hooper, N. M. Amyloid-β receptors: the good, the bad, and the prion protein. *J. Biol. Chem.* **291,** 3174–3183 (2016).
32. Morkuniene, R. *et al.* Small Aβ1-42 oligomer-induced membrane depolarization of neuronal and microglial cells: role of *N*-methyl-D-aspartate receptors. *J. Neurosci. Res.* **93,** 475–486 (2015).
33. Um, J. W. *et al.* Metabotropic glutamate receptor 5 is a coreceptor for Alzheimer aβ oligomer bound to cellular prion protein. *Neuron* **79,** 887–902 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements These studies were funded by Biogen. The authors thank the patients and their family members participating in the aducanumab studies, and the PRIME investigators (Supplementary Information) and staff conducting these studies. Medical writing support, under direction of the authors, was provided by A. Smith at Complete Medical Communications, and was funded by Biogen. We thank N. Pederson, J. Dolnikova and E. Garber for help in generating the recombinant antibodies, D. Fahrer, C. Quigley, M. Themeles, X. Zhang and P. Auluck for help in generating the histological data, and K. Mack for editorial support and coordination of the authors in combining the preclinical and clinical work in this manuscript.

Author Contributions T.B., P.H.W., M.M., T.E., K.R., J.G. and R.M.N. designed the preclinical studies, and J.S., Y.L., J.G., J.F., C.H., R.M.N. and A.S. designed the clinical study. P.C. led the imaging implementation for the clinical study. T.C. and J.O. were clinical study statisticians. T.B., P.H.W., M.M., R.D., F.Q., M.A., M.L., S.C., M.S.B., O.Q.-M., R.H.S., H.M.A., T.E., J.G. and R.M.N. generated, analysed, and/or interpreted data from preclinical studies. T.B., P.H.W., M.M., R.D., F.Q., M.A., M.L., S.C., M.S.B., O.Q.-M., R.H.S., H.M.A., T.E., K.R., J.G., C.H., R.M.N. and A.S. critically reviewed preclinical sections of the manuscript. J.S., P.C., L.W., S.S., T.C., Y.L., J.O., J.F., Y.H., A.M., J.G., C.H., R.M.N. and A.S. analysed and interpreted clinical study data and critically reviewed clinical sections of the manuscript. All authors approved the final version of the manuscript for submission. Biogen and Neurimmune reviewed and provided feedback on the paper. The authors had full editorial control of the paper, and provided their final approval of all content.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S. (alfred.sandrock@biogen.com).

Reviewer Information *Nature* thanks L. Lannfelt, R. Thomas and the other anonymous reviewer(s) for their contribution to the peer review of this work.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

METHODS

Clinical study subjects. Patients were screened for inclusion in three stages. First, patients were evaluated on demographic, and clinical and laboratory criteria, including being between 50–90 years of age, and meeting clinical criteria for either prodromal or mild AD, as determined by the investigator. The criteria for prodromal AD were: MMSE score between 24–30 (inclusive), a spontaneous memory complaint, objective memory loss defined as a free recall score of ≤ 27 on the FCSRT[34], a global CDR score of 0.5, absence of significant levels of impairment in other cognitive domains and essentially preserved activities of daily living, and an absence of dementia[35]. The criteria for mild AD were: MMSE score between 20–26 (inclusive), a global CDR of 0.5 or 1.0, and meeting the National Institute on Aging–Alzheimer's Association core clinical criteria for probable AD[36]. Second, patients who remained eligible underwent MRI to exclude those with confounding pathology, including acute or sub-acute micro- or macro-haemorrhage, prior macro-haemorrhage, >4 micro-haemorrhages, superficial siderosis or any finding that might be a contributing cause of the patient's dementia, pose a risk to the patient, or prevent a satisfactory MRI assessment for safety monitoring. Third, remaining eligible patients underwent a florbetapir PET scan, and those with a positive scan based on a visual assessment, as determined by a qualified reader, were eligible. The Aβ PET screening process has been described in a separate publication[9]. Stable use of most concomitant background medications was permitted and, in the case of cholinesterase inhibitors and/or memantine, patients were required to be on a stable dose for a minimum of 4 weeks before screening with no adjustment of dosing during the double-blind phase of the study. Patients were excluded if they had a medical condition that might be a contributing cause of cognitive impairment.

Clinical study design. This was a multicentre, randomized, 12-month, double-blind, placebo-controlled, multiple-dose study of aducanumab followed by a 42-month, dose-blinded LTE study in patients with either prodromal or mild AD who were Aβ PET-positive (ClinicalTrials.gov identifier NCT01677572). The primary objective was to evaluate the safety and tolerability of multiple doses of aducanumab in patients with prodromal AD or mild AD dementia. The secondary objectives were to: (i) assess the effect on cerebral Aβ plaque content as measured by ^{18}F-florbetapir PET imaging at week 26; (ii) assess the multiple-dose serum concentrations of aducanumab; and (iii) evaluate the immunogenicity of aducanumab after multiple-dose administration. The key exploratory objectives were assessments of the effect of aducanumab on the following: the clinical progression of AD as measured by change from baseline on the CDR-SB, a NTB, and the FCSRT; disease-related biomarkers in blood, cerebral Aβ plaque content as measured by ^{18}F-florbetapir PET imaging at week 54; and cerebral Aβ plaque content by ApoE ε4 carrier status (carrier/non-carrier). Other exploratory endpoints were change from baseline on the Neuropsychiatric Inventory Questionnaire, Cognitive Drug Research computerized test battery, volumetric MRI, and, in a subset of patients, glucose metabolism as measured by fluorodeoxyglucose PET, functional connectivity by task-free functional MRI, cerebral blood flow by arterial spin labelling MRI, and disease-related biomarkers in cerebrospinal fluid. MMSE was another exploratory assessment.

During the 12-month, double-blind, placebo-controlled phase, patients received aducanumab or placebo by IV infusion once every 4 weeks for 52 weeks. In a staggered, parallel-group design, the treatment arms were enrolled as follows: first Arms 1–3 (aducanumab 1 mg kg^{-1} ($n=30$); aducanumab 3 mg kg^{-1} ($n=30$); placebo ($n=20$), respectively) in parallel. Once enrolment was open, Arms 4 and 5 (aducanumab up to 10 mg kg^{-1} ($n=30$) (actual dose 10 mg kg^{-1}); placebo ($n=10$), respectively) were enrolled in parallel with Arms 1–3. Once enrolment in Arms 1–5 was complete, enrolment in Arms 6 and 7 (aducanumab up to 30 mg kg^{-1} ($n=30$) (actual dose 6 mg kg^{-1}); placebo ($n=10$), respectively) began. The trial was initially designed to dose up to 30 mg kg^{-1}, but when ARIA were detected at 10 mg kg^{-1} it was decided not to proceed to doses higher than 10 mg kg^{-1} with repeated infusions. Dose escalation in Arms 4 and 5, and then Arms 6 and 7, was based on review of existing safety, tolerability, and pharmacokinetic data, and recommendation of the external Data Monitoring Committee. Patients were randomized (using a centralized interactive voice and web response System (IXRS)) to a treatment group within Arms 1–3, 4 and 5, or 6 and 7, stratified by ApoE ε4 status (carrier or non-carrier). Patient enrolment was monitored so that the ratio of ApoE ε4 carriers to non-carriers was no more than 2:1 and no less than 1:2. During the overlap in enrolment of Arms 1–3 and Arms 4 and 5, patients were randomized using a minimization algorithm. Patients who discontinued study treatment for any reason were encouraged to remain in the study and complete all assessments during the double-blind period. Patients completing the double-blind period and meeting certain eligibility criteria entered the LTE. After enrolment on Arms 6 and 7 were completed, the protocol was amended to include a titration arm and a corresponding placebo group—Arms 8 and 9. Both the LTE and Arms 8 and 9 are ongoing and were not part of this interim analysis.

Investigators, study site staff (except for a designated pharmacist/technician), and study patients were blinded to the patients' randomized treatment assignment for the placebo-controlled period. Only the designated pharmacist/technician at each site was aware of the assigned treatment for each patient. Aducanumab was supplied as a sterile clear-to-yellow solution for IV infusion at a dose of 200 mg in 4 ml. For patients randomized to receive aducanumab, undiluted aducanumab (required volume based on patient weight) was added to a 100 ml 0.9% saline bag to reach the assigned dose (an equivalent amount of saline was first withdrawn so that the final total volume of all IV bags was identical). All IV bags (active and placebo (100 ml 0.9% saline)) were covered with a sealed brown light-protective bag to maintain blinding with a label including protocol and patient randomization number.

Cases of ARIA were managed in accordance with protocol-defined rules using centrally read MRI findings coupled with clinical symptoms, if present. The rules were consistent with the guidelines published by the Alzheimer Association Research Roundtable Working Group[18]. Briefly, patients developing mild ARIA-E or ARIA-H (≤ 4 incident micro-haemorrhages) without clinical symptoms could continue at the same dose; patients developing moderate or severe ARIA-E without clinical symptoms, or those with ARIA-E accompanied by mild clinical symptoms, could suspend treatment and resume at the next lower dose level once ARIA (and symptoms, if any) resolved. Patients who developed ARIA-E or ARIA-H (≤ 4 incident micro-haemorrhages) accompanied by moderate, severe, or serious clinical symptoms, >4 incident micro-haemorrhages, any incident macro-haemorrhage, or >1 incident haemosiderosis at any time during the study were to permanently discontinue treatment.

The study was conducted in accordance with the Declaration of Helsinki, and the International Conference on Harmonisation and Good Clinical Practice guidelines, and had ethics committee approval at each participating site. All patients provided written informed consent.

Clinical study assessments. Amyloid plaque content, as measured by florbetapir PET imaging, was assessed at screening, and at weeks 26 and 54. Detailed PET scanning protocols have been described in a separate publication[9]. Briefly, for each florbetapir scan, a dose of 370 MBq was injected intravenously, with PET scanning starting around 50 min later and continuing for approximately 20 min.

Visual reads, the basis for meeting the inclusion criterion of a positive Aβ PET scan, were based upon PET image data, with the registered MRI and fused PET/MRI data providing supplementary anatomical information. Scans were independently interpreted by two board-certified neuroradiologists who, in accordance with the Amyvid Prescribing Information[37], had successfully completed a training programme (provided by the manufacturer using either an in-person tutorial or an electronic process). Images were designated as positive or negative, following guidelines described in the Amyvid Prescribing Information[37].

A composite cortical SUVR was computed using a volume-weighted average across six brain regions of interest (frontal, parietal, lateral temporal and sensorimotor, anterior, and posterior cingulate cortices), as previously described[16], normalized to whole cerebellar activity[10,38].

Clinical tests including the CDR and an NTB (comprising RAVLT Immediate and Delayed Recall, Wechsler Memory Scale Verbal Pair Associate Learning Test Immediate and Delayed Recall, Delis–Kaplan Executive Function System Verbal Fluency Conditions 1 and 2, and the Wechsler Adult Intelligence Scale Fourth Edition Symbol Search and Coding Subsets) were performed during screening and at weeks 26 and 54. The FCSRT was performed at screening and at week 52. These clinical tests were administered by a trained, certified clinician or rater experienced in the assessment of patients with cognitive deficits. When possible, the same rater would administer a given test across all visits. In order to maintain blinding to adverse events, raters were not permitted to perform other clinical assessments, and were blinded to other clinical and safety assessments. The rater who conducted the CDR for a patient could not complete any other rating scales for that same patient, and was blinded to the results of all other cognitive scales.

The following safety assessments were performed at regular intervals: physical examination, neurological examination, vital signs, electrocardiogram, and laboratory safety assessments. During the placebo-controlled period, brain MRI was performed at screening and at weeks 6, 18, 30, 42, and 54, and end of study or termination. The MMSE was completed at screening, and at weeks 24, 52, and end of study or termination, and, in patients who developed ARIA, at every scheduled visit until ARIA resolved.

The concentrations of aducanumab in serum and presence of anti-aducanumab antibodies were determined using validated ELISA techniques (Supplementary Information).

Statistical analysis in the clinical study. This interim analysis included all patients randomized to a fixed-dose regimen and completing the double-blind period of the study (data cut-off February 2015). For all analyses, all patients assigned to placebo were treated as a single group. The safety population was defined as all patients who were randomized and received at least one dose of study treatment. Adverse events were coded using the Medical Dictionary for Regulatory Activities

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

classification. The pharmacodynamic and pharmacokinetic populations were defined as all patients who were randomized, received at least one dose of study treatment, and had at least one post-baseline assessment of the pharmacodynamic parameter or at least one measurable aducanumab concentration in serum, respectively.

The primary analysis of the pharmacodynamic and efficacy data was based on Analysis of Covariance (ANCOVA), adjusting for baseline and ApoE ε4 status (carrier and non-carrier) using observed data. No imputation was performed for missing data. For each time point, adjusted means for each treatment, pairwise adjusted differences with placebo, 95% confidence intervals for the pairwise differences, and associated nominal *P* values for comparison were calculated. No adjustments were made for multiple comparisons/multiple interim analyses. Dose–response was tested using a linear contrast from the ANCOVA model. The linear contrast test is sensitive to a variety of positive dose–response shapes, including a linear dose–response relationship. This served as the primary analysis for the cognition analyses. To account for missing data, a MMRM was used as a sensitivity analysis for the longitudinal data change from baseline data, adjusting for baseline and ApoE ε4 status (carrier and non-carrier). Visit and treatment group were treated as categorical variables in the model along with their interactions. An unstructured covariance matrix was assumed to model the within-patient variability. This model imposes no assumptions on mean trend and correlation structure, and is considered robust.

Subgroup analyses were performed for change from baseline Aβ PET and change from baseline for cognition measures (CDR-SB and MMSE) for baseline clinical stage and ApoE ε4 status (carrier and non-carrier). The subgroup analysis of the pharmacodynamic and efficacy data was based on ANCOVA, adjusting for baseline and ApoE ε4 status (carrier and non-carrier) (for baseline clinical stage only) using observed data.

Serum pharmacokinetics were determined by nonlinear mixed effects model (NONMEM) approach. Sparse samples in the multiple-ascending-dose study and intensive samples from an earlier single-ascending-dose study[8] were combined to construct a population pharmacokinetic model. The model was built in NONMEM software using the first-order conditional estimation with interaction method. Cumulative AUC up to month 12 was estimated for each patient. The plasma terminal elimination half-life was estimated in the pharmacokinetic analysis population. The analysis population for the primary immunogenicity analysis was defined as all patients who were randomized, received study treatment, and had at least one post-dose immunogenicity sample evaluated for immunogenicity.

Interim analyses were specified in the protocol for the purpose of planning future studies; no changes were to be made for this study based on the interim analysis results.

A sample size of 30 patients per treatment group would provide more than 90% power to detect a treatment difference of 1 standard deviation with respect to the reduction of Aβ from baseline, based on comparison of each aducanumab group with placebo, at a two-sided significance level of 0.05, and assuming a dropout rate of 20%.

Transgenic mouse studies. Penetration of aducanumab into the brain and target engagement were assessed in 22-month-old female Tg2576 mice following a single dose of aducanumab at 30 mg kg^{-1} administered i.p. ('Target engagement study'; $n = 4$–5 per time point). The ability of aducanumab to reduce Aβ burden was assessed following chronic treatment of 9-month-old male and female Tg2576 transgenic mice dosed weekly i.p. for 6 months with PBS or 0.3, 1, 3, 10, or 30 mg kg^{-1} of the murine chimaeric variant ^{ch}aducanumab ('Chronic efficacy study'; $n = 20$–55 per treatment group). An additional dosing study ('Chronic efficacy study with Agly'; $n = 12$–14 per treatment group) comparing the plaque clearing ability of ^{ch}aducanumab to that of an effector function-impaired variant (^{ch}aducanumab-Agly) was conducted using a similar study design (chronic treatment of 9.5-month-old Tg2576 transgenic mice dosed weekly i.p. for 6 months with PBS or 3 mg kg^{-1} of ^{ch}aducanumab or ^{ch}aducanumab-Agly).

Mice were killed following anaesthesia with ketamine/xylazine (100/10 mg kg^{-1} i.p.). Blood was collected by cardiac puncture, and mice were perfused with ice-cold heparinized saline (0.9%) using a peristaltic pump. The brain was removed and halved along the medio-sagittal line. The right hemisphere was frozen on dry ice and stored at −80 °C for biochemical analysis. The left hemisphere was fixed by immersion in 10% neutral buffered formalin.

Size of the treatment groups was determined to take into account natural mortality (10–20%) and high inter-animal variability specific to the Tg2576 strain of mice. No animals were excluded from the analyses, unless the animal died prematurely. '*n*' reported in the manuscript represents the number of animals in each group that were euthanized as scheduled at the end of the study. The allocation of animals to treatment groups took into account date of birth, gender, and weight at baseline. Each treatment group was balanced for mean age, gender, and mean weight. Dosing solutions were coded with letters so that all experimenters were blinded to the treatment. The labelling of the samples collected did not reflect treatment group, so that experimenters processing and analysing the samples were still blinded. Codes were broken once all analyses were completed, including statistical analysis.

All in-life procedures were conducted in strict accordance with protocols approved by Biogen's Institutional Animal Care and Use Committee.

Biochemical measurements. Please see Supplementary Information.

Histological assessment. Please see Supplementary Information.

Preparation of different Aβ peptide conformations. Synthetic Aβ1–42 (Aβ_{42}) peptide (AnaSpec, Fremont, California, USA) was reconstituted in hexafluoroisopropanol at a concentration of 1 mg/ml, aliquoted, air-dried, and vacuum-concentrated to form a film, and dissolved in dimethyl sulfoxide (DMSO) at a concentration of 5 mg/ml. Aβ_{42} oligomers and Aβ_{42} fibrils were prepared by diluting DMSO-reconstituted monomeric into PBS at a concentration of 100 μg/ml and incubating at 37 °C for at least 3 days and 1 week, respectively. The solution was centrifuged at 14,000*g* for 15 min at 4 °C, and oligomers were recovered from the supernatant following the shorter incubation, whereas fibrils were recovered from the pellet following the longer incubation. For details on the biophysical characterization of high molecular weight Aβ_{42} aggregates, please see Supplementary Information.

In immunoprecipitation experiments, samples of freshly prepared monomeric, soluble oligomeric, or insoluble fibrillar Aβ_{42} were immunoprecipitated with ^{ch}aducanumab, 3D6 or a murine IgG2a control antibody (P1.17), dot-blotted onto a nitrocellulose membrane, and detected with biotinylated pan-Aβ antibody 6E10. Similar results were observed for ^{ch}aducanumab when immunoblotted with 3D6.

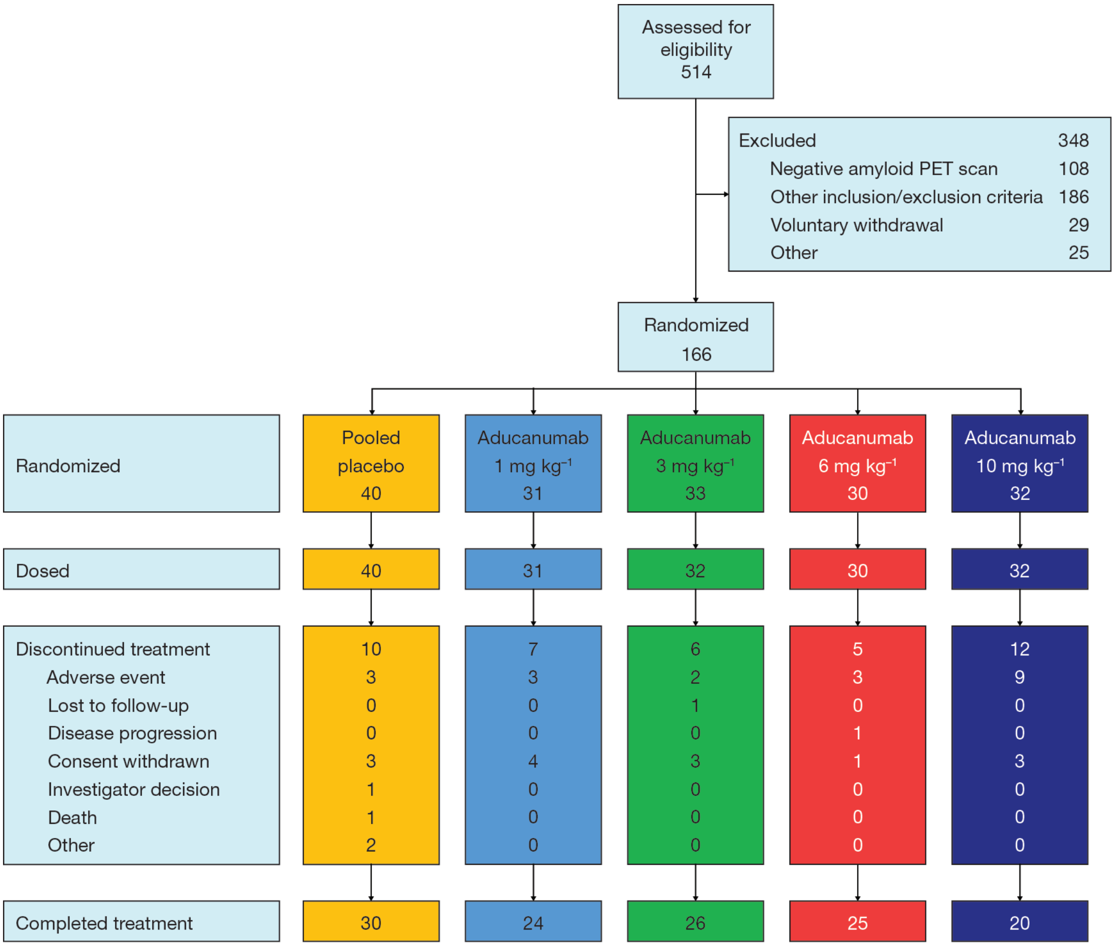
ELISA. Please see Supplementary Information.

Antibody generation using reverse translational medicine. Aducanumab was derived from a de-identified blood lymphocyte library collected from healthy elderly subjects with no signs of cognitive impairment and cognitively impaired elderly subjects with unusually slow cognitive decline. Memory B cells, isolated from peripheral blood lymphocyte preparations by anti-CD22-mediated sorting were cultured on gamma-irradiated human peripheral blood mononuclear cell feeder layers. Supernatants from isolated B cells were screened for their ability to stain Aβ plaques on brain tissue sections, from either patients with AD or aged APP transgenic mice[39], and for their binding to aggregated forms of Aβ_{40} and Aβ_{42} *in vitro*. Positive hits meeting the above criteria were counter-screened to exclude clones cross-reacting with full-length APP expressed on stably transfected HEK293 cells (provided by U. Konietzko, University of Zurich, Switzerland; tested negative for mycoplasma contamination; not independently authenticated). Selected Aβ-reactive B-cell clones were subjected to cDNA cloning of IgG heavy and κ or λ light chain variable region sequences, and sub-cloned in expression constructs using Ig-framework specific primers for human variable heavy and light chain families in combination with human J-H segment-specific primers. Aducanumab was engineered to incorporate glycosylated human IgG1 heavy and human κ light chain constant domain sequences. A murine chimaeric IgG2a/κ version of aducanumab (^{ch}aducanumab) was generated for use in chronic efficacy studies in APP transgenic mice. An aglycosylated variant of ^{ch}aducanumab (^{ch}aducanumab-Agly), incorporating a single point mutation (N297Q, using standard Kabat EU numbering) which eliminates *N*-glycosylation of the Fc region and severely reduces FcγR binding[40], was generated to test for Fc-related activities. The recombinant mouse IgG2b monoclonal antibody 3D6[41] was used as a comparator in some studies.

***Ex vivo* phagocytosis assay.** Please see Supplementary Information.

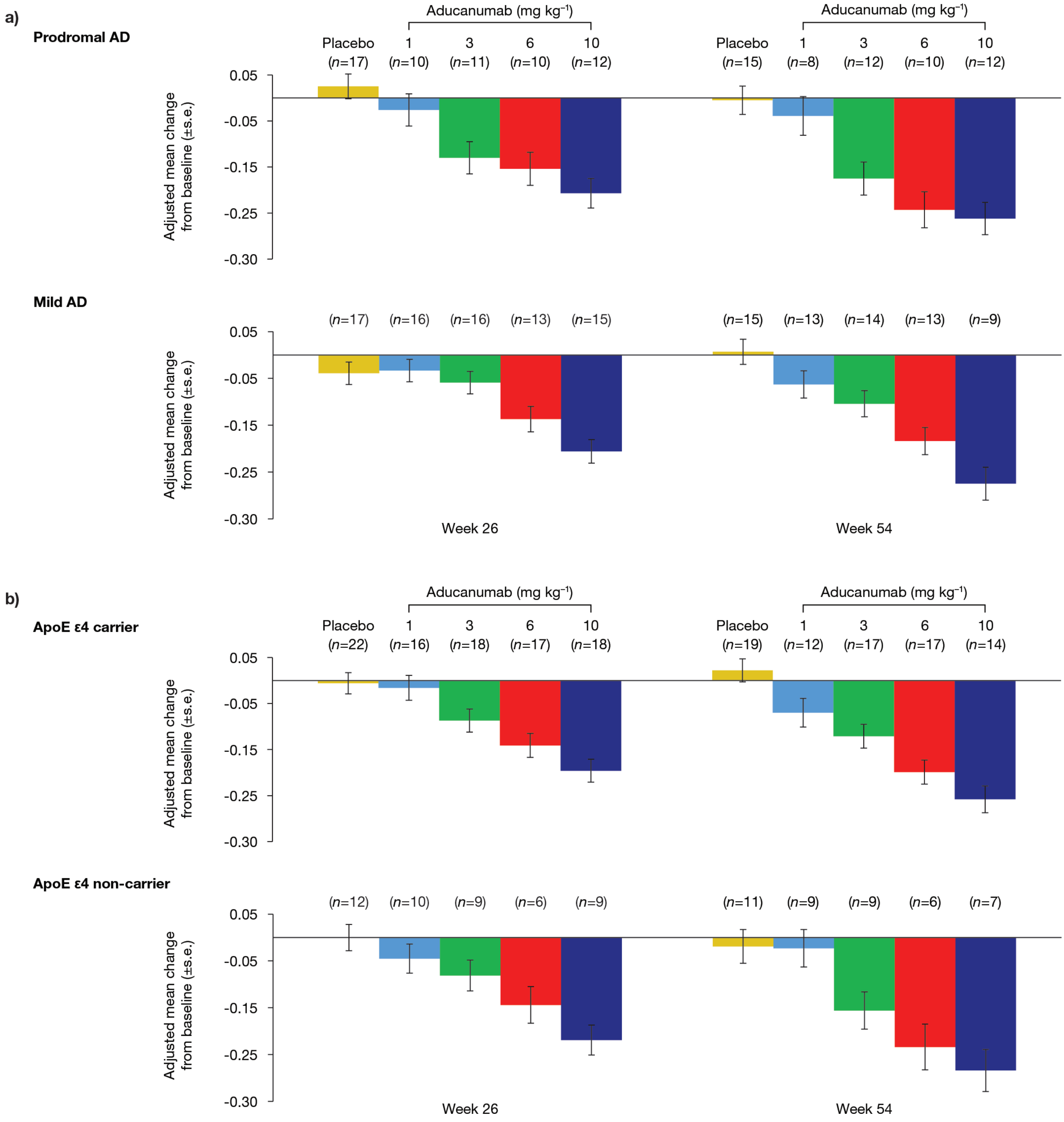
34. Derby, C. A. *et al.* Screening for predementia AD: time-dependent operating characteristics of episodic memory tests. *Neurology* **80,** 1307–1314 (2013).
35. Dubois, B. *et al.* Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol.* **9,** 1118–1127 (2010).
36. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* **7,** 263–269 (2011).
37. Eli Lilly and Company. Amyvid Prescribing Information. http://www.lilly.com. (2013).
38. Clark, C. M. *et al.* Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *Lancet Neurol.* **11,** 669–678 (2012).
39. Hock, C. *et al.* Generation of antibodies specific for beta-amyloid by vaccination of patients with Alzheimer disease. *Nat. Med.* **8,** 1270–1275 (2002).
40. Tao, M. H. & Morrison, S. L. Studies of aglycosylated chimeric mouse-human IgG. Role of carbohydrate in the structure and effector functions mediated by the human IgG constant region. *J. Immunol.* **143,** 2595–2601 (1989).
41. Johnson-Wood, K. *et al.* Amyloid precursor protein processing and A beta42 deposition in a transgenic mouse model of Alzheimer disease. *Proc. Natl Acad. Sci. USA* **94,** 1550–1555 (1997).

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.


Assessed for eligibility
514
Excluded 348
Negative amyloid PET scan 108
Other inclusion/exclusion criteria 186
Voluntary withdrawal 29
Other 25
Randomized
166
Randomized
Pooled placebo 40
Aducanumab 1 mg kg^{-1} 31
Aducanumab 3 mg kg^{-1} 33
Aducanumab 6 mg kg^{-1} 30
Aducanumab 10 mg kg^{-1} 32
Dosed
40
31
32
30
32
Discontinued treatment
Adverse event
Lost to follow-up
Disease progression
Consent withdrawn
Investigator decision
Death
Other
10 3 0 0 3 1 1 2
7 3 0 0 4 0 0 0
6 2 1 0 3 0 0 0
5 3 0 1 1 0 0 0
12 9 0 0 3 0 0 0
Completed treatment
30
24
26
25
20

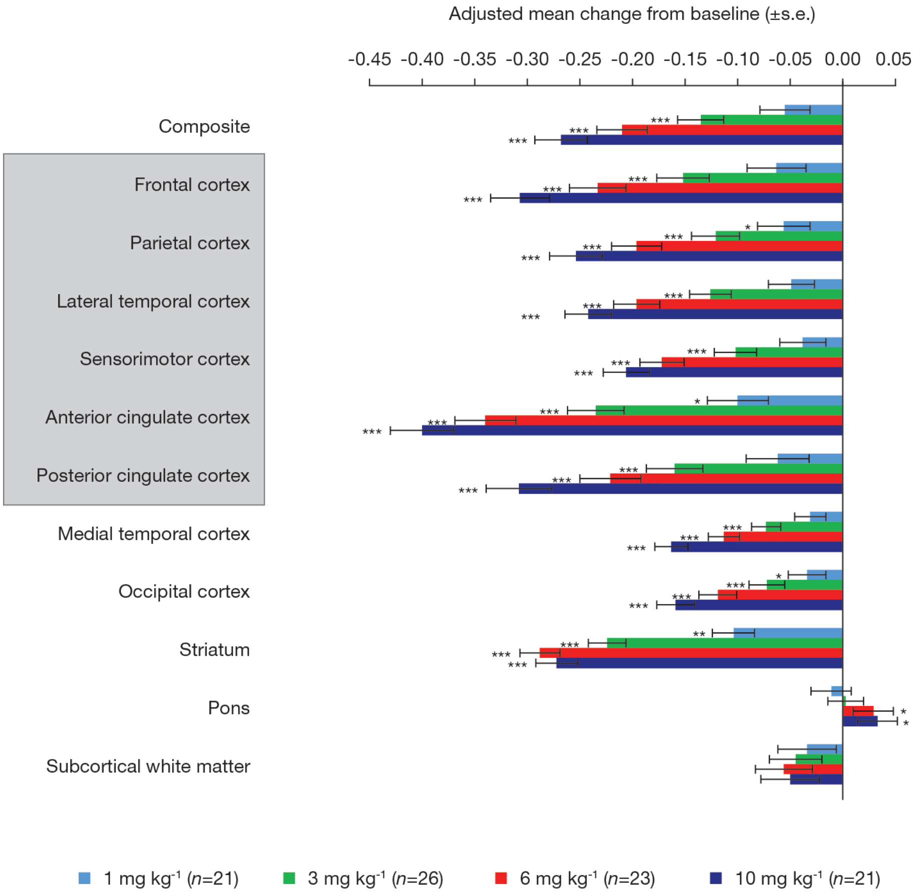
Extended Data Figure 1 | Participant accounting. PET, positron emission tomography.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



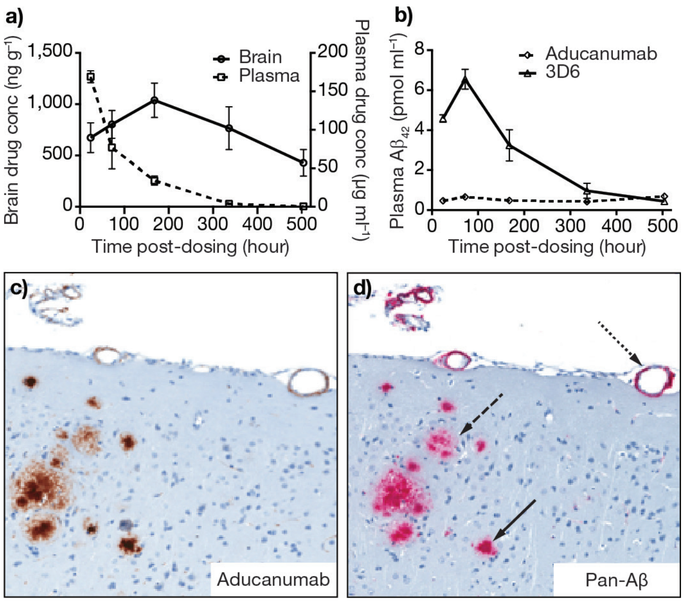
Extended Data Figure 2 | Amyloid plaque reduction with aducanumab by baseline clinical stage and baseline ApoE ε4 status. a, **b**, Analyses by baseline clinical stage were performed using ANCOVA for change from baseline with factors of: treatment, ApoE ε4 status (carrier and non-carrier) and baseline composite SUVR (**a**), and for analyses by ApoE ε4 status, using treatment and baseline composite SUVR (**b**). Adjusted mean ± s.e. ApoE ε4, apolipoprotein E ε4 allele; SUVR, standard uptake value ratio.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



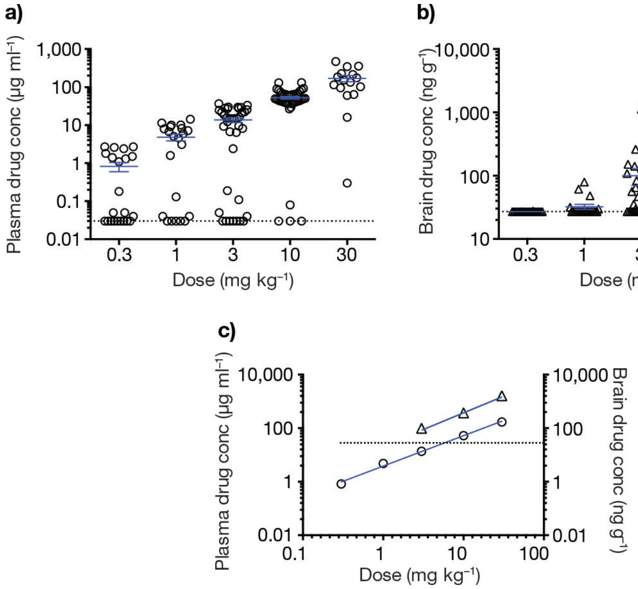
Extended Data Figure 3 | Amyloid plaque reduction: regional analysis SUVR at week 54. The boxed area indicates the six regions included in the composite score. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$ versus placebo; two-sided tests with no adjustments for multiple comparisons. Adjusted mean ± s.e. Analyses using ANCOVA. SUVR, standard uptake value ratio.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



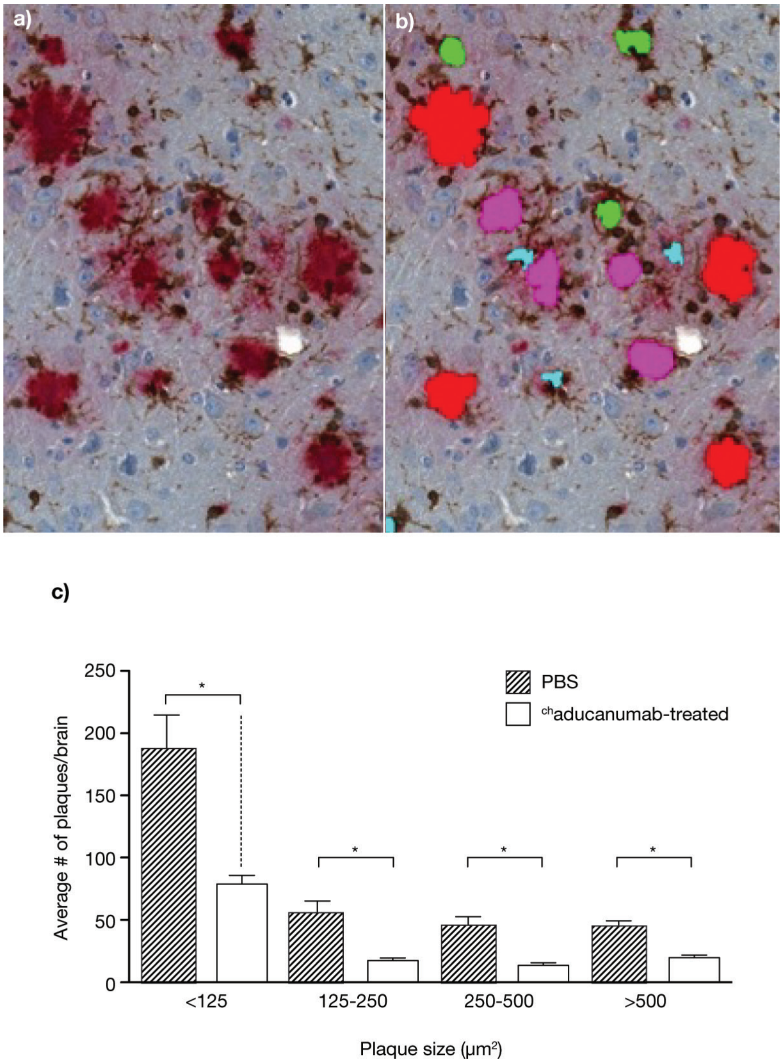
Extended Data Figure 4 | Brain penetration of aducanumab after a single intraperitoneal administration in 22-month-old Tg2576 transgenic mice. a, **b**, Aducanumab levels in plasma and brain (**a**), and plasma Aβ levels after a single dose (**b**; $n = 4$–5; mean ± s.e.). **c**, **d**, *In vivo* binding of aducanumab to amyloid deposits detected using a human IgG-specific secondary antibody (**c**), and *ex vivo* immunostaining with a pan-Aβ antibody on consecutive section (**d**). Examples of a compact Aβ plaque (solid arrow), diffuse Aβ deposit (dashed arrow), and CAA lesion (dotted arrow). CAA, cerebral amyloid angiopathy.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



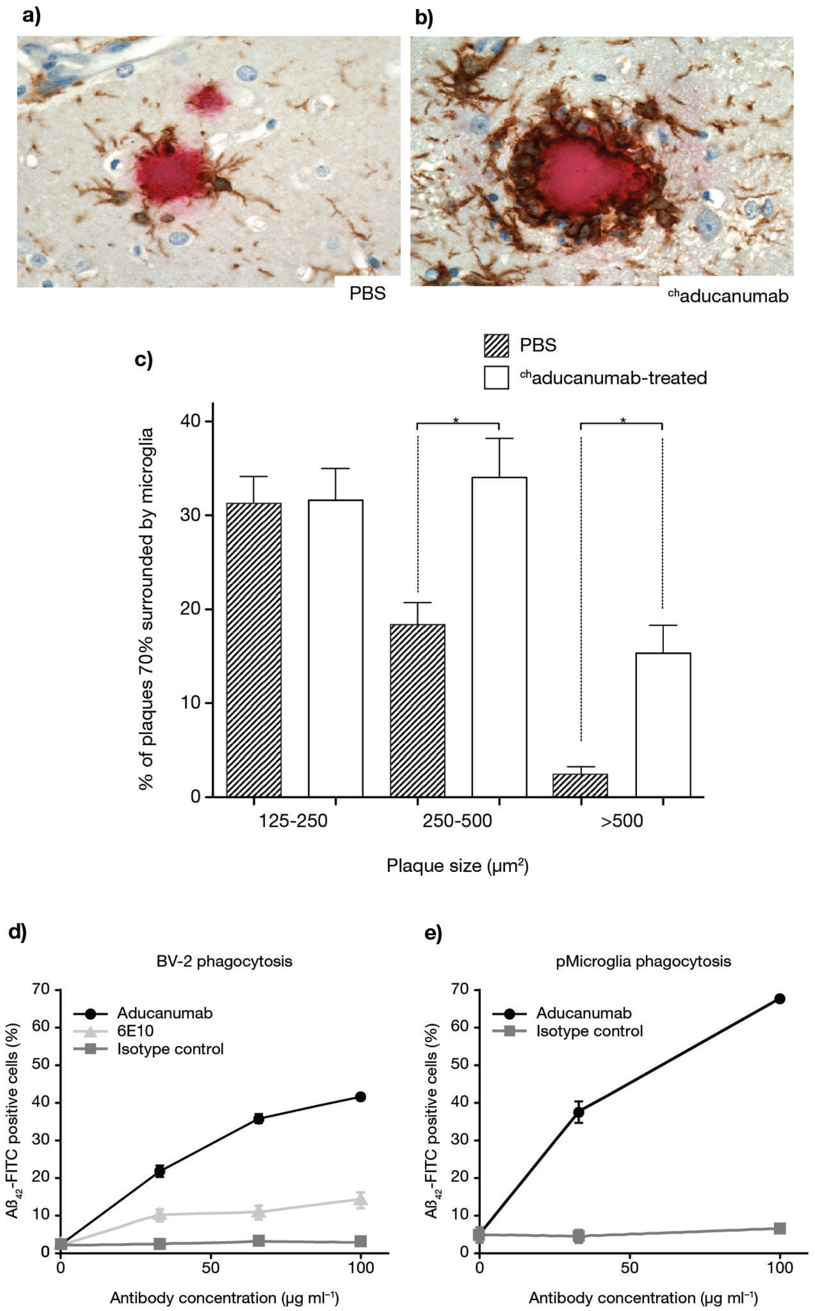
Extended Data Figure 5 | Exposure following weekly dosing with ^{ch}aducanumab in 9.5- to 15.5-month-old Tg2576 transgenic mice. **a**, **b**, ^{ch}aducanumab concentrations in plasma (**a**), or DEA-soluble brain extract (**b**) were measured in samples collected 24 h after the last dose in the 'Chronic efficacy study'. Mean ± s.e. Dotted lines represent the limits of quantitation of each assay. **c**, Correlations of drug concentrations in plasma (open circles) or brain (open triangles) with administered dose. The average brain concentrations in the two groups receiving the lowest dose were below the limit of quantitation for that assay, which is indicated by a dotted line on the figure.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



Extended Data Figure 6 | Treatment with ^{ch}aducanumab affects plaques of all sizes. a, Following weekly dosing of ^{ch}aducanumab in Tg2576 from 9.5–15.5 months of age, amyloid plaques were stained with 6E10 and quantified using Visiopharm software. **b**, Plaque size was defined by area, and coloured as follows: <125 μm^2 (cyan), 125–250 μm^2 (green), 250–500 μm^2 (pink), and >500 μm^2 (red). **c**, ^{ch}aducanumab treatment was associated with a significant decrease in plaque number in all size ranges relative to vehicle-treated controls, with reductions of 58%, 68%, 68%, and 53% in the number of plaques for the <125 μm^2, 125–250 μm^2, 250–500 μm^2, and >500 μm^2 groups size, respectively. Mean ± s.e.; statistically significant differences from vehicle for each size range are indicated with asterisks; $^*P < 0.05$, Mann–Whitney test.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



Extended Data Figure 7 | Enhanced recruitment of microglia to amyloid plaques following $^{\text{ch}}$aducanumab treatment and engagement of Fcγ receptors. a, **b**, Brain sections from either PBS- or $^{\text{ch}}$aducanumab-treated mice (‘Chronic efficacy study’; 3 mg kg^{-1} group) were immunostained for Aβ (6E10; red) and a marker of microglia (Iba1; brown). **c**, The area of individual amyloid plaques was measured, and Iba1-stained microglia were grouped into two categories, either associated with plaques (within 25 µm of a plaque) or not associated with plaques (>25 µm from a plaque). Plaques with circumferences ≥ 70% surrounded by microglia were quantified and stratified based on plaque size. The fraction of plaques that were at least 70% surrounded by microglia was significantly greater in the $^{\text{ch}}$aducanumab-treated group (white bars) compared with the PBS control group (grey bars), for plaques ≥250 μm^2. Mean ± s.e.; statistically significant differences from vehicle for each size range are indicated with asterisks; $*P < 0.05$, Bonferroni’s post hoc test following one-way analysis of variance. All quantifications were done using the Visiopharm software. **d**, **e**, FITC-labelled $\text{A}\beta_{42}$ fibrils were incubated with different concentrations of the antibodies before adding to BV-2 microglia cell line (**d**), or primary microglia (**e**) for phagocytosis experiment measuring uptake of $\text{A}\beta_{42}$ fibrils into the cells by FACS analysis. Mean ± s.d.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 1 | Change from baseline in amyloid PET SUVR values (a secondary endpoint at 6 months), and in exploratory clinical endpoints at the end of the placebo-controlled period (6-month data also shown for amyloid PET)

Adjusted mean ± SE change from baseline for:	Placebo	Aducanumab 1 mg kg^{-1}	 3 mg kg^{-1}	 6 mg kg^{-1}	 10 mg kg^{-1}	p-value (dose-response)
Amyloid PET SUVR values						
At 6 months	(*n*=34)	(*n*=26)	(*n*=27)	(*n*=23)	(*n*=27)	
	–0.005 ± 0.018	–0.030 ± 0.020	–0.087 ± 0.020**	–0.143 ± 0.022***	–0.205 ± 0.020***	<0.0001
At 1 year[†]	(*n*=30)	(*n*=21)	(*n*=26)	(*n*=23)	(*n*=21)	
	0.003 ± 0.021	–0.055 ± 0.024	–0.135 ± 0.022***	–0.210 ± 0.024***	–0.268 ± 0.025***	<0.0001
CDR-SB[†]	(*n*=31)	(*n*=23)	(*n*=27)	(*n*=26)	(*n*=23)	
	1.87 ± 0.41	1.72 ± 0.46	1.37 ± 0.43	1.11 ± 0.44	0.63 ± 0.47*	<0.05
MMSE[‡]	(*n*=32)	(*n*=25)	(*n*=26)	(*n*=26)	(*n*=25)	
	–2.81 ± 0.67	–2.18 ± 0.75	–0.70 ± 0.75*	–1.96 ± 0.75	–0.56 ± 0.76*	<0.05
NTB overall Z score[†]	(*n*=29)	(*n*=23)	(*n*=26)	(*n*=24)	(*n*=24)	
	–0.11 ± 0.08	–0.25 ± 0.09	–0.13 ± 0.08	–0.19 ± 0.09	–0.10 ± 0.09	NS
FCSRT: sum of free recall score[‡]	(*n*=31)	(*n*=23)	(*n*=25)	(*n*=25)	(*n*=25)	
	–2.33 ± 1.07	–1.63 ± 1.24	–1.25 ± 1.20	–4.04 ± 1.21	–0.69 ± 1.20	NS

$^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ versus placebo; two-sided tests with no adjustments for multiple comparisons.

[†] At week 54.

[‡] At week 52.

Analyses using ANCOVA. ApoE ε4, apolipoprotein E ε4 allele; CDR-SB, Clinical Dementia Rating—Sum of Boxes; FCSRT, Free and Cued Selective Reminding Test; MMSE, Mini-Mental State Examination; NS, not significant; NTB, neuropsychological test battery; SE, standard error; SUVR, standard uptake value ratio.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 2 | Incidence of ARIA based on MRI data and ARIA-E patient disposition

		Aducanumab			
	Placebo	1 mg kg^{-1}	3 mg kg^{-1}	6 mg kg^{-1}	10 mg kg^{-1}
Number of dosed subjects with at least one post-baseline MRI	38	31	32	30	32
ApoE ε4 carrier	24	19	21	21	20
ApoE ε4 non-carrier	14	12	11	9	12
ARIA-E, n (%)	0	1 (3)	2 (6)	11 (37)	13 (41)
By ApoE ε4					
ApoE ε4 carrier	0	1 (5)	1 (5)	9 (43)	11 (55)
ApoE ε4 non-carrier	0	0	1 (9)	2 (22)	2 (17)
ARIA-E and:					
Continued treatment	0	0	2 (6)	8 (27)	5 (16)
Same dose	0	0	0	2 (7)	0
Dose reduced	0	0	2 (6)	6 (20)	5 (16)
Discontinued treatment	0	1 (3)	0	3 (10)	8 (25)
ApoE ε4 carrier	0	1 (5)	0	2 (10)	7 (35)
ApoE ε4 non-carrier	0	0	0	1 (11)	1 (8)
Isolated ARIA-H, n (%)	2 (5)	2 (6)	3 (9)	0	2 (6)
ApoE ε4 carrier	2 (8)	1 (5)	2 (10)	0	2 (10)
ApoE ε4 non-carrier	0	1 (8)	1 (9)	0	0
ARIA-E and ARIA-H, n (%)	0	1 (3)	1 (3)	5 (17)	8 (25)
ApoE ε4 carrier	0	1 (5)	1 (5)	5 (24)	7 (35)
ApoE ε4 non-carrier	0	0	0	0	1 (8)

ApoE ε4, apolipoprotein E ε4 allele; ARIA, amyloid-related imaging abnormalities; ARIA-E (oedema); ARIA-H (micro-haemorrhages, macro-haemorrhages, or superficial siderosis); MRI, magnetic resonance imaging.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 3 | Pharmacokinetic data

	Aducanumab			
	1 mg kg^{-1}	3 mg kg^{-1}	6 mg kg^{-1}	10 mg kg^{-1}
PK analysis population (intent-to-treat)*	*n*=31	*n*=32	*n*=30	*n*=32
Cumulative AUC (µg.h/mL, mean ± SD)	47,078 ± 17,555	143,395 ± 59,986	251,535 ± 122,883	346,163 ± 198,603
Subjects who received all 14 planned doses	*n*=18	*n*=18/19†	*n*=16	*n*=14
$C_{max\ ss}$ (µg/mL, mean ± SD)‡	21.2 ± 3.7	59.6 ± 19.6	123.8 ± 42.5	250.8 ± 33.5
Cumulative AUC (µg.h/mL, mean ± SD)	55,223 ± 11,529	169,457 ± 41,775	315,352 ± 76,300	524,511 ± 95,622

*Data include patients who missed doses.
†A total of 19 patients received all 14 doses but 1 patient missed the concentration measurement at Week 40 and so $n = 18$ for $C_{max,ss}$ at 3 mg kg^{-1} aducanumab.
‡The observed post-infusion concentrations at Week 40 were reported as steady-state C_{max}.
AUC, area under the concentration curve; $C_{max,ss}$, maximum concentration at steady state; PK, pharmacokinetic; SD, standard deviation.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 4 | Change from baseline in amyloid PET SUVR values, CDR-SB, and MMSE at the end of the placebo-controlled period by absence/presence[*] of ARIA-E

	Treatment group (# without ARIA-E, # with ARIA-E)				
		Aducanumab			
Adjusted mean ± SE for:	Placebo	1 mg kg^{-1}	3 mg kg^{-1}	6 mg kg^{-1}	10 mg kg^{-1}
Amyloid PET SUVR values[†]	(30, 0)	(21, 0)	(24, 2)	(17, 6)	(13, 8)
ARIA-E					
Absence	0.003 ± 0.020	–0.056 ± 0.024	–0.141 ± 0.023	–0.243 ± 0.027	–0.278 ± 0.031
Presence	0.001 ± 0.020	–	–0.069 ± 0.075	–0.114 ± 0.049	–0.263 ± 0.040
CDR-SB[†]	(31, 0)	(23, 0)	(25, 2)	(18, 8)	(14, 9)
ARIA-E					
Absence	1.84 ± 0.42	1.72 ± 0.48	1.33 ± 0.47	1.11 ± 0.54	0.78 ± 0.61
Presence	1.95 ± 0.35	–	2.04 ± 1.38	1.18 ± 0.73	0.67 ± 0.67
MMSE[‡]	(32, 0)	(25, 0)	(24, 2)	(18, 8)	(16, 9)
ARIA-E					
Absence	–2.86 ± 0.69	–2.20 ± 0.77	–0.47 ± 0.80	–1.82 ± 0.91	–1.05 ± 0.96
Presence	–2.60 ± 0.69	–	–3.41 ± 2.69	–1.95 ± 1.42	0.83 ± 1.35

[*] Since there were no ARIA-E events in the placebo group, the overall placebo group was used as the comparator in the subgroup analysis for presence of ARIA-E.

[†] At week 54.

[‡] At week 52.

Analyses based on observed data. Adjusted mean change and standard errors are based on an ANCOVA model for change from baseline with factors of treatment, laboratory ApoE ε4 status (carrier and non-carrier), and baseline composite SUVR, CDR-SB, or MMSE, respectively. ARIA-E, amyloid-related imaging abnormalities (oedema); CDR-SB, Clinical Dementia Rating—Sum of Boxes; MMSE, Mini-Mental State Examination; PET, positron emission tomography; SE, standard error; SUVR, standard uptake value ratio.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

ADDENDUM

doi:10.1038/nature22809

Addendum: The antibody aducanumab reduces Aβ plaques in Alzheimer's disease

Jeff Sevigny, Ping Chiao, Thierry Bussière, Paul H. Weinreb, Leslie Williams, Marcel Maier, Robert Dunstan, Stephen Salloway, Tianle Chen, Yan Ling, John O'Gorman, Fang Qian, Mahin Arastu, Mingwei Li, Sowmya Chollate, Melanie S. Brennan, Omar Quintero-Monzon, Robert H. Scannevin, H. Moore Arnold, Thomas Engber, Kenneth Rhodes, James Ferrero, Yaming Hang, Alvydas Mikulskis, Jan Grimm, Christoph Hock, Roger M. Nitsch & Alfred Sandrock

Nature **537**, 50–56 (2016); doi:10.1038/nature19323

Figure 1 of our original Article illustrated that treatment with aducanumab reduced human brain amyloid-β plaques in a dose-dependent fashion as measured by florbetapir positron emission tomography (PET) imaging. The figure gave the visual appearance of standard uptake value ratio (SUVR) reduction in subcortical white matter as well as cortical regions, although statistically validated evidence of dose-dependent SUVR reduction was demonstrated only in cortical regions. We provide an updated figure (Fig. 1 of this Addendum), which includes colour bars and difference images to aid in the understanding and interpretation of the representative florbetapir PET images. An additional panel on the right illustrates the differences between baseline and week 54 images, computed by simple subtraction of the baseline from follow-up images, after co-registration to a common coordinate system. The difference images show that the SUVR reduction (which is unitless) occurs primarily in the cortical regions (highlighted in red) in patients treated with aducanumab.

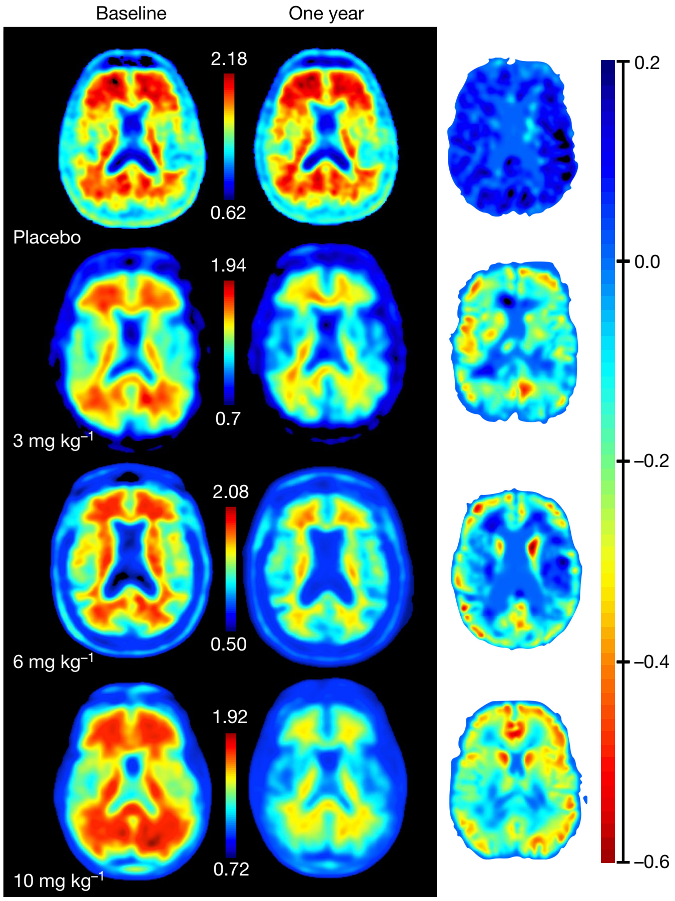


Figure 1 | This is the updated Fig. 1 of the original Article.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Journal of Alzheimer's Disease 44 (2015) 549–560
DOI 10.3233/JAD-141626
IOS Press

A Clinical Study of Lupron Depot in the Treatment of Women with Alzheimer's Disease: Preservation of Cognitive Function in Patients Taking an Acetylcholinesterase Inhibitor and Treated with High Dose Lupron Over 48 Weeks

Richard L. Bowen[a,*], George Perry[b,c], Chengjie Xiong[d], Mark A. Smith[c,†] and Craig S. Atwood[e,f,g,*]
[a] *OTB Research, Charleston, SC, USA*
[b] *UTSA Neurosciences Institute and Department of Biology, University of Texas at San Antonio, San Antonio, TX, USA*
[c] *Department of Pathology, School of Medicine, Case Western Reserve University, Cleveland, OH, USA*
[d] *Department of Biostatistics, Washington University School of Medicine, St. Louis, MO, USA*
[e] *Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI, USA*
[f] *Geriatric Research, Education and Clinical Center, Veterans Administration Hospital, Madison, WI, USA*
[g] *School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, WA, Australia*

Handling Editor: Massimo Tabaton

Accepted 8 September 2014

Abstract. To test the efficacy and safety of leuprolide acetate (Lupron Depot®) in the treatment of Alzheimer's disease (AD), we conducted a 48-week, double-blind, placebo-controlled, dose-ranging study in women aged 65 years or older with mild to moderate AD. A total of 109 women with mild to moderate AD and a Mini-Mental State Examination score between 12 and 24 inclusive were randomized to low dose Lupron Depot® (11.25 mg leuprolide acetate), high dose Lupron Depot® (22.5 mg leuprolide acetate), or placebo injections every 12 weeks. There were no statistically significant differences in primary efficacy parameters (ADAS-Cog and ADCS-CGIC), although there was a non-statistically significant trend in favor of the high dose Lupron group on the ADAS-Cog. There were no statistically significant differences in secondary efficacy parameters (NPI, ADCS-ADL, BI, and ADCS-Severity Rating). However, in the *a priori* designated subgroup analysis of patients taking an

---

[†] Deceased.

[*] Correspondence to: Richard L. Bowen, MD, OTB Research, 217 Calhoun St, Unit 1, Charleston, SC 29401, USA. Tel.: +1 843 480 2273; Fax: +1 843 501 7698; E-mail: richard.bowen@yahoo.com. and Craig S. Atwood, PhD, University of Wisconsin-Madison School of Medicine and Public Health, William S. Middleton Memorial VA (GRECC 11 G), 2500 Overlook Terrace, Madison, WI 53705, USA. Tel.: +1 608 256 1901/Ext. 11664; Fax: +1 608 280 7291; E-mail: csa@medicine.wisc.edu.

ISSN 1387-2877/15/$27.50 © 2015 – IOS Press and the authors. All rights reserved

acetylcholinesterase inhibitor (AChEI), there was a statistically significant benefit in the high dose group compared to both the low dose and placebo groups as determined by ADAS-Cog (mean decline: 0.18, 4.21, and 3.30), ADCS-CGIC (% subjects experiencing decline: 38, 82, and 63), and ADCS-ADL (mean decline: −0.54, −8.00, and −6.85), respectively. No differences between treatment groups were seen on the NPI, ADCS-CGI Severity Rating, or the BI in the subgroup analysis. These data indicate that cognitive function is preserved in patients treated with high dose Lupron who were already using AChEIs. The positive interaction between Lupron and AChEIs warrants further investigation for the treatment of AD.

Keywords: 17β-estradiol, acetylcholinesterase inhibitor, Alzheimer's disease, apolipoprotein E, clinical trial, cognitive testing, gonadotropin-releasing hormone, Lupron, luteinizing hormone, women

INTRODUCTION

Age-related changes in hormones of the hypothalamic-pituitary-gonadal axis have been suggested as a major etiological factor in Alzheimer's disease (AD) [1–4]. In addition to the age-related decline in circulating sex steroids, there is evidence to suggest that simultaneous elevations in the circulating concentrations of gonadotropins and gonadotropin-releasing hormone (GnRH) at this time play a role in AD [5–8]. Evidence for suppressing GnRH and gonadotropin signaling in the treatment of AD comes from a growing number of epidemiological, preclinical and biological studies.

Compelling epidemiological data suggest that Lupron Depot® (otherwise referred to as Lupron) treatment decreases the risk for AD in men. The most frequent use of Lupron is for the treatment of prostate cancer. A study, utilizing the Medicare inpatient database, of men who underwent prostatectomy for prostate cancer ($n = 115{,}789$) found that the incidence of dementia within 5 years of the procedure date was ∼ 34–55% that of age-matched men undergoing a similar surgical procedure ($n = 433{,}736$) ([9] and Beaird, Bowen, Perry, Atwood et al., unpublished data). That GnRH agonist treatment was the cause of this dramatic decrease in AD incidence was verified by D'Amico and colleagues [10] who demonstrated a significant 55% reduction in the risk of death from AD in men with prostate cancer treated with a GnRH agonist compared with untreated patients.

Preclinical evidence for the use of Lupron in the treatment of AD comes from studies of both normal and amyloid-β protein precursor (AβPP)-transgenic models of AD. Suppression of gonadotropins with Lupron improves cognitive performance in aged AβPP-transgenic mice [2] while increases in luteinizing hormone (LH)/human chorionic gonadotropin (hCG) have been attributed to cognitive decline in ovariectomized rats [11], LHβ-transgenic mice [12], and ovariectomized C57/Bl6 mice [13]. Moreover, Lupron treatment has been shown to decrease amyloid-β (Aβ) production in C57/Bl6 mice [8] and Aβ load in aged AβPP-transgenic mice [2]. The role of LH in mediating AβPP processing was confirmed in a bigenic mouse model that expresses $\text{A}\beta\text{PPsw}^{+}$ in the background of a LH receptor (*Lhr*) knockout ($\text{A}\beta\text{PPsw}^{+}/\text{Lhr}^{-/-}$; [14]). Despite the ∼10-fold elevation in AβPP/Aβ production by $A\beta PPsw^{+}$ mice [15], genetic ablation of *Lhr* significantly reduced amyloid load and the total number of Aβ plaques in the hippocampus and cerebral cortex of male and female mice. Genetic ablation of *Lhr* in $A\beta PPsw^{+}$ mice also decreases tau phosphorylation by ∼50% that induced by AβPP overexpression in these mice [14].

Pathological and biochemical studies support the role of gonadotropins in amyloidosis and neurofibrillary tangle formation. LH/hCG promotes the processing of AβPP toward the amyloidogenic pathway *in vitro* [16]. LH induced an increase in the generation and secretion of Aβ, coupled with decreased secretion of AβPP and increased AβPPCT100 production in human neuroblastoma cells [8].

This clinical study was conducted as a dose-ranging study designed to investigate the efficacy and safety of Lupron in the treatment of individuals with AD. In order to minimize any effects due to the loss of sex steroids, it was decided to make this a woman only study since women in this age group are postmenopausal and have little if any endogenous sex steroid production. The study design, patient selection criteria, and outcome measures were guided by regulatory standards in clinical studies. We find that Lupron treatment in combination with acetylcholinesterase inhibitor (AChEI) halts or slows the progression of cognitive decline in women with mild-moderate AD.

METHODS

The study was conducted from April 16, 2003 through December 16, 2004. Participants were recruited from five U.S. sites. The institutional review board at each site (Baumel-Eisner Neuromedical Institute – three sites; Sun Health Research Institute; Meridien Research) reviewed and approved the study protocol. 109 patients were enrolled who met all of the following criteria: had given their consent by signing the Informed Consent Form and the responsible caregiver also had signed the consent form; or, if the patient was judged by the investigator to be unable to give consent, the legally authorized representative gave consent by signing the consent form and the patient gave assent, in accord with local regulations; were female; were 65 years of age or older; had a diagnosis of probable AD according to the National Institute of Neurological and Communicative Disorders-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria, and the investigator ascertained that the condition had been present at least 6 months prior to screening; were either presently taking a AChEI, and had begun taking it at least 90 days prior to baseline and, in the investigator's opinion, the dosage would likely remain stable throughout the study or they had never taken AChEIs or stopped taking such medication at least 90 days prior to baseline and would likely remain off AChEIs throughout the study; if they were taking other drugs or substances that have purported cognition enhancing properties such as *Ginkgo biloba* and vitamin E, they had begun taking it at least 60 days prior to baseline and, in the investigator's opinion, the dosage would remain stable throughout the study; had scored no lower than 12 or higher than 24 on the Mini-Mental State Examation (MMSE) administered at the screening visit; had a brain imaging study (CT scan, MRI, or PET) performed at the time of their initial diagnosis of AD or after that time, and the findings had been consistent with a diagnosis of probable AD (if a brain imaging study had not been performed, one was performed during the screening process); had a Rosen Modified Hachinski score of 4 or lower at the screening visit, supporting the investigator's clinical judgment that the patient's dementia was "probable AD" and not a dementia of vascular origin; were fluent in English or Spanish and had completed at least 6 years of education; lived at home or in a congregate living facility for requirements other than skilled nursing care, and had a caregiver who saw the patient at least three times a week for a total of at least 10 hours and could sign the consent form, provide information pertinent to the patient's cognitive status, accompany the patient on clinic visits, and participate in the evaluations; hormone replacement therapy, if any, had been stable for at least 60 days prior to baseline, and was not expected to change during the course of the study; scored less than 15 on the Hamilton Depression Scale (17-item version) administered as part of the screening evaluation; values on their screening laboratory tests did not indicate significant medical conditions that would have interfered with their participation in, and completion of, the study.

Exclusion criteria were: The presence of a significant neurological disease affecting the brain, or psychiatric disease other than AD, such as major depression, schizophrenia, epilepsy, Parkinson's disease, or stroke; current significant systemic illness or symptoms of organ failure; a screening electrocardiogram (ECG) that showed evidence of a serious and/or unstable condition or a recent (within 6 months) myocardial infarction; a history of cancer within the last 5 years, except for basal cell or squamous cell cancer, or cervical carcinoma *in situ*; receiving Coumadin or anti-Parkinsonian medications; receiving other investigational drugs within 30 days or 5 half-lives prior to randomization, whichever was longer; taking other medications known to affect serum gonadotropin concentrations, such as gosorelin or danazol, except for estrogen and/or progesterone; had a history of bone fracture secondary to low bone mineral density; had a history of osteoporosis/osteopenia, unless they were receiving therapy for osteoporosis/osteopenia for at least 3 weeks prior to baseline, and the treatment regimen was expected to remain stable; abuse or dependence on alcohol or other substances satisfied criteria for DSM-IV categories 303.9 or 305; had donated blood within 30 days of baseline or were likely to do so during the course of the study.

Intervention

The study was a 48-week, double-blind, placebo-controlled, stratified, parallel-group study conducted in a group of women aged 65 years or older with mild to moderate AD at five sites in the United States. Those whose screening assessments showed that they were eligible to enter the study were assigned to receive either: An 11.25 mg formulation (marketed by TAP Pharmaceuticals Inc. of Lake Forest, Illinois, as Lupron Depot®-3 Month 11.25 mg) given as intramuscular injections; a 22.5 mg formulation (marketed by TAP Pharmaceuticals Inc. of Lake Forest, Illinois, as Lupron Depot® -3 Month 22.5 mg) given as intramuscular injections; or a placebo (physiologic saline) injection. Patients received intramuscular injections of study drug at Day 0 (baseline visit), week 12 (visit 5), week 24 (visit 7), and week 36 (visit 10) (see Table 1

Table 1
Schedule of assessments

	Screening	Baseline	Post-baseline Visits									
Visit Number	1	2	3[1]	4	5	6[1]	7	8	9[1]	10	11	12
Weeks Post-baseline	≤ −6	0	1	4	12	18	24	26	30	36	42	48
Informed Consent	X											
Medical & Social History	X											
Diagnosis of Probable AD[2,3]	X											
MMSE	X											
Rosen Modified HIS	X											
ECG	X											
Ham-D (17 item version)	X	X						X				X
Review Inclusion/Exclusion Criteria	X	X										
AE Assessment		X	X	X	X	X	X	X	X	X	X	X
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	X
Physical Examination	X											
Height	X											
Hematology & Chemistry	X	X		X	X		X	X		X	X	X
Urinalysis	X											
DEXA scan	X											X
APOE Assay		X										
TSH	X											
FSH, LH, Estradiol Assays[4]		X		X	X		X	X		X	X	X
Weight & Vital Signs	X	X		X	X		X	X		X	X	X
Randomization		X										
Administer Study Drug		X			X		X			X		
Phone Contact			X			X			X			
ADAS-Cog		X		X	X		X	X		X	X	X
ADCS-CGIC				X	X		X	X		X	X	X
NPI		X						X				X
BI		X		X	X		X	X		X	X	X
ADCS-ADL		X		X	X		X	X		X	X	X
ADCS-CGI Severity Rating		X										X

[1] Patients and caregivers were contacted by phone for assessments of safety and concomitant medications. [2] Defined in the NINCDS-ADRDA, including neuro imaging, history or cognitive and memory loss, and examinations to exclude other causes of dementia. [3] Brain imaging was obtained during screening period if not previously obtained after onset of symptoms of AD. [4] Optional blood samples were to be collected only if patients had consented to them in the Informed Consent Form. AD, Alzheimer's disease; MMSE, Mini-Mental State Examination; HIS, Hachinski Ischemic Score; ECG, electrocardiography; Ham-D, Hamilton Depression Rating Scale; AE, adverse event; DEXA, dual-energy x-ray absorptiometry; *APOE*, apolipoprotein E; TSH, thyroid-stimulating hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; ADAS-Cog, Alzheimer's Disease Assessment Scale - Cognitive Subscale; ADCS-CGIC, Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change; NPI, neuropsychiatric inventory; BI, burden interview; ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory.

for Schedule of Assessment). The study randomization was stratified so that the number of patients with and without evidence of osteoporosis or osteopenia, based on dual-energy x-ray absorptiometry (DEXA) scan findings, was balanced among the three treatment groups. Lupron Depot® is composed of leuprolide acetate, an analogue of the endogenous decapeptide GnRH. It has a substitution of a D-amino acid for glycine at position 6 and deletion of glycine at position 10 with the insertion of ethylamide, causing it to have a longer half-life and much higher affinity for the GnRH receptor than endogenous GnRH [17]. Once administered, it elicits an initial surge in LH and subsequently sex steroids, but within 2 weeks, GnRH receptors are down regulated resulting in very low levels of LH and follicle-stimulating hormone (FSH) [18] (Table 2).

Outcome measures

Outcome and safety measures were evaluated at baseline and weeks 4, 12, 24, 26, 36, 42, and 48. Additional telephone assessments were performed at weeks 1, 18, and 30. The primary efficacy parameters were the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) and the Alzheimer's Disease Cooperative Study Clinical Global Impression of Change (ADCS-CGIC). Secondary efficacy parameters were the Neuropsychiatric Inventory (NPI),

Table 2
Serum luteinizing hormone (LH) and follicle-stimulating hormone (FSH) concentrations is per protocol patients

Serum LH in patients treated with placebo or lupron (mean ± SD; mIU/mL)			
Study week	11.25 mg $n = 24$	22.5 mg $n = 21$	Placebo $n = 24$
Baseline	27.7 ± 11.2*	30.9 ± 19.1*	24.1 ± 14.5
Week 4	2.3 ± 0.9	2.6 ± 1.8	25.7 ± 14.7
Week 24	0.7 ± 0.3	0.3 ± 0.2	26.5 ± 15.4
Week 26	0.4 ± 0.2	0.3 ± 0.2	25.8 ± 15.1
Week 48	0.8 ± 0.4	0.4 ± 0.2	25.9 ± 15.3

Serum FSH in patients treated with placebo or lupron (mean ± SD; mIU/mL)			
Study week	11.25 mg $n = 24$	22.5 mg $n = 21$	Placebo $n = 24$
Baseline	52.3 ± 20.9*	48.5 ± 31.6*	46.3 ± 19.9
Week 4	5.0 ± 1.9	3.8 ± 2.4	47.4 ± 20.1
Week 24	7.3 ± 2.6	5.1 ± 3.3	50.8 ± 20.7
Week 26	4.9 ± 1.9	4.7 ± 3.1	50.5 ± 20.9
Week 48	6.8 ± 2.9	5.3 ± 3.5	50.1 ± 20.8

*Statistical difference ($p < 0.001$) between Baseline and Weeks 4, 24, 26, and 48 for low and high dose Lupron for serum LH and FSH.

Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory (ADCS-ADL), Burden Interview (BI), and ADCS-Severity Rating.

Safety was assessed by reviews of treatment-emergent adverse events and post-baseline changes in vital signs, physical examinations, clinical laboratory measures, and bone mineral density.

Bone mineral density

Bone mineral density was measured by means of DEXA scans of the lumbar vertebrae and a hip (including femoral neck). A DEXA scan was performed at screening and the end of study (week 48). The final DEXA scan was performed within 2 weeks before or after the final visit.

APOE genotyping

Direct sequencing of *APOE* genotype was performed by the Michigan State University DNA Diagnostic Program, East Lansing, MI.

Hormonal analyses

Serum LH, FSH and 17β-estradiol concentrations were measured at Quest Diagnostics, Miramar, FL.

Statistical analyses

All groups were analyzed for primary and secondary efficacy endpoints. In addition, pre-defined subgroup analyses included AChEI use and *APOE* status.

Primary efficacy analyses

The primary efficacy analyses were defined as comparisons between treatment groups for scores on the ADAS-Cog and ADCS-CGIC and were performed on the Intent-to-Treat population. The Intent-to-Treat population was defined as patients who received at least one dose of randomized drug and who had at least one post-baseline assessment of at least one primary efficacy variable.

ADAS-Cog: The efficacy analysis of the ADAS-Cog score for both treatment groups (low and high doses of Lupron Depot®) and the placebo group were analyzed by the method of analysis of variance and analysis of covariance. The primary analysis was the two-way analysis of variance model containing the main effects for both the treatment groups and the study sites along with their possible interaction. The final analysis was carried out on the 48-week endpoint by using the change in ADAS-Cog score from baseline.

ADCS-CGIC: The primary efficacy comparisons of the ADCS-CGIC score for both active treatment groups and the placebo group were analyzed by the Cochran-Mantel-Haenszel test which treats study sites as strata. In order to adjust for the other covariate effects, similar tests were based on the strata according to the levels of covariates. These covariates included the baseline osteoporosis/osteopenia status, the *APOE* genotype status, and the education level. If there was a significant association between the treatment groups and the ADCS-CGIC score, the common odds ratios were estimated by the Mantel-Haenszel estimator and the corresponding confidence interval determined across strata that ADCS-CGIC improved (or at least stabilized) over time between each active treatment group and the placebo group. The final analysis was carried out on the 48-week endpoint for ADCS-CGIC score.

Secondary efficacy analyses

The secondary efficacy analyses were the comparisons between treatment groups in scores on the ADCS-ADL, NPI (degree of behavioral disturbances associated with AD), BI (the impact of the patient's illness on the caregiver), and ADCS-CGI Severity Rating. Methods of statistical analysis similar to those used for ADAS-Cog score were used to analyze the change from baseline for the ADCS-ADL, NPI, and BI. The change from baseline in ADCS-ADL, NPI, and BI were analyzed by ANOVA and ANCOVA with the incorporation of important covariates such as the baseline age, the baseline osteoporosis/osteopenia status, the *APOE* status, and the education level. The effects of treatment on the change in ADCS-ADL, NPI, and BI were assessed using the appropriate hypotheses tests and confidence interval estimations.

In addition, the ADCS-CGI severity rating was summarized descriptively using frequency and percentage for each level of the rating at baseline, and using continuous statistics in the change from baseline at week 48.

In the use of all of these techniques in efficacy analysis, a variety of technical assumptions were required for each type of analysis. In order to assure that the reported results were not simply artifacts of the particular method of analysis, different analyses with a variety of analytic techniques that have slightly differing theoretical assumptions were carried out and compared. In order to control the Type I error rate for the final analysis, Bonferroni's method was used to adjust for the multiple comparisons made between each of the two active treatment groups and the placebo group. However, no adjustment was made for multiple analyses in the *a priori* subgroup analysis of patients taking AChEIs.

RESULTS

Demographic and clinical characteristics

The demographics and baseline characteristics of each treatment group are listed in Table 3. Each group was comparable for all demographic and clinical characteristics ($p > 0.05$) which included: age, race, height, weight, education level, *APOE* genotype, AChEI usage, MMSE score, Rosen Modified Hachinski Ischemic Score, Hamilton Psychiatric Rating Scale for Depression, abnormal physical exam findings at screening, abnormal ECG findings at screening, and 17β-estradiol, LH and FSH concentrations.

Of the 109 patients who entered the study, 37 were assigned to low dose, 36 to high dose, and 36 to placebo. There was no significant difference in completion rates between the groups: 72 patients (66%) completed the study; 25 (68%) in the low dose group, 22 (61%) in the high dose group, and 25 (69%) in the placebo group (Supplementary Table 1).

Primary outcomes

In the primary analysis there was a trend, although not statistically significant, in favor of the high dose Lupron group on the ADAS-Cog. The mean decline in the ADAS-Cog scores after 48 weeks of treatment was 1.7 points in the high dose group compared to 2.4 points in the placebo group and 4.9 points in the low dose group (Fig. 1A). A similar, although not as pronounced trend, was observed for ADCS-CGIC scores with 39% of patients in the high dose group exhibiting decline compared to 54% in the placebo group and 72% in the low dose group (Fig. 1B).

However, in the *a priori* designated subgroup analysis of patients taking AChEIs, there was a statistically significant benefit to subjects as determined by the ADAS-Cog and the ADCS-CGIC in the high dose Lupron group compared to both the placebo and low dose groups (Fig. 2). The mean decline in the ADAS-Cog scores after 48 weeks of treatment was 0.18 points in the high dose group compared to 3.30 points in the placebo group and 4.21 points in the low dose groups (Fig. 2A). Similarly, 9% of patients in the high dose group exhibited decline on ADCS-CGIC scores after 48 weeks of treatment compared to 63% in the placebo group and 82% in the low dose group (Fig. 2B). In patients not taking AChEIs, there was no significant difference by the ADAS-Cog and the ADCS-CGIC between individuals in the high dose Lupron, low dose Lupron, or placebo groups (see Supplementary Figure 1).

Secondary outcomes

In the primary analysis, there was no statistically significant difference on any of the secondary outcome measures, which included the ADCS-ADL (Fig. 3), the NPI, the ADCS-CGI Severity Rating, and the BI.

However, in the *a priori* subgroup analysis, patients taking high dose Lupron showed a statistically significant benefit seen on the ADCS-ADL. The mean decline in the high dose group was 0.54 points compared to 6.9 points in the placebo group and 8.0 points in the low dose group (Fig. 3). No differences between treatment groups were seen on the NPI, ADCS-CGI Severity Rating, or the BI in the subgroup analysis. In patients not taking AChEIs, there was no significant difference by the ADCS-ADL between individuals in the high dose Lupron, low dose Lupron, or placebo groups (see Supplementary Material).

It is known that patients who are homozygous for *APOE* $\varepsilon 4$ allele have an increased risk of AD.

Table 3
Demographics and baseline characteristics of each treatment group

Characteristic	Category	Lupron 11.25 mg (*n* = 36)	Lupron 22.5 mg (*n* = 36)	Placebo (*n* = 36)	*p-value*
Age (years)	Mean ± Std	78.75 ± 6.25	78.25 ± 6.01	76.97 ± 5.54	0.461[1]
	Median	80.0	80.0	77.5	
	Interquartile Range	73.5 – 83.0	73.5 – 83.0	74.0 – 80.0	
	Min-Max	67–93	67–88	65–88	
Race	Caucasian	30 (83.3%)	26 (72.2%)	27 (75%)	0.514[2]
	African-American	1 (2.8%)	2 (5.6%)	0	
	Hispanic	5 (13.9%)	8 (22.2%)	9 (25.0%)	
Height (inches)	Mean ± Std	60.95 ± 1.94	60.97 ± 2.88	61.58 ± 2.53	0.508[1]
	Median	61	61.5	61.3	
	Interquartile Range	59.0 – 62.0	59.0 – 62.8	60.0 – 64.0	
	Min-Max	57.0 – 64.5	55.0 – 67.0	56.0 – 67.0	
Weight (pounds)	Mean ± Std	131.9 ± 27.3	139.4 ± 20.9	140.4 ± 25.1	0.373[1]
	Median	132.0	134.8	136.5	
	Interquartile Range	113.0 – 147.0	127.5 – 146.0	123.0 – 152.3	
	Min-Max	92.0 – 220.0	108.0 – 193.0	95.0 – 223.0	0.817[2]
Education	Grade 6	6 (16.7%)	6 (16.7%)	6 (16.7%)	
	High school Grad	20 (55.6%)	21 (58.3%)	23 (63.9%)	
	Some College	5 (13.9%)	4 (11.1%)	3 (8.3%)	
	College Grad	5 (13.9%)	4 (11.1%)	2 (5.6%)	
	Post-Grad	0	1 (2.8%)	2 (5.6%)	
APOE Genotype	2/3	2 (5.6%)	2 (5.6%)	0	0.484[2]
	2/4	0	3 (8.8%)	1 (2.8%)	
	3/3	15 (41.7%)	16 (44.4%)	12 (33.3%)	
	3/4	16 (44.4%)	12 (33.3%)	18 (50.0%)	
	4/4	1 (2.8%)	2 (5.6%)	2 (5.6%)	
AChEI	Yes	28 (77.8%)	23 (63.9%)	26 (72.2%)	0.462[2]
	No	8 (22.2%)	13 (36.1%)	10 (27.8%)	
Estrogen supplementation	Yes	4	1	3	
Serum 17β-estradiol (pg/mL)	Mean±Std	73.8 ± 214.9 (*n* = 36)	21.1 ± 25.0 (*n* = 35)	32.3 ± 48.3 (*n* = 35)	>0.05[3]
	Median	15.5	15.0	17.0	
	Interquartile Range	11.0 – 25.0	11.0 – 20.0	10.0 – 22.0	
	Min-Max	10.0 – 1170.0	10.0 – 155.0	10.0 – 214.0	
FSH (mIU/mL)	Mean±Std	48.2 ± 22.2	52.4 ± 28.8	48.8 ± 21.8	
	Median	45.1	48.0	43.8	>0.05[3]
	Interquartile Range	32.5 – 63.7	30.9 – 70.8	32.7 – 65.4	
	Min-Max	7.4 – 105.0	13.3 – 145.0	15.7 – 106.0	
LH (mIU/mL)	Mean±Std	27.7 ± 15.0	33.6 ± 22.6	27.6 ± 14.8	>0.05[3]
	Median	26.3	30.5	23.5	
	Interquartile Range	22.0 – 32.3	18.9 – 40.9	17.6 – 36.1	
	Min-Max	4.3 – 84.5	3.9 – 130.1	3.3 – 71.9	
ADAS-Cog	Overall	19.73 ± 6.41	20.14 ± 9.36	21.90 ± 9.50	>0.29[3]
	Sub-group analysis (AChEI users)	20.73 ± 5.94	20.31 ± 9.03	24.29 ± 9.93	>0.06[4]
ADCS-ADL	Sub-group analysis (AChEI users)	59.2 ± 7.8	55.8 ± 12.7	55.6 ± 13.9	
NPI	Overall	8.9 ± 11.8	8.8 ± 9.6	9.1 ± 8.5	1.00[5]
MMSE	Overall	18.2 ± 3.3	18.6 ± 3.5	17.9 ± 3.3	0.33[6]
Rosen Modified HIS	Overall	0.72 ± 0.74	0.50 ± 0.56	0.72 ± 0.88	0.82[6]
Ham-D	Overall	3.3 ± 3.0	4.1 ± 3.6	4.6 ± 3.2	0.12[6]
Abnormal physical findings	Overall	29 (81%)	32 (89%)	26 (72)	0.13[7]
Abnormal ECG findings	Overall	20 (56%)	30 (83%)	27 (75)	0.008[7]

[1] *p-values* for treatment comparisons using a two-way analysis of variance test with factors of treatment and site (if the assumptions of ANOVA are satisfied) or using Friedman's test if these assumptions are not satisfied. [2] *p-values* for treatment comparisons using the Cochran-Mantel-Haenszel test for general association, adjusted for site. [3] *p-values* for baseline serum hormone concentrations. [4] *p-value* for placebo versus high dose group. [5] *p-values* and confidence intervals for treatment comparisons from analysis of variance with treatment and site as factors. [6] *p-values* for treatment comparisons using Friedman's test with factors of treatment and site. [7] *p-values* for treatment comparisons using Cochran-Mantel-Haenszel test for general association, adjusted for site. MMSE, Mini-Mental State Examination; HIS, Hachinski Ischemic Score; ECG, electrocardiography; Ham-D, Hamilton Depression Rating Scale; *APOE*, apolipoprotein E; FSH, follicle-stimulating hormone; LH, luteinizing hormone; ADAS-Cog, Alzheimer's Disease Assessment Scale - Cognitive Subscale; NPI, neuropsychiatric inventory; ADCS-ADL, Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory; AChEI, acetylcholinesterase inhibitor.

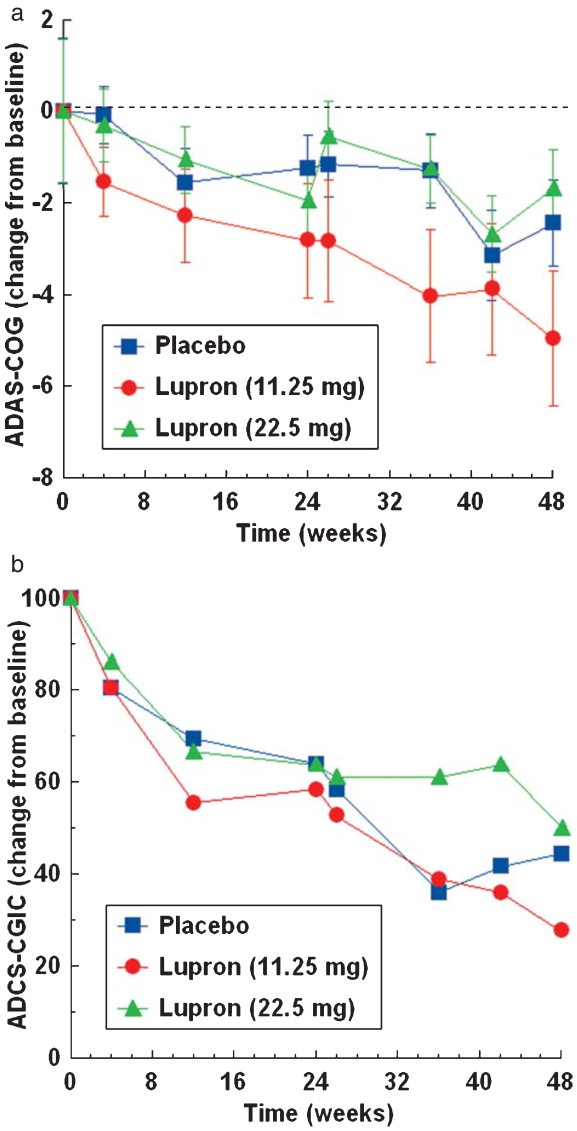


Fig. 1. Changes in cognitive performance over 48 weeks in individuals treated with placebo, low dose Lupron, or high dose Lupron as determined with (A) ADAS-Cog and (B) ADCS-CGIC ($n = 36$ in each group). For ADAS-Cog and ADCS-CGIC, there were no significant differences for placebo versus high dose, placebo versus low dose, and low dose versus high dose at 26 and 48 weeks.

Sub-analyses were performed for efficacy endpoints based upon patients' *APOE* status. No statistical differences were found.

Safety

The safety profile of Lupron at doses similar to those used in this study has been established in other indications such as the treatment of advanced prostate cancer,

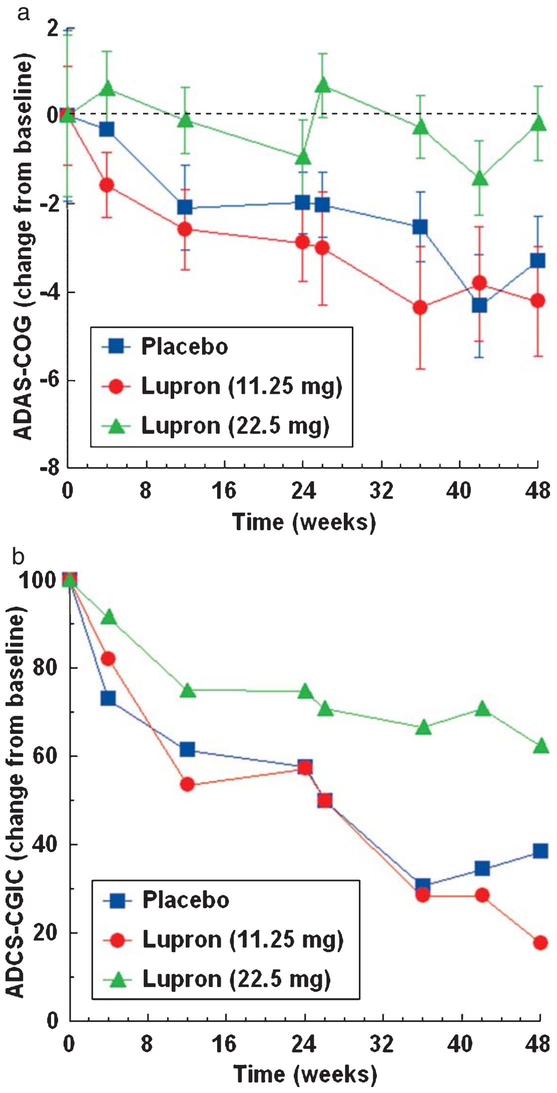


Fig. 2. Changes in cognitive performance over 48 weeks in individuals taking AChEIs and treated with placebo ($n = 26$), low dose ($n = 28$), or high dose ($n = 24$) Lupron as determined with (A) ADAS-Cog and (B) ADCS-CGIC. A) Statistical differences, ADAS-Cog - Adjusted for multiple comparisons: *p-value* and 95% confidence interval for the high dose group versus placebo with treatment and site as factors using ANOVA = 0.037 (−5.78, −0.18) and with treatment and site as factors and baseline ADAS-Cog score as covariate using ANCOVA = 0.057 (−5.67, 0.08) at week 26. *p-value* and 95% confidence interval for the high dose group versus placebo with treatment and site as factors using ANOVA = 0.042 (−6.14, −0.11) and with treatment and site as factors and baseline ADAS-Cog score as covariate using ANCOVA = 0.060 (−6.08, 0.12) at week 48. ADAS-Cog - Unadjusted for multiple comparisons: The p values, unadjusted for multiple analyses for high dose and placebo were = 0.0009 and 0.026 at weeks 26 and 48, respectively. B) Statistical differences, ADCS-CGIC - Adjusted for multiple comparisons: There were no statistical differences between any treatment groups. ADCS-CGIC - Unadjusted for multiple comparisons: The *p-values*, unadjusted for multiple analyses for high dose and placebo, were 0.223 and 0.031 at weeks 26 and 48, respectively.

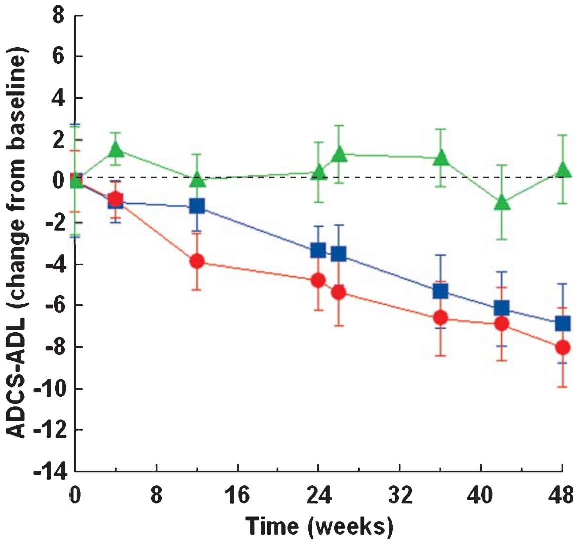


Fig. 3. Changes in cognitive performance as determined by ADCS-ADLs over 48-weeks in individuals treated with placebo ($n = 26$), low dose ($n = 28$), or high dose ($n = 24$) with AChEIs. The *p-values*, unadjusted for multiple analyses for high dose and placebo, were = 0.016 and 0.015 at weeks 26 and 48, respectively.

children with central precocious puberty, endometriosis, and uterine fibroids. However, the safety of Lupron treatment in patients with AD has not previously been described. The majority of patients (77 of 109 patients or 71%) experienced at least one adverse event (AE) (Table 4 and Supplementary Table 2). These were mostly mild or moderate in severity and the ADCS safety monitoring committee regarded these as mainly unrelated to study drug. The most common AEs reported were consistent with the known safety profile of Lupron (Supplementary Table 3). There were 8 patient discontinuations due to AEs.

Twenty serious AEs were reported in 18 patients including two deaths. One death was attributed to respiratory failure in the high dose group, and one death was attributed to cerebral hemorrhage in the placebo group. The ADCS safety monitoring committee categorized three AEs as possibly related to study drug: Two cases (upper gastrointestinal hemorrhage and a syncopal episode) occurred in the low dose group and one case (deep vein thrombosis) occurred in the placebo group. Therefore, based on this data, there were no unexpected safety concerns in this patient population. This is consistent with the well-established safety profile of Lupron in the treatment of other conditions.

The extended use of Lupron is known to cause the loss of bone mineral density in men [19–22] and pre-menopausal women [23, 24]. Whether any similar effect occurs in postmenopausal women has not been studied. In our study, there were no significant differences in bone mineral density between any of the treatment groups over 48 weeks of treatment.

DISCUSSION

This Phase II dose-ranging study demonstrated that high dose Lupron in combination with AChEIs halted the progression of cognitive decline in women with mild to moderate AD over a 48-week period (Figs. 2 and 3). A similar effect was not observed in the low dose Lupron group taking AChEIs or in the placebo group taking AChEIs (Figs. 2 and 3). This combination treatment was safe and well tolerated at both dose levels (Table 4) and AEs were consistent with the current product labels for other indications.

In the primary analysis there was a trend, although not statistically significant, in favor of the high dose Lupron group on the ADAS-Cog (Fig. 1). In the *a priori* designated subgroup analysis of patients already taking AChEIs, there was a clear statistically significant benefit demonstrated on the ADAS-cog, ADCS-CGIC, and the ADCS-ADL in the high dose Lupron group compared to the low dose and placebo groups (Figs. 2 and 3).

All drugs currently approved for the treatment of AD confer an initial improvement in cognitive function followed by a decline whose rate is similar to placebo [25]. In contrast to these treatments, there was no initial improvement in cognitive function following initiation of Lupron treatment but most importantly, there was no decline in cognitive performance in the high dose/AChEI group. These findings together with biological and epidemiological evidence suggest that the effects seen with high dose Lupron are one of potential disease modification rather than symptomatic improvement [1–9, 11–14, 26].

The mechanism by which Lupron acts with AChEI to improve cognitive performance is unclear. It is known that the AChEI rivastigmine can reduce the lipopolysaccharide-induced decreases in GnRH and LH, and perhaps stimulate GnRH/LH secretion [27]. In this connection, the modulation of GnRH release has been suggested to be mediated via cholinergic (and GABAergic) neurotransmission [28]. Thus, one possible additive mechanism of action might involve the further downregulation of GnRH receptor signaling and LH expression/signaling. Alternatively, since GnRH mediates neurotransmission itself [29, 30], Lupron might act directly to improve cognitive per-

Table 4
Summary of adverse events (AEs)

	Treatment			*p-value*[1]		
	Leuprolide 11.25 mg ($n = 37$)	Leuprolide 22.5 mg ($n = 36$)	Placebo ($n = 36$)	11.25 versus placebo	22.5 versus placebo	11.25 versus 22.5
Patients with at least 1 AE	27 (72.9%)	27 (75.0%)	23 (63.9%)	0.40	0.31	0.84
Not related	13 (35.1%)	12 (33.3%)	8 (22.2%)			
Probably not related	6 (16.2%)	10 (27.7%)	6 (16.7%)			
Possibly related	7 (18.9%)	3 (8.3%)	8 (22.2%)			
Probably related	1 (2.7%)	2 (5.5%)	1 (2.8%)			
Related	0	0	0			
Patients with serious AE	10 (27.0%)	4 (11.1%)	5 (13.8%)	0.17	1.0	0.08
Not related	7 (18.9%)	1 (2.8%)	2 (5.6%)			
Probably not related	1 (2.7%)	3 (8.3%)	2 (5.6%)			
Possibly related	2 (5.4%)	0	1 (2.8%)			
Probably related	0	0	0			
Related	0	0	0			
Patients with AEs that led to discontinuation	4 (10.8%)	2 (5.5%)	0	0.12	0.49	0.67
Patients with AEs resulting in death	0	1 (2.8%)	1 (2.8%)	0.49	1.0	0.49

At each level of summarization each patient is only counted once. [1]*p-values* for treatment comparisons from Pearson's Chi-square test or Fisher's exact test if appropriate.

formance. Another possibility is that Lupron acts to halt any further neurodegeneration thereby allowing AChEIs to act on remaining neurons to maintain cholinergic function.

The dose effect seen in this study suggests that Lupron's action is not solely due to its suppression of peripheral circulating concentrations of gonadotropins, which were similarly suppressed in low dose and high dose groups (Table 2). Therefore, Lupron's actions might also be due to 9 direct effect on GnRH receptor signaling within the brain [31]. GnRH receptors are expressed throughout the brain and their expression correlates to those areas with AD neuropathology [31]. In this connection, we recently identified the existence of autocrine/paracrine feedback loops within the brain, in essence a feedback loop similar to the hypothalamic-pituitary-gonadal axis that regulates neurohormone production [32]. Since GnRH receptor mediates neuronal LH expression and LH receptor signaling, high doses of Lupron might suppress the neuroautocrine production of LH, which we have previously demonstrated is elevated in expression and colocalizes with AD neuropathology [33], while low doses might stimulate LH production.

This dose effect might also explain some conflicting preclinical results. Most researchers have found that lowering LH signaling with the GnRH agonist Lupron decreases Aβ levels and improves cognitive performance in wild-type mice [8, 34] and AβPP-transgenic mice [2]. However, a decrease in brain Aβ and improvement in cognition following leuprolide acetate treatment was not observed in the overexpressing AβPP(Swt), PS1(M146 V), and tau(P301L) (triple) transgenic mice [35]. Whether this is a dose effect (or an artifact of the 3xTg mice) is not clear since multiple doses have not been evaluated. Future animal studies are warranted to help understand the dose effect and the synergism with AChEIs.

In conclusion, our data demonstrate that cognitive function was preserved in patients treated with high dose Lupron who were already using AChEIs. Caution should be used in the interpretation of the results due to: The small sample size, which did not allow determination of whether this treatment is best suited to early or later phases of the disease; the fact that baseline demographics were not compared for the subgroup; and non-adjustment for multiple analyses. The results of this study should however encourage further investigation of GnRH agonist therapy for the treatment of AD. Future clinical studies should be conducted with Lupron at doses providing systemic exposure at least equivalent to those provided by Lupron 22.5 mg every 12 weeks. Such studies could be expanded to include the use of GnRH antagonists.

ACKNOWLEDGMENTS

The authors thank John Stone for his unwavering support and encouragement. The opinions expressed herein are those of the authors. The contents do not represent the views of the Department of Veterans Affairs or the US Government.

Authors' disclosures available online (http://www.j-alz.com/disclosures/view.php?id=2545).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD141626.

REFERENCES

[1] Bowen RL, Atwood CS (2004) Living and dying for sex. A theory of aging based on the modulation of cell cycle signaling by reproductive hormones. *Gerontology* **50**, 265-290.

[2] Casadesus G, Webber KM, Atwood CS, Pappolla MA, Perry G, Bowen RL, Smith MA (2006) Luteinizing hormone modulates cognition and amyloid-beta deposition in Alzheimer APP transgenic mice. *Biochim Biophys Acta* **1762**, 447-452.

[3] Atwood CS, Meethal SV, Liu T, Wilson AC, Gallego M, Smith MA, Bowen RL (2005) Dysregulation of the hypothalamic-pituitary-gonadal axis with menopause and andropause promotes neurodegenerative senescence. *J Neuropathol Exp Neurol* **64**, 93-103.

[4] Atwood CS, Bowen RL (2011) The reproductive-cell cycle theory of aging: An update. *Exp Gerontol* **46**, 100-107.

[5] Bowen RL, Isley JP, Atkinson RL (2000) An association of elevated serum gonadotropin concentrations and Alzheimer disease? *J Neuroendocrinol* **12**, 351-354.

[6] Short RA, Bowen RL, O'Brien PC, Graff-Radford NR (2001) Elevated gonadotropin levels in patients with Alzheimer disease. *Mayo Clin Proc* **76**, 906-909.

[7] Bowen RL, Smith MA, Harris PL, Kubat Z, Martins RN, Castellani RJ, Perry G, Atwood CS (2002) Elevated luteinizing hormone expression colocalizes with neurons vulnerable to Alzheimer's disease pathology. *J Neurosci Res* **70**, 514-518.

[8] Bowen RL, Verdile G, Liu T, Parlow AF, Perry G, Smith MA, Martins RN, Atwood CS (2004) Luteinizing hormone, a reproductive regulator that modulates the processing of amyloid-beta precursor protein and amyloid-beta deposition. *J Biol Chem* **279**, 20539-20545.

[9] Bowen RL, Beaird H, Atwood CS, Smith MA, Rimm AA. (2004) Men treated for prostate cancer have a decreased incidence of dementia. *9th International Conference on Alzheimer's Disease and Related Disorders*, 17–22 July 2004, Philadelphia, Pennsylvania, USA.

[10] D'Amico AV, Braccioforte MH, Moran BJ, Chen MH (2010) Luteinizing-hormone releasing hormone therapy and the risk of death from Alzheimer disease. *Alzheimer Dis Assoc Disord* **24**, 85-89.

[11] Berry A, Tomidokoro Y, Ghiso J, Thornton J (2008) Human chorionic gonadotropin (a luteinizing hormone homologue) decreases spatial memory and increases brain amyloid-beta levels in female rats. *Horm Behav* **54**, 143-152.

[12] Casadesus G, Milliken EL, Webber KM, Bowen RL, Lei Z, Rao CV, Perry G, Keri RA, Smith MA (2007) Increases in luteinizing hormone are associated with declines in cognitive performance. *Mol Cell Endocrinol* **269**, 107-111.

[13] Bryan KJ, Mudd JC, Richardson SL, Chang J, Lee HG, Zhu X, Smith MA, Casadesus G Down-regulation of serum gonadotropins is as effective as estrogen replacement at improving menopause-associated cognitive deficits. *J Neurochem* **112**, 870-881.

[14] Lin J, Li X, Yuan F, Lin L, Cook CL, Rao Ch V, Lei Z (2010) Genetic ablation of luteinizing hormone receptor improves the amyloid pathology in a mouse model of Alzheimer disease. *J Neuropathol Exp Neurol* **69**, 253-261.

[15] Hsiao K, Chapman P, Nilsen S, Eckman C, Harigaya Y, Younkin S, Yang F, Cole G (1996) Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science* **274**, 99-102.

[16] Bowen RL, Verdile G, Liu T, Parlow AF, Perry G, Smith MA, Martins RN, Atwood CS (2004) Luteinizing hormone, a reproductive regulator that modulates the processing of amyloid-beta precursor protein and amyloid-beta deposition. *J Biol Chem* **279**, 20539-20545.

[17] Chrisp P, Sorkin EM (1991) Leuprorelin. A review of its pharmacology and therapeutic use in prostatic disorders. *Drugs Aging* **1**, 487-509.

[18] Periti P, Mazzei T, Mini E (2002) Clinical pharmacokinetics of depot leuprorelin. *Clin Pharmacokinet* **41**, 485-504.

[19] Ziaran S, Goncalves FM, Sn JB (2013) Complex metabolic and skeletal changes in men taking long-term androgen deprivation therapy. *Clin Genitourin Cancer* **11**, 33-38.

[20] Yu EY, Kuo KF, Gulati R, Chen S, Gambol TE, Hall SP, Jiang PY, Pitzel P, Higano CS (2012) Long-term dynamics of bone mineral density during intermittent androgen deprivation for men with nonmetastatic, hormone-sensitive prostate cancer. *J Clin Oncol* **30**, 1864-1870.

[21] Spry NA, Galvao DA, Davies R, La Bianca S, Joseph D, Davidson A, Prince R (2009) Long-term effects of intermittent androgen suppression on testosterone recovery and bone mineral density: Results of a 33-month observational study. *BJU Int* **104**, 806-812.

[22] Khosla S, Melton LJ, 3rd, Riggs BL (2001) Estrogens and bone health in men. *Calcif Tissue Int* **69**, 189-192.

[23] Palomba S, Morelli M, Di Carlo C, Noia R, Pellicano M, Zullo F (2002) Bone metabolism in postmenopausal women who were treated with a gonadotropin-releasing hormone agonist and tibolone. *Feril Infertil* **78**, 63-68.

[24] Mitwally MF, Gotlieb L, Casper RF (2002) Prevention of bone loss and hypoestrogenic symptoms by estrogen and interrupted progestogen add-back in long-term GnRH-agonist down-regulated patients with endometriosis and premenstrual syndrome. *Menopause* **9**, 236-241.

[25] Howard R, McShane R, Lindesay J, Ritchie C, Baldwin A, Barber R, Burns A, Dening T, Findlay D, Holmes C, Hughes A, Jacoby R, Jones R, McKeith I, Macharouthu A, O'Brien J, Passmore P, Sheehan B, Juszczak E, Katona C, Hills R, Knapp M, Ballard C, Brown R, Banerjee S, Onions C, Griffin M, Adams J, Gray R, Johnson T, Bentham P, Phillips P (2012) Donepezil and memantine for moderate-to-severe Alzheimer's disease. *N Engl J Med* **366**, 893-903.

[26] D'Amico AV, Braccioforte MH, Moran BJ, Chen MH Luteinizing-hormone releasing hormone therapy and the risk of death from Alzheimer disease. *Alzheimer Dis Assoc Disord* **24**, 85-89.

[27] Herman AP, Krawczynska A, Bochenek J, Haziak K, Romanowicz K, Misztal T, Antushevich H, Herman A, Tomaszewska-Zaremba D (2013) The effect of rivastigmine on the LPS-induced suppression of GnRH/LH secretion during the follicular phase of the estrous cycle in ewes. *Anim Reprod Sci* **138**, 203-212.

[28] Kaur G (2001) Role of cholinergic and GABAergic neurotransmission in the opioids-mediated GnRH release mechanism of EBP-primed OVX rats. *Mol Cell Biochem* **219**, 13-19.

[29] Lu F, Yang JM, Wu JN, Chen YC, Kao YH, Tung CS, Yang SN (1999) Activation of gonadotropin-releasing hormone receptors produces neuronal excitation in the rat hippocampus. *Chin J Physiol* **42**, 67-71.

[30] Yang SN, Lu F, Wu JN, Liu DD, Hsieh WY (1999) Activation of gonadotropin-releasing hormone receptors induces a long-term enhancement of excitatory postsynaptic currents mediated by ionotropic glutamate receptors in the rat hippocampus. *Neurosci Lett* **260**, 33-36.

[31] Wilson AC, Salamat MS, Haasl RJ, Roche KM, Karande A, Meethal SV, Terasawa E, Bowen RL, Atwood CS (2006) Human neurons express type I GnRH receptor and respond to GnRH I by increasing luteinizing hormone expression. *J Endocrinol* **191**, 651-663.

[32] Meethal SV, Liu T, Chan HW, Ginsburg E, Wilson AC, Gray DN, Bowen RL, Vonderhaar BK, Atwood CS (2009) Identification of a regulatory loop for the synthesis of neurosteroids: A steroidogenic acute regulatory protein-dependent mechanism involving hypothalamic-pituitary-gonadal axis receptors. *J Neurochem* **110**, 1014-1027.

[33] Bowen RL, Smith MA, Harris PL, Kubat Z, Martins RN, Castellani RJ, Perry G, Atwood CS (2002) Elevated luteinizing hormone expression colocalizes with neurons vulnerable to Alzheimer's disease pathology. *J Neurosci Res* **70**, 514-518.

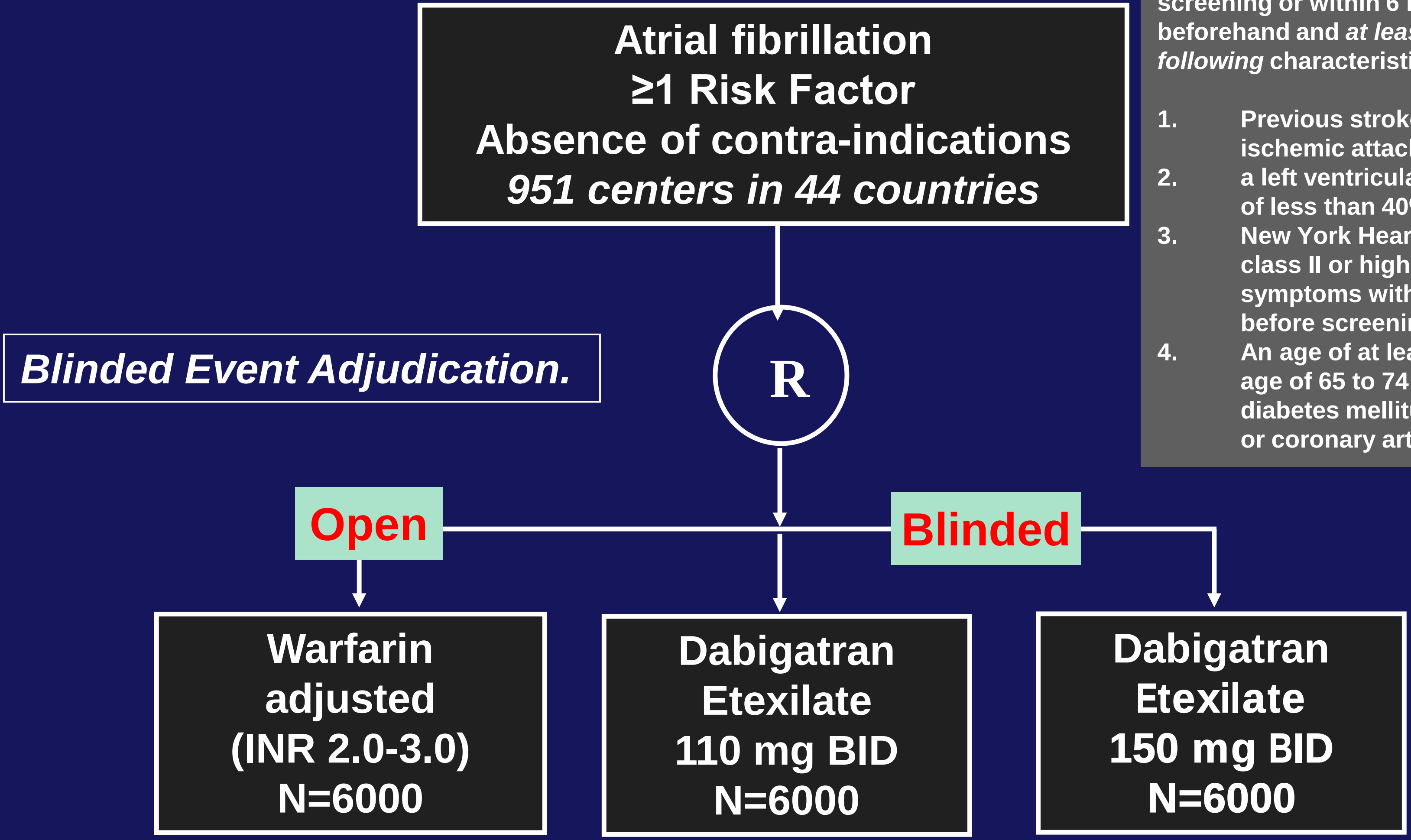
[34] Bryan KJ, Mudd JC, Richardson SL, Chang J, Lee HG, Zhu X, Smith MA, Casadesus G (2010) Down-regulation of serum gonadotropins is as effective as estrogen replacement at improving menopause-associated cognitive deficits. *J Neurochem* **112**, 870-881.

[35] Rosario ER, Carroll JC, Pike CJ (2012) Evaluation of the effects of testosterone and luteinizing hormone on regulation of beta-amyloid in male 3xTg-AD mice. *Brain Res* **1466**, 137-145.

A Comparison of RE-LY and ROCKET AF Trial Designs

C. Michael Gibson, M.S., MD.

RE-LY: Study Design



Atrial fibrillation
≥1 Risk Factor
Absence of contra-indications
951 centers in 44 countries

Blinded Event Adjudication.

R

Open

Blinded

Warfarin adjusted (INR 2.0-3.0)
N=6000

Dabigatran Etexilate 110 mg BID
N=6000

Dabigatran Etexilate 150 mg BID
N=6000

Patients were eligible if they had atrial fibrillation documented on electrocardiography performed at screening or within 6 months beforehand and *at least one of the following* characteristics:

1. **Previous stroke or transient ischemic attack**
2. **a left ventricular ejection fraction of less than 40%**
3. **New York Heart Association class II or higher heart-failure symptoms within 6 months before screening**
4. **An age of at least 75 years or an age of 65 to 74 years plus diabetes mellitus, hypertension, or coronary artery disease.**

Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Rocket AF Study Design

Risk Factors
- CHF
- Hypertension
- Age ≥ 75
- Diabetes

At least 2 or 3 required*

OR
- Stroke, TIA or Systemic embolus

Atrial Fibrillation

Randomize
Double Blind /
Double Dummy
(n ~ 14,000)

Rivaroxaban

20 mg daily
15 mg for Cr Cl 30-49 ml/min

Warfarin

INR target - 2.5
(2.0-3.0 inclusive)

Monthly Monitoring
Adherence to standard of care guidelines

Primary Endpoint: Stroke or non-CNS Systemic Embolism

* Enrollment of patients without prior Stroke, TIA or systemic embolism and only 2 factors capped at 10%

Rocket AF Investigators, AHA 2010

Comparison of Study Designs

- Both had non-inferiority to warfarin as primary endpoint

- Rocket AF required 2 risk factors for entry, RE-LY 1 risk factor

- Rocket AF capped $CHADS_2$ = 2 early in the trial unless a patient scored two points by having a prior stroke/TIA. This may account for the high rate of prior stroke in Rocket AF.

- Both randomized trials

- Rocket AF administered warfarin in a blinded fashion, RE-LY did not

- There was a dose adjustment for impaired CrCl in Rocket AF

- INR target range 2-3 in both

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Comparison of Study Designs in Other Trials

Trial	Inclusion	Design	Start date	Duration (mo)	Current / Goal Enrollment	# sites
ROCKET AF	CHADS ≥ 3 or stroke/TIA (15% CHADS 2)	Sham INR	12/06	15	14,264	~1200
ARISTOTLE	CHADS ≥ 1 (50% VKA naïve)	Sham INR	1/07	15	~15,000	~937
RE-LY	CHADS ≥ 1 (30% VKA naïve)	Open label	12/2005	26	18,113	706
AMADEUS	CHADS ≥ 1	Open label	9/2003	23	4576	165
SPORTIF V	CHADS ≥ 1	Sham INR	8/2000	17	3922	409
ENGAGE	CHADS ≥ 2	Sham INR	10/2008	24	20,500	~1400

C. Michael Gibson, M.S., M.D.

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Statistical Methods: Efficacy

RELY:

Primary Efficacy Evaluation: Stroke or non-CNS Embolism

- Non-Inferiority: Intention-to-treat
- Superiority: Intention-to-treat

Rocket AF:

Primary Efficacy Evaluation: Stroke or non-CNS Embolism

- Non-Inferiority: Protocol Compliant on treatment
- Superiority: On Treatment, then by Intent-to-Treat

RE-LY used *Intention to treat* for both non-inferiority and superiority testing; Rocket AF used *on treatment analysis* for first tests of non-inferiority and superiority

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Statistical Methods: Safety

RELY:

- Primary Safety Evaluation: Major bleeding

Rocket AF:

- Primary Safety Evaluation: Major or non-Major Clinically Relevant Bleeding

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

RE-LY Definitions of Stroke

- Stroke was defined as the sudden onset of a focal neurologic deficit in a location consistent with the territory of a major cerebral artery and categorized as ischemic, hemorrhagic, or unspecified.
- Hemorrhagic transformation of ischemic stroke was not considered to be hemorrhagic stroke.
- Intracranial hemorrhage consisted of hemorrhagic stroke and subdural or subarachnoid hemorrhage.
- Systemic embolism was defined as an acute vascular occlusion of an extremity or organ, documented by means of imaging, surgery, or autopsy.

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Rocket AF Definitions of Stroke

- The primary efficacy outcome is the composite of stroke
  - Stroke is defined as a new, sudden, focal neurological deficit resulting from a presumed cerebrovascular cause that is not reversible within 24 hours and not due to a readily identifiable cause such as a tumor or seizure
  - All strokes will be classified as primary ischemic or primary hemorrhagic
- And non-CNS systemic embolism
  - Non-CNS systemic embolism is defined as abrupt vascular insufficiency associated with clinical or radiological evidence of arterial occlusion in the absence of other likely mechanisms, (e.g., trauma, atherosclerosis, instrumentation)

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

RE-LY: Baseline Characteristics

Characteristic	Dabigatran 110 mg	Dabigatran 150 mg	Warfarin
Randomized	6015	6076	6022
Mean age (years)	71.4	71.5	71.6
Male (%)	64.3	63.2	63.3
CHADS2 score (mean) 0-1 (%) 2 (%) 3+ (%)	2.1 32.6 34.7 32.7	2.2 32.2 35.2 32.6	2.1 30.9 37.0 32.1
Prior stroke/TIA (%)	19.9	20.3	19.8
Prior MI (%)	16.8	16.9	16.1
CHF (%)	32.2	31.8	31.9
Baseline ASA (%)	40.0	38.7	40.6
Warfarin Naïve (%)	49.9	49.8	51.4

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Rocket AF: Baseline Demographics

	Rivaroxaban (N=7081)	Warfarin (N=7090)
CHADS$_2$ Score (mean)	3.48	3.46
2 (%)	13	13
3 (%)	43	44
4 (%)	29	28
5 (%)	13	12
6 (%)	2	2
Prior VKA Use (%)	62	63
Congestive Heart Failure (%)	63	62
Hypertension (%)	90	91
Diabetes Mellitus (%)	40	39
Prior Stroke/TIA/Embolism (%)	55	55
Prior Myocardial Infarction (%)	17	18

Based on Intention-to-Treat Population

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Rocket AF was a Higher Risk Patient Population

- Whereas 32.4% of patients in RE-LY were low risk CHADS 0-1, there were *none of these patients in Rocket AF*

- Whereas just over 32% of patients in RE-LY were high risk CHADS score of 3 or more, *over 85% of Rocket AF patients had a CHADS score of 3 or more*

- RE-LY patients were about 71.5 years old, and Rocket AF patients were 73 years old

- Prior stroke TIA embolism was about 20% in RE-LY and was 55% in Rocket AF

- About half of RE-LY patients were warfarin naïve, whereas on 37.5% of Rocket AF patients were warfarin naive

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Impact of Enrolling Higher CHADs Score Patients

Higher CHADs scores are associated with:

1. **Higher rates of major bleeding**
2. **Lower TTRs**

Personal communication RE-LY Investigators. Connolly SJ, et al. *N Engl J Med.* 2009;361:1139-1151

Comparison of Trial Metrics

	RE-LY	Rocket AF
Countries	44	45
Patients	18,113	14,264
Median Duration of Follow-Up	2 years (about 730 days)	589 days of exposure, 707 days including period off drug during follow-up
Time in Therapeutic Range (TTR)	64% 67% warfarin-experienced 61% warfarin-naïve	57.8%

C. Michael Gibson, M.S., M.D.

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

Rates of Drug Discontinuation

RE-LY

1 Year:

Dabigatran 110 mg:	14.5%
Dabigatran 150 mg:	15.5%
Warfarin:	10.2%

2 Years:

Dabigatran 110 mg:	20.7%
Dabigatran 150 mg:	21.2%,
Warfarin:	16.6%

Rocket AF

Rivaroxaban:	23.9%
Warfarin:	22.4%

Rocket AF Investigators, AHA 2010; Connolly SJ, et al. *N Engl J Med*. 2009;361:1139-1151

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 8, 2011 VOL. 365 NO. 10

Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation

Manesh R. Patel, M.D., Kenneth W. Mahaffey, M.D., Jyotsna Garg, M.S., Guohua Pan, Ph.D., Daniel E. Singer, M.D., Werner Hacke, M.D., Ph.D., Günter Breithardt, M.D., Jonathan L. Halperin, M.D., Graeme J. Hankey, M.D., Jonathan P. Piccini, M.D., Richard C. Becker, M.D., Christopher C. Nessel, M.D., John F. Paolini, M.D., Ph.D., Scott D. Berkowitz, M.D., Keith A.A. Fox, M.B., Ch.B., Robert M. Califf, M.D., and the ROCKET AF Steering Committee, for the ROCKET AF Investigators*

ABSTRACT

BACKGROUND

The use of warfarin reduces the rate of ischemic stroke in patients with atrial fibrillation but requires frequent monitoring and dose adjustment. Rivaroxaban, an oral factor Xa inhibitor, may provide more consistent and predictable anticoagulation than warfarin.

METHODS

In a double-blind trial, we randomly assigned 14,264 patients with nonvalvular atrial fibrillation who were at increased risk for stroke to receive either rivaroxaban (at a daily dose of 20 mg) or dose-adjusted warfarin. The per-protocol, as-treated primary analysis was designed to determine whether rivaroxaban was noninferior to warfarin for the primary end point of stroke or systemic embolism.

RESULTS

In the primary analysis, the primary end point occurred in 188 patients in the rivaroxaban group (1.7% per year) and in 241 in the warfarin group (2.2% per year) (hazard ratio in the rivaroxaban group, 0.79; 95% confidence interval [CI], 0.66 to 0.96; P<0.001 for noninferiority). In the intention-to-treat analysis, the primary end point occurred in 269 patients in the rivaroxaban group (2.1% per year) and in 306 patients in the warfarin group (2.4% per year) (hazard ratio, 0.88; 95% CI, 0.74 to 1.03; P<0.001 for noninferiority; P=0.12 for superiority). Major and nonmajor clinically relevant bleeding occurred in 1475 patients in the rivaroxaban group (14.9% per year) and in 1449 in the warfarin group (14.5% per year) (hazard ratio, 1.03; 95% CI, 0.96 to 1.11; P=0.44), with significant reductions in intracranial hemorrhage (0.5% vs. 0.7%, P=0.02) and fatal bleeding (0.2% vs. 0.5%, P=0.003) in the rivaroxaban group.

CONCLUSIONS

In patients with atrial fibrillation, rivaroxaban was noninferior to warfarin for the prevention of stroke or systemic embolism. There was no significant between-group difference in the risk of major bleeding, although intracranial and fatal bleeding occurred less frequently in the rivaroxaban group. (Funded by Johnson & Johnson and Bayer; ROCKET AF ClinicalTrials.gov number, NCT00403767.)

From the Duke Clinical Research Institute (M.R.P., K.W.M., J.G., J.P.P., R.C.B.) and Duke Translational Medicine Institute (R.M.C.), Duke University Medical Center, Durham, NC; Johnson & Johnson Pharmaceutical Research and Development, Raritan (G.P., C.C.N.), and Bayer HealthCare Pharmaceuticals, Montville (J.F.P., S.D.B.) — both in New Jersey; Massachusetts General Hospital and Harvard Medical School — both in Boston (D.E.S.); Ruprecht-Karls-University, Heidelberg (W.H.), and Hospital of the University of Münster, Münster (G.B.) — both in Germany; the Cardiovascular Institute, Mount Sinai Medical Center, New York (J.L.H.); Royal Perth Hospital, Perth, WA, Australia (G.J.H.); and the University of Edinburgh and Royal Infirmary of Edinburgh — both in Edinburgh (K.A.A.F.). Address reprint requests to Dr. Patel at Duke Clinical Research Institute, Duke University Medical Center, Rm. 0311 Terrace Level, 2400 Pratt St., Durham, NC 27705, or at manesh.patel@duke.edu.

*A complete listing of the steering committee members and trial investigators in the Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF) is provided in the Supplementary Appendix, available at NEJM.org.

This article (10.1056/NEJMoa1009638) was published on August 10, 2011, at NEJM.org.

N Engl J Med 2011;365:883-91.

Copyright © 2011 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

ATRIAL FIBRILLATION IS ASSOCIATED with an increase in the risk of ischemic stroke by a factor of four to five[1] and accounts for up to 15% of strokes in persons of all ages and 30% in persons over the age of 80 years.[2] The use of vitamin K antagonists is highly effective for stroke prevention in patients with nonvalvular atrial fibrillation and is recommended for persons at increased risk.[3-5] However, food and drug interactions necessitate frequent coagulation monitoring and dose adjustments, requirements that make it difficult for many patients to use such drugs in clinical practice.[6-8]

Rivaroxaban is a direct factor Xa inhibitor that may provide more consistent and predictable anticoagulation than warfarin.[9,10] It has been reported to prevent venous thromboembolism more effectively than enoxaparin in patients undergoing orthopedic surgery[11,12] and was noninferior to enoxaparin followed by warfarin in a study involving patients with established venous thrombosis.[13] This trial was designed to compare once-daily oral rivaroxaban with dose-adjusted warfarin for the prevention of stroke and systemic embolism in patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke.[14]

METHODS

STUDY DESIGN AND OVERSIGHT

The Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF) was a multicenter, randomized, double-blind, double-dummy, event-driven trial that was conducted at 1178 participating sites in 45 countries.[14] The study was supported by Johnson & Johnson Pharmaceutical Research and Development and Bayer HealthCare. The Duke Clinical Research Institute coordinated the trial, managed the database, and performed the primary analyses independently of the sponsors. Pertinent national regulatory authorities and ethics committees at participating centers approved the protocol, which is available with the full text of this article at NEJM.org. The members of an international executive committee designed the trial, were responsible for overseeing the study's conduct, retained the ability to independently analyze and present the data, made the decision to submit the manuscript for publication, and take responsibility for the accuracy and completeness of the data and all analyses. The first academic author wrote the initial draft of the manuscript.

STUDY PARTICIPANTS

We recruited patients with nonvalvular atrial fibrillation, as documented on electrocardiography, who were at moderate-to-high risk for stroke. Elevated risk was indicated by a history of stroke, transient ischemic attack, or systemic embolism or at least two of the following risk factors: heart failure or a left ventricular ejection fraction of 35% or less, hypertension, an age of 75 years or more, or the presence of diabetes mellitus (i.e., a $CHADS_2$ score of 2 or more, on a scale ranging from 1 to 6, with higher scores indicating a greater risk of stroke). According to the protocol, the proportion of patients who had not had a previous ischemic stroke, transient ischemic attack, or systemic embolism and who had no more than two risk factors was limited to 10% of the cohort for each region; the remainder of patients were required to have had either previous thromboembolism or three or more risk factors. Complete inclusion and exclusion criteria are provided in the Supplementary Appendix, available at NEJM.org. All patients provided written informed consent.

STUDY TREATMENT

Patients were randomly assigned to receive fixed-dose rivaroxaban (20 mg daily or 15 mg daily in patients with a creatinine clearance of 30 to 49 ml per minute) or adjusted-dose warfarin (target international normalized ratio [INR], 2.0 to 3.0). Patients in each group also received a placebo tablet in order to maintain blinding. Randomization was performed with the use of a central 24-hour, computerized, automated voice-response system. A point-of-care device was used to generate encrypted values that were sent to an independent study monitor, who provided sites with either real INR values (for patients in the warfarin group in order to adjust the dose) or sham values (for patients in the rivaroxaban group receiving placebo warfarin) during the course of the trial. Sham INR results were generated by means of a validated algorithm reflecting the distribution of values in warfarin-treated patients with characteristics similar to those in the study population.[15]

It was intended that patients would continue to take the assigned therapy throughout the course of the trial, unless discontinuation was considered to be clinically indicated. Follow-up

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

procedures and restrictions on concomitant medications are summarized in the Supplementary Appendix.

OUTCOMES

The primary efficacy end point was the composite of stroke (ischemic or hemorrhagic) and systemic embolism. Brain imaging was recommended to distinguish hemorrhagic from ischemic stroke. In the presence of atherosclerotic peripheral arterial disease, the diagnosis of embolism required angiographic demonstration of abrupt arterial occlusion.

Secondary efficacy end points included a composite of stroke, systemic embolism, or death from cardiovascular causes; a composite of stroke, systemic embolism, death from cardiovascular causes, or myocardial infarction; and individual components of the composite end points. The principal safety end point was a composite of major and nonmajor clinically relevant bleeding events. Bleeding events involving the central nervous system that met the definition of stroke were adjudicated as hemorrhagic strokes and included in both the primary efficacy and safety end points. Other overt bleeding episodes that did not meet the criteria for major or clinically relevant nonmajor bleeding were classified as minor episodes.

An independent clinical end-point committee applied protocol definitions to adjudicate all suspected cases of stroke, systemic embolism, myocardial infarction, death, and bleeding events that contributed to the prespecified end points. Detailed definitions of the end-point events are provided in the Supplementary Appendix.

STATISTICAL ANALYSIS

The primary hypothesis was that rivaroxaban would be noninferior to warfarin for the prevention of stroke or systemic embolism. The primary analysis was prespecified to be performed in the per-protocol population, which included all patients who received at least one dose of a study drug, did not have a major protocol violation, and were followed for events while receiving a study drug or within 2 days after discontinuation (group A in Fig. 1 in the Supplementary Appendix).[16-19]

For the primary analysis, we determined that a minimum of 363 events would provide a power of 95% to calculate a noninferiority margin of 1.46 with a one-sided alpha level of 0.025. However, 405 events were selected as the prespecified target to ensure a robust statistical result. On the basis of a projected event rate of 2.3% per 100 patient-years in the warfarin group and a projected 14% rate of annual attrition, it was estimated that approximately 14,000 patients would need to be randomly assigned to a study group.

If noninferiority was achieved in the primary analysis, a closed testing procedure was to be conducted for superiority in the safety population during treatment, which included patients who received at least one dose of a study drug and were followed for events, regardless of adherence to the protocol, while they were receiving the assigned study drug or within 2 days after discontinuation (group B in Fig. 1 in the Supplementary Appendix). Key secondary efficacy end points were also tested for superiority in the as-treated safety population.[20] Testing for noninferiority and superiority was also performed in the intention-to-treat population, which included all patients who underwent randomization and were followed for events during treatment or after premature discontinuation (group C in Fig. 1 in the Supplementary Appendix).

In addition, we performed post hoc analyses of events in the intention-to-treat population and events occurring during the end-of-study transition to open-label treatment with conventional anticoagulant agents. In the warfarin group, we used the method of Rosendaal et al.[21] to calculate the overall time that INR values fell within the therapeutic range. Comparative analyses of treatment efficacy were performed according to quartiles of time that INR values fell within the therapeutic range at the participating clinical sites.

Event rates per 100 patient-years are presented as proportions of patients per year. Hazard ratios, confidence intervals, and P values were calculated with the use of Cox proportional-hazards models with treatment as the only covariate. Testing for noninferiority was based on a one-sided significance level of 0.025; testing for superiority was based on a two-sided significance level of 0.05.

RESULTS

RECRUITMENT AND FOLLOW-UP

From December 18, 2006, through June 17, 2009, a total of 14,264 patients underwent randomization (Fig. 1 in the Supplementary Appendix). The study was terminated on May 28, 2010. The proportions of patients who permanently stopped

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

Table 1. Characteristics of the Intention-to-Treat Population at Baseline.

Characteristic	Rivaroxaban (N=7131)	Warfarin (N=7133)
Age — yr		
Median	73	73
Interquartile range	65–78	65–78
Female sex — no. (%)	2831 (39.7)	2832 (39.7)
Body-mass index*		
Median	28.3	28.1
Interquartile range	25.2–32.1	25.1–31.8
Blood pressure — mm Hg		
Systolic		
Median	130	130
Interquartile range	120–140	120–140
Diastolic		
Median	80	80
Interquartile range	70–85	70–85
Type of atrial fibrillation — no. (%)		
Persistent	5786 (81.1)	5762 (80.8)
Paroxysmal	1245 (17.5)	1269 (17.8)
Newly diagnosed or new onset	100 (1.4)	102 (1.4)
Previous medication use — no. (%)		
Aspirin	2586 (36.3)	2619 (36.7)
Vitamin K antagonist	4443 (62.3)	4461 (62.5)
CHADS$_2$ risk of stroke†		
Mean score (±SD)	3.48±0.94	3.46±0.95
Score — no. (%)		
2	925 (13.0)	934 (13.1)
3	3058 (42.9)	3158 (44.3)
4	2092 (29.3)	1999 (28.0)
5	932 (13.1)	881 (12.4)
6‡	123 (1.7)	159 (2.2)
Coexisting condition — no. (%)		
Previous stroke, systemic embolism, or transient ischemic attack	3916 (54.9)	3895 (54.6)
Congestive heart failure	4467 (62.6)	4441 (62.3)
Hypertension	6436 (90.3)	6474 (90.8)
Diabetes mellitus	2878 (40.4)	2817 (39.5)
Previous myocardial infarction‡	1182 (16.6)	1286 (18.0)
Peripheral vascular disease	401 (5.6)	438 (6.1)
Chronic obstructive pulmonary disease	754 (10.6)	743 (10.4)
Creatinine clearance — ml/min§		
Median	67	67
Interquartile range	52–88	52–86

* The body-mass index is the weight in kilograms divided by the square of the height in meters.

† The CHADS$_2$ score for the risk of stroke ranges from 1 to 6, with higher scores indicating an increased risk. Three patients (one in the rivaroxaban group and two in the warfarin group) had a CHADS$_2$ score of 1.

‡ P<0.05 for the between-group comparison.

§ Creatinine clearance was calculated with the use of the Cockcroft–Gault formula.

their assigned therapy before an end-point event and before the termination date were 23.7% in the rivaroxaban group and 22.2% in the warfarin group. The median duration of treatment exposure was 590 days; the median follow-up period was 707 days. Only 32 patients were lost to follow-up. Because of violations in Good Clinical Practice guidelines at one site that made the data unreliable, 93 patients (50 in the rivaroxaban group and 43 in the warfarin group) were excluded from all efficacy analyses before unblinding. An additional issue with data quality was raised at another trial site, but this issue was resolved without the exclusion of the patients from the analysis (for details, see the Supplementary Appendix).

PATIENT CHARACTERISTICS AND TREATMENTS

Key clinical characteristics of the patients who underwent randomization are shown in Table 1. The median age was 73 years (a quarter of the patients were 78 years of age or older), and 39.7% of the patients were women. The patients had substantial rates of coexisting illnesses: 90.5% had hypertension, 62.5% had heart failure, and 40.0% had diabetes; 54.8% of the patients had had a previous stroke, systemic embolism, or transient ischemic attack. The mean and median CHADS$_2$ scores were 3.5 and 3.0, respectively. Data on medication use at baseline are provided in Table 1 in the Supplementary Appendix. Previous use of vitamin K antagonists was reported by 62.4% of patients. At some time during the study, 34.9% of patients in the rivaroxaban group and 36.2% of those in the warfarin group took aspirin concurrently with the assigned study drug. Among patients in the warfarin group, INR values were within the therapeutic range (2.0 to 3.0) a mean of 55% of the time (median, 58%; interquartile range, 43 to 71).

PRIMARY OUTCOME

In the per-protocol population (the patients included in the primary efficacy analysis), stroke or systemic embolism occurred in 188 patients in the rivaroxaban group (1.7% per year) and in 241 patients in the warfarin group (2.2% per year) (hazard ratio in the rivaroxaban group, 0.79; 95% confidence interval [CI], 0.66 to 0.96; P<0.001 for noninferiority) (Table 2 and Fig. 1A). In the as-treated safety population, primary events occurred in 189 patients in the rivaroxaban group (1.7% per year) and in 243 patients in the warfarin

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

Table 2. Primary End Point of Stroke or Systemic Embolism.*

Study Population	Rivaroxaban			Warfarin			Hazard Ratio (95% CI)†	P Value	
	No. of Patients	No. of Events	Event Rate	No. of Patients	No. of Events	Event Rate		Noninferiority	Superiority
			no./100 patient-yr			*no./100 patient-yr*			
Per-protocol, as-treated population‡	6958	188	1.7	7004	241	2.2	0.79 (0.66–0.96)	<0.001	
Safety, as-treated population	7061	189	1.7	7082	243	2.2	0.79 (0.65–0.95)		0.02
Intention-to-treat population§	7081	269	2.1	7090	306	2.4	0.88 (0.75–1.03)	<0.001	0.12
During treatment		188	1.7		240	2.2	0.79 (0.66–0.96)		0.02
After discontinuation		81	4.7		66	4.3	1.10 (0.79–1.52)		0.58

* The median follow-up period was 590 days for the per-protocol, as-treated population during treatment; 590 days for the safety, as-treated population during treatment; and 707 days for the intention-to-treat population.

† Hazard ratios are for the rivaroxaban group as compared with the warfarin group.

‡ The primary analysis was performed in the as-treated, per-protocol population during treatment.

§ Follow-up in the intention-to-treat population continued until notification of study termination.

group (2.2% per year) (hazard ratio, 0.79; 95% CI, 0.65 to 0.95; P=0.01 for superiority). Among all randomized patients in the intention-to-treat analysis, primary events occurred in 269 patients in the rivaroxaban group (2.1% per year) and in 306 patients in the warfarin group (2.4% per year) (hazard ratio, 0.88; 95% CI, 0.74 to 1.03; P<0.001 for noninferiority; P=0.12 for superiority) (Fig. 1B).

During treatment in the intention-to-treat population, patients in the rivaroxaban group had a lower rate of stroke or systemic embolism (188 events, 1.7% per year) than those in the warfarin group (240 events, 2.2% per year) (P=0.02) (Table 2 and Fig. 2). Among patients who stopped taking the assigned study drug before the end of the study, during a median of 117 days of follow-up after discontinuation, primary events occurred in 81 patients in the rivaroxaban group (4.7% per year) and in 66 patients in the warfarin group (4.3% per year) (P=0.58). (Details regarding the time to events in patients who completed the study and were switched to standard medical therapy are provided in Fig. 2 in the Supplementary Appendix.)

BLEEDING OUTCOMES

Major and clinically relevant nonmajor bleeding occurred in 1475 patients in the rivaroxaban group and in 1449 patients in the warfarin group (14.9% and 14.5% per year, respectively; hazard ratio in the rivaroxaban group, 1.03; 95% CI, 0.96 to 1.11; P=0.44) (Table 3). Rates of major bleeding were similar in the rivaroxaban and warfarin groups (3.6% and 3.4%, respectively; P=0.58). Decreases in hemoglobin levels of 2 g per deciliter or more and transfusions were more common among patients in the rivaroxaban group, whereas fatal bleeding and bleeding at critical anatomical sites were less frequent. Rates of intracranial hemorrhage were significantly lower in the rivaroxaban group than in the warfarin group (0.5% vs. 0.7% per year; hazard ratio, 0.67; 95% CI, 0.47 to 0.93; P=0.02). Major bleeding from a gastrointestinal site was more common in the rivaroxaban group, with 224 bleeding events (3.2%), as compared with 154 events in the warfarin group (2.2%, P<0.001) (Table 2 in the Supplementary Appendix). (Data on nonhemorrhagic adverse events are provided in Table 3 in the Supplementary Appendix.)

SECONDARY EFFICACY OUTCOMES

The rates of secondary efficacy outcomes in the as-treated safety population are presented in Table 4 in the Supplementary Appendix. During treatment, myocardial infarction occurred in 101 patients in the rivaroxaban group and in 126 patients in the warfarin group (0.9% and 1.1% per year, respectively; hazard ratio in the rivaroxaban group, 0.81; 95% CI, 0.63 to 1.06; P=0.12). In the same analysis population, there were 208 deaths in the rivaroxaban group and 250 deaths in the warfarin group (1.9% and 2.2% per year, respectively; hazard ratio, 0.85; 95% CI, 0.70 to 1.02;

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

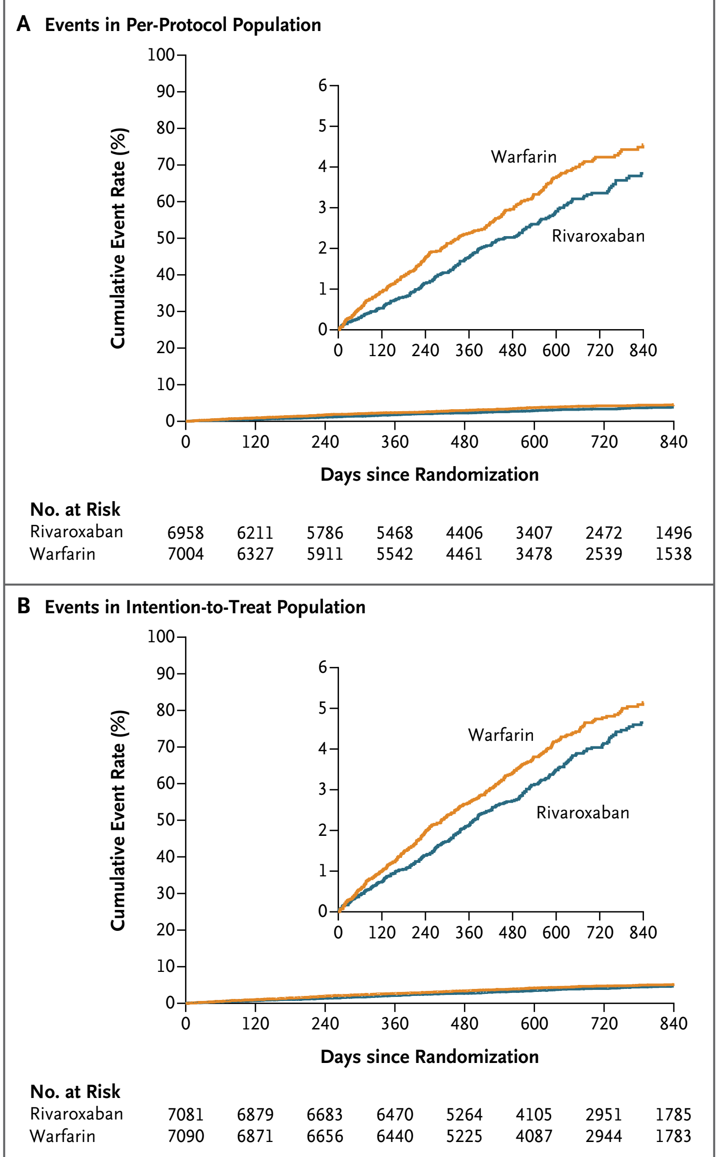


Figure 1. **Cumulative Rates of the Primary End Point (Stroke or Systemic Embolism) in the Per-Protocol Population and in the Intention-to-Treat Population.**

P=0.07). In addition, in the intention-to-treat analysis throughout the trial, there were 582 deaths in the rivaroxaban group and 632 deaths in the warfarin group (4.5% and 4.9% per year, respectively; hazard ratio, 0.92; 95% CI, 0.82 to 1.03; P=0.15).

SELECTED SUBGROUP ANALYSES

The effect of rivaroxaban, as compared with warfarin, in both efficacy and safety analyses was consistent across all prespecified subgroups (Fig. 3, 4, and 5 in the Supplementary Appendix). Furthermore, the effect of rivaroxaban did not differ across quartiles of the duration of time that INR values were within the therapeutic range according to study center (P=0.74 for interaction) (Table 5 in the Supplementary Appendix). Within the highest quartile according to center, the hazard ratio with rivaroxaban versus warfarin was 0.74 (95% CI, 0.49 to 1.12).

DISCUSSION

In this randomized trial, we compared rivaroxaban with warfarin for the prevention of stroke or systemic embolism among patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke. In both the primary analysis, which included patients in the per-protocol population, and in the intention-to-treat analysis, we found that rivaroxaban was noninferior to warfarin. In the primary safety analysis, there was no significant difference between rivaroxaban and warfarin with respect to rates of major or nonmajor clinically relevant bleeding.

As prespecified in the statistical-analysis plan, we analyzed the trial data in a variety of ways because we anticipated that some patients would discontinue the study treatment and we wished to evaluate both noninferiority and superiority. Although an intention-to-treat analysis is the standard method for assessing superiority in a randomized trial, noninferiority is best established when patients are actually taking the randomized treatment.[16-19] Thus, the primary analysis was performed in the per-protocol population during receipt of the randomly assigned therapy. In the intention-to-treat population, we found no significant between-group difference in a conventional superiority analysis. In contrast, in the analyses of patients receiving at least one dose of a study drug who were followed for events during treatment, we found that rivaroxaban was superior to warfarin. The difference between these results reflects the fact that among patients who discontinued therapy before the conclusion of the trial, no significant difference in outcomes would have been anticipated, and none was seen.

The most worrisome complication of anticoagulation is bleeding. Rates of major and nonmajor clinically relevant bleeding, the main measure of treatment safety, were similar in the rivaroxaban and warfarin groups. Bleeding that proved fatal or involved a critical anatomical site occurred less frequently in the rivaroxaban group, mainly

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

because of lower rates of hemorrhagic stroke and other intracranial bleeding. In contrast, bleeding from gastrointestinal sites, including upper, lower, and rectal sites, occurred more frequently in the rivaroxaban group, as did bleeding that led to a drop in the hemoglobin level or bleeding that required transfusion. Even though patients in our trial were at increased risk for bleeding events, rates of major bleeding were similar to those in other recent studies involving patients with atrial fibrillation.[4,15,22,23]

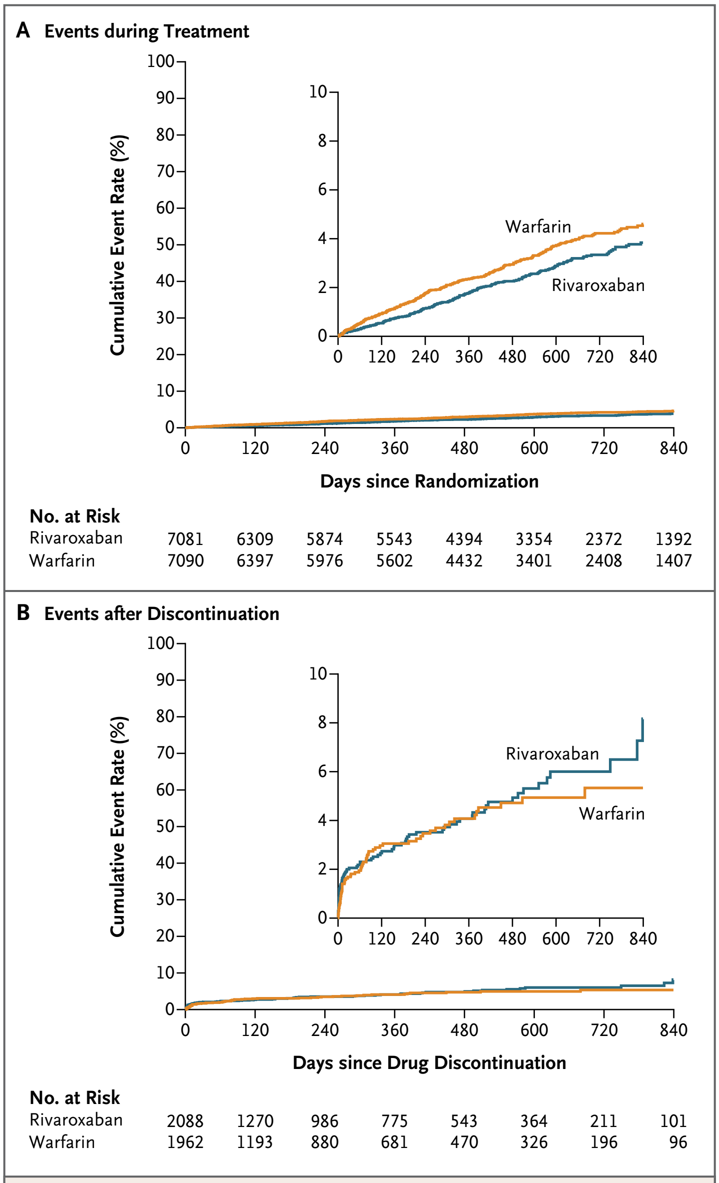


Figure 2. **Cumulative Rates of the Primary End Point during Treatment and after Discontinuation in the Intention-to-Treat Population.**

Among patients in our study who survived and did not reach the primary end point, the rate of premature, permanent cessation of randomized treatment (14.3% in year 1) was slightly higher than in other studies (average, 11%).[15,23] This may have been a consequence of the trial's double-blind design or the inclusion of patients with more coexisting illnesses. Among patients who permanently discontinued their assigned treatment before the end of the study, only about half were treated thereafter with a vitamin K antagonist. This observation suggests that for at least some of the patients who participated in the trial, the risks of open-label therapy with currently available anticoagulants were ultimately judged to outweigh the risk of stroke or systemic embolism. Event rates were similar at 30 days and 1 year after withdrawal, suggesting that the mechanism of events did not involve hypercoagulability early after withdrawal of rivaroxaban. Events occurring at the end of the study were probably related to increased difficulty in achieving the transition from blinded trial therapy to the open-label use of a vitamin K antagonist when the patient had previously been assigned to the rivaroxaban group, since presumably many patients who had previously been assigned to the warfarin group would have already had a therapeutic INR.

Among patients in the warfarin group, the proportion of time in which the intensity of anticoagulation was in the therapeutic range (mean, 55%), which was calculated from all INR values during the study and for 7 days after warfarin interruptions, was lower than in previous studies of other new anticoagulants in patients with atrial fibrillation (range, 64 to 68%). Among these trials, the only study of blinded treatment was limited to North American sites, which may have facilitated trial compliance.[15] Most earlier trials of warfarin included fewer high-risk patients,[3] and no previous studies addressed patient populations with overall levels of coexisting illnesses and geographic diversity that were similar to those of the patients in our study.[24] Significant variations in the duration of time in the therapeutic range may reflect regional differences and differential skill in managing warfarin.[25] In a recent analysis of anticoagulation management involving more than 120,000 patients in the Veterans Affairs health care system, the mean proportion of time in the therapeutic range was 58%, with significant variation across sites.[24] The efficacy of rivaroxaban, as compared with warfarin, was as favorable in centers with the best INR control as in those with poorer control.

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

Table 3. Rates of Bleeding Events.*

Variable	Rivaroxaban (N=7111)		Warfarin (N=7125)		Hazard Ratio (95% CI)†	P Value‡
	Events	Event Rate	Events	Event Rate		
	no. (%)	*no./100 patient-yr*	*no. (%)*	*no./100 patient-yr*		
Principal safety end point: major and nonmajor clinically relevant bleeding§	1475 (20.7)	14.9	1449 (20.3)	14.5	1.03 (0.96–1.11)	0.44
Major bleeding						
Any	395 (5.6)	3.6	386 (5.4)	3.4	1.04 (0.90–1.20)	0.58
Decrease in hemoglobin ≥2 g/dl	305 (4.3)	2.8	254 (3.6)	2.3	1.22 (1.03–1.44)	0.02
Transfusion	183 (2.6)	1.6	149 (2.1)	1.3	1.25 (1.01–1.55)	0.04
Critical bleeding¶	91 (1.3)	0.8	133 (1.9)	1.2	0.69 (0.53–0.91)	0.007
Fatal bleeding	27 (0.4)	0.2	55 (0.8)	0.5	0.50 (0.31–0.79)	0.003
Intracranial hemorrhage	55 (0.8)	0.5	84 (1.2)	0.7	0.67 (0.47–0.93)	0.02
Nonmajor clinically relevant bleeding	1185 (16.7)	11.8	1151 (16.2)	11.4	1.04 (0.96–1.13)	0.35

* All analyses of rates of bleeding are based on the first event in the safety population during treatment.

† Hazard ratios are for the rivaroxaban group as compared with the warfarin group and were calculated with the use of Cox proportional-hazards models with the study group as a covariate.

‡ Two-sided P values are for superiority in the rivaroxaban group as compared with the warfarin group.

§ Minimal bleeding events were not included in the principal safety end point.

¶ Bleeding events were considered to be critical if they occurred in intracranial, intraspinal, intraocular, pericardial, intraarticular, intramuscular (with compartment syndrome), or retroperitoneal sites.

In conclusion, in this trial comparing a once-daily, fixed dose of rivaroxaban with adjusted-dose warfarin in patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke, rivaroxaban was noninferior to warfarin in the prevention of subsequent stroke or systemic embolism. There were no significant differences in rates of major and clinically relevant nonmajor bleeding between the two study groups, although intracranial and fatal bleeding occurred less frequently in the rivaroxaban group.

Supported by Johnson & Johnson Pharmaceutical Research and Development and Bayer HealthCare.

Dr. Patel reports receiving consulting fees from Ortho McNeil Janssen and Bayer HealthCare and serving on an advisory board for Genzyme; Dr. Mahaffey, receiving consulting fees from Adolor, Alexion, Amgen, Argolyn Bioscience, AstraZeneca, Bayer, Bristol-Myers Squibb, Daiichi Sankyo, Eli Lilly, Elsevier, Forest Labs, Genentech, GlaxoSmithKline, Guidant, Ikarja, Johnson & Johnson, Merck, Novartis, Pfizer, Proctor & Gamble, Sanofi Aventis, Schering Plough, Scios, and WebMD, grant support from Abbott Vascular, Amgen, Amylin, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, CardioKinetix, Cierra, Cordis, Edwards LifeSciences, Eli Lilly, Genentech, GlaxoSmithKline, Guidant, Innocoll Pharmaceuticals, Johnson & Johnson, KCI Medical, Luitpold Pharmaceutical, Medtronic, Merck, Momenta Pharmaceuticals, Novartis, Portola Pharmaceutical, Pozen, Regado Biotechnologies, Sanofi Aventis, Schering Plough, and the Medicines Company, and lecture fees from Johnson & Johnson; Dr. Pan, being an employee of Johnson & Johnson; Dr. Singer, receiving consulting fees from Johnson & Johnson, Bayer HealthCare, Boehringer Ingelheim, Daiichi Sankyo, Merck, Sanofi Aventis, Medtronic, and St. Jude Medical, grant support from Daiichi Sankyo, and lecture fees from Bristol-Myers Squibb and Pfizer; Dr. Hacke, serving as a board member for Sygnis Pharma Germany and receiving consulting fees, grant support, lecture fees, and payments for the development of educational presentations from Boehringer Ingelheim, consulting fees from Photothera USA and Codman USA, and lecture fees from Bayer and Photothera USA; Dr. Breithardt, receiving consulting fees from Bayer HealthCare, Sanofi Aventis, Bristol-Myers Squibb, Boehringer Ingelheim, Boston Scientific, and Otsuka Pharmaceutical, grant support from Sanofi Aventis, St. Jude, and Meda Pharma, and lecture fees from Sanofi Aventis, Boehringer Ingelheim, Bayer HealthCare, and Boston Scientific; Dr. Halperin, receiving consulting fees from Astellas Pharma, Boehringer Ingelheim, Bristol-Myers Squibb, Daiichi Sankyo, Pfizer, and Sanofi Aventis; Dr. Hankey, receiving consulting fees from Sanofi Aventis, Schering Plough, and Boehringer Ingelheim and lecture fees from Sanofi Aventis and Pfizer; Dr. Becker, receiving grant support from Johnson & Johnson, Bayer, Regado Biosciences, and Bristol-Myers Squibb, consulting fees from Boehringer Ingelheim, and lecture fees from Merck and AstraZeneca; Dr. Nessel, being an employee of and having an equity interest in Johnson & Johnson; Drs. Paolini and Berkowitz, being employees of Bayer HealthCare, and Dr. Paolini, having an equity interest in Bayer HealthCare; Dr. Fox, receiving grant support and lecture fees from Eli Lilly and lecture fees from Sanofi Aventis and AstraZeneca; Dr. Califf, receiving consulting fees from Kowa, Nile, Orexigen, Sanofi Aventis, Novartis, and Xoma and grant support from Novartis, Merck, and Amilyn/Lilly and having an equity interest in Nitrox. No other potential conflict of interest relevant to this article was reported. Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Lisa G. Berdan, Karen Hannan, and Kimberly Schwabe for their operational oversight of the trial and Elizabeth Cook for providing editorial support in the preparation of the manuscript.

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

REFERENCES

1. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983-8.

2. *Idem*. Atrial fibrillation: a major contributor to stroke in the elderly: the Framingham Study. Arch Intern Med 1987;147:1561-4.

3. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med 1999;131:492-501.

4. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007;146:857-67.

5. Singer DE, Albers GW, Dalen JE, Go AS, Halperin JL, Manning WJ. Antithrombotic therapy in atrial fibrillation: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126:Suppl:429S-456S.

6. Albers GW, Yim JM, Belew KM, et al. Status of antithrombotic therapy for patients with atrial fibrillation in university hospitals. Arch Intern Med 1996;156:2311-6.

7. Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV, Singer DE. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. Ann Intern Med 1999;131:927-34.

8. Piccini JP, Hernandez AF, Zhao X, et al. Quality of care for atrial fibrillation among patients hospitalized for heart failure. J Am Coll Cardiol 2009;54:1280-9.

9. Kubitza D, Becka M, Wensing G, Voith B, Zuehlsdorf M. Safety, pharmacodynamics, and pharmacokinetics of BAY 59-7939 — an oral, direct Factor Xa inhibitor — after multiple dosing in healthy male subjects. Eur J Clin Pharmacol 2005;61:873-80.

10. Kubitza D, Becka M, Roth A, Mueck W. Dose-escalation study of the pharmacokinetics and pharmacodynamics of rivaroxaban in healthy elderly subjects. Curr Med Res Opin 2008;24:2757-65.

11. Eriksson BI, Borris LC, Friedman RJ, et al. Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. N Engl J Med 2008;358:2765-75.

12. Lassen MR, Ageno W, Borris LC, et al. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty. N Engl J Med 2008;358:2776-86.

13. The EINSTEIN Investigators. Oral rivaroxaban for symptomatic venous thromboembolism. N Engl J Med 2010;363:2499-510.

14. ROCKET AF Study Investigators. Rivaroxaban — once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and Embolism Trial in Atrial Fibrillation: rationale and design of the ROCKET AF study. Am Heart J 2010;159(3):340.e1-347.e1.

15. Albers GW, Diener HC, Frison L, et al. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: a randomized trial. JAMA 2005;293:690-8.

16. Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: per protocol or intent-to-treat. Stat Med 2006;25:1169-81.

17. Brittain E, Lin D. A comparison of intent-to-treat and per-protocol results in antibiotic non-inferiority trials. Stat Med 2005;24:1-10.

18. Draft: Guidance for Industry: non-inferiority clinical trials. Silver Spring, MD: Food and Drug Administration, 2010. (http://www.fda.gov/Drugs/Guidance ComplianceRegulatoryInformation/Guidances/default.htm.)

19. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA 2006;295:1152-60. [Erratum, JAMA 2006;296:1842.]

20. Soulakova JN. Comparison of several testing strategies for combination drug efficacy trials based on the closure principle. Stat Med 2009;28:260-73.

21. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993;69:236-9.

22. Connolly S, Pogue J, Hart R, et al. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet 2006;367:1903-12.

23. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009;361:1139-51. [Erratum, N Engl J Med 2010;363:1877.]

24. Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Risk-adjusted percent time in therapeutic range as a quality indicator for outpatient oral anticoagulation: results of the Veterans Affairs Study to Improve Anticoagulation (VARIA). Circ Cardiovasc Qual Outcomes 2011;4:22-9.

25. Wallentin L, Yusuf S, Ezekowitz MD, et al. Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. Lancet 2010;376:975-83.

Copyright © 2011 Massachusetts Medical Society.

RECEIVE IMMEDIATE NOTIFICATION WHEN AN ARTICLE IS PUBLISHED ONLINE FIRST

To be notified by e-mail when *Journal* articles are published Online First, sign up at **NEJM.org.**

The New England Journal of Medicine
Downloaded from nejm.org by KAREEM ROYES on September 30, 2011. For personal use only. No other uses without permission.
Copyright © 2011 Massachusetts Medical Society. All rights reserved.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 17, 2009 VOL. 361 NO. 12

Dabigatran versus Warfarin in Patients with Atrial Fibrillation

Stuart J. Connolly, M.D., Michael D. Ezekowitz, M.B., Ch.B., D.Phil., Salim Yusuf, F.R.C.P.C., D.Phil., John Eikelboom, M.D., Jonas Oldgren, M.D., Ph.D., Amit Parekh, M.D., Janice Pogue, M.Sc., Paul A. Reilly, Ph.D., Ellison Themeles, B.A., Jeanne Varrone, M.D., Susan Wang, Ph.D., Marco Alings, M.D., Ph.D., Denis Xavier, M.D., Jun Zhu, M.D., Rafael Diaz, M.D., Basil S. Lewis, M.D., Harald Darius, M.D., Hans-Christoph Diener, M.D., Ph.D., Campbell D. Joyner, M.D., Lars Wallentin, M.D., Ph.D., and the RE-LY Steering Committee and Investigators*

ABSTRACT

BACKGROUND

Warfarin reduces the risk of stroke in patients with atrial fibrillation but increases the risk of hemorrhage and is difficult to use. Dabigatran is a new oral direct thrombin inhibitor.

METHODS

In this noninferiority trial, we randomly assigned 18,113 patients who had atrial fibrillation and a risk of stroke to receive, in a blinded fashion, fixed doses of dabigatran — 110 mg or 150 mg twice daily — or, in an unblinded fashion, adjusted-dose warfarin. The median duration of the follow-up period was 2.0 years. The primary outcome was stroke or systemic embolism.

RESULTS

Rates of the primary outcome were 1.69% per year in the warfarin group, as compared with 1.53% per year in the group that received 110 mg of dabigatran (relative risk with dabigatran, 0.91; 95% confidence interval [CI], 0.74 to 1.11; $P<0.001$ for noninferiority) and 1.11% per year in the group that received 150 mg of dabigatran (relative risk, 0.66; 95% CI, 0.53 to 0.82; $P<0.001$ for superiority). The rate of major bleeding was 3.36% per year in the warfarin group, as compared with 2.71% per year in the group receiving 110 mg of dabigatran ($P=0.003$) and 3.11% per year in the group receiving 150 mg of dabigatran ($P=0.31$). The rate of hemorrhagic stroke was 0.38% per year in the warfarin group, as compared with 0.12% per year with 110 mg of dabigatran ($P<0.001$) and 0.10% per year with 150 mg of dabigatran ($P<0.001$). The mortality rate was 4.13% per year in the warfarin group, as compared with 3.75% per year with 110 mg of dabigatran ($P=0.13$) and 3.64% per year with 150 mg of dabigatran ($P=0.051$).

CONCLUSIONS

In patients with atrial fibrillation, dabigatran given at a dose of 110 mg was associated with rates of stroke and systemic embolism that were similar to those associated with warfarin, as well as lower rates of major hemorrhage. Dabigatran administered at a dose of 150 mg, as compared with warfarin, was associated with lower rates of stroke and systemic embolism but similar rates of major hemorrhage. (ClinicalTrials.gov number, NCT00262600.)

From the Population Health Research Institute, McMaster University and Hamilton Health Sciences, Hamilton, ON, Canada (S.J.C., S.Y., J.E., J.P., E.T.); Lankenau Institute for Medical Research and the Heart Center, Wynnewood, PA (M.D.E., A.P.); Uppsala Clinical Research Center, Uppsala, Sweden (J.O., L.W.); Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT (P.A.R., J.V., S.W.); Working Group on Cardiovascular Research the Netherlands, Utrecht, the Netherlands (M.A.); St. John's National Academy of Health Sciences, Bangalore, India (D.X.); FuWai Hospital, Beijing (J.Z.); Estudios Clínicos Latinoamérica, Rosario, Argentina (R.D.); Lady Davis Carmel Medical Center, Haifa, Israel (B.S.L.); Vivantes Klinikum Neukölln, Berlin (H.D.); University Duisburg-Essen, Essen, Germany (H.-C.D.); and Sunnybrook Health Sciences Centre, Toronto (C.D.J.). Address reprint requests to Dr. Connolly at the Population Health Research Institute, 237 Barton St. E., Hamilton, ON L8L 2X2, Canada, or at connostu@phri.ca.

*Members of the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) Study Group are listed in the Appendix and the Supplementary Appendix, available with the full text of this article at NEJM.org.

Drs. Connolly, Ezekowitz, Yusuf, and Wallentin contributed equally to this article.

This article (10.1056/NEJMoa0905561) was published on August 30, 2009, and updated on September 16, 2009, at NEJM.org.

N Engl J Med 2009;361:1139-51.

Copyright © 2009 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Atrial fibrillation increases the risks of stroke and death. Vitamin K antagonists, such as warfarin, reduce the risks of stroke and death but increase the risk of hemorrhage as compared with control therapy.[1] Therefore, warfarin is recommended for patients who have atrial fibrillation and are at risk for stroke.[2]

Vitamin K antagonists are cumbersome to use, because of their multiple interactions with food and drugs, and they require frequent laboratory monitoring. Therefore, they are often not used, and when they are, rates of discontinuation are high.[3,4] Many patients receiving warfarin still have inadequate anticoagulation.[5] Thus, there is a need for new anticoagulant agents that are effective, safe, and convenient to use.

Dabigatran etexilate is an oral prodrug that is rapidly converted by a serum esterase to dabigatran, a potent, direct, competitive inhibitor of thrombin. It has an absolute bioavailability of 6.5%, 80% of the given dose is excreted by the kidneys, its serum half-life is 12 to 17 hours, and it does not require regular monitoring.[6] Dabigatran has been evaluated in a pilot trial involving patients with atrial fibrillation and in a study for the prevention of venous thromboembolism, in which doses of 150 mg twice daily and 220 mg once daily, respectively, were promising.[7,8] We performed a large, randomized trial comparing the use of dabigatran, at doses of 110 mg twice daily and 150 mg twice daily, with warfarin.

METHODS

TRIAL DESIGN

The Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) was a randomized trial designed to compare two fixed doses of dabigatran, each administered in a blinded manner, with open-label use of warfarin in patients who had atrial fibrillation and were at increased risk for stroke. The design of this study has been described previously.[9]

The study was funded by Boehringer Ingelheim and was coordinated by the Population Health Research Institute (Hamilton, ON, Canada), which independently managed the database and performed the primary data analyses. An operations committee, with assistance from an international steering committee and with participation by the sponsor, was responsible for the design, conduct, and reporting of the study. The study was approved by all appropriate national regulatory authorities and ethics committees of the participating centers. All the authors vouch for the accuracy and completeness of the data and the analyses.

STUDY PARTICIPANTS

Patients were recruited from 951 clinical centers in 44 countries. In brief, patients were eligible if they had atrial fibrillation documented on electrocardiography performed at screening or within 6 months beforehand and at least one of the following characteristics: previous stroke or transient ischemic attack, a left ventricular ejection fraction of less than 40%, New York Heart Association class II or higher heart-failure symptoms within 6 months before screening, and an age of at least 75 years or an age of 65 to 74 years plus diabetes mellitus, hypertension, or coronary artery disease. Reasons for exclusion were the presence of a severe heart-valve disorder, stroke within 14 days or severe stroke within 6 months before screening, a condition that increased the risk of hemorrhage, a creatinine clearance of less than 30 ml per minute, active liver disease, and pregnancy. (Detailed inclusion and exclusion criteria are available in Tables 1 and 2 of the Supplementary Appendix, available with the full text of this article at NEJM.org.)

PROCEDURES

After providing written informed consent, all trial participants were randomly assigned to receive one of two doses of dabigatran, or to receive warfarin, by means of a central, interactive, automated telephone system. Dabigatran was administered, in a blinded fashion, in capsules containing either 110 mg or 150 mg of the drug, to be taken twice daily. Warfarin was administered, in an unblinded fashion, in tablets of 1, 3, or 5 mg and was adjusted locally to an international normalized ratio (INR) of 2.0 to 3.0, with the INR measured at least monthly. The time that the INR was within the therapeutic range was calculated with the use of the method of Rosendaal et al.,[10] excluding INRs from the first week and after discontinuation of the study drug. These data were reported back to the participating centers with advice for optimal INR control. Concomitant use of aspirin (at a dose of <100 mg per day) or other antiplatelet agents was permitted. Quinidine use was

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

permitted until 2 years after the trial started, when the protocol was amended to prohibit its use, because of its potential to interact with dabigatran.

Follow-up visits occurred 14 days after randomization, at 1 and 3 months, every 3 months thereafter in the first year, and then every 4 months until the study ended. Liver-function testing was performed monthly during the first year of the follow-up period. On the basis of a prespecified evaluation of liver-function tests in at least 6000 patients in the dabigatran group after they had been followed for 6 months or more, the data safety monitoring board recommended that the frequency of liver-function testing be reduced, with such testing performed only at the regular visits.

OUTCOMES

The primary study outcome was stroke or systemic embolism. The primary safety outcome was major hemorrhage. Secondary outcomes were stroke, systemic embolism, and death. Other outcomes were myocardial infarction, pulmonary embolism, transient ischemic attack, and hospitalization. The primary net clinical benefit outcome was the composite of stroke, systemic embolism, pulmonary embolism, myocardial infarction, death, or major hemorrhage. Stroke was defined as the sudden onset of a focal neurologic deficit in a location consistent with the territory of a major cerebral artery and categorized as ischemic, hemorrhagic, or unspecified. Hemorrhagic transformation of ischemic stroke was not considered to be hemorrhagic stroke. Intracranial hemorrhage consisted of hemorrhagic stroke and subdural or subarachnoid hemorrhage. Systemic embolism was defined as an acute vascular occlusion of an extremity or organ, documented by means of imaging, surgery, or autopsy. Major bleeding was defined as a reduction in the hemoglobin level of at least 20 g per liter, transfusion of at least 2 units of blood, or symptomatic bleeding in a critical area or organ. Life-threatening bleeding was a subcategory of major bleeding that consisted of fatal bleeding, symptomatic intracranial bleeding, bleeding with a decrease in the hemoglobin level of at least 50 g per liter, or bleeding requiring transfusion of at least 4 units of blood or inotropic agents or necessitating surgery. All other bleeding was considered minor.

An international team of adjudicators reviewed documents in local languages after blinding, or documents were translated by an independent group and were centrally blinded. Each primary and secondary outcome event was adjudicated by two independent investigators who were unaware of the treatment assignments. All transient ischemic attacks were reviewed to ensure that strokes had not been missed. To detect possible unreported events, symptom questionnaires were regularly administered to patients, and adverse-event and hospitalization reports were scrutinized for unreported primary or secondary outcomes.

STATISTICAL ANALYSIS

The primary analysis was designed to test whether either dose of dabigatran was noninferior to warfarin, as evaluated with the use of Cox proportional-hazards modeling. To satisfy the noninferiority hypothesis, the upper bound of the one-sided 97.5% confidence interval for the relative risk of an outcome with dabigatran as compared with warfarin needed to fall below 1.46. This noninferiority margin was derived from a meta-analysis of trials of vitamin K antagonists as compared with control therapy in patients with atrial fibrillation, with the margin defined according to the upper bound of the 95% confidence interval for the relative risk of the primary outcome in the control group versus the warfarin group.[11] The margin of 1.46 represents half the 95% confidence interval of the estimated effect of control therapy over warfarin. To account for testing of both dabigatran doses against warfarin, we planned to determine whether the higher of the two one-sided P values for the two doses was less than 0.025, in which case both treatments would be declared to be noninferior. If the higher of the two one-sided P values was 0.025 or greater, the lower of the two was required to be less than 0.0125 to permit a claim of statistical significance. All analyses were based on the intention-to-treat principle. After noninferiority of both doses of dabigatran was established, all subsequent P values were reported for two-tailed tests of superiority. Cox regression was used to calculate relative risks, confidence intervals, and P values. Chi-square testing was used to compare rates of medication discontinuation and adverse events.

We planned to enroll 15,000 patients, an enrollment that we estimated would provide 84% power to evaluate the noninferiority of each dose

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

of dabigatran. Two protocol changes were made by the operations committee during the enrollment period, without knowledge of emerging treatment effects. These were the enforcement of balanced enrollment of patients who had not received long-term therapy with a vitamin K antagonist (i.e., had a total lifetime use of <61 days) and those who had (i.e., had a total lifetime use of ≥61 days), and an increase in the sample size to 18,000 patients to maintain the statistical power in case of a low event rate. An independent data safety monitoring board reviewed the unblinded study data and performed two prespecified interim analyses of efficacy, with a plan to recommend study termination if the benefit of dabigatran exceeded 3 SD from unity of the parameter estimate and if that benefit persisted on repeat analysis 3 months later.

RESULTS

CHARACTERISTICS OF THE STUDY PATIENTS

A total of 18,113 patients were enrolled between December 22, 2005, and December 15, 2007. The three treatment groups were well balanced with respect to baseline characteristics (Table 1). The mean age of the patients was 71 years, and 63.6% were men. Half the patients had received long-term therapy with vitamin K antagonists. The mean $CHADS_2$ score was 2.1 (Table 1).

FOLLOW-UP DATA

Final follow-up visits occurred between December 15, 2008, and March 15, 2009. The median duration of the follow-up period was 2.0 years, and complete follow-up was achieved in 99.9% of patients, with 20 patients lost to follow-up. The rates of discontinuation for 110 mg of dabigatran, 150 mg of dabigatran, and warfarin were 14.5%, 15.5%, and 10.2%, respectively, at 1 year and 20.7%, 21.2%, and 16.6% at 2 years. Aspirin was used continuously during the treatment period in 21.1%, 19.6%, and 20.8% of patients receiving 110 mg of dabigatran, 150 mg of dabigatran, and warfarin, respectively. In the warfarin group, the mean percentage of the study period during which the INR was within the therapeutic range was 64%.

PRIMARY OUTCOME

Stroke or systemic embolism occurred in 182 patients receiving 110 mg of dabigatran (1.53% per year), 134 patients receiving 150 mg of dabigatran (1.11% per year), and 199 patients receiving warfarin (1.69% per year) (Table 2 and Fig. 1). Both doses of dabigatran were noninferior to warfarin (P<0.001). The 150-mg dose of dabigatran was also superior to warfarin (relative risk, 0.66; 95% confidence interval [CI], 0.53 to 0.82; P<0.001), but the 110-mg dose was not (relative risk, 0.91; 95% CI, 0.74 to 1.11; P=0.34). Rates of hemorrhagic stroke were 0.38% per year in the warfarin group, as compared with 0.12% per year in the group that received 110 mg of dabigatran (relative risk with dabigatran, 0.31; 95% CI, 0.17 to 0.56; P<0.001) and 0.10% per year in the group that received 150 mg of dabigatran (relative risk, 0.26; 95% CI, 0.14 to 0.49; P<0.001).

OTHER OUTCOMES

Rates of death from any cause were 4.13% per year with warfarin, as compared with 3.75% per year with 110 mg of dabigatran (relative risk with dabigatran, 0.91; 95% CI 0.80 to 1.03; P=0.13) and 3.64% per year with 150 mg of dabigatran (relative risk, 0.88; 95% CI, 0.77 to 1.00; P=0.051). The rate of myocardial infarction was 0.53% per year with warfarin and was higher with dabigatran: 0.72% per year in the 110-mg group (relative risk, 1.35; 95% CI, 0.98 to 1.87; P=0.07) and 0.74% per year in the 150-mg group (relative risk, 1.38, 95% CI, 1.00 to 1.91; P=0.048).

BLEEDING

The rate of major bleeding was 3.36% per year in the warfarin group, as compared with 2.71% per year in the group that received 110 mg of dabigatran (relative risk with dabigatran, 0.80; 95% CI, 0.69 to 0.93; P=0.003) and 3.11% per year in the group that received 150 mg of dabigatran (relative risk, 0.93; 95% CI, 0.81 to 1.07; P=0.31) (Table 3). Rates of life-threatening bleeding, intracranial bleeding, and major or minor bleeding were higher with warfarin (1.80%, 0.74%, and 18.15%, respectively) than with either the 110-mg dose of dabigatran (1.22%, 0.23%, and 14.62%, respectively) or the 150-mg dose of dabigatran (1.45%, 0.30%, and 16.42%, respectively) (P<0.05 for all comparisons of dabigatran with warfarin). There was a significantly higher rate of major gastrointestinal bleeding with dabigatran at the 150-mg dose than with warfarin.

The net clinical benefit outcome consisted of major vascular events, major bleeding, and death. The rates of this combined outcome were 7.64% per year with warfarin and 7.09% per year with 110 mg of dabigatran (relative risk with dabiga-

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Table 1. **Baseline Characteristics of the Study Participants, According to Treatment Group.***

Characteristic	Dabigatran, 110 mg	Dabigatran, 150 mg	Warfarin
Age — yr	71.4±8.6	71.5±8.8	71.6±8.6
Weight — kg	82.9±19.9	82.5±19.4	82.7±19.7
Blood pressure — mm Hg			
Systolic	130.8±17.5	131.0±17.6	131.2±17.4
Diastolic	77.0±10.6	77.0±10.6	77.1±10.4
Male sex — no./total no. (%)	3865/6015 (64.3)	3840/6076 (63.2)	3809/6022 (63.3)
Type of atrial fibrillation — no./total no. (%)			
Persistent	1950/6011 (32.4)	1909/6075 (31.4)	1930/6021 (32.0)
Paroxysmal	1929/6011 (32.1)	1978/6075 (32.6)	2036/6021 (33.8)
Permanent	2132/6011 (35.4)	2188/6075 (36.0)	2055/6021 (34.1)
$CHADS_2$ score†	2.1±1.1	2.2±1.2	2.1±1.1
0 or 1 — no./total no. (%)	1958/6014 (32.6)	1958/6076 (32.2)	1859/6022 (30.9)
2 — no./total no. (%)	2088/6014 (34.7)	2137/6076 (35.2)	2230/6022 (37.0)
3–6 — no./total no. (%)	1968/6014 (32.7)	1981/6076 (32.6)	1933/6022 (32.1)
Previous stroke or transient ischemic attack — no./total no. (%)	1195/6015 (19.9)	1233/6076 (20.3)	1195/6022 (19.8)
Prior myocardial infarction — no./total no. (%)	1008/6015 (16.8)	1029/6076 (16.9)	968/6022 (16.1)
Heart failure — no./total no. (%)	1937/6015 (32.2)	1934/6076 (31.8)	1922/6022 (31.9)
Diabetes mellitus — no./total no. (%)	1409/6015 (23.4)	1402/6076 (23.1)	1410/6022 (23.4)
Hypertension — no./total no. (%)	4738/6015 (78.8)	4795/6076 (78.9)	4750/6022 (78.9)
Medications in use at baseline — no./total no. (%)			
Aspirin	2404/6013 (40.0)	2352/6075 (38.7)	2442/6017 (40.6)
ARB or ACE inhibitor	3987/6013 (66.3)	4053/6075 (66.7)	3939/6017 (65.5)
Beta-blocker	3784/6013 (62.9)	3872/6075 (63.7)	3719/6017 (61.8)
Amiodarone	624/6013 (10.4)	665/6075 (10.9)	644/6017 (10.7)
Statin‡	2698/6013 (44.9)	2667/6075 (43.9)	2673/6017 (44.4)
Proton-pump inhibitor	812/6013 (13.5)	847/6075 (13.9)	832/6017 (13.8)
H_2-receptor antagonist	225/6013 (3.7)	241/6075 (4.0)	256/6017 (4.3)
Long-term VKA therapy	3011/6015 (50.1)	3049/6076 (50.2)	2929/6022 (48.6)

* Plus-minus values are means ±SD. ARB denotes angiotensin-receptor blocker, and ACE angiotensin-converting enzyme.

† The $CHADS_2$ score is a measure of the risk of stroke in which congestive heart failure, hypertension, an age of 75 years or older, and diabetes mellitus are each assigned 1 point and previous stroke or transient ischemic attack is assigned 2 points; the score is calculated by summing all the points for a given patient.[12]

‡ Statins are defined here as 3-hydroxy-3-methylglutaryl–coenzyme A reductase inhibitors.

§ Long-term therapy with a vitamin K antagonist (VKA) denotes a total lifetime use of a VKA of 61 or more days.

tran, 0.92; 95% CI, 0.84 to 1.02; P=0.10) and 6.91% per year with 150 mg of dabigatran (relative risk, 0.91; 95% CI, 0.82 to 1.00; P=0.04).

COMPARISON OF DABIGATRAN DOSES

As compared with the 110-mg dose, administration of the 150-mg dose of dabigatran reduced the risk of stroke or systemic embolism (P=0.005). This difference was driven mostly by a decrease in the rate of stroke with ischemic or unspecified cause, whereas rates of hemorrhagic stroke were similar in the two dabigatran groups. There was no significant difference in the rates of death from either vascular causes or any cause between the two doses. On the other hand, as compared with the 110-mg dose, the 150-mg dose of dabigatran

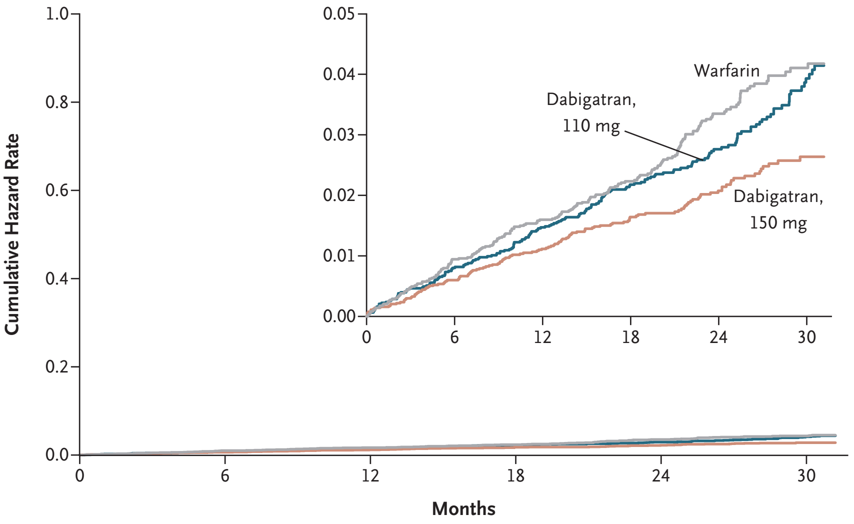
The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Table 2. Efficacy Outcomes, According to Treatment Group.

Event	Dabigatran, 110 mg (N=6015)		Dabigatran, 150 mg (N=6076)		Warfarin (N=6022)		Dabigatran, 110 mg, vs. Warfarin		Dabigatran, 150 mg, vs. Warfarin		Dabigatran, 150 mg vs. 110 mg	
							Relative Risk (95% CI)	P Value	Relative Risk (95% CI)	P Value	Relative Risk (95% CI)	P Value
	no. of patients	*%/yr*	*no. of patients*	*%/yr*	*no. of patients*	*%/yr*						
Stroke or systemic embolism*	182	1.53	134	1.11	199	1.69	0.91 (0.74–1.11)	<0.001 for noninferiority, 0.34	0.66 (0.53–0.82)	<0.001 for noninferiority, <0.001	0.73 (0.58–0.91)	0.005
Stroke	171	1.44	122	1.01	185	1.57	0.92 (0.74–1.13)	0.41	0.64 (0.51–0.81)	<0.001	0.70 (0.56–0.89)	0.003
Hemorrhagic	14	0.12	12	0.10	45	0.38	0.31 (0.17–0.56)	<0.001	0.26 (0.14–0.49)	<0.001	0.85 (0.39–1.83)	0.67
Ischemic or unspecified	159	1.34	111	0.92	142	1.20	1.11 (0.89–1.40)	0.35	0.76 (0.60–0.98)	0.03	0.69 (0.54–0.88)	0.002
Nondisabling stroke	60	0.50	44	0.37	69	0.58	0.86 (0.61–1.22)	0.40	0.62 (0.43–0.91)	0.01	0.72 (0.49–1.07)	0.10
Disabling or fatal stroke	112	0.94	80	0.66	118	1.00	0.94 (0.73–1.22)	0.65	0.66 (0.50–0.88)	0.005	0.70 (0.53–0.94)	0.02
Myocardial infarction	86	0.72	89	0.74	63	0.53	1.35 (0.98–1.87)	0.07	1.38 (1.00–1.91)	0.048	1.02 (0.76–1.38)	0.88
Pulmonary embolism	14	0.12	18	0.15	11	0.09	1.26 (0.57–2.78)	0.56	1.61 (0.76–3.42)	0.21	1.27 (0.63–2.56)	0.50
Hospitalization	2311	19.4	2430	20.2	2458	20.8	0.92 (0.87–0.97)	0.003	0.97 (0.92–1.03)	0.34	1.06 (1.00–1.12)	0.04
Death from vascular causes	289	2.43	274	2.28	317	2.69	0.90 (0.77–1.06)	0.21	0.85 (0.72–0.99)	0.04	0.94 (0.79–1.11)	0.44
Death from any cause	446	3.75	438	3.64	487	4.13	0.91 (0.80–1.03)	0.13	0.88 (0.77–1.00)	0.051	0.97 (0.85–1.11)	0.66

* Data are shown for all patients who had at least one event. All analyses were based on the time to the first event. P values are for superiority, unless otherwise indicated. The modified Rankin scale (on which scores can range from 0 [no neurologic disability] to 5 [severe disability], with 6 indicating a fatal stroke) was used to categorize stroke: nondisabling stroke was defined by a score of 0 to 2, and disabling or fatal stroke, a score of 3 to 6.

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.



No. at Risk						
Warfarin	6022	5862	5718	4593	2890	1322
Dabigatran, 110 mg	6015	5862	5710	4593	2945	1385
Dabigatran, 150 mg	6076	5939	5779	4682	3044	1429

Figure 1. Cumulative Hazard Rates for the Primary Outcome of Stroke or Systemic Embolism, According to Treatment Group.

was associated with a trend toward an increased risk of major bleeding (P=0.052) and also with increased risks of gastrointestinal, minor, and any bleeding. The net clinical benefit was almost identical for the two doses.

ADVERSE EVENTS AND LIVER FUNCTION

The only adverse effect that was significantly more common with dabigatran than with warfarin was dyspepsia (Table 4). Dyspepsia occurred in 348 patients (5.8%) in the warfarin group and in 707 patients (11.8%) and 688 patients (11.3%) in the 110-mg and 150-mg dabigatran groups, respectively (P<0.001 for both comparisons) (Table 4). Elevations in the serum aspartate aminotransferase or alanine aminotransferase level of more than 3 times the upper limit of the normal range did not occur more frequently with dabigatran, at either dose, than with warfarin.

SUBGROUP ANALYSES

For the subgroups shown in Figure 2, no significant interaction was seen with the treatment effect of dabigatran (at either dose). There was no significant interaction between the treatment effect of dabigatran and presence or absence of long-term therapy with a vitamin K antagonist. Although 80% of the dabigatran dose is renally excreted, there was no significant interaction in the treatment effect of dabigatran across levels of the baseline calculated creatinine clearance.

DISCUSSION

We compared two fixed-dose regimens of dabigatran (110 mg twice daily and 150 mg twice daily), administered in a blinded fashion, with adjusted-dose warfarin, administered in an unblinded fashion, in patients who had atrial fibrillation and were at risk for stroke. Both dabigatran doses were noninferior to warfarin with respect to the primary efficacy outcome of stroke or systemic embolism. In addition, the 150-mg dose of dabigatran was superior to warfarin with respect to stroke or systemic embolism, and the 110-mg dose was superior to warfarin with respect to major bleeding.

Previous studies seeking to identify a safe and effective alternative to warfarin for patients with atrial fibrillation have all had specific limitations. The combination of clopidogrel and aspirin was more effective than aspirin alone[13] but less effective than warfarin.[14] Subcutaneous idraparinux was more effective than warfarin but was associated with a substantially higher risk of bleeding.[15]

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Table 3. Safety Outcomes, According to Treatment Group.*

Event	Dabigatran, 110 mg		Dabigatran, 150 mg		Warfarin		Dabigatran, 110 mg, vs. Warfarin		Dabigatran, 150 mg, vs. Warfarin		Dabigatran, 150 mg vs. 110 mg	
							Relative Risk (95% CI)	P Value	Relative Risk (95% CI)	P Value	Relative Risk (95% CI)	P Value
	no. of patients	%/yr	*no. of patients*	%/yr	*no. of patients*	%/yr						
Major bleeding	322	2.71	375	3.11	397	3.36	0.80 (0.69–0.93)	0.003	0.93 (0.81–1.07)	0.31	1.16 (1.00–1.34)	0.052
Life threatening	145	1.22	175	1.45	212	1.80	0.68 (0.55–0.83)	<0.001	0.81 (0.66–0.99)	0.04	1.19 (0.96–1.49)	0.11
Non–life threatening	198	1.66	226	1.88	208	1.76	0.94 (0.78–1.15)	0.56	1.07 (0.89–1.29)	0.47	1.14 (0.95–1.39)	0.17
Gastrointestinal†	133	1.12	182	1.51	120	1.02	1.10 (0.86–1.41)	0.43	1.50 (1.19–1.89)	<0.001	1.36 (1.09–1.70)	0.007
Minor bleeding	1566	13.16	1787	14.84	1931	16.37	0.79 (0.74–0.84)	<0.001	0.91 (0.85–0.97)	0.005	1.16 (1.08–1.24)	<0.001
Major or minor bleeding	1740	14.62	1977	16.42	2142	18.15	0.78 (0.74–0.83)	<0.001	0.91 (0.86–0.97)	0.002	1.16 (1.09–1.23)	<0.001
Intracranial bleeding	27	0.23	36	0.30	87	0.74	0.31 (0.20–0.47)	<0.001	0.40 (0.27–0.60)	<0.001	1.32 (0.80–2.17)	0.28
Extracranial bleeding	299	2.51	342	2.84	315	2.67	0.94 (0.80–1.10)	0.45	1.07 (0.92–1.25)	0.38	1.14 (0.97–1.33)	0.11
Net clinical benefit outcome‡	844	7.09	832	6.91	901	7.64	0.92 (0.84–1.02)	0.10	0.91 (0.82–1.00)	0.04	0.98 (0.89–1.08)	0.66

* Data are shown for all patients who had at least one event. All analyses were based on the time to the first event. Hemorrhagic stroke was a subcategory of stroke in the efficacy analysis and in the safety analysis is also counted as major, life-threatening bleeding and as part of intracranial bleeding.

† Gastrointestinal bleeding could be life threatening or non–life threatening.

‡ The net clinical benefit outcome was the composite of stroke, systemic embolism, pulmonary embolism, myocardial infarction, death, or major bleeding.

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Table 4. **Discontinuation of the Study Drug, Adverse Events, and Liver Function According to Treatment Group.***

Variable	Dabigatran, 110 mg (N=6015)	Dabigatran, 150 mg (N=6076)	Warfarin (N=6022)
		number of patients (percent)	
Study-drug discontinuation			
Discontinued at 1 yr†	862 (15)	935 (16)	608 (10)
Discontinued at 2 yr†	1161 (21)	1211 (21)	902 (17)
Reason for discontinuation			
Patient's decision	440 (7.3)	474 (7.8)	375 (6.2)
Outcome event	192 (3.2)	164 (2.7)	130 (2.2)
Serious adverse event‡	163 (2.7)	166 (2.7)	105 (1.7)
Gastrointestinal symptoms§	134 (2.2)	130 (2.1)	38 (0.6)
Gastrointestinal bleeding	58 (1.0)	80 (1.3)	54 (0.9)
Adverse events¶			
Dyspepsia‡‖	707 (11.8)	688 (11.3)	348 (5.8)
Dizziness	486 (8.1)	506 (8.3)	568 (9.4)
Dyspnea	557 (9.3)	580 (9.5)	586 (9.7)
Peripheral edema	473 (7.9)	478 (7.9)	468 (7.8)
Fatigue	399 (6.6)	401 (6.6)	372 (6.2)
Cough	344 (5.7)	348 (5.7)	364 (6.0)
Chest pain	312 (5.2)	377 (6.2)	357 (5.9)
Back pain	316 (5.3)	314 (5.2)	337 (5.6)
Arthralgia	270 (4.5)	335 (5.5)	346 (5.7)
Nasopharyngitis	337 (5.6)	330 (5.4)	336 (5.6)
Diarrhea	377 (6.3)	397 (6.5)	346 (5.7)
Atrial fibrillation	330 (5.5)	357 (5.9)	349 (5.8)
Urinary tract infection	273 (4.5)	289 (4.8)	335 (5.6)
Upper respiratory tract infection	288 (4.8)	285 (4.7)	313 (5.2)
Liver function			
ALT or AST >3× ULN	124 (2.1)	117 (1.9)	132 (2.2)
ALT or AST >3× ULN with concurrent bilirubin >2× ULN	13 (0.2)	13 (0.2)	21 (0.3)
Hepatobiliary disorder**			
Serious adverse event	33 (0.5)	34 (0.6)	33 (0.5)
Non–serious adverse event	101 (1.7)	109 (1.8)	112 (1.9)

* ALT denotes alanine aminotransferase, AST aspartate aminotransferase, and ULN upper limit of the normal range.

† Rates of discontinuation at 1 and 2 years were higher with dabigatran than with warfarin (P<0.001). The rates are based on Kaplan–Meier estimates.

‡ P<0.001 for the comparison of either dose of dabigatran with warfarin.

§ Gastrointestinal disorders included pain, vomiting, and diarrhea.

¶ The adverse events listed are those that were reported in more than 5% of patients in any of the three treatment groups.

‖ Dyspepsia was defined to include the coding terms abdominal pain upper, abdominal pain, abdominal discomfort, and dyspepsia.

** Hepatobiliary disorders were classified as serious adverse events if they consisted of clinical or biochemical liver dysfunction requiring hospitalization, most frequently cholelithiasis or cholecystitis. Hepatobiliary disorders classified as adverse events were most frequently cholelithiasis, cholecystitis, abnormal hepatic function, and jaundice.

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Figure 2 (facing page). Relative Risk of the Primary Outcome of Stroke or Systemic Embolism with Dabigatran versus Warfarin, According to Subgroup.

Ethnic group was self-reported. Long-term therapy with a vitamin K antagonist (VKA) denotes a total lifetime use of a VKA of 61 days or more. The body-mass index is the weight in kilograms divided by the square of the height in meters. The $CHADS_2$ score is a measure of the risk of stroke in which congestive heart failure, hypertension, an age of 75 years or older, and diabetes mellitus are each assigned 1 point and previous stroke or transient ischemic attack is assigned 2 points; the score is calculated by summing all the points for a given patient.[12] Creatinine clearance was calculated according to the Cockcroft–Gault method. The squares with horizontal lines are hazard ratios and corresponding 95% confidence intervals; the sizes of squares are proportional to the sizes of the subgroups. PPI denotes proton-pump inhibitor.

Ximelagatran, an earlier direct thrombin inhibitor, appeared to be similar to warfarin with respect to efficacy and safety but was found to be hepatotoxic.[16] In our serial measurement of liver function, we did not find evidence of hepatotoxicity with dabigatran.

The rate of myocardial infarction was higher with both doses of dabigatran than with warfarin. An explanation might be that warfarin provides better protection against coronary ischemic events than dabigatran, and warfarin is known to reduce the risk of myocardial infarction.[17] However, rates of myocardial infarction were similar between patients with atrial fibrillation who were receiving warfarin and those who were receiving ximelagatran, another direct thrombin inhibitor.[16] The explanation for this finding is therefore uncertain.

The most devastating complication of warfarin therapy is intracranial hemorrhage, especially hemorrhagic stroke. As compared with aspirin, warfarin doubles the risk of intracranial hemorrhage.[1] Thus, our finding that the rate of this complication with both doses of dabigatran was less than one third the rate with warfarin, without a reduction in the efficacy against ischemic stroke, suggests an important advantage of dabigatran. The rate of major bleeding with warfarin was higher in our study than in some previous trials.[11,13,14] This is partly explained by the more inclusive definition of major bleeding in our study. There was an increase in the rate of gastrointestinal bleeding with the higher dabigatran dose, despite the overall lower rates of bleeding at other sites. To enhance absorption of dabigatran, a low pH is required. Therefore, dabigatran capsules contain dabigatran-coated pellets with a tartaric acid core. This acidity may partly explain the increased incidence of dyspeptic symptoms with both dabigatran doses and the increased risk of gastrointestinal bleeding with the 150-mg dose.

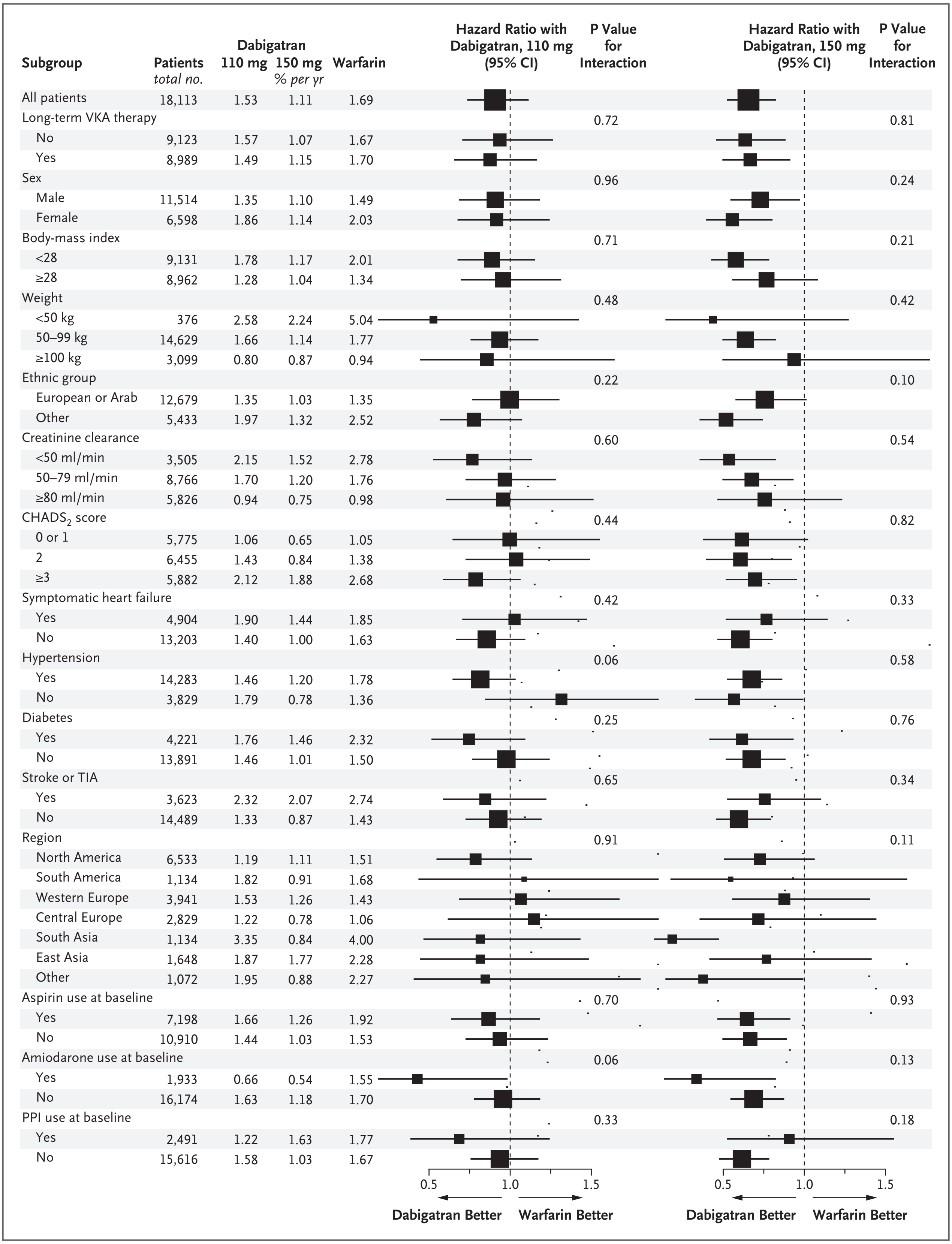
The benefit of dabigatran may be explained in part by the twice-daily dosing regimen. Since dabigatran has an elimination half-life of 12 to 17 hours, twice-daily dosing reduces variability in the anticoagulation effect, especially as compared with the anticoagulation effect of warfarin, which is difficult to control. Warfarin broadly inhibits coagulation (inhibiting factors II, VII, IX, and X and proteins C and S). By selectively inhibiting only thrombin, dabigatran may have antithrombotic efficacy while preserving some other hemostatic mechanisms in the coagulation system and thus potentially mitigating the risk of bleeding.

The use of open-label warfarin could have introduced a bias in the reporting or adjudication of events. This risk was reduced by the implementation of several validated procedures, including blinded evaluation of outcome events. The unexpectedly different rates of myocardial infarction and gastrointestinal bleeding among the three treatment groups support an absence of bias. Control of anticoagulation with warfarin in our study was similar to that in previous international clinical trials, even though half our patients had not previously had extensive treatment with warfarin.[10,17]

The net clinical benefit outcome, which is a measure of the overall benefit and risk, was similar between the two doses of dabigatran, owing to the lower risk of ischemia with the 150-mg dose and the lower risk of hemorrhage with the 110-mg dose. These findings suggest that the dose of dabigatran could potentially be tailored to take into consideration the risk characteristics of a specific patient, although this concept was not specifically tested in our trial.

In conclusion, we compared two doses of dabigatran with warfarin in patients who had atrial fibrillation and who were at risk for stroke. As compared with warfarin, the 110-mg dose of dabigatran was associated with similar rates of stroke and systemic embolism and lower rates of major

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.



Subgroup	Patients *total no.*	Dabigatran 110 mg	Dabigatran 150 mg *% per yr*	Warfarin	P Value for Interaction (Dabigatran, 110 mg)	P Value for Interaction (Dabigatran, 150 mg)
All patients	18,113	1.53	1.11	1.69		
Long-term VKA therapy					0.72	0.81
No	9,123	1.57	1.07	1.67		
Yes	8,989	1.49	1.15	1.70		
Sex					0.96	0.24
Male	11,514	1.35	1.10	1.49		
Female	6,598	1.86	1.14	2.03		
Body-mass index					0.71	0.21
<28	9,131	1.78	1.17	2.01		
≥28	8,962	1.28	1.04	1.34		
Weight					0.48	0.42
<50 kg	376	2.58	2.24	5.04		
50–99 kg	14,629	1.66	1.14	1.77		
≥100 kg	3,099	0.80	0.87	0.94		
Ethnic group					0.22	0.10
European or Arab	12,679	1.35	1.03	1.35		
Other	5,433	1.97	1.32	2.52		
Creatinine clearance					0.60	0.54
<50 ml/min	3,505	2.15	1.52	2.78		
50–79 ml/min	8,766	1.70	1.20	1.76		
≥80 ml/min	5,826	0.94	0.75	0.98		
CHADS$_2$ score					0.44	0.82
0 or 1	5,775	1.06	0.65	1.05		
2	6,455	1.43	0.84	1.38		
≥3	5,882	2.12	1.88	2.68		
Symptomatic heart failure					0.42	0.33
Yes	4,904	1.90	1.44	1.85		
No	13,203	1.40	1.00	1.63		
Hypertension					0.06	0.58
Yes	14,283	1.46	1.20	1.78		
No	3,829	1.79	0.78	1.36		
Diabetes					0.25	0.76
Yes	4,221	1.76	1.46	2.32		
No	13,891	1.46	1.01	1.50		
Stroke or TIA					0.65	0.34
Yes	3,623	2.32	2.07	2.74		
No	14,489	1.33	0.87	1.43		
Region					0.91	0.11
North America	6,533	1.19	1.11	1.51		
South America	1,134	1.82	0.91	1.68		
Western Europe	3,941	1.53	1.26	1.43		
Central Europe	2,829	1.22	0.78	1.06		
South Asia	1,134	3.35	0.84	4.00		
East Asia	1,648	1.87	1.77	2.28		
Other	1,072	1.95	0.88	2.27		
Aspirin use at baseline					0.70	0.93
Yes	7,198	1.66	1.26	1.92		
No	10,910	1.44	1.03	1.53		
Amiodarone use at baseline					0.06	0.13
Yes	1,933	0.66	0.54	1.55		
No	16,174	1.63	1.18	1.70		
PPI use at baseline					0.33	0.18
Yes	2,491	1.22	1.63	1.77		
No	15,616	1.58	1.03	1.67		

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

hemorrhage; the 150-mg dose of dabigatran was associated with lower rates of stroke and systemic embolism but with a similar rate of major hemorrhage.

Supported by a grant from Boehringer Ingelheim.

Dr. Connolly reports receiving consulting fees, lecture fees, and grant support from Boehringer Ingelheim; Dr. Ezekowitz, consulting fees, lecture fees, and grant support from Boehringer Ingelheim and Aryx Therapeutics, consulting fees from Sanofi-Aventis, and lecture fees and grant support from Portola Pharmaceuticals; Dr. Yusuf, consulting fees, lecture fees and grant support from Boehringer Ingelheim and consulting fees from AstraZeneca, Bristol-Myers Squibb, and Sanofi-Aventis; Dr. Eikelboom, consulting fees, lecture fees, and grant support from Boehringer Ingelheim, AstraZeneca, Sanofi-Aventis, and GlaxoSmithKline, consulting fees and lecture fees from Eisai Pharmaceuticals, Eli Lilly, and McNeil, and consulting fees from Bristol-Myers Squibb, Corgenix Medical Corporation, and Daiichi-Sankyo; Dr. Oldgren, consulting fees, lecture fees, and grant support from Boehringer Ingelheim and lecture fees from AstraZeneca; and Drs. Parekh and Xavier, grant support from Boehringer Ingelheim. Drs. Reilly, Varrone, and Wang report being employees of Boehringer Ingelheim. Drs. Alings and Zhu report receiving consulting fees and grant support from Boehringer Ingelheim; Dr. Diaz, consulting fees from GlaxoSmithKline, lecture fees from Sanofi-Aventis, GlaxoSmithKline, and Boehringer Ingelheim, and grant support from Boehringer Ingelheim; Dr. Lewis, consulting fees from Sanofi-Aventis, Bristol-Myers Squibb, and Boehringer Ingelheim and grant support from Boehringer Ingelheim; Dr. Darius, consulting fees, lecture fees, and grant support from Boehringer Ingelheim, consulting fees from Sanofi-Aventis and Bayer Schering Pharma, and lecture fees from the Medicines Company and Eli Lilly; Dr. Diener, consulting fees and lecture fees from Boehringer Ingelheim, Abbott, AstraZeneca, Bayer, Bristol-Myers Squibb, CoAxia, D-Pharm, Fresenius, GlaxoSmithKline, Janssen Cilag, Merck Sharp and Dohme, MindFrame, Neurobiological Technologies, Novartis, Novo-Nordisk, Paion, Parke-Davis, Pfizer, Sanofi-Aventis, Sankyo, Servier, Solvay, Thrombogenics, Wyeth and Yamaguchi and grant support from Boehringer Ingelheim, AstraZeneca, GlaxoSmithKline, Novartis, Janssen-Cilag, and Sanofi-Aventis; Dr. Joyner, grant support from Boehringer Ingelheim, AstraZeneca, Sanofi-Aventis, and Bristol-Myers Squibb; and Dr. Wallentin, consulting fees, lecture fees, and grant support from Boehringer Ingelheim, consulting fees from Regado and Athera, lecture fees from Boehringer Ingelheim, AstraZeneca, and Eli Lilly, and grant support from AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, and Schering Plough. No other potential conflict of interest relevant to this article was reported.

APPENDIX

Members of the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) Study Group are as follows (a full list of investigators is included in the Supplementary Appendix): **Operations Committee:** S.J. Connolly (principal investigator), M.D. Ezekowitz (principal investigator), S. Yusuf (chair, steering committee), J. Eikelboom, J. Oldgren, A. Parekh, P.A. Reilly, E. Themeles, J. Varrone, S. Wang, E. Palmcrantz-Graf, M. Haehl, L. Wallentin (chair, steering committee). **Steering Committee:** A.M.W. Alings, J.V. Amerena, A. Avezum, I. Baumgartner, J. Brugada, A. Budaj, V. Caicedo, L. Ceremuzynski, J.H. Chen, P.J. Commerford, S.J. Connolly, A.L. Dans, H. Darius, G. Di Pasquale, R. Diaz, C. Erol, M.D. Ezekowitz, J. Ferreira, G.C. Flaker, M.D. Flather, M.G. Franzosi, R. Gamboa, S.P. Golitsyn, J.A. Gonzalez Hermosillo, D. Halon, H. Heidbuchel, S.H. Hohnloser, M. Hori, K. Huber, P. Jansky, G. Kamensky, M. Keltai, S. Kim, C.P. Lau, J.Y.F. Le Heuzey, B.S. Lewis, L.S. Liu, J. Nanas, J. Oldgren, P.S. Pais, A.N. Parkhomenko, K.E. Pedersen, L.S. Piegas, D. Raev, O. Razali, T.A. Simmers, P.J. Smith, M. Talajic, R.S. Tan, S. Tanomsup, L. Toivonen, D. Vinereanu, L. Wallentin, D. Xavier, S. Yusuf, J. Zhu. **Core Adjudication–Stroke Advisory Committee:** H.C. Diener (cochair), C.D. Joyner (cochair), A. Diehl, G. Ford, M. Robinson (core adjudication only), J. Silva. **Data Safety Monitoring Board:** P. Sleight (cochair), D.G. Wyse (cochair), J. Collier, D. De Mets, J. Hirsh, E. Lesaffre, L. Ryden, P. Sandercock. **Investigators who recruited at least 12 patients:** *Argentina:* A. Caccavo, L. Cartasegna, C.A. Cuneo, M.V. Elizari, A.J. Gabito, M.A. Hominal, A.D. Hrabar, I.J. Mackinnon, M.A. Rodriguez, A.S. Sanchez, D.R. Vogel; *Australia:* D.M. Colquhoun, R.G. Lehman, B.B. Singh, J.H. Waites; *Austria:* F. Gazo, A. Podczeck, P. Siostrzonek; *Belgium:* J.L. Boland, A. de Meester, L. De Wolf, D. Dilling-Boer, M. Goethals, P. Goethals, O. Gurne, S. Hellemans, B. Ivan, M. Jottrand, I. Kersschot, F. Provenier, W. van Mieghem, Y. Vandekerckhove; *Brazil:* J.C.F. Braga, J.F. Kerr Saraiva, A.M. Lorga-Filho, L.N. Maia, C.S. Melo, D.B. Precoma, S. Rassi, P.R. Rossi; *Bulgaria:* E.S. Baldjiev, I.G. Filibev, N.N. Gotcheva, A.R. Goudev, D.T. Guenova, E.I. Manov, A.L. Radoslavov, D.A. Spirova, V.M. Tsanova, M.L. Tzekova; *Canada:* A.I. Bakbak, J.C. Berlingieri, K. Blackburn, A.W. Booth, S. Bose, P. Boucher, C. Constance, P. Costi, B. Coutu, M.S. Green, J. Habot, S. Kouz, A.V. Lalani, A.S. Lam, P.T.S. Ma, L.B. Mitchell, C.A. Morillo, G. O'Hara, A.S. Pandey, Y. Pesant, C.N. Powell, T.M. Rebane, D. Savard, S. Schulman, M.J. Sehl, D. Shu, L.D. Sterns, R. St-Hilaire, G.S. Syan, P. Talbot, I. Teitelbaum, S.A. Vizel; *China:* X.J. Bai, W. Gao, Z.S. He, Q. Hua, L. Li, W.M. Li, G.P. Lu, S. Lv, N.L. Sun, N.F. Wang, Y.M. Yang, F. Zhang, L. Zhang; *Czech Republic:* J. Belohlavek, M. Cernohous, M. Choura, V. Dedek, B. Filipensky, K. Kovarova, Z. Poklopova, H. Skalicka, J. Spinar; *Denmark:* K.K. Dodt, K. Egstrup, S. Husted, K.K. Klarlund, S. Lind Rasmussen, T.M. Melchior, M.E. Olsen, N. Ralfkiaer, L.H. Rasmussen, K. Skagen; *Finland:* K.E. Airaksinen, H.V. Huikuri, M. Mänttäri, J.H. Melin, K. Peuhkurinen, V.K. Virtanen, A. Ylitalo; *France:* L. Boucher, A. Boye, P. Igigabel, C. Magnani, M. Martelet, J.E. Poulard; *Germany:* K.F. Appel, C.E.H. Dempfle, A. Dormann, J. Harenberg, W.L. Haverkamp, T. Horacek, D. Koudonas, J. Kreuzer, A. Müegge, T.F. Munzel, P.B. Salbach, W. Schoels, R. Veltkamp, E. Von Hodenberg; *Greece:* M.I. Anastasiou-Nana, K. Karamitsos, C. Lappas, A. Manolis; *Hong Kong (China):* W.K. Chan, H.F. Tse, P.T. Tsui, C.M. Yu; *Hungary:* A. Jánosi, A. Kadar, P. Karpati, K. Keltai, J. Rapi, L.I. Regos, I. Szakal; *India:* R.K. Aggarwal, A. Bharani, J.S. Bhuvaneswaran, R.B. Byrapaneni, J.B. Gupta, K.K. Haridas, J. Hiremath, A.S. Jain, J. Joseph, A.M. Naik, R.B. Panwar, J.P.S. Sawhney, N. Sinha, A. Srinivas, P.S. Vadagenalli; *Israel:* J. Benhorin, N.M. Bornstein, B. Brenner, A. Butnaru, A. Caspi, M. Elias, E. Grossman, S. Hamoud, R. Ilia, E.I. Klainman, M. Lishner, C. Lotan, A. Marmor, M. Motro, L.H. Reisen, E. Schwammenthal, M. Shochat, B. Strasberg, Y. Turgeman, A.T. Weiss, D.H. Wexler, R. Wolff, D. Zeltser, R. Zimlichman; *Italy:* G. Argiolas, M. Barbiero, A. Fraticelli, M. Mennuni, L. Moretti, L. Mos, S. Pirelli; *Japan:* J. Furuya, S. Hara, M. Hiroe, S. Kakinoki, N. Miyamoto, M. Yokochi; *Malaysia:* K.H. Chee, A.F.Y. Fong, O. Ismail, S. Jeyaindran; *Mexico:* J. Carrillo, P.A. Fernandez Bonetti; *the Netherlands:* G.L. Bartels, T.A. Bruning, R. Ciampricotti, L. Cozijnsen, H.J. Crijns, M.C.G. Daniels, D.E.P. de Waard, F.R. den Hartog, W.F. Heesen, P.A. Hoogslag, A. Huizenga, J.A. Kragten, G.C. Linssen, D.J. Lok, H.R. Michels, J. Plomp, L. Pos, P.G. Postema, R. Salomonsz, I. Stoel, J.G. Tans, H.J. Thijssen, A.J.M. Timmermans, H.A. van de Klippe, C. van der Zwaan, I.C. van Gelder, L.H. van Kempen, H.A. van Kesteren, P. van Rossum; *Norway:* O. Breder, P.A. Sirnes, A. Tveit; *Peru:* W. Cabrera, J.M. Heredia, M.E. Horna, P.M. Salazar; *Philippines:* M.T.B. Abola, J.C. Anonuevo, D.D. Morales, G.G. Rogelio, A.A. Roxas, Jr.; *Poland:* B. Bacior, J. Grodecki, Z.F. Kalarus, P. Miekus, F. Monies, J. Rekosz, M. Szpajer, A. Wasilewska-Piepiorka, B. Zaborska; *Portugal:* L. Cunha, L.A. Providencia; *Romania:* M.A. Cinteza; *Russia:* I.G. Gordeev, A.Y. Ivleva, T.N. Novikova, E.P. Panchenko, E.V. Shlyakhto, S.B. Shustov, B.A. Sidorenko, V. Sulimov, A.L. Syrkin, D.A. Zateyshchikov; *Singapore:* D. Foo; *Slovak Republic:* V. Bugan, J. Dúbrava, G. Kaliska, M. Masarovicova, D. Pella, J. Sedlák, R. Uhliar; *South Africa:* J.M. Engelbrecht, D. Jankelow,

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

R.J. Routier, F.A. Snyders, H.D. Theron; *South Korea:* T.J. Cha, J.G. Cho, I.S. Choi, K.J. Choi, K.S. Kim, Y.N. Kim, M.H. Lee, M.Y. Lee, D.J. Oh, T.H. Rho; *Spain:* L. Cano, A. Martinez-Rubio, L. Mont; *Sweden:* S. Bandh, C.M. Blomstrom Lundqvist, P. Cherfan, E. Fengsrud, J. Herlitz, T. Juhlin, F. Maru, O.R. Nilsson, F. Ronn, M. Rosenqvist; *Taiwan:* H.H. Hu, C.J. Jack, W.T. Lai, C.H. Liu, H.L. Po, S.J. Ryu, C.D. Tsai, C.D. Tseng, J.H. Wang, S.P. Yang; *Thailand:* S. Kiatchoosakun, R. Krittayaphong, S. Kuanprasert, T. Simtharakaew; *Turkey:* A. Usal; *Ukraine:* K.M. Amosova, O.I. Karpenko, O. Kuryata, L. Martynova, S.S. Pavlyk, L.V. Rudenko, V. Tseluyko, N. Usan, L.G. Voronkov, O.J. Zharinov; *United Kingdom:* J.G. Cleland, A.T. Cohen, P. Davey, J. Davies, H.H. Kadr, G.H. Lip, G.T. McInnes, A.J. Moriarty, M. Pye, S.H.L. Thomas, P.R. Wilkinson; *United States:* Z.A. Abbud, K.V. Adams, I.S. Ahmed, R. Arora, G.R. Aycock, A.J. Bartkowiak, Jr., B. Bean, N.W. Bedwell, A.D. Belber, D.R. Bensimhon, R.E. Benton, O. Ben-Yehuda, B.D. Bertolet, S.D. Bilazarian, D.E. Bolster, J. Bomba, C.R. Bricker, R.I. Brock, K.F. Browne, F. Cardona, T. Carlson, K.W. Carr, W.R. Cashion, Y.S. Chandrashekar, J.H. Chappell, Y. Chen, J.H. Cieszkowski, D.M. Clark, J.F. Cole, S.F. Cossu, D.M. Denny, V.S. Desai, N.J. Deumite, L. Dewey, D.W. Dunning, C. East, W.A. Edmiston, H.S. Ellison, M. Elshahawy, G. Emlein, B. Ewing, P.W. Farrell, P.G. Fattal, J.A. Fialkow, R.H. Fields, M.S. Finkel, G.J. Fishbein, F.J. Fleischhauer, M.A. Frais, D.C. Friedman, M.D. Gelernt, H.C. Genovely, E.L. Gillespie, R.K. Goldberg, D.A. Goldscher, M. Goldstein, A.J. Greenspon, T.C. Hack, S.W. Halpern, C. Hamburg, G.S. Hamroff, G.D. Hanovich, E.J. Haskel, W.H. Haught, S.E. Hearne, J.A. Hemphill, C.H. Henes, P.R. Hermany, W.R. Herzog, T.C. Hilton, M.B. Honan, D.A. Hotchkiss, V.N. Howard, R.K. Ison, A.K. Jacobson, N.F. Jarmukli, D.B. Joyce, P.L. Judson, S. Kaatz, K.J. Kaplan, H.B. Karunaratne, R.N. Khant, M.J. Koren, E.J. Kosinski, P. Kosolcharoen, J.H. Kramer, S.J. Kulback, A. Kumar, J.S. Landzberg, D.T. Lang, T.K. Lau, J.K. Lee, R.M. Levy, D. Linzer, L. Maletz, M.F. McGough, W.P. McGuinn, M.D. McMillen, C.A. McPherson, M.E. Meengs, W.L. Meengs, A.W. Meholick, M.B. Melucci, S.K. Meymandi, R.H. Miller, S.T. Minor, M. Modi, J.F. Moloney, S.K. Mukherjee, V.K. Nadar, M. Nallasivan, J.P. Navas, E.N. Nsah, J.L. Nunamaker, D.J. O'Dea, B. Olliff, P.G. O'Neill, J.R. Onufer, R.C. Orchard, P. Pandey, S. Pezzella, D. Phillips, S.G. Pollock, S.D. Promisloff, J.T. Rittelmeyer, T.A. Rocco, Jr., D. Rosenbaum, D.M. Salerno, K.A. Saxman, R.M. Schneider, K.H. Sheikh, N.K. Shemonsky, W.G. Short, N. Singh, R. Sperling, A.D. Steljes, D.P. Suresh, S.M. Teague, J.M. Teixeira, R.W. Terry, J.A. Trippi, J.E. Usedom, R.M. Vicari, N. Vijay, K.N. Vora, R.B. Vranian, J.L. Walker, R.L. Walsh, S. Weiner, R.J. Weiss, W. Wera-Archakul, J.H. Whitaker, R.H. White, J.N. Whitehill, M.L. Williams, V.E. Wilson, W.W. Wilson, C.W. Wulff, K.A. Yousuf, B.G. Zakhary, P. Zimetbaum, R. Zoble, P.L. Zwerner.

REFERENCES

1. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007;146:857-67.

2. Fuster V, Rydén LE, Cannom DS, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation — executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients with Arial Fibrillation). J Am Coll Cardiol 2006;48:854-906. [Erratum, J Am Coll Cardiol 2007;50:562.]

3. Birman-Deych E, Radford MJ, Nilasena DS, Gage BF. Use and effectiveness of warfarin in Medicare beneficiaries with atrial fibrillation. Stroke 2006;37:1070-4.

4. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115:2689-96.

5. Connolly SJ, Pogue J, Eikelboom J, et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. Circulation 2008;118:2029-37.

6. Stangier J. Clinical pharmacokinetics and pharmacodynamics of the oral direct thrombin inhibitor dabigatran etexilate. Clin Pharmacokinet 2008;47:285-95.

7. Ezekowitz MD, Reilly PA, Nehmiz G, et al. Dabigatran with or without concomitant aspirin compared with warfarin alone in patients with nonvalvular atrial fibrillation (PETRO Study). Am J Cardiol 2007;100:1419-26.

8. Eriksson BI, Dahl OE, Rosencher N, et al. Dabigatran etexilate versus enoxaparin for prevention of venous thromboembolism after total hip replacement: a randomized, double-blind, non-inferiority trial. Lancet 2007;370:949-56. [Erratum, Lancet 2007;370:2004.]

9. Ezekowitz MD, Connolly SJ, Parekh A, et al. Rationale and design of RE-LY: randomized evaluation of long-term anticoagulant therapy, warfarin, compared with dabigatran. Am Heart J 2009;157:805-10.

10. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993;69:236-9.

11. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med 1999;131:492-501.

12. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA 2001;285:2864-70.

13. The ACTIVE Investigators. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. N Engl J Med 2009;360:2066-78.

14. ACTIVE Writing Group of the ACTIVE Investigators. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet 2006;367:1903-12.

15. Amadeus Investigators, Bousser MG, Bouthier J, et al. Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomised, open-label, non-inferiority trial. Lancet 2008;371:315-21. [Erratum, Lancet 2008;372:2022.]

16. Deiner HC, Executive Steering Committee of the SPORTIFF III and V Investigators. Stroke prevention using the oral direct thrombin inhibitor ximelagatran in patients with non-valvular atrial fibrillation: pooled analysis from the SPORTIF III and V studies. Cerebrovasc Dis 2006;21:279-93.

17. Hurlen M, Abdelnoor M, Smith P, Erikssen J, Arnesen H. Warfarin, aspirin, or both after myocardial infarction. N Engl J Med 2002;347:969-74.

Copyright © 2009 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on October 18, 2011. For personal use only. No other uses without permission.
Copyright © 2009 Massachusetts Medical Society. All rights reserved.

Supplementary Appendix

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

Supplement to: Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009;361. DOI: 10.1056/NEJMoa0905561.

(PDF last updated on November 29, 2010.)

Dabigatran Compared to Warfarin in Patients with Atrial Fibrillation

Stuart J. Connolly, M.D.[a], Michael D. Ezekowitz, M.B., Ch.B., D.Phil.[b], Salim Yusuf, F.R.C.P.C., D.Phil.[a], John Eikelboom, M.D.[a], Jonas Oldgren, M.D., Ph.D.[d], Amit Parekh, M.D.[b], Janice Pogue, M.Sc.[a], Paul A. Reilly, Ph.D.[c], Ellison Themeles[a], Jeanne Varrone, M.D.[c], Susan Wang, Ph.D.[c], Marco Alings, M.D., Ph.D.[e], Denis Xavier, M.D.[f], Jun Zhu, M.D.[g], Rafael Diaz, M.D.[h], Basil Lewis, M.D.[i], Harald Darius, M.D.[j], Hans-Christoph Diener, M.D., Ph.D.[k], Campbell D. Joyner, M.D.[l], Lars Wallentin, M.D., Ph.D.[d], the RELY Steering Committee and RELY investigators

[a]Population Health Research Institute, McMaster University and Hamilton Health Sciences, Hamilton, Canada; [b]Lankenau Institute for Medical Research and The Heart Center, Wynnewood, USA; [c]Boehringer-Ingelheim Pharmaceuticals Inc., Ridgefield, USA; [d]Uppsala Clinical Research Centre, Uppsala, Sweden; [e]Working group on Cardiovascular Research the Netherlands, Utrecht, The Netherlands; [f]St. John's National Academy of Health Sciences, Bangalore, India; [g]FuWai Hospital, Beijing, China; [h]Estudios Clínicos Latinoamérica, Rosario, Argentina; [i]Lady Davis Carmel Medical Center, Haifa, Israel; [j]Vivantes Klinikum Neukölln, Berlin, Germany; [k]University Duisburg-Essen, Essen, Germany; [l]Sunnybrook Health Sciences Centre, Toronto, Canada

Correspondence to :
Stuart J. Connolly MD FRCPC
Population Health Research Institute
McMaster university
Hamilton Health Sciences,
McMaster University
237Barton St. E.
Hamilton, Ontario L8L 2X2 Canada
connostu@phri.ca

Online Supplementary Appendix Table 1: RE-LY Inclusion Criteria

1. Atrial fibrillation documented as follows:
    a. There is atrial fibrillation documented by electrocardiogram on the day of screening or randomization
    b. The patient has had a symptomatic episode of paroxysmal or persistent atrial fibrillation documented by 12-lead electrocardiogram within 6 months before randomization
    c. There is documentation of symptomatic or asymptomatic paroxysmal or persistent atrial fibrillation on 2 separate occasions, at least 1 day apart, one of which is within 6 months before randomization. In this case, atrial fibrillation may be documented by 12-lead electrocardiogram, rhythm strip, pacemaker/ICD electrogram, or Holter electrocardiogram. The duration of atrial fibrillation should be at least 30 seconds. Electrograms (not marker channels or mode switch episodes) from pacemakers and defibrillators can be used to document only 1 episode of paroxysmal or persistent atrial fibrillation.
2. In addition to documented atrial fibrillation, patients must have one of the following:
    a. History of previous stroke, transient ischemic attack, or systemic embolism
    b. Ejection fraction less than 40% documented by echocardiogram, radionuclide or contrast angiogram in the last 6 months
    c. Symptomatic heart failure, New York Heart Association class 2 or higher in the last 6 months
    d. Age at least 75 years
    e. Age at least 65 years and one of the following:
        i. Diabetes mellitus on treatment
        ii. Documented coronary artery disease (any of: prior myocardial infarction, positive stress test, positive nuclear perfusion study, prior CABG surgery or PCI, angiogram showing at least 75% stenosis in a major coronary artery)
        iii. Hypertension requiring medical treatment
3. Age at least 18 years at study entry
4. Written, informed consent

Online Supplementary Appendix Table 2: RE-LY Exclusion Criteria

1. History of heart valve disorder (i.e., prosthetic valve or hemodynamically relevant valve disease)
2. Severe, disabling stroke within the previous 6 months, or any stroke within the previous 14 days
3. Conditions associated with an increased risk of bleeding
    a. Major surgery within the previous month
    b. Planned surgery or intervention within the next 3 months
    c. History of intracranial, intraocular, spinal, retroperitoneal or atraumatic intra-articular bleeding
    d. Gastrointestinal hemorrhage within the past year
    e. Symptomatic or endoscopically documented gastroduodenal ulcer disease in the previous 30 days
    f. Hemorrhagic disorder or bleeding diathesis
    g. Need for anticoagulant treatment of disorders other than atrial fibrillation
    h. Fibrinolytic agents within 48 hours of study entry
    i. Uncontrolled hypertension (systolic blood pressure greater than 180 mm Hg and/or diastolic blood pressure greater than 100 mm Hg)
    j. Recent malignancy or radiation therapy (within 6 months) and not expected to survive 3 years
4. Contraindication to warfarin treatment
5. Reversible causes of atrial fibrillation (e.g., cardiac surgery, pulmonary embolism, untreated hyperthyroidism)
6. Plan to perform a pulmonary vein ablation or surgery for cure of the atrial fibrillation
7. Severe renal impairment (estimated creatinine clearance 30 mL/min or less)
8. Active infective endocarditis
9. Active liver disease, including but not limited to
    a. Persistent ALT, AST, Alk Phos greater than twice the upper limit of the normal range
    b. Active hepatitis C (positive HCV RNA)
    c. Active hepatitis B (HBs antigen +, anti HBc IgM +)
    d. Active hepatitis A
10. Women who are pregnant or of childbearing potential who refuse to use a medically acceptable form of contraception throughout the study

11. Anemia (hemoglobin level less than 100 g/L) or thrombocytopenia (platelet count less than 100 X 10^9/L)
12. Patients who have developed transaminase elevations upon exposure to ximelagatran
13. Patients who have received an investigational drug in the past 30 days
14. Patients considered unreliable by the investigator, or having a life expectancy less than the expected duration of the trial because of concomitant disease, or having any condition which, in the opinion of the investigator, would not allox safe participation in the study (e.g., drug addiction, alcohol abuse)

The RELY Study Group

Operations Committee: S.J. Connolly, M.D. Ezekowitz, S. Yusuf, J. Eikelboom, J. Oldgren, A. Parekh, P.A. Reilly, E. Themeles, J. Varrone, S. Wang, E. Palmcrantz-Graf, M. Haehl, L. Wallentin; **Steering Committee:** S.J. Connolly, M.D. Ezekowitz, S. Yusuf, J. Oldgren, L. Wallentin, A.M.W. Alings, J.V. Amerena, A. Avezum, I. Baumgartner, J. Brugada, A. Budaj, V. Caicedo, L. Ceremuzynski, J.H. Chen, P.J. Commerford, A.L. Dans, H. Darius, G. Di Pasquale, R. Diaz, C. Erol, J. Ferreira, G.C. Flaker, M.D. Flather, M.G. Franzosi, R. Gamboa, S.P. Golitsyn, J.A. Gonzalez Hermosillo, D. Halon, H. Heidbuchel, S.H. Hohnloser, M. Hori, K. Huber, P. Jansky, G. Kamensky, M. Keltai, S. Kim, C.P. Lau, J.Y.F. Le Heuzey, B.S. Lewis, Liu L.S., J. Nanas, O. Razali, P.S. Pais, A.N. Parkhomenko, K.E. Pedersen, Piegas, LS., D. Raev, T.A. Simmers, P.J. Smith, M. Talajic, R.S. Tan, S. Tanomsup, L. Toivonen, D. Vinereanu, D. Xavier, J. Zhu; **Core Adjudication Committee:** H.C. Diener (co-chair), C.D. Joyner (co-chair), A. Diehl, G. Ford, M. Robinson, J. Silva; **Stroke Advisory Committee:** H.C. Diener (co-chair), C.D. Joyner (co-chair), A. Diehl, G. Ford, J. Silva; **Data Safety Monitoring Board:** P. Sleight (co-chair), D.G. Wyse (co-chair), J. Collier, D. De Mets, J. Hirsh, E. Lesaffre, L. Ryden, P. Sandercock; **Event Adjudicators:** M.I. Anastasiou-Nana, G. Andersen, B.H. Annex, M. Atra, A. Avezum, N.M. Bornstein, G. Boysen, P.J.A.M. Brouwers, M. Buerke, L.M. Burrell, Y.K. Chan, W.H. Chen, R.T.F. Cheung, P.J. Commerford, S. Divakaramenon, G.A. Donnan, G.Z. Duray, H. Dvorakova, J. Eikelboom, J. Fiedler, E. Gardinale, P.C. Gates, E.G. Goshev, S. Goto, B. Gross, H.P. Guimaraes, O. Gulkevych, R.L. Haberl, D. Halon, G. Hankey, J. Hartikainen, J. Healey, S.H. Hohnloser, A.M. Iliesiu, O. Irkin, P. Jansky, T. Jaxa-Chamiec, S. Jolly, G. Kamensky, K.A.M. Kaste, M. Keltai, B. Kies, K.D. Kostov, K.S. Kristensen, A.J. Labovitz, R.P.T. Lassila, K.L.F. Lee, B.S. Lewis, Piegas, LS., Y.M. Lutay, P. Magloire, K.H. Mak, A. Meijer, L. Mihov, C.A. Morillo, L.E. Morillo, G.M. Nair, B. Norrving, A. Ntalianis, M. Ntsekhe, L. Olah, P.M.D. Pasco, A. Peeters, V. Perovic, I. Petrov, G. Pizzolato, F. Rafti, N.R. Rey, S. Ribas, M. Rokoss, I.J. Sarembock, T. Sheth, A. Shuaib, T.A. Simmers, E. Sitkei, E. Sorokin, M. Šrámek, E. Strozynska, D. Tanne, V.N.S. Thijs, A. Tomek, F. Turazza, G. Vanhooren, S.A. Vizel, J. Vos, N. Wahlgren, R. Weachter, B. Zaborska, J. Zaborski, R. Zimlichman; **Boehringer-Ingelheim Core Project Team:** J. Cong, K. Fendt, S. Muldoon, P.A. Reilly, J. Varrone, S. Wang; **Population Health Research Institute Core Project Team:** S. Bajkor, A. Grinvalds, M. Malvaso, J. Pogue, M. Robinson, K. Simek, E. Themeles, S. Yang; **Investigative sites who recruited at least 1 patient (each site represented by a maximum of 2 site representatives):** *Argentina:* M.F. Alzogaray, J.O. Bono, A. Caccavo, L. Cartasegna, W.P. Casali, J.L. Cuello, C.A. Cuneo, M.V. Elizari, A.A. Fernandez, A.E. Ferrari, A.J. Gabito, R.F. Goicoechea, V.M. Gorosito, A. Hirschson, M.A. Hominal, A.D. Hrabar, A. Liberman, I.J. Mackinnon, R.D. Manzano, C.A. Muratore, S.A. Nemi, M.A. Rodriguez, A.S. Sanchez, J. Secchi, D.R. Vogel; *Australia:* J.V. Amerena, D.M. Colquhoun, D.S. Crimmins, A.M. Dart, S.M. Davis, P.J. Hand, P.A. Kubler, R.G. Lehman, G. McBain, H.C. Morrison, G. New, B.B. Singh, C.Z. Spence, J.H. Waites; *Austria:* J. Auer, L. Doweik, F. Freihoff, G. Gaul, F. Gazo, H. Geiger, G. Giacomini, G.W. Huber, K. Huber, I. Jukic, G. Lamm, H. Niessner, A. Podczeck, J. Schuh, P. Siostrzonek, C. Steger, B. Vogel, R. Watzak, H.S. Weber, W. Weihs; *Belgium:* I. Blankoff, J.L. Boland, C. Brike, M. Carlier, F. Cools, A. de Meester, H.J. De Raedt, L. De Wolf, G.M. Dhooghe, D. Dilling-Boer, S.R. Elshot, S. Fasseaux, M. Goethals, P. Goethals, O. Gurne, H. Heidbuchel, S. Hellemans, B. Ivan, M. Jottrand, I. Kersschot, E. Lecoq, O. Marcovitch, D. Melon, H. Miljoen, L. Missault, L.A. Pierard, F. Provenier, M.F. Rousseau, D. Stockman, E. Tran-Ngoc, W. van Mieghem, Y. Vandekerckhove, P. Vandervoort, J. Verrostte, J. Vijgen; *Brazil:* D. Armaganijan, C. Braga, J.C.F. Braga, R. Cipullo, C.L.P. Cunha, A. de Paola, M.I. Delmonaco, F.V. Guimaraes, L. Herek, J.F. Kerr Saraiva, Maia L.N., A.M. Lorga, A.M. Lorga-Filho, R.L. Marino, C.S. Melo, O.M. Mouco, V.C. Pereira, D.B. Precoma, W. Rabelo, S. Rassi, P.R. Rossi, J.M. Rossi Neto, F.M. Silva, M.H.

Vidotti, S.L. Zimmermann; *Bulgaria:* E.D. Anev, T.A. Balabanov, E.S. Baldjiev, E.S. Bogusheva, M.A. Chaneva, I.G. Filibev, N.N. Gotcheva, A.R. Goudev, I.T. Gruev, D.T. Guenova, Z.A. Kamenova, E.I. Manov, I.A. Panov, Z.I. Parvanova, M.B. Pehlivanova, P.T. Penchev, N.Y. Penkov, A.L. Radoslavov, D. Raev, K.N. Ramshev, N.M. Runev, M.N. Sindzhielieva, D.A. Spirova, V.M. Tsanova, M.L. Tzekova, G.K. Yaramov; *Canada:* R. Aggarwal, A.I. Bakbak, K. Bayly, J.C. Berlingieri, K. Blackburn, C. Bobbie, A.W. Booth, D. Borts, S. Bose, P. Boucher, K. Brown, J.M. Burstein, J.C. Butt, B.D. Carlson, R. Chetty, J.D. Chiasson, C. Constance, P. Costi, B. Coutu, I. Deneufbourg, D. Dion, P. Dorian, J.D. Douketis, J. Eikelboom, S. Farukh, N.G. Filipchuk, B.A. Fox, H.I. Fox, C.B. Gailey, M. Gauthier, A. Glanz, M.S. Green, J. Habot, H. Hink, S. Jolly, C. Kearon, S. Kouz, C. Lai, K. Lai, A.V. Lalani, A.S. Lam, L.A. Lapointe, R.A. Leather, P.T.S. Ma, E. MacKay, I. Mangat, S. Mansour, E. Melton, L.B. Mitchell, C.A. Morillo, A.L. Morris, G.M. Nair, W.A. Nisker, M.J. O'Donnell, G. O'Hara, L.M. Omichinski, A.S. Pandey, R. Parkash, Y. Pesant, C. Pilon, K.J. Pistawka, C.N. Powell, J.B. Price, S. Prieur, T.M. Rebane, A.J. Ricci, J. Roberge, M. Roy, J.L. Sapp, D. Savard, S. Schulman, M.J. Sehl, F. Sestier, R. Shandera, D. Shu, L.D. Sterns, R. St-Hilaire, G.S. Syan, M. Talajic, P. Talbot, I. Teitelbaum, R.H. Tytus, S.A. Vizel, L. Winkler, R. Zadra, B.S. Zidel; *China:* X.J. Bai, W. Gao, X. Gao, D.M. Guan, Z.S. He, Q. Hua, H. Li, L. Li, W.M. Li, G.P. Lu, S. Lv, K. Meng, H.Y. Niu, D.G. Qi, S.Y. Qi, F. Qian, N.L. Sun, H.Y. Wang, N.F. Wang, Y.M. Yang, H. Zeng, F. Zhang, F.R. Zhang, L. Zhang; *Colombia:* R. Bohorquez, J.F. Rosas, L. Saent, M. Vacca, V.M. Velasco; *Czech Republic:* J. Belohlavek, M. Cernohous, M. Choura, V. Dedek, B. Filipensky, L. Hemzsky, P. Jansky, I. Karel, I. Kopeckova, K. Kovarova, R. Labrova, T. Madr, Z. Poklopova, D. Rucka, J. Simon, H. Skalicka, M. Smidova, J. Spinar; *Denmark:* K.K. Dodt, K. Egstrup, J. Friberg, D. Haar, S. Husted, G.V. Jensen, A.M. Joensen, K.K. Klarlund, S. Lind Rasmussen, T.M. Melchior, M.E. Olsen, K.E. Pedersen, M.K. Poulsen, N. Ralfkiaer, L.H. Rasmussen, K. Skagen; *Finland:* K.E. Airaksinen, H.V. Huikuri, E.J. Hussi, P. Kettunen, M. Mänttäri, J.H. Melin, J. Mikkelsson, K. Peuhkurinen, L. Toivonen, V.K. Virtanen, A. Ylitalo; *France:* B. Agraou, L. Boucher, J.M. Bouvier, A. Boye, B. Boye, E.M. Decoulx, P. Defaye, M. Delay, H. Desrues, K. Gacem, P. Igigabel, P. Jacon, J.Y.F. Le Heuzey, S. Leparree, C. Magnani, M. Martelet, J. Movallem, T. Olive, J.E. Poulard, B. Tiam; *Germany:* K.F. Appel, S. Appel, L. Bansemir, M. Borggrefe, J. Brachmann, N. Bulut-Streich, K. Busch, H. Darius, C.E.H. Dempfle, M. Desaga, V. Desaga, A. Dormann, G.Z. Duray, I. Fechner, S. Genth-Zotz, W.G. Haberbosch, J. Harenberg, W.L. Haverkamp, R. Henzgen, H. Heuer, S.H. Hohnloser, T. Horacek, H.B. Huttner, U. Janssens, H.J. Jantke, V. Klauss, D. Koudonas, J. Kreuzer, H. Kuckuck, A. Maselli, A. Müegge, T.F. Munzel, K. Nitsche, A. Nledegjen, A. Parwani, M. Pluemer-Schmidt, B.W. Pollock, B.I. Salbach, P.B. Salbach, T. Schaufele, W. Schoels, S. Schwab, U. Siegmund, R. Veltkamp, E. Von Hodenberg, R. Weber, M. Zechmeister; *Greece:* M.I. Anastasiou-Nana, A.A. Anastasopoulous, V.O. Foulidis, E. Kaldara, K. Karamitsos, J. Karantzis, H. Kirpizidis, C. Kokkinakis, A. Krommydas, C. Lappas, G.I. Lappas, A. Manolis, A.S. Manolis, Z. Orfanidis, M. Papamichalis, L. Peltekis, S. Savvas, E.A. Skoumpourdis, D.A. Stakos, I. Styliadis, F. Triposkiadis, D. Tsounis, D.N. Tziakas, T. Zafiridis, J.H. Zarifis; *Hong Kong (China):* G.C.P. Chan, W.K. Chan, W.S. Chan, C.L. Lau, H.F. Tse, P.T. Tsui, C.M. Yu, C.S. Yue; *Hungary:* K. Fugedi, B. Garai, A. Jánosi, A. Kadar, P. Karpati, K. Keltai, I. Kosa, I. Kovacs, Z. Laszlo, L. Mezei, J. Rapi, L.I. Regos, E. Sitkei, I. Szakal, I. Szigyarto, K. Toth, A. Zsa'ry; *India:* D.K. Agarwal, R.K. Aggarwal, Arulvenkatesh R., A. Bharani, J.S. Bhuvaneswaran, R.B. Byrapaneni, P. Chandwani, S. Chopra, N. Desai, V. Deshpande, N.P. Golla, J.B. Gupta, K.K. Haridas, J. Hiremath, A.S. Jain, M. Jain, D.A. Jhala, J. Joseph, M. Kaila, A. Kannaiyan, S. Kumar, P. Kuruvila, V.K. Mahorkar, A. Metha, A.M. Naik, S. Narayanan, P.S. Pais, R.B. Panwar, C. Reddy, J.P.S. Sawhney, S.M. Shah, S. Sharma, G.S. Shetty, N. Sinha, N.N. Sontakke, A. Srinivas, M.R. Trivedi, P.S. Vadagenalli, M.

Vijayakumar; *Israel:* Y. Ben-Aharon, J. Benhorin, N. Bogomolny, N.M. Bornstein, S. Botwin-Shimko, I. Bova, B. Brenner, M. Burstein, A. Butnaru, A. Caspi, H.D. Danenberg, M. Dayan, M. Eldar, D. Elian, M. Elias, A. Elis, G. Esanu, I. Genin, L.H. Goldstein, E. Grossman, S. Hamoud, T. Hayek, N. Ilani, R. Ilia, E.I. Klainman, A. Leibowitz, D. Leibowitz, I. Levin, M. Lishner, C. Lotan, A. Mahagney, A. Marmor, M. Motro, D. Peres, T. Plaev, L.H. Reisen, O. Rogowski, E. Schwammenthal, Y. Schwammenthal, M. Shechter, M. Shochat, A. Shotan, B. Strasberg, E. Sucher, D. Tanne, G. Telman, Y. Turgeman, I. Tzoran, A.T. Weiss, T. Weitsman, B. Weller, D.H. Wexler, R. Wolff, D. Yarnitsky, D. Zeltser, R. Zimlichman; *Italy:* G. Argiolas, F. Arteni, M. Barbiero, R. Bazzucco, D. Bernardi, L. Bianconi, D. Bicego, R. Brandini, B. Bresciani, F. Busoni, E. Carbonieri, M. Carini, A. Catalano, C. Cavallini, G. D'Angelo, R. De Caterina, M. Di Niro, G. Di Pasquale, F. Filigheddu, A. Fraticelli, R. Marconi, M. Mennuni, L. Moretti, L. Mos, L.G. Pancaldi, S. Pirelli, G. Renda, M. Santini, I. Tavarozzi, P. Terrosu, F. Uneddu, M. Viccione, R. Zanini, G. Zingarini; *Japan:* T. Aoyagi, H. Eguma, K. Fujii, M. Fukuchi, M. Fukunami, Y. Furukawa, J. Furuya, K. Haneda, S. Hara, M. Hiroe, Y. Iesaka, T. Iijima, Y. Ishibashi, K. Iwade, T. Kajiya, S. Kakinoki, S. Kamakura, Y. Katayama, Y. Kihara, K. Kimura, S. Kobayashi, K. Kono, Y. Koretsune, N. Marui, T. Matsuyama, H. Meno, N. Miyamoto, S. Morikawa, K. Myojin, T. Nakamura, Y. Nishi, T. Ogawa, H. Onaka, T. Sakakibara, S. Sakurai, Y. Sasaki, H. Sato, M. Sugii, K. Sumii, S. Suzuki, M. Takagi, T. Takenaka, K. Takeuchi, S. Tanaka, J. Tanouchi, K. Ueda, Y. Ueyama, T. Ujihira, M. Usui, M. Yagi, T. Yamada, H. Yamamoto, M. Yokochi, E. Zen; *Malaysia:* A.K. Abd Ghaphar, C.K. Ang, K.H. Chee, A.F.Y. Fong, O. Ismail, S. Jeyaindran, S. Kaur, T.C. Lee, O. Razali, R.S. Sandhu, R.P. Shah, S. Suganthi, S. Zainal Abidin; *Mexico:* R. Alvarado-Ruiz, J. Carrillo, E. Delgado, P.A. Fernandez Bonetti, J.A. Gonzalez Hermosillo, J.L. Leiva, A. Meaney, R. Olvera, R. Olvera Jr., R. Peralta-Heredia, I. Rodriguez, C.M. Ruiz Rabasa, G. Solache, E. Villeda Espinosa; *Netherlands:* S. Ahmed, A.M.W. Alings, E. Badings, G.L. Bartels, M. Beganovic, T.A. Bruning, R. Ciampricotti, L. Cozijnsen, H.J. Crijns, M.C.G. Daniels, D.E.P. de Waard, F.R. den Hartog, A. Dirkali, B.E. Groenemeijer, W.F. Heesen, J.A. Heijmeriks, P.A. Hoogslag, A. Huizenga, H.H. Idzerda, J.A. Kragten, K. Krasznai, T. Lenderink, A.H. Liem, G.C. Linssen, D.J. Lok, J.G. Meeder, H.R. Michels, J. Plomp, L. Pos, J.L. Posma, P.G. Postema, R. Salomonsz, T.A. Simmers, I. Stoel, J.G. Tans, H.J. Thijssen, A.J.M. Timmermans, R.G. Tteleman, P.F.M.M. van Bergen, H.A. van de Klippe, C. van der Zwaan, J.W.M. van Eck, A.J.J. van Es, I.C. van Gelder, L.H. van Kempen, H.A. van Kesteren, P. van Rossum, S. Veldmeyer, A.A.M. Wilde; *Norway:* H. Arnesen, D. Atar, O. Breder, H. Istad, Z. Radunovic, D.E. Rykke, P.A. Sirnes, P.J. Smith, A. Tveit, S.R. Ulimoen; *Peru:* W. Cabrera, R. Duenas, R. Gamboa, J.M. Heredia, M.E. Horna, Y. Hurtado, P.M. Salazar; *Philippines:* M.T.B. Abola, J.C. Anonuevo, R.S. Arellano, C. Dioquino, D.R. Maranon, D.D. Morales, E.B. Reyes, G.G. Rogelio, A.A. Roxas Jr., D.J.V. Sulit; *Poland:* B. Bacior, E. Dulak, J. Gniot, J. Goncikowski, J. Grodecki, Z.F. Kalarus, K. Kawecka-Jaszcz, P. Miekus, F. Monies, M. Piepiorka, E. Pilichowska, E. Plizio, J. Rekosz, A. Rybicka-Musialik, W.A. Streb, M. Styczkiewicz, M. Szpajer, M. Trusz-Gluza, A. Wasilewska-Piepiorka, B. Zaborska; *Portugal:* P. Adragao, V. Branco, P. Canhão, L. Cunha, F. Falcão, G. Lopes, C. Machado, J. Martinez-Marcos, P.F. Monteiro, Parreira L., Pinto A.N., L.A. Providencia, Salgado A.V., Santos J.F., A.T. Timoteo; *Romania:* R. Capalneanu, M. Cinteza, A.D. Margulescu, D.S. Turdeanu, D. Vinereanu, V.D. Vintila; *Russia:* V.L. Baranov, E.R. Berngardt, T.R. Dzhordzhikiya, S.P. Golitsyn, I.G. Gordeev, Y.V. Grigoryev, M.U. Isaeva, A.Y. Ivleva, V.A. Kokorin, A.L. Komarov, O.K. Maximenko, E.B. Maykov, N. Novikova, T.N. Novikova, E.P. Panchenko, M.G. Poltavskaya, Y.N. Popova, S.A. Pronina, A.Sh. Revishvili, E.V. Shlyakhto, S.B. Shustov, B.A. Sidorenko, A.I. Sinopalnikov, V. Sulimov, A.L. Syrkin, A.Y. Titkov, Y.S. Titkov, D.A. Zateyshchikov, O.P. Zavaritskaya; *Singapore:* P.L. Chia, D. Foo, K.L. Sim, R.S. Tan; *Slovak Republic:* V. Bugan, I. Buganova, J. Dúbrava, G.

Kaliska, G. Kamensky, M. Masarovicova, P. Mikes, Z. Mikes, J. Murin, D. Pella, R. Rybar, J. Sedlák, M. Skamla, P. Spurný, J. Strbova, R. Uhliar; *South Africa:* P.J. Commerford, L.J. Disler, J.M. Engelbrecht, D. Jankelow, J. King, E.Q. Klug, M. Munnick, A.M. Okreglicki, R.J. Routier, F.A. Snyders, H.D. Theron, H. Wittmer; *South Korea:* T.J. Cha, J.G. Cho, I.S. Choi, J.I. Choi, K.J. Choi, K.R. Han, J.H. Heo, S.W. Jang, T.S. Kang, H.S. Kim, K.S. Kim, S.J. Kim, S.S. Kim, Y.H. Kim, Y.N. Kim, M.H. Lee, M.Y. Lee, G.B. Nam, D.J. Oh, H.W. Park, J.S. Park, T.H. Rho, D.G. Shin, E.K. Shin; *Spain:* J.J. Alonso, L. Cano, N.P. Castellano, A.J. Criado-Millan, A. Curcio, P. Egea, J.M. Escudier, A. Grande, J.M. Grande, G. Gusi-Tragant, I.F. Lozano, A.M. Martin, A. Martinez-Rubio, L. Mont, J. Perez-Villacastin, L. Sosa; *Sweden:* M. Ali, T. Andersson, S. Bandh, C.M. Blomstrom Lundqvist, P. Cherfan, E. Fengsrud, C. Fluur, J. Herlitz, Z. Hijazi, N. Hoglund, B. Hojeberg, J. Jabro, T. Juhlin, B. Kjellman, I. Lonnberg, F. Maru, L. Morlid, O.R. Nilsson, J. Oldgren, F. Ronn, M. Rosenqvist, H. Walfridsson; *Switzerland:* I. Baumgartner, S.T. Engelter, A. Gallino, P.A. Lyrer, T. Moccetti, I. Petrova; *Taiwan:* Y.J. Chang, C.H. Chen, J.H. Chen, M.Y.C. Chen, J.J. Cheng, T.R. Chiang, W.T. Chung, C.H. Hsia, C.Y. Hsu, H.H. Hu, J.S. Jeng, W.T. Lai, L.M. Lien, K.H. Lin, C.H. Liu, H.S. Lo, G.S. Peng, H.L. Po, S.J. Ryu, C.D. Tsai, L.M. Tsai, C.D. Tseng, J.H. Wang, S.F. Wang, S.P. Yang; *Thailand:* Kiatchoosakun S., R. Krittayaphong, S. Kuanprasert, T. Ngarmukos, T. Simtharakaew, B. Sukanandachai, A. Sukonthasam, A. Suwanagool, S. Tanomsup, Tatsanavivat P.; *Turkey:* Y. Atmaca, N. Baris, B. Boyaci, M. Demir, C. Erol, S. Guneri, A. Usal, R. Yalcin; *Ukraine:* K.M. Amosova, O.P. Beregova, YE.M. Besaga, M.R. Ikorkin, K. Karapetyan, O.I. Karpenko, L. Kononenko, O. Kuryata, L. Martynova, T. Motylevska, O. Okhryamkina, A.N. Parkhomenko, S.S. Pavlyk, M.V. Perepelytsya, L.V. Rudenko, O.A. Skarzhevsky, L.A. Tkachenko, V. Tseluyko, N. Usan, L.G. Voronkov, K.V. Yshchenko, O.J. Zharinov; *United Kingdom:* V.G. Bryson, R. Butler, R.I. Cargill, N.S. Chahal, J.G. Cleland, A.T. Cohen, E.M. Cruddas, P. Davey, J. Davies, M.D. Flather, S.L. Ford, K. Griffith, R. Haynes, S. Hill, M. Javed, H.H. Kadr, G.H. Lip, J. Machin, D.J. McEneaney, G.T. McInnes, A.J. McNeill, A.J. Moriarty, S. Muir, J. O'Callaghan, J.A. Purvis, M. Pye, R. Senior, D.A. Sutton, S.H.L. Thomas, P.R. Wilkinson, R. Wilmott, M.J. Wrigley; *United States:* R. Abadier, Z.A. Abbud, K.V. Adams, S.W. Adler, S. Agarwal, A.M. Ahmed, I.S. Ahmed, M.A. Aiuto, T.D. Albrittun, P. Aliyar, J.J. Allan, D.P. Allen, S.L. Allen, A. Altschuller, M. Amin, I.S. Anand, A.B. Antolick, R. Arora, A.J. Arouni, C.L. Arslanian, R.W. Asinger, G.R. Aycock, R.J. Bariciano, S.B. Baron, M.A. Barr, A.J. Bartkowiak Jr., L. Baruch, C. Basignani, M.L. Bass, B. Bean, N.W. Bedwell, A.D. Belber, K. Belew, Y.C. Bell, R.L. Bellinger, W.T. Bennett, D.R. Bensimhon, R. Benton, R.E. Benton, O. Ben-Yehuda, B.D. Bertolet, A.S. Betkowski, S.D. Bilazarian, J.K. Bissette, M.B. Bobade, D.E. Bolster, J. Bomba, D.M. Book, J.A. Boscia, A. Bouchard, L.M. Bowman, A.J. Bradley III, H.D. Brandt, C.R. Bricker, T.L. Brobyn, R.I. Brock, T.M. Broderick, K. Broedlin, A.M. Brown, K.F. Browne, S.W. Burke, M.E. Burton, G.A. Buser, M.K. Capasso, W.E. Caplan, J. Cappelli, C. Cardona, F. Cardona, T. Carlson, K.W. Carr, T. Casey, W.R. Cashion, D.T. Cass, Y.S. Chandrashekar, M. Changlani, P.G. Chapla, J.H. Chappell, C. Chen, Y. Chen, N.R. Cho, J.H. Cieszkowski, D.M. Clark, R. Clayton, C.W. Clogston, D.J. Cockrell, A.I. Cohen, T.J. Cohen, J.F. Cole, G. Conway, V.R. Cook, A.L. Cornish, S.F. Cossu, D.L. Costello, D.J. Courtade, H.C. Covelli, B.S. Crenshaw, L.A. Crews, G.H. Crossley, S.C. Culp, B.M. Curtis, K. Darrow, R.E. De Raad, M. DeGregorio, A.A. DelNegro, D.M. Denny, V.S. Desai, N.J. Deumite, L. Dewey, R.N. Dharawat, B. Dobbs, S.M. Donahue, B. Downey, J. Downing, M.A. Drehobl, W.A. Drewes, M.N. Drucker, R. Duff, M. Duggal, S.H. Dunlap, D.W. Dunning, V. DuThinh, G.T. Dykstra, C. East, M.C. Eblaghie, J. Edelstein, W.A. Edmiston, H.J. Eisen, S.J. Eisenberg, J.R. Ellis, H.S. Ellison, S. Ellsworth, M. Elshahawy, G. Emlein, M. Entcheva, L.K. Essandoh, A.Q. Estrada, B. Ewing, M.D. Ezekowitz, R.T. Faillace, A. Fanelli, P.W. Farrell, S.W. Farris, P.G. Fattal, D.Y. Feigenblum, G.J. Feldman, J.A. Fialkow, K.M. Fiddler,

The RELY Study Group

R.H. Fields, M.S. Finkel, C. Finn, T.A. Fischell, M. Fishbach, G.J. Fishbein, M.M. Fisher, G.C. Flaker, F.J. Fleischhauer, T.G. Folk, S.D. Folkerth, R.R. Fortman, M.A. Frais, D.C. Friedman, G. Fuchs, F. Fuller, G. Garibian, F.H. Gee, M.D. Gelernt, H.C. Genovely, J.R. Gerber, J.J. Germano, J.J. Giardina, J.M. Gilbert, E.L. Gillespie, E.M. Gilman, B. Gitler, D.H. Givens, R. Glover, H.S. Gogia, D.C. Gohn, R.K. Goldberg, J.J. Goldberger, D.A. Goldscher, M. Goldstein, T. Goraya, D.F. Gordon, D. Gottlieb, H.L. Grafner, M. Graham, M.W. Graves, M. Graziano, S.N. Greco, M.L. Greenberg, A.J. Greenspon, G.S. Greer, D.D. Griffin, E.W. Grogan, V.L. Groo, T. Guarnieri, A. Gupta, J. Gupta, T.C. Hack, B. Hall, O. Hallak, S.W. Halpern, C. Hamburg, G.S. Hamroff, J. Han, F. Handel, S.R. Hankins, G.D. Hanovich, J.A. Hanrahan, I.U. Haque, J.L. Hargrove, P.E. Harnick, J.L. Harris, P.A. Hartley, E.J. Haskel, D. Hatch, W.H. Haught, S. Hearne, S.E. Hearne, J.A. Hemphill, D.A. Henderson, C.H. Henes, T. Hengerer-Yates, P.R. Hermany, W.R. Herzog, K. Hickey, T.C. Hilton, E.S. Hockstad, P. Hodnett, R. Hoffmeister, J. Holland, L. Hollenweger, M.B. Honan, D.A. Hoopes, A.R. Hordes, D.A. Hotchkiss, M.A. Howard, V.N. Howard, A.R. Hulyalkar, P. Hurst, L.C. Hutchison, J. Ingram, T. Isakov, R.K. Ison, C.N. Israel, B.K. Jackson, K.N. Jackson, A.K. Jacobson, S. Jain, N.F. Jarmukli, I. Joffe, L.E. Johnson, S.A. Johnson, S.L. Johnson, A.A. Jones, D.B. Joyce, P.L. Judson, S.S. Juk Jr., S. Kaatz, R.M. Kaddaha, K.J. Kaplan, H.B. Karunaratne, J.D. Kennett, D.M. Kenton, J.A. Kettunen, M.A. Khan, R.N. Khant, S.S. Kim, M.D. Kirkwood, B.P. Knight, P.O. Knight, T.J. Knutson, J.F. Kobayashi, A. Kogan, A.D. Kogan, M.J. Koren, E.J. Kosinski, P. Kosolcharoen, J.B. Kostis, J.H. Kramer, S.D. Kramer, J. Kron, C.R. Kuchenrither, S.J. Kulback, A. Kumar, D. Kushner, A. Kutscher, A.J. Labovitz, C.K. Lai, J.B. Lam, C. Landau, J.S. Landzberg, D.T. Lang, J.M. Lang, C.J. Lanzarotti, D.L. Lascewski, T.K. Lau, J.K. Lee, S. Lee, W.N. Leimbach, A.M. LePine, M.F. Lesser, S.H. Leuchak, R.M. Levy, W.R. Lewis, T.L. Lincoln, W.M. Lingerfelt, M. Liston, Z.G. Liu, R.L. Lloret, L. Lohrbauer, D.C. Longoria, B.M. Lott, D.R. Louder, K.L. Loukinen, J. Lovell, S. Lue, J.A. Mackall, L. Maletz, L. Marlow, R.C. Martin, M. Matsumura, M.J. McCartney, D. McDuffie, M.F. McGough, F.A. McGrew III, Wm.P. McGuinn, M.D. McMillen, J. McNeff, C.A. McPherson, M.E. Meengs, W.L. Meengs, A.W. Meholick, J.S. Meisner, M.B. Melucci, A. Mercando, J.D. Merlino, S.K. Meymandi, M.B. Miele, R.H. Miller, S.H. Miller, S.T. Minor, M.R. Mitchell, M. Modi, F.V. Mody, C.L. Moeller, J.F. Moloney, J.E. Moran, N.C. Morcos, A. Morgan, S.K. Mukherjee, K. Mullinax, A.L. Murphy, A.J. Mustin, G.I. Myers, G.V. Naccarelli, V.K. Nadar, M. Nallasivan, J.P. Navas, I.K. Niazi, E.N. Nsah, J.L. Nunamaker, T.B. Ochalek, D.J. O'Dea, P.D. Ogilvie, B. Olliff, A.K. Omalley, P.G. O'Neill, J.R. Onufer, R.C. Orchard, L.A. Orihuela, E.C. Ortiz, M.T. O'Sullivan, B.J. Padanilam, P. Pandey, A. Parekh, D.V. Patel, R.J. Patel, V.B. Patel, R.R. Patlola, G.D. Pennock, R. Perlman, P.H. Peters Jr., A.V. Petrillo, S. Pezzella, D. Phillips, J.R. Pierre-Louis, G. Pilcher Jr., C. Pillai, S.G. Pollock, M.S. Pond, J.K. Porterfield, L. Presant, J. Pressler, A.H. Pribble, S.D. Promisloff, K.K. Pudi, D.L. Putnam, J. Quartner, J.C. Quinn, C.M. Quinnell, G.L. Raad, L.A. Rasmussen, C. Ray, J.A. Reiffel, S. Reynertson, J.W. Richardson, C.P. Riley, J.S. Rippy, J.T. Rittelmeyer, D.M. Roberts, R. Robertson, V.J.B. Robinson, T.A. Rocco Jr., D. Rosenbaum, E.M. Roth, J.N. Rottman, R.R. Rough, J.J. Rubenstein, A.M. Sakkal, T. Saleem, D.M. Salerno, M.L. Samendinger, S. Sandeno, T.M. Santilli, P. Santucci, P. Sattar, K.A. Saxman, S. Schaefer, J. Schmidt, R.M. Schneider, D.D. Schocken, M.K. Schrader, B.A. Schramm, R.W. Schultz, A.E. Schussheim, E.F. Schwarz, M.C. Seamon, J.D. Sestero, M.P. Shah, R. Shah, A. Shalaby, J.G. Shanes, G.L. Sheftel, K.H. Sheikh, A.B. Shein, N.K. Shemonsky, A. Shepler, E. Sheridan, T.M. Shipwash, R.I. Shopnick, W.G. Short, M.F. Shoukfeh, R.S. Sibia, T.M. Siler, J.A. Silva, G.R. Simons, A.G. Simpson, H.R. Simpson, V.J. Simpson, B.N. Singh, N. Singh, V.N. Singh, C.J. Sitz, L. Skatrud, J. Sklar, D.J. Slotwiner, P.F. Smith, P.N. Smith, R.H. Smith, J.E. Smith III, B.C. Sodowick, A.J. Solomon, E.A. Soltero, A.F. Sonel, R. Sperling, C. Spiller, B.Z. Spink, L.W. Sprinkle, A.C.

Spyropoulos, T.D. Stamos, A.D. Steljes, M.E. Stillabower, T. Stover, J.E. Strain, T.L. Strickland, D.P. Suresh, T.S. Takata, J.S. Taylor, M. Taylor, S.M. Teague, J.M. Teixeia, E.A. Telfer, P.S. Terry, R.W. Terry, H.M. Thai, M. Thalin, V.N. Thomas, C.A. Thompson, M.A. Thompson, J.W. Thornton, R.E. Tidman, B.S. Toler, M.I. Traina, J.A. Trippi, D.L. Ujiiye, J.E. Usedom, E. Van De Graaff, L.R. Van De Wall, J.W. Vaughn, D. Ver Steeg, R.M. Vicari, N. Vijay, C.B. Vitale, A.G. Vlastaris, J. Voda, K.N. Vora, W.F. Voyles, R.B. Vranian, P.S. Vrooman, P. Waack, A.L. Waldo, J.L. Walker, M.A. Wallace, E.A. Walsh, R.L. Walsh, A. Walton, M. Washam, P.S. Wehner, J.Y. Wei, S. Weiner, R.J. Weiss, D.M. Wells, W. Wera-Archakul, J.H. Wertheimer, S.A. West, J.H. Whitaker, M.L. White, R.H. White, J.N. Whitehill, P.J. Wiegman, J. Wiesel, J. Williams, L.E. Williams, M.L. Williams, V.K. Williamson, V.E. Wilson, W.W. Wilson, S.L. Woodfield, C.W. Wulff, S.W. Yates, K.A. Yousuf, B.G. Zakhary, R. Zambrano, P. Zimetbaum, R. Zoble, A.R. Zopo, P.L. Zwerner.

American Pain Society
RESEARCH
EDUCATION
TREATMENT
ADVOCACY

PUBLISHED BY
ELSEVIER

The Journal of Pain, Vol 17, No 1 (January), 2016: pp 76-89
Available online at www.jpain.org and www.sciencedirect.com

Cognitive-Behavioral–Based Physical Therapy for Patients With Chronic Pain Undergoing Lumbar Spine Surgery: A Randomized Controlled Trial

Kristin R. Archer,[*,†] Clinton J. Devin,[*] Susan W. Vanston,[*] Tatsuki Koyama,[‡]
Sharon E. Phillips,[‡] Steven Z. George,[§] Matthew J. McGirt,[¶] Dan M. Spengler,[*]
Oran S. Aaronson,[‖] Joseph S. Cheng,[‖] and Stephen T. Wegener[**]

[*]Department of Orthopaedic Surgery, [†]Department of Physical Medicine and Rehabilitation, [‡]Department of Biostatistics, and [‖]Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee.
[§]Department of Physical Therapy, University of Florida, Gainesville, Florida.
[¶]Carolina Neurosurgery and Spine Associates, Charlotte, North Carolina.
[**]Department of Physical Medicine and Rehabilitation, Johns Hopkins Medicine, Baltimore, Maryland.

Abstract: The purpose of this study was to determine the efficacy of a cognitive-behavioral–based physical therapy (CBPT) program for improving outcomes in patients after lumbar spine surgery. A randomized controlled trial was conducted on 86 adults undergoing a laminectomy with or without arthrodesis for a lumbar degenerative condition. Patients were screened preoperatively for high fear of movement using the Tampa Scale for Kinesiophobia. Randomization to either CBPT or an education program occurred at 6 weeks after surgery. Assessments were completed pretreatment, posttreatment and at 3-month follow-up. The primary outcomes were pain and disability measured by the Brief Pain Inventory and Oswestry Disability Index. Secondary outcomes included general health (SF-12) and performance-based tests (5-Chair Stand, Timed Up and Go, 10-Meter Walk). Multivariable linear regression analyses found that CBPT participants had significantly greater decreases in pain and disability and increases in general health and physical performance compared with the education group at the 3-month follow-up. Results suggest a targeted CBPT program may result in significant and clinically meaningful improvement in postoperative outcomes. CBPT has the potential to be an evidence-based program that clinicians can recommend for patients at risk for poor recovery after spine surgery.

Perspective:* *This study investigated a targeted cognitive-behavioral–based physical therapy program for patients after lumbar spine surgery. Findings lend support to the hypothesis that incorporating cognitive-behavioral strategies into postoperative physical therapy may address psychosocial risk factors and improve pain, disability, general health, and physical performance outcomes.

© 2016 by the American Pain Society

Key words: *Lumbar degenerative disease, cognitive-behavioral therapy, randomized controlled trial, postoperative rehabilitation, lumbar spinal fusion.*

Degenerative lumbar conditions, such as spinal stenosis, lead to chronic pain, physical impairment, and reduced quality of life.[2] The prevalence in the general population ranges from 20 to 25% and increases to above 45% in individuals greater than 60 years of age.[31,34,35] Lumbar spinal stenosis is one of the most common diagnoses associated with spine surgery.[2,18,72] The surgical technique for lumbar

---

Received May 11, 2015; Revised September 2, 2015; Accepted September 29, 2015.

The research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number R21AR062880 and the Magistro Family Foundation grant through the Foundation for Physical Therapy. This study used REDCap as the secure database which was supported by CTSA award No. UL1TR000445 from the National Center for Advancing Translational Sciences. The authors declare no conflict of interest in the preparation of this manuscript.

Address reprint requests to Kristin R. Archer, PhD, DPT, Department of Orthopaedic Surgery, Vanderbilt University, School of Medicine, Medical Center East – South Tower, Suite 4200, Nashville, TN, 37232. E-mail: kristin.archer@vanderbilt.edu

1526-5900/$36.00

© 2016 by the American Pain Society

http://dx.doi.org/10.1016/j.jpain.2015.09.013

degenerative conditions is well established and studies have reported on the benefits of surgery compared with nonoperative management.[24,39] The Spine Patient Outcomes Research Trial (SPORT), using as-treated analysis, found that surgery for lumbar stenosis had a significant advantage over nonoperative treatment at 2 and 4 years after surgery.[84] However, as-treated SPORT findings demonstrated that the advantage of surgery was no longer significant after 5 years.[45]

The estimated percentage of people over 60 years of age is expected to increase steadily toward 2050.[77] An increased number of people will experience age-associated degenerative conditions and chronic pain; spine surgery rates will continue to increase.[4] Despite surgical advances, adults who have undergone lumbar spine surgery continue to have poorer physical and mental health outcomes compared with the general population.[50,83] Studies have found persistent pain, functional disability, and poor quality of life in up to 40% of individuals after spine surgery for lumbar degenerative conditions.[9,32,48,84] The reoperation rate has been reported to range from 18 to 23% at 8 to 10 years after surgery.[45]

Archer et al.[5,6,8] and others have found that fear of movement, avoidance coping, positive affect, and depression are independently associated with persistent pain and disability and decreased physical function after lumbar spine surgery.[17,29,47,69] Despite the literature recommending a biopsychosocial approach to postoperative care,[52,88] physical therapy programs after spine surgery continue to focus on trunk and lower extremity strengthening, flexibility, range of motion, and education on posture and proper body mechanics. Randomized trials to date have found no significant difference between traditional physical therapy and either no treatment, an educational booklet, or advice to keep active.[1,46,51,52] These results suggest that an alternative approach to postoperative rehabilitation may be needed to address the psychosocial factors often associated with poor spinal surgery outcomes.

The purpose of this study was to incorporate cognitive-behavioral strategies into physical therapy to improve outcomes in patients with chronic pain undergoing lumbar spine surgery. Individuals with high fear of movement were targeted in order to focus on adults at risk for poor postoperative recovery.[8,17,29,47] The Changing Behavior through Physical Therapy (CBPT) program was designed to decrease fear of movement and increase self-efficacy[7] and be delivered by physical therapists. Because clinic-based rehabilitation can be impractical for many older adults, a telephone-delivery model was used to allow individuals with financial, geographic, and mobility constraints to participate in the study. We hypothesized that CBPT participants would have greater improvement in patient-reported pain, disability, and general health and performance-based tests compared with education participants at 6 months after lumbar spine surgery for degenerative conditions.

Methods

Trial Design

This study was a randomized controlled trial. Participants were recruited from a single academic medical center and randomized to either CBPT or an education program during a routine postoperative clinic visit at 6 weeks after surgery. At this visit, all participants also received standard care, which may include having lifting and/or driving restrictions removed and referral to traditional physical therapy. The education program was chosen as a comparison to control for the time and attention of the therapist and for normal healing that occurs from 6 weeks to 3 months after surgery.

The investigators, participating surgeons, research personnel conducting the assessments, and patients were blinded to group assignment. Participants were informed that they would be randomly assigned to 1 of 2 different educational treatments and were asked not to discuss study procedures with their treating surgeon, medical staff, and research personnel. The study physical therapist was blinded to the aims and hypotheses of the study.

The overall study design included a clinic screening visit, preoperative assessment, pretreatment assessment (6 weeks after surgery), treatment phase, posttreatment assessment (3 months after surgery), and 3-month follow-up assessment (6 months after surgery) (see ClinicalTrials.gov and NCT01131611). The Institutional Review Board at the participating site approved the study and all patients provided informed consent before study enrollment and data collection.

Sample Size and Power

The number of study participants was based on a sample size calculation for a comparison of treatment groups on change in the outcomes of pain intensity and interference measured by the Brief Pain Inventory (BPI), disability measured by the Oswestry Disability Index (ODI), and general health measured by the 12-Item Short-Form Health Survey. Power was estimated by generating simulated data from available pilot data, then using the simulated data to estimate the original model parameters. A sample size of 80 was chosen to be able to detect minimum clinically important differences (MCID) in pain intensity of 1.2 to 2.0 points, pain interference of 1.6 to 2.2 points, disability of 10 to 12.8 points, and general health of 4.9 to 6.2 points during the postoperative period, with an 80% power while controlling type I error rate at 5%. These MCIDs were based on studies conducted in patients following lumbar spine surgery.

Participants

Participants for this study were recruited from 499 individuals, between March 2012 and April 2013, undergoing a laminectomy with or without arthrodesis for a lumbar degenerative condition (spinal stenosis, spondylosis with or without myelopathy, and degenerative spondylolisthesis). The following inclusion criteria were used for recruitment purposes: 1) 21 years of age or older; 2) English speaking; 3) back and/or lower extremity pain for greater than 6 months; 4) no history of a neurological movement disorder; and 5) no presence of psychotic disease in the medical record. Participants also needed to report high fear of movement, based on a score of 39 or greater on the Tampa Scale for Kinesiophobia (TSK). A cut-off of 39 on the TSK has been found to identify individuals who have a high probability of dysfunctional pain beliefs and poor outcomes after spine surgery.[5,6,57,79,80,85]

Study exclusion criteria included 1) spinal deformity as the primary indication for surgery; 2) surgery for pseudarthrosis, trauma, infection, or tumor; and 3) having microsurgical techniques as the primary procedure.

Study Procedures and Randomization

Eligible participants were approached for consent before surgery and completed a screening questionnaire to determine high fear of movement. Individuals who remained eligible completed an intake assessment, a battery of validated questionnaires that assessed pain, disability, general health, pain self-efficacy, depression, and a series of performance-based tests. Participants returned to the clinic at 6 weeks after surgery for a standard postoperative visit. At this clinic visit, participants completed a pretreatment assessment for the study and the first treatment session (CBPT or education).

Randomization was administered through the Research Electronic Data Capture (REDCap) system[28] and occurred immediately after the baseline assessment in order to initiate treatment. A computer-generated scheme randomized patients to either CBPT or education in a 1:1 ratio in blocks of assignments. Because preliminary data demonstrated that surgery type and fear of movement influenced patient-reported outcomes, these assignments were frequency matched on type of surgery (fusion or no fusion) and screening score on the TSK (39–45, 46–49, 50–68), resulting in 6 strata.

Participants returned for an in-person posttreatment assessment and a 3-month follow-up assessment at 3 and 6 months after surgery, respectively. All assessments included a self-report questionnaire that measured psychosocial characteristics (fear of movement and pain self-efficacy) and pain, disability, and general health outcomes as well as use of physical therapy and other health care services. Performance-based tests were completed to assess lower extremity strength, functional mobility, and gait speed. Participants who were unable to return to the clinic for follow-up visits were asked to complete questionnaires at home and return them in self-addressed stamped envelopes. See Fig 1 for a CONSORT flow diagram.

Participants were reimbursed $25 for their time in completing the baseline assessment and $100 for each of the in-person follow-up assessments.

Demographic and Depressive Symptoms

A preoperative intake assessment collected demographic and health information pertaining to age, sex, race, education, employment, smoking status, height and weight, comorbid conditions, narcotic use, history of spinal surgery, and expectations of a successful surgery. Participants also provided information on depressive symptoms by completing the 9-item Patient Health Questionnaire (PHQ-9).[40] A total score on the PHQ-9 can range from 0 to 27 with higher numbers indicating severe depressive symptoms. Participants rate each item on a 4-point Likert scale with scoring that ranges from “not at all” to “nearly every day.” In a psychometric study of the PHQ-9 compared with independent diagnoses made by mental health professionals, the instrument was both sensitive (.75) and specific (.90) for the diagnosis of major depression.[40,41]

Psychosocial Measures

Fear of movement was assessed with the 17-item TSK.[38] A total score can range from 17 to 68. Participants rate items on a 4-point Likert scale with scoring alternatives ranging from “strongly disagree” to “strongly agree.” The MCID for the TSK has been reported to be 4 points in patients with back pain.[89] The TSK has good internal consistency and test-retest reliability in surgical patients and patients with various musculoskeletal conditions.[22,67]

The Pain Self-Efficacy Questionnaire (PSEQ) was used to measure the strength and generality of a person’s belief in his/her ability to accomplish a range of activities despite pain.[58] Participants rate how confident they are on a 7-point scale from “not at all confident” to “completely confident.” Scores range from 0 to 60, with a score greater than 40 indicating high self-efficacy.[55] The PSEQ has been found to have excellent internal consistency, good test-retest reliability, and construct validity through correlations with depression, anxiety, coping strategies, pain ratings, and work-related tasks in patients with chronic pain.[58]

Primary Outcome Measures

Pain Intensity and Pain Interference

The BPI was used to measure both pain intensity and pain interference with daily activity.[12] The pain intensity scale includes 4 pain items assessing current, worst, least, and average pain (0, no pain at all; 10, as bad as you can imagine). The pain interference scale is a 7-item scale measuring the degree to which pain interferes with areas of daily life: general activity, mood, walking, work, relations with others, sleep, and enjoyment of life (0, does not interfere; 10, completely interferes). The BPI has proven reliable (Cronbach’s alpha > .80) and valid (highly correlated with the SF-36 brief pain scale, the Roland Disability Questionnaire, the McGill Pain Questionnaire, and the Visual Analog Scale for pain) in both surgical patients and patients with chronic low back pain.[36,53,92] The MCID for back and leg pain has been found to be 1.2 and 1.6, respectively, in patients after lumbar spine surgery.[14,25]

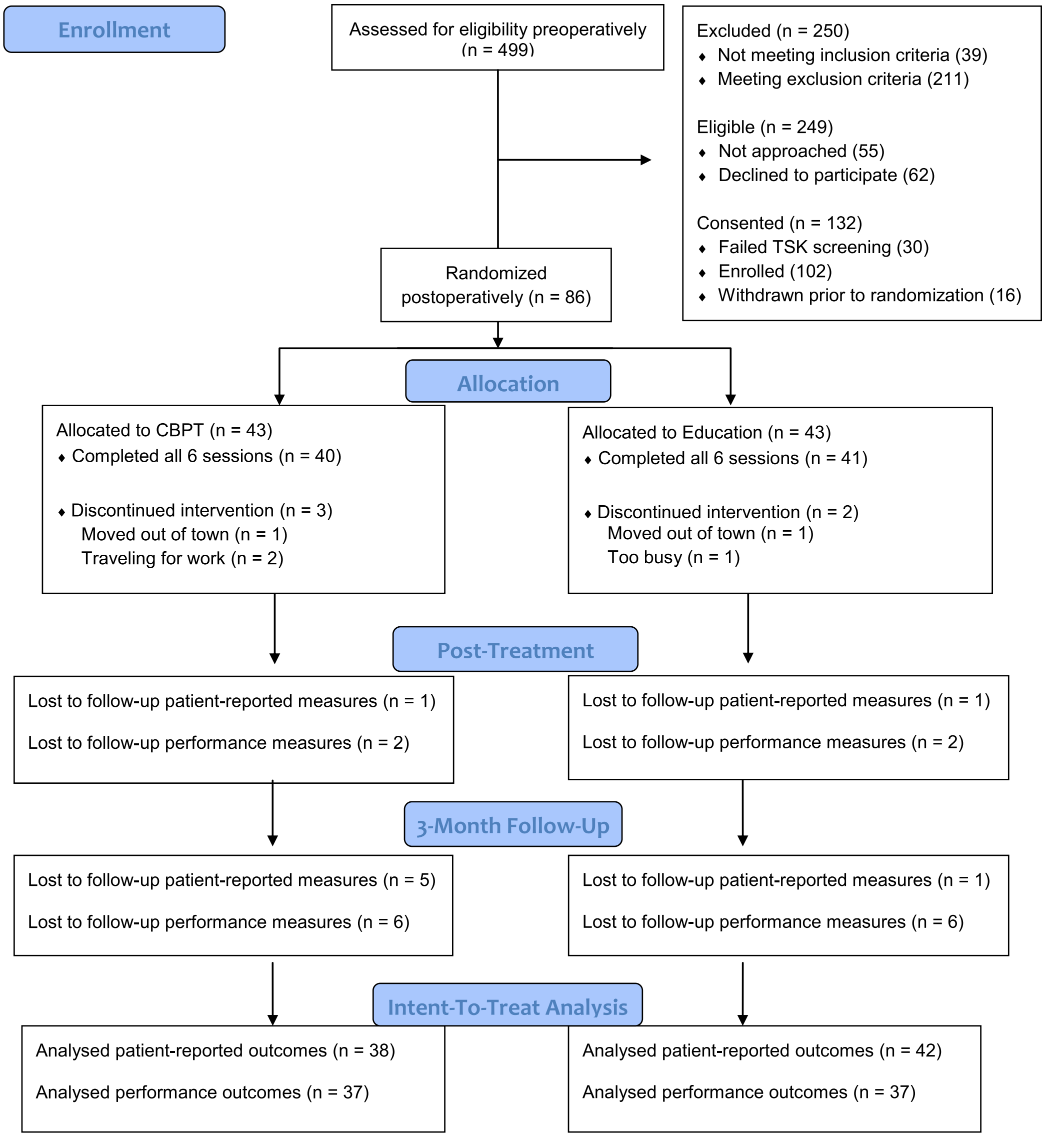

Enrollment
Assessed for eligibility preoperatively (n = 499)
Excluded (n = 250)
Not meeting inclusion criteria (39)
Meeting exclusion criteria (211)
Eligible (n = 249)
Not approached (55)
Declined to participate (62)
Consented (n = 132)
Failed TSK screening (30)
Enrolled (102)
Withdrawn prior to randomization (16)
Randomized postoperatively (n = 86)
Allocation
Allocated to CBPT (n = 43)
Completed all 6 sessions (n = 40)
Discontinued intervention (n = 3)
Moved out of town (n = 1)
Traveling for work (n = 2)
Allocated to Education (n = 43)
Completed all 6 sessions (n = 41)
Discontinued intervention (n = 2)
Moved out of town (n = 1)
Too busy (n = 1)
Post-Treatment
Lost to follow-up patient-reported measures (n = 1)
Lost to follow-up performance measures (n = 2)
Lost to follow-up patient-reported measures (n = 1)
Lost to follow-up performance measures (n = 2)
3-Month Follow-Up
Lost to follow-up patient-reported measures (n = 5)
Lost to follow-up performance measures (n = 6)
Lost to follow-up patient-reported measures (n = 1)
Lost to follow-up performance measures (n = 6)
Intent-To-Treat Analysis
Analysed patient-reported outcomes (n = 38)
Analysed performance outcomes (n = 37)
Analysed patient-reported outcomes (n = 42)
Analysed performance outcomes (n = 37)

Figure 1. Flowchart of recruitment and follow-up.

Disability

Low back disability was measured using the ODI.[21] The 10-item ODI assesses 10 aspects of daily living: pain intensity, lifting, sitting, standing, walking, sleeping, hygiene, traveling, social life, and sex life. Ratings for each item are from 0 (high functioning) to 5 (low functioning). Total item scores are divided by the total possible score and multiplied by 100 to create a percentage of disability. Disability categories include 0 to 20% (minimal disability), 21 to 40% (moderate disability), 41 to 60% (severe disability), 61 to 80% (crippled), and 81 to 100% (bed bound or exaggerated symptoms). The ODI has demonstrated excellent test-retest reliability (Pearson's r >.80), adequate internal consistency (Cronbach's alpha >.70), and validity with moderately high correlations with other disability measures.[16,65] The MCID has been found to range from 10 to 12.8 points in patients after lumbar spine surgery.[14,25,61,62]

Secondary Outcome Measures

General Physical and Mental Health

General physical and mental health was measured with the physical and mental component scales of the 12-Item Short-Form Health Survey (SF-12).[82] The physical component scale (PCS) assesses the 4 subdomains of physical functioning, role-physical, bodily pain, and general health, and the mental component scale (MCS) assesses the 4 subdomains of vitality, social functioning, role-emotional, and mental health. Total subscale scores range from 0 to 100, and higher scores represent better health status. The PCS and MCS of the SF-12 have demonstrated responsiveness, good test-retest reliability, good internal consistency, and validity in both generalized and various patient populations.[33,82] The MCID for the PCS and MCS has been estimated at 4.9 points in patients after lumbar arthrodesis.[25]

Performance-Based Function

The 5-Chair Stand test[26] was used to assess lower extremity strength. Participants were asked to fold their arms across their chest and stand up from and sit down on a standard chair. If able to perform 1 time successfully, patients were asked to stand up and sit down 5 times as fast as possible starting in the sitting position and stopping after the fifth rise. Performance on the 5-Chair Stand test was measured in seconds. The 5-Chair Stand test has demonstrated good test-retest reliability and validity, with significant correlations with other measures of physical performance and self-reported disability.[26] The MCID for the 5-Chair Stand test has been estimated as a reduction of 2.3 seconds in patients with balance and vestibular disorders.[54]

The Timed Up and Go (TUG) test[64] was used to assess functional mobility. Participants were asked to stand from a chair, walk 3 m, turn around, walk back, and sit down, and the time to complete was recorded in seconds. The TUG has been shown to have excellent test-retest reliability and be a valid and responsive performance measure in older individuals.[10,64] A major clinically important improvement for the TUG has been reported as a reduction in time ranging from 1.2 to 1.4 seconds in older adults with osteoarthritis.[91]

The 10-Meter Walk test[27] was used to assess gait speed. Patients were given a 2-m warm-up distance preceding the 10-m distance and 2 m beyond the 10 m to continue walking. The time that it took to traverse the 10 m at a comfortable pace was recorded. Two trials were conducted, with a brief rest as needed between trials. Measurements for both trials were averaged. Excellent interrater and intrarater reliability and good test-retest reliability for self-paced timed walking speed tests using a stopwatch have been reported.[49] Validity for walking speed tests has been determined by significant correlations with measures of function and mortality in older adults.[27,49] The MCID for gait speed has been estimated to be .16 m/s in patients with subacute stroke and substantial meaningful change has been found to be .10 m/s in older adults.[63,73]

Treatments

Therapist Training

One physical therapist with no previous experience delivering cognitive-behavioral strategies participated in a training program for both the CBPT and education programs. Formal training included 8 hours of didactic and 16 hours of experiential session-by-session training with a clinical psychologist (S.T.W.) and 8 hours of training with a physical therapist who designed the programs (K.R.A.). Knowledge and skills competence was determined through a written test after the first 2-day session and a skills test after the second 2-day session (ie, scores needed to be >85). After training, the CBPT and education programs were implemented with research personnel and a pretest occurred with 2 patients in each group. All sessions during the pretest were audiotaped and reviewed with the physical therapist to evaluate adherence to the CBPT and education treatment protocols and cognitive and behavioral competencies specific to the CBPT treatment.[78]

CBPT Program

The CBPT program is a cognitive-behavioral–based approach to rehabilitation (see www.spine-surgery-recovery.com for more information). Brief cognitive-behavioral therapy (CBT) programs for pain developed by Woods and Asmundson,[90] Williams and McCracken,[87] and Turner et al,[75] and a self-management program developed for older adults by Lorig[44] provided the basis for the CBPT program. Specific cognitive-behavioral strategies were selected from these evidence-based CBT programs and adapted for use by physical therapists. The main goal of the CBPT program was to reduce pain and disability through reductions in fear of movement and increases in self-efficacy. Patients received weekly sessions with a study physical therapist for 6 weeks. The first session was conducted in person and participants were given a manual to follow along with the study therapist. The remaining sessions were delivered over the telephone. All sessions were 30 minutes in length, except the first session, which was approximately 1 hour.

The CBPT program focused on empirically supported behavioral self-management, problem solving, cognitive restructuring, and relaxation training.[74,87,90] The main components of the program include education on the relationship between the body, mind, and one's activity level, a graded activity plan (ie, a comprehensive list of activities ordered from least to most difficult based on fear or pain) and weekly activity and walking goals. Goals were rated by patients on a scale from 0 to 10 (completely confident), and scores of 8 or greater indicated a realistic goal. A cognitive or behavioral strategy was introduced in each session, with the therapist helping patients identify enjoyable activities (ie, distraction), replace negative thinking with positive thoughts, find a balance between rest and activity, and manage setbacks by recognizing high-risk situations and negative thoughts. Details of the CBPT intervention were published previously.[7]

Education Program

The education program focused on postoperative recovery and consisted of topics commonly covered by physical therapists during outpatient treatment sessions. Sessions addressed the benefits of physical therapy, proper biomechanics after surgery, importance of daily exercise, and ways to promote healing. Information on stress reduction, sleep hygiene, energy management, communication with health providers, and preventing future injury were also provided. Patients received weekly sessions with a study physical therapist for 6 weeks. The first session was conducted in person and participants were given a manual to follow along with the study therapist. The remaining sessions were delivered over the telephone. All sessions were 30 minutes in length, except the first session, which was approximately 1 hour.

Treatment Fidelity

The study physical therapist's adherence to the CBPT and education manuals was assessed by digitally recording all sessions and randomly selecting 30% of all sessions (balanced evenly across the sessions) to review. A clinical psychologist (S.T.W.) and a physical therapist (K.R.A.) with expertise in the programs rated the CBPT and education sessions for treatment integrity and potential contamination using a standardized checklist. The study therapist also completed a checklist of all the components delivered during each CBPT or education session and made note of any protocol deviations. A therapist adherence score was determined for each session using a scale from 0 (completely nonadherent) to 100 (completely adherent).

Treatment Acceptability

Acceptability of the CBPT and education programs was assessed after treatment. Participants were asked to rate how helpful the program was to their recovery and how likely they were to recommend the program to a friend. These items were scored using an 11-point numeric rating scale with 0 being "not at all helpful or likely" and 10 being "extremely helpful or likely." Participants were also asked to rate the overall benefit of the program taking into account the effort put into it, the importance of changes in pain and activity due to the program, and the importance of the program compared with other services on a 5-point Likert scale. Finally, participants were asked through open-ended questions to comment on the strengths and weaknesses of the program.

Data Analysis

Descriptive statistics were used to summarize the mean scores and standard deviations (SD) or frequency of demographic, clinical, and psychosocial characteristics as well as outcomes measures. Group means and corresponding confidence intervals or frequency for preoperative variables and baseline measures were compared using Student t-tests or χ^2 tests to confirm balance between groups. The characteristics of the patients who were lost to follow-up were compared with those who completed the follow-up assessments. Missing items were less than 5% for the completed psychosocial and outcome scales and imputed based on a single mean imputation.

All analyses were intent to treat. The mean change from pretreatment to posttreatment and 3-month follow-up was calculated for the primary and secondary outcome measures and psychosocial characteristics. Between-group differences of mean change from baseline to each follow-up time point were compared using repeated-measures analysis of variance. Standardized mean effect size differences of the programs were assessed with Cohen's d and $d = .20$ indicated a small effect, $d = .50$ a medium effect, $d = .80$ a large effect, and $d = 1.3$ a very large effect.[13,68] Separate multivariable linear regression models were then conducted for the outcomes at the 3-month follow-up, controlling for a priori variables of the pretreatment score of the outcome of interest, age, education, presence of comorbid conditions, and number of physical therapy visits since the baseline visit. Potential interactions between treatment and age and type of surgery were tested. Stata software (StataCorp, 2011, College Station, TX) was used to analyze the data. The level of significance was set at $P < .05$.

Results

Of the 194 eligible participants who were approached about the study, 132 (68%) consented and 102 passed screening and were enrolled (Fig 1). Eight-six participants were randomized. Sixteen participants (16%) were not treated surgically for a lumbar degenerative condition and were withdrawn from the study before randomization. The dropout rate for the CBPT and education programs was 7% and 5%, respectively. For the 5 individuals who did not complete all 6 sessions, reasons provided were moving out of town, traveling for work, and other time commitments. The follow-up rate for the patient-reported and performance-based outcomes after treatment was 98% and 95% and at the 3-month follow-up was 93% and 86%, respectively. There were no significant differences between patients with and without complete follow-up data on demographic, clinical, and psychosocial variables.

Participant demographic and clinical variables are presented in Table 1. No significant differences were noted across groups. Seventy-eight participants (91%) received clinic-based physical therapy during the treatment phase between 6 weeks and 3 months after surgery. The average number of physical therapy visits for the CBPT group was 8.6 (SD = 4.9) and 8.0 (SD = 4.2) for the education group ($P = .55$). Forty patients (47%) continued with physical therapy between the posttreatment and 3 month follow-up time point; the CBPT participants had an average of 6.5 visits (SD = 7.5) and the education group 6.6 visits (SD = 10.3; $P = .94$).

Treatment Fidelity and Acceptability

Adherence to the CBPT and education programs was high, with no statistical differences between groups (97.7 vs 98; $P = .41$). Both groups reported that the program was extremely helpful and it was extremely likely they would recommend the program to a friend (Table 2). The majority of CBPT participants (59.5%) reported that the benefits of the program far outweighed the effort compared with 45.2% of education participants ($P = .33$). Significant differences were noted for the questions on the importance of changes in pain and activity, with 54.8% and 76.2% of CBPT participants and 21.4% and 33.3% of education participants noting that their pain decreased and activity increased a meaningful amount, respectively ($P < .01$). No significant difference was found between groups on whether the program was more important than other services since leaving the hospital (CBPT 45.2% vs education 38.1%; $P = .11$).

Table 1. Demographic and Clinical Characteristics of Study Participants (N = 86)

Characteristic	Total	CBPT N = 43	Education N = 43
Demographic			
Age in y, mean ± SD	57.6 ± 12.2	56.9 ± 11.1	58.4 ± 13.3
Female sex, n (%)	48 (55.8)	25 (58.1)	23 (53.5)
Self-report white race, n (%)	69 (80.2)	36 (83.7)	33 (76.7)
More than high school education, n (%)	62 (72.1)	30 (69.8)	32 (74.4)
Married, n (%)	61 (70.9)	32 (74.4)	29 (67.4)
Obese BMI category, n (%)	44 (51.1)	23 (53.5)	21 (48.8)
Employed before surgery, n (%)			
Not working	29 (33.7)	14 (32.6)	15 (34.9)
Working	39 (45.3)	21 (48.8)	18 (41.9)
Retired	18 (21.0)	8 (18.6)	10 (23.2)
Current smoker, n (%)	17 (19.8)	10 (23.3)	7 (16.3)
Comorbid conditions, n (%)			
0	6 (7.0)	4 (9.3)	2 (4.7)
1–2	66 (76.7)	32 (74.4)	34 (79.0)
>2	14 (16.3)	7 (16.3)	7 (16.3)
Clinical			
Fusion surgery, n (%)	60 (69.8)	29 (67.4)	31 (72.1)
Previous spine surgery, n (%)	34 (39.5)	17 (39.5)	17 (39.5)
Duration of preoperative pain, mean ± SD	24.1 ± 27.4	25.1 ± 30.2	23.1 ± 24.5
Taking narcotics before surgery, n (%)	47 (54.6)	23 (53.5)	24 (55.8)
Expectations of successful surgery, mean ± SD	8.9 (1.7)	8.7 ± 2.1	9.2 ± 1.1
Preoperative depression, PHQ-9 mean ± SD	10.3 (5.8)	11 (5.6)	9.6 (6.0)
Preoperative fear of movement, TSK mean ± SD	43.3 (5.3)	43.5 (5)	43.2 (5.6)
Preoperative pain self-efficacy, PSEQ mean ± SD	26.6 (11.3)	25.5 (10.6)	27.7 (12.1)
Preoperative back pain, BPI mean ± SD	6.7 (2.1)	6.8 (1.9)	6.5 (2.3)
Preoperative leg pain, BPI mean ± SD	7.1 (2.4)	7.0 (2.6)	7.1 (2.2)
Preoperative disability, ODI mean ± SD	49.1 (13.3)	49.2 (13.7)	49.0 (13.1)
Preoperative physical health, SF-12 mean ± SD	25.8 (5.8)	25.4 (5.7)	26.2 (6.1)
Preoperative mental health, SF-12 mean ± SD	46.8 (11.7)	46 (11)	47.7 (12.4)
5-Chair Stand score, mean s ± SD	39.3 (21.5)	38.0 (21.7)	40.6 (21.5)
TUG score, mean s ± SD	20.0 (10.5)	18.7 (9.8)	21.3 (11.2)
10-Meter Walk score, mean m/s ± SD	.80 (.32)	.79 (.29)	.81 (.35)

Abbreviation: BMI, body mass index.

The main strength of both programs from the patients' perspective was the telephone-delivery format. Additional strengths noted were that the sessions provided encouragement/motivation/confidence to engage in the recovery process and that the sessions increased activity. Specific to the CBPT program, 40% reported that the program made them feel accountable to someone, 30% felt that the program made them more aware of what they could do about their condition, and 25% learned to discuss things more openly with their doctor or feel more connected to their medical staff. Weaknesses noted by participants included the following: 1) programs were not long enough; 2) more information was needed on the healing process and restrictions; 3) guidelines for recovery were needed; and 4) the programs needed to start closer to discharge from the hospital.

Primary Outcomes

Average primary outcome scores for the CBPT group demonstrated an improvement in back and leg pain, pain interference, and disability over time (Table 3). The education group scores for leg pain and disability improved; however, average back pain and pain interference scores remain unchanged from posttreatment to 3-month follow-up.

Group differences in pain and disability were statistically significant at 3-month follow-up ($P < .05$), but not posttreatment (Table 3). The mean change from pretreatment to 3-month follow-up for the CBPT group was above MCID for the BPI pain interference score (−1.7 points; 95% CI = −2.4 to −1.1) and ODI score (−17.3; 95% CI = −20.3 to −14.4). The effect size for back and leg pain was .62, pain interference was .72, and disability was .79.

Multivariable linear regression analyses controlling for the pretreatment score of the outcome of interest, age, education, comorbid conditions, and number of physical therapy visits found that CBPT participants had BPI back pain scores that were −.85 points lower (95% CI = −1.4 to −.25; $P = .006$), BPI leg pain scores were −1.1 points lower (95% CI = −1.9 to −.27; $P = .009$), BPI pain interference scores were −1.3 points lower (95% CI = −2.1 to −.40; $P = .005$), and ODI scores were −9.4 points lower (95% CI = −14.9 to −4.0; $P = .001$) than education participants at the 3-month follow-up. The regression models accounted for 64% and 44% of the variance for back and

Table 2. **Acceptability of CBPT and Education Programs to Study Participants (N = 84)**

Measure	*CBPT (N = 42)*	*Education (N = 42)*
1. Helpful (0–10), mean (SD)	8.9 (1.7)	8.1 (2.1)
2. Likely to recommend (0–10), mean (SD)*	9.3 (1.6)	8.3 (2.7)
3. Overall benefit, taking into account the effort put into it, n (%)		
Benefits far outweighed the effort	25 (59.5)	19 (45.2)
Benefits somewhat outweighed the effort	5 (11.9)	4 (9.5)
Benefits equaled the effort	11 (26.2)	14 (33.3)
Effort somewhat outweighed the benefits	1 (2.4)	2 (4.8)
Effort far outweighed the benefits	0 (0)	3 (7.2)
4. Importance of changes in pain, n (%)*		
Pain decreased a meaningful amount	23 (54.8)	9 (21.4)
Some decrease in pain, but not enough to be meaningful	6 (14.3)	6 (14.3)
No change in pain	13 (30.9)	27 (64.3)
Some increase in pain, but not enough to be meaningful	0 (0)	0 (0)
Pain increased a meaningful amount	0 (0)	0 (0)
5. Importance of changes in activity, n (%)*		
Activity increased a meaningful amount	32 (76.2)	14 (33.3)
Some increase in activity, but not enough to be meaningful	4 (9.5)	7 (16.7)
No change in activity	6 (14.3)	20 (47.6)
Some decrease in activity, but not enough to be meaningful	0 (0)	1 (2.4)
Activity decreased a meaningful amount	0 (0)	0 (0)
6. Compared with other services, the importance of the program to recovery, n (%)		
Much more important	12 (28.5)	6 (14.3)
Somewhat more important	7 (16.7)	10 (23.8)
As important	21 (50)	17 (40.5)
Somewhat less important	1 (2.4)	5 (11.9)
Much less important	1 (2.4)	4 (9.5)

**P* < .05 for significant differences across groups.

leg pain, 49% of the variance for pain interference, and 59% of the variance for disability.

Secondary Outcomes: General Health

Average general health outcome scores for the CBPT group demonstrated an improvement in physical and mental health over time (Table 4). The education group scores for physical health improved; however, average mental health scores remained relatively unchanged.

Group differences in general physical health were statistically significant at the 3-month follow-up and in mental health at the posttreatment follow-up and the 3-month follow-up (*P* < .05; Table 4). The change scores for the SF-12 PCS of 6.6 and 13.4 and SF-12 MCS of 7.4 and 12.5 at 3 and 6 months after surgery in the CBPT

Table 3. **Primary Outcome Scores and Change from Pretreatment to Posttreatment and 3-Month Follow-Up by Group**

			Mean Change from Pretreatment		*Between-Group Difference*	
Variable	*CBPT Mean (SD)*	*Education Mean (SD)*	*CBPT*	*Education*	*CBPT vs Education*	P *Value*
BPI: back pain						
Pretreatment	3.0 (2.2)	2.8 (2.0)				
Posttreatment	2.9 (2.6)	2.5 (2.0)	−.08 (−.65 to .49)	−.3 (−.68 to .08)	.22 (−.46 to .9)	.52
3 mo	1.9 (2.0)	2.5 (2.4)	−1.1 (−1.5 to −.74)	−.26 (−.76 to .23)	−.88 (−1.5 to −.25)	.007
BPI: leg pain						
Pretreatment	2.5 (2.6)	2.2 (2.1)				
Posttreatment	2.1 (2.2)	2.3 (2.2)	−.48 (−.91 to −.06)	.05 (−.34 to .44)	−.53 (−1.1 to .04)	.07
3 mo	1.3 (2.1)	2.1 (2.6)	−1.3 (−1.9 to −.72)	−.1 (−.75 to .55)	−1.2 (−2.1 to −.34)	.007
BPI: interference						
Pretreatment	3.8 (3.0)	3.1 (2.6)				
Posttreatment	3.2 (3.2)	2.8 (2.9)	−.65 (−1.16 to −.14)	−.3 (−.84 to .24)	−.35 (−1.1 to .38)	.34
3 mo	2.1 (2.5)	2.8 (2.8)	−1.7 (−2.4 to −1.1)	−.26 (−.89 to .38)	−1.5 (−2.4 to −.57)	.002
ODI score						
Pretreatment	38.8 (17.3)	34.0 (16.7)				
Posttreatment	28.6 (17.6)	27.9 (19.4)	−9.8 (−12.1 to −7.5)	−6.1 (−10.5 to −1.7)	−3.7 (−8.6 to 1.2)	.14
3 mo	21.1 (16.7)	26.5 (20.5)	−17.3 (−20.3 to −14.4)	−7.5 (−12.1 to −2.9)	−9.8 (−15.3 to −4.4)	<.001

Table 4. Secondary Outcome Scores and Change from Pretreatment to Posttreatment and 3-Month Follow-Up by Group

			Mean Change from Pretreatment		Between-Group Difference	
Variable	CBPT Mean (SD)	Education Mean (SD)	CBPT	Education	CBPT vs Education	P Value
SF-12: PCS						
Pretreatment	29.8 (.5)	32.0 (9.8)				
Posttreatment	36.5 (10.8)	36.9 (11.6)	6.6 (4.2–8.9)	4.8 (2.1–7.6)	1.7 (−1.9 to 5.3)	.34
3 mo	43.0 (10.9)	38.3 (11.4)	13.4 (10.4–16.4)	6.3 (3.3–9.3)	7.1 (2.9–11.3)	.001
SF-12: MCS						
Pretreatment	43.6 (11.9)	53.9 (10.1)				
Posttreatment	50.9 (10.0)	53.7 (12.2)	7.4 (5.3–9.5)	−.2 (−3.0 to 2.6)	7.6 (4.2–11.1)	<.001
3 mo	56.6 (8.1)	53.4 (12.0)	12.5 (9.6–15.4)	−.5 (−3.7 to 2.7)	13.0 (8.7–17.2)	<.001
5-Chair Stand, s						
Pretreatment	24.7 (18.9)	20.3 (15.4)				
Posttreatment	22.6 (18.4)	21.0 (16.2)	−2.7 (−5.4 to .01)	.4 (−3.2 to 4)	−3.1 (−7.5 to 1.4)	.17
3 mo	13.6 (5.1)	16.7 (11.7)	−11.6 (−17.3 to −5.9)	−4.6 (−8.2 to −.91)	−7 (−13.7 to −.37)	.04
TUG test, s						
Pretreatment	11.5 (5.0)	11.6 (5.9)				
Posttreatment	10.0 (2.7)	12.1 (6.4)	−1.7 (−2.8 to −.55)	.33 (−1.2 to 1.9)	−2.0 (−3.9 to −.11)	.04
3 mo	9.6 (4.2)	10.9 (6.2)	−2.1 (−3.3 to −.9)	−.54 (−1.9 to .78)	−1.6 (−3.3 to .19)	.08
10-Meter Walk test, m/s						
Pretreatment	1.02 (.26)	1.07 (.29)				
Posttreatment	1.19 (.20)	1.08 (.30)	.10 (.05 to .15)	.01 (−.08 to .09)	.09 (−.01 to .19)	.07
3 mo	1.21 (.22)	1.17 (.26)	.20 (.13 to .26)	.10 (.004 to .19)	.10 (−.14 to .21)	.08

group, respectively, were above the MCID value of 4.9 points. Physical and mental health effects sizes were .75 and 1.35, respectively.

Multivariable analyses found that SF-12 PCS and MCS scores were 6.4 points higher (95% CI = 2.3 to 10.6; *P* = .003) and 8.6 points higher (95% CI = 4.5 to 12.7; *P* < .001), respectively, in the CBPT group compared with the education group, indicating better overall health for CBPT participants at the 3-month follow-up. The regression models accounted for 59% and 35% of the variance for physical and mental health.

Secondary Outcomes: Performance-Based Tests

Average physical performance outcome scores for the CBPT group demonstrated an improvement in performance-based function over time (Table 4). The education group scores for the 10-Meter Walk test improved; however, the average time it took participants to complete the 5-Chair Stand and TUG tests increased from the pretreatment to the posttreatment time point.

The CBPT group had a greater mean improvement in the 5-Chair Stand and TUG tests at 3-month follow-up and posttreatment, respectively (*P* < .05; Table 4). Change in seconds for the 5-Chair Stand and TUG tests were clinically significant with values greater than the MCID of 2.3 and 1.4, respectively. Effect sizes for the performance-based tests ranged from .41 to .49, with the largest effect found for the 5-Chair Stand test.

In multivariable analyses, CBPT participants had greater improvement in performance-based tests scores than the education group at the 3 month follow-up, with scores 4.3 seconds lower (95% CI = −7.7 to −.82; *P* = .02) for the 5-Chair Stand test, 1.8 seconds lower (95% CI = −3.2 to −.16; *P* = .02) for the TUG test, and .09 m/s higher (95% CI = −.008 to .18; *P* = .07) for the 10-Meter Walk test. Treatment effects were significant for the 5-Chair Stand and TUG tests. The regression models accounted for 52% of the variance for 5-Chair Stand test, 62% of the variance for TUG test, and 33% of the variance for 10-Meter Walk test.

Psychosocial Characteristics

Average psychosocial scores for the CBPT and education groups demonstrated an improvement in fear of movement and pain self-efficacy over time (Table 5).

Group differences in fear of movement and pain self-efficacy were statistically significant at the 3-month follow-up (*P* < .05) but not posttreatment (Table 4). The 5.9-point decrease from pretreatment to 3-month follow-up for the TSK was greater than the MCID of 4 points. The effect size for the TSK and PSEQ were .59 and .85, respectively.

Discussion

This trial was conducted to determine whether a CBPT program would lead to greater improvement in postoperative outcomes compared with an education program in patients with chronic pain undergoing lumbar spine surgery for degenerative conditions. A targeted CBT-based rehabilitation approach decreased fear of movement and increased self-efficacy as well as improved patient-reported and performance-based outcomes at 6 months after surgery.

CBPT participants demonstrated greater improvement in back and leg pain and pain interference with activity. Change scores in the CBPT group

Table 5. Psychosocial Scores and Change from Pretreatment to Posttreatment and 3-Month Follow-Up by Group

			Mean Change from Pretreatment		Between-Group Difference	
Variable	CBPT Mean (SD)	Education Mean (SD)	CBPT	Education	CBPT vs Education	P Value
TSK score (17–68)						
Pretreatment	40.2 (7.3)	38.7 (6.5)				
Posttreatment	37.5 (7.0)	36.3 (7.2)	2.6 (−4 to −1.2)	−2.2 (−3.9 to −.48)	−.4 (−2.6 to 1.8)	.70
3 mo	33.9 (8.1)	36.2 (8.4)	−5.9 (−7.7 to −4.2)	−2.3 (−4.4 to −.16)	−3.6 (−6.3 to −.93)	.009
PSEQ score (0–60)						
Pretreatment	36.0 (16.8)	41.5 (15.2)				
Posttreatment	42.7 (17.3)	44.1 (16.5)	6.5 (3.8–9.3)	2.6 (−1.4 to 6.6)	3.9 (−.86 to 8.7)	.11
3 mo	48.8 (14.3)	43.5 (16.3)	11.9 (8.7–15.1)	1.9 (−2.2 to 6)	10.0 (4.8–15.1)	<.001

between an early postoperative time point and 6 months after surgery are consistent with studies testing a psychologist-delivered CBT program (60-minute sessions twice a week for 8 weeks) and a group behavioral physical therapy intervention (90-minute sessions; 3 times over 8 weeks) in patients recovering from surgery for a degenerative condition.[11,56] The minimal changes in back and leg pain for our education program are similar to findings in trials examining traditional postoperative physical therapy and prospective cohort studies.[1,8,39] Group differences at 3-month follow-up also support work by Abbott et al[3] that found a greater improvement in back pain after lumbar fusion with a 3-session psychomotor therapy program compared with exercise training.

Additional research is needed to determine whether larger and clinically meaningful improvements in pain can be obtained through a CBT-based approach. Improvement in back and leg intensity was statistically significant in our study but not clinically meaningful, which may be due to the relatively low pain scores after surgery. The average back and leg pain scores pretreatment (6 weeks after surgery) were 2.9 and 2.4, respectively. A more time-intensive or in-person CBPT program may be needed to achieve MCID and substantial clinical benefit thresholds ranging from 1.2- to 2.5-point net improvement.[25]

The hypothesis that CBPT participants would have greater improvement in disability and general health was supported. Disability and general health improvement in the CBPT group at the 3-month follow-up was both statistically and clinically significant based on published MCID values. Our disability findings are consistent with trials by Abbott et al[3] and Monticone et al[56] comparing psychomotor therapy and CBT, respectively, with exercise training in patients after lumbar fusion. However, Abbott et al[3] did not find a significant difference in general health at 6 months after surgery, which may be accounted for by the short duration of the program (3 sessions) or early initiation (first 3 months after surgery). The large and clinically relevant changes noted in our study for disability and general physical and mental health may be due to the CBPT intervention's focus on decreasing barriers to functional activity and walking rather than focusing solely on resolution of pain symptoms.

Significant differences between the CBPT and education groups were found at the 3-month follow-up, but not posttreatment, for the pain, disability, and physical health outcomes. The non-significant findings posttreatment may be due to the rapid improvement in pain and disability that occurs during the initial 3-month postoperative period, regardless of the postoperative management strategy.[46,51,60] Another explanation may be that additional time is needed for patients to practice the cognitive and behavioral strategies presented in the CBPT program in order for improvements in pain and disability to occur.

Overall, these findings suggest that the CBPT program has the potential to be more effective for improving patient-reported outcomes than education and traditional clinic-based rehabilitation. Moderate to large effect sizes were found for the CBPT program during the postoperative recovery period (ie, the pretreatment measure was 6 weeks after surgery). It is important to note that previous studies have used preoperative scores as the baseline measure; thus, they may have overestimated treatment effects by capitalizing on the benefits of surgery. This study makes an important contribution by documenting that a CBT-based approach has the potential to make significant and clinically meaningful differences in outcomes beyond improvements that can be attributed to surgery.

The hypothesis regarding the performance-based tests was partially supported. Multivariable regression analyses found greater improvement in the 5-Chair Stand and TUG scores for CBPT participants at the 3 month follow-up but not for the 10-Meter Walk test. The high prevalence of comorbidities or fear of falling in this patient population may negatively affect gait speed to a greater extent than strength and mobility. Additional follow-up time may also be needed to detect meaningful change in 10-Meter Walk scores. The CBPT program produced a mean change in seconds for the 5-Chair Stand and TUG tests that can be considered clinically meaningful based on MCID values from previous studies.[54,91] This is the first study to assess the effects of a CBT-based intervention on physical performance in patients after spine surgery.

Clinical Implications

Research supports a comprehensive biopsychosocial approach to postoperative spine management.[59,70] Brief and telephone-administered CBT has been shown to improve pain and function in patients with chronic pain.[42,43,75,81,86] However, rehabilitation in surgical populations has not traditionally focused on CBT. This study along with others supports the use of CBT-based interventions for patients having surgery for chronic musculoskeletal conditions.[3,7,11,56,66]

This study also has clinical implications with regard to targeted rehabilitation interventions. The findings are consistent with studies in low back pain and whiplash populations that have found improvement in pain and disability after programs that specifically target the psychosocial variables contributing to poor outcomes.[23,71] The literature recommends identifying the variables that mediate the effects of CBT in order to provide more effective and clinically relevant treatments.[20]

The CBPT program broadens the availability of effective CBT strategies by expanding the implementation from traditional providers (ie, psychologists) to physical therapists. Several studies have reported on the benefits of CBT-based interventions when delivered by dental hygienists and nurses.[15,19,76] Our work and that of others demonstrates that physical therapists can learn and successfully implement the cognitive-behavioral techniques needed to make meaningful differences in pain-related outcomes.[7,23,30,37,71]

Limitations

This study is limited by incomplete follow-up at 3 months after treatment, especially for the performance-based tests. However, this concern is minimized by our finding that there were no significant differences in baseline characteristics between patients with and without complete follow-up data. The sample size was underpowered to detect small to medium group differences in performance-based tests. It is important to consider clinical significance as well as statistical significance when interpreting these results. The education group reported similar outcomes to those found in previous usual care arms and prospective cohort studies,[2,8,39] but additional research is needed to determine the effectiveness of CBPT alone compared with usual care, which typically consists of traditional clinic-based physical therapy. The treatments were delivered by a single physical therapist at a single center and to patients screened for high fear of movement, which limits the generalizability of our results. The long-term effectiveness of CBPT after spine surgery remains to be determined. Studies on CBT-based interventions after spine surgery have found inconsistent results. Monticone et al[56] reported positive findings at 1 year for a CBT program focusing on fear of movement and pain catastrophizing, whereas Abbott et al[3] and Christensen et al[11] found no group differences in back pain at 1 and 2 years after lumbar fusion. Additional research is needed to determine the optimal time for delivery (ie, addition of sessions preoperatively and/or immediately after surgery) and whether in-person administration would strengthen the effect of the CBPT program.

Conclusions

This randomized trial demonstrates that screening patients for fear of movement and using a targeted CBPT program results in significant and clinically meaningful improvement in pain, disability, general health, and physical performance after spine surgery for degenerative conditions. The CBPT program, delivered by physical therapists over the telephone, has the potential to be an evidence-based program that clinicians can recommend for patients at risk for poor postoperative outcomes.

References

1. Aalto TJ, Leinonen V, Herno A, Alen M, Kröger H, Turunen V, Savolainen S, Saari T, Airaksinen O: Postoperative rehabilitation does not improve functional outcome in lumbar spinal stenosis: a prospective study with 2-year postoperative follow-up. Eur Spine J 20: 1331-1340, 2011

2. Abbas J, Hamoud K, May H, Peled N, Sarig R, Stein D, Alperovitch-Najemson D, Hershkovitz I: Socioeconomic and physical characteristics of individuals with degenerative lumbar spinal stenosis. Spine 38:E554-E561, 2013

3. Abbott AD, Tyni-Lenné R, Hedlund R: Early rehabilitation targeting cognition, behavior, and motor function after lumbar fusion: a randomized controlled trial. Spine 35: 848-857, 2010

4. Anderson T, Bunger C, Sogaard R: Long-term health care utilization and costs of spinal fusion in elderly patients. Eur Spine J 22:977-984, 2013

5. Archer KR, Wegener ST, Seebach C, Song Y, Skolasky RL, Thornton C, Khanna AJ, Riley LH 3rd: The effect of fear of movement beliefs on pain and disability after surgery for lumbar and cervical degenerative conditions. Spine 36: 1554-1562, 2011

6. Archer KR, Phelps KD, Seebach CL, Song Y, Riley LH 3rd, Wegener ST: Comparative study of short forms of the Tampa Scale for Kinesiophobia: fear of movement in a surgical spine population. Arch Phys Med Rehabil 93:1460-1462, 2012

7. Archer KR, Motzny N, Abraham CM, Yaffe D, Seebach CL, Devin CJ, Spengler DM, McGirt MJ, Aaronson OS, Cheng JS, Wegener ST: Cognitive-behavioral based physical therapy to improve surgical spine outcomes: a case series. Phys Ther 83: 1140-1149, 2013

8. Archer KR, Seebach CL, Mathis S, Riley LH, Wegener ST: Early postoperative fear of movement predicts pain, disability, and physical health 6 months after spinal surgery for degenerative conditions. Spine J 14:759-767, 2014

9. Atlas SJ, Keller RB, Wu YA, Deyo RA, Sinder DE: Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the Maine lumbar spine study. Spine 30:936-943, 2005

10. Brooks D, Davis AM, Naglie G: Validity of 3 physical performance measures in inpatient geriatric rehabilitation. Arch Phys Med Rehabil 87:105-110, 2006

11. Christensen FB, Laurberg I, Bünger CE: Importance of the back-café concept to rehabilitation after lumbar spinal fusion: a randomized clinical study with a 2-year follow-up. Spine 28:2561-2569, 2003

12. Cleeland CS, Ryan KM: Pain assessment: global use of the brief pain inventory. Ann Acad Med Singapore 23:129-138, 1994

13. Cohen J: Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, Lawrence Erlbaum, 1988

14. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY: Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. Spine J 8:968-974, 2008

15. Dalton JA, Keefe FJ, Carlson J, Youngblood R: Tailoring cognitive-behavioral treatment for cancer pain. Pain Manag Nurs 5:3-18, 2004

16. Davidson M, Keating JL: A comparison of 5 low back disability questionnaires: reliability and responsiveness. Phys Ther 82:8-24, 2002

17. den Boer JJ, Oostendorp R, Beems T, Munneke M, Evers A: Continued disability and pain after lumbar disc surgery. Pain 123:45-52, 2006

18. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG: Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. J Am Med Assoc 303:1259-1265, 2010

19. Dworkin SF, Huggins KH, Wilson L, Mancl L, Turner J, Massoth D, LeResche L, Truelove E: A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-Axis II to target clinic cases for a tailored self-care TMD treatment program. J Orofac Pain 16:48-63, 2002

20. Ehde DM, Dillworth TM, Turner JA: Cognitive-behavioral therapy for individuals with chronic pain: efficacy, innovations, and directions for research. Am Psychol 69:152-166, 2014

21. Fairbank JC, Pynsent PB: The Oswestry Disability Index. Spine 25:2940-2952, 2000

22. French DJ, France CR, Vigneau F, French JA, Evans RT: Fear of movement/(re)injury in chronic pain: a psychometric assessment of the original English version of the Tampa scale for kinesiophobia (TSK). Pain 127:42-51, 2007

23. George SZ, Fritz JM, Bialosky JE, Donald DA: The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. Spine 28:2551-2560, 2003

24. Gibson JNA, Waddell G: Surgery for degenerative lumbar spondylosis. Cochrane Database Syst Rev (4), 2005. Art. No.: CD001352

25. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY: Defining substantial clinical benefit following lumbar spine arthrodesis. J Bone Joint Surg Am 90:1839-1847, 2008

26. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB: A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 49:M85-94, 1994

27. Hardy SE, Perera S, Roumani YF, Chandler JM, Studenski SA: Improvement in usual gait speed predicts better survival in older adults. J Am Geriatr Soc 55:1727-1734, 2007

28. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42:377-381, 2009

29. Hasenbring MI, Plaas H, Fischbein B, Willburger R: The relationship between activity and pain in patients 6 months after lumbar disc surgery: do pain-related coping modes act as moderator variables? Eur J Pain 10:701-709, 2006

30. Hay E, Lewis M, Vohara K, Main CJ, Watson P, Dziedzic KS, Sim J, Minns Lowe C, Croft PR: Comparison of physical treatments versus a brief pain-management programme for back pain in primary care: a randomized clinical trial in physiotherapy practice. Lancet 365:2024-2030, 2005

31. Issak PS, Cunningham ME, Pumberger M, Hughes AP, Cammisa FP: Degenerative lumbar spinal stenosis: evaluation and management. J Am Acad Orthop Surg 20:527-535, 2012

32. Jansson K, Németh G, Granath F, Jönsson B, Blomqvist P: Health-related quality of life (EQ-5D) before and 1 year after surgery for lumbar spinal stenosis. J Bone J Surg B 91:210-216, 2009

33. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, Stradling J: A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? J Public Health Med 19:179-186, 1997

34. Kalichman L, Cole R, Kim Dh, Li L, Suri P, Guermazi A, Hunter DJ: Spinal stenosis prevalence and association with symptoms: the Framingham Study. Spine J 9:545-550, 1990

35. Kalichman L, Kim Dh, Li L, Guermazi A, Hunter DJ: Computed tomography-evaluated features of spinal degeneration: prevalence, intercorrelation, and association with self-reported low back pain. Spine J 10:200-208, 2010

36. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS: Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. Clin J Pain 20:309-318, 2004

37. Klaber Moffett JA, Jackson DA, Richmond S, Hahn S, Coulton S, Farrin A, Manca A, Torgerson DJ: Randomized trial of brief physiotherapy intervention compared with usual physiotherapy for neck pain patients: outcomes and patients' preference. BMJ 330:75-80, 2005

38. Kori SH, Miller RP, Todd DD: Kinesiophobia: a new region of chronic pain behavior. Pain Manag 3:35-43, 1990

39. Kreiner DS, Shaffer WO, Baisden JL, Gilbert TJ, Summers JT, Toton JF, Hwang SW, Mendel RC, Reitman CA: An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. Spine J 13:734-743, 2013

40. Kroenke K, Spitzer RL, Williams JB: The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16:606-613, 2001

41. Kroenke K, Spitzer RL, Williams JB, Lowe B: The patient health questionnaire somatic, anxiety, and depressive

symptom scales: a systematic review. Gen Hosp Psychiatry 32:345-359, 2010

42. Litt MD, Shafer DM, Kreutzer DL: Brief cognitive-behavioral treatment for TMD pain: long-term outcomes and moderators of treatment. Pain 151:110-116, 2010

43. Linton SJ, Andersson T: Can chronic disability be prevented? A randomized trial of a cognitive behavioral intervention and 2 forms of information for patients with spinal pain. Spine 25:2825-2831, 2000

44. Lorig KR, Ritter P, Stewart AL, Sobel DS, Brown BW Jr, Bandura A, Gonzalez VM, Laurent DD, Holman HR: Chronic disease self-management program: 2-year health status and health care utilization outcomes. Med Care 39:1217-1223, 2001

45. Lurie JD, Tosteson TD, Tosteson A, Abdu WA, Zhao W, Morgan TS, Weinstein JN: Long-term outcomes of lumbar spinal stenosis: 8-year results of the Spine Patient Outcomes Research Trial (SPORT). Spine 40:63-76, 2015

46. Mannion AF, Denzler R, Dvorak J, Muntener M, Grob D: A randomized controlled trial of post-operative rehabilitation after surgical compression of the lumbar spine. Eur Spine J 16:1101-1107, 2007

47. Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Dvorak J, Jacobshagen N, Semmer NK, Boos N: Predictors of multidimensional outcome after spinal surgery. Eur Spine J 16:777-786, 2007

48. Mannion AF, Denzler R, Dvorak J, Grob D: Five-year outcome of surgical decompression of the lumbar spine without fusion. Eur Spine J 19:1883-1891, 2010

49. Marks R: Reliability and validity of selfpaced walking time measures for knee osteoarthritis. Arthritis Care Res 7:50-53, 1994

50. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, Sullivan SD: Expenditures and health status among adults with back and neck problems. JAMA 299:656-664, 2008

51. McGregor AH, Dore CJ, Morris TP, Morris S, Jamrozik K: Function after spinal treatment, exercise, and rehabilitation (FASTER). Spine 36:1711-1720, 2011

52. McGregor AH, Probyn K, Cro S, Doré CJ, Burton AK, Balagué F, Pincus T, Fairbank J: A Cochrane review. Rehabilitation following surgery for lumbar spinal stenosis. Spine 39:1044-1054, 2014

53. Mendoza TR, Chen C, Brugger A, Hubbard R, Snabes M, Palmer SN, Zhang Q, Cleeland CS: The utility and validity of the modified brief pain inventory in a multiple-dose postoperative analgesic trial. Clin J Pain 20:357-362, 2004

54. Meretta BM, Whitney SL, Marchetti GF, Sparto PJ, Muirhead RJ: The five times sit to stand test: responsiveness to change and concurrent validity in adults undergoing vestibular rehabilitation. J Vestib Res 16:233-243, 2006

55. Miles CL, Pincus T, Carnes D, Taylor SJ, Underwood M: Measuring pain self-efficacy. Clin J Pain 27:461-470, 2011

56. Monticone M, Ferrante S, Teli M, Rocca B, Foti C, Lovi A, Brayda Bruno M: Management of catastrophising and kinesiophobia improves rehabilitation after fusion for lumbar spondylolisthesis and stenosis. A randomised controlled trial. Eur Spine J 23:87-95, 2014

57. Nederhand MJ, Ijzerman M, Hermens HJ, Turk DC, Zilvold G: Predictive value of fear avoidance in developing chronic neck pain disability: consequences in clinical decision making. Arch Phys Med Rehabil 85:496-501, 2004

58. Nicholas MK: The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 11:153-163, 2007

59. Ostelo RW, de Vet HC, Waddell G, Kerckhoffs MR, Leffers P, van Tulder M: Rehabilitation following first-time lumbar disc surgery: a systematic review within the framework of the Cochrane Collaboration. Spine 28:209-218, 2003

60. Oestergaard LG, Nielsen CV, Bünger CE, Sogaard R, Fruensgaard S, Helmig P, Christensen FB: The effect of early initiation of rehabilitation after lumbar spinal fusion: a randomized clinical study. Spine 37:1803-1809, 2012

61. Parker SL, Adogwa O, Mendenhall SK, Shau DN, Anderson WN, Cheng JS, Devin CJ, McGirt MJ: Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. J Neurosurg Spine 16:61-67, 2012

62. Parker SL, Adogwa O, Paul AR, Anderson WN, Aaronson O, Cheng JS, McGirt MJ: Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. J Neurosurg Spine 14:598-604, 2011

63. Perera S, Mody SH, Woodman RC, Studenski SA: Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 54:743-749, 2006

64. Podsiadlo D, Richardson S: The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 39:142-148, 1991

65. Pratt RK, Fairbank JC, Virr A: The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. Spine 27:84-91, 2002

66. Riddle DL, Keefe FJ, Nay WT, McKee D, Attarian DE, Jensen MP: Pain coping skills training for patients with elevated pain catastrophizing who are scheduled for knee arthroplasty: a quasi-experimental study. Arch Phys Med Rehabil 92:859-865, 2011

67. Roelofs J, Goubert L, Peters ML, Vlaeyen JW, Crombez G: The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain 8:495-502, 2004

68. Rosenthal JA: Qualitative descriptors of strength of association and effect size. J Soc Serv Res 21:37-59, 1996

69. Seebach CL, Kirkhart M, Lating JM, Wegener ST, Song Y, Riley LH 3rd, Archer KR: Examining the role of positive and negative affect in recovery from spine surgery. Pain 153:518-525, 2012

70. Soegaard R, Christensen FB, Lauerberg I, Bünger CE: Lumbar spinal fusion patients' demands to the primary health sector: evaluation of three rehabilitation protocols. A prospective randomized study. Eur Spine J 15:648-656, 2006

71. Sullivan MJ, Adams H, Rhodenizer T, Stanish WD: A psychosocial risk factor-targeted intervention for the prevention of chronic pain and disability following whiplash injury. Phys Ther 86:8-18, 2006

72. Szpalski M, Gunzburg R: Lumbar spinal stenosis in the elderly: an overview. Eur Spine J 12(Suppl 2):S170-S175, 2013

73. Tilson JK, Sullivan KJ, Cen SY, Rose DK, Koradia CH, Azen SP, Duncan PW: Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. Phys Ther 90:196-208, 2010

74. Turner JA, Mancl L, Aaron LA: Brief cognitive-behavioral therapy for temporomandibular disorder pain: effects on daily electronic outcome and process measures. Pain 117: 377-387, 2005

75. Turner JA, Mancl L, Aaron LA: Short- and long-term efficacy of brief cognitive-behavioral therapy for patients with chronic temporomandibular disorder pain: a randomized controlled trial. Pain 121:181-194, 2006

76. Turner JA, Mancl L, Huggins KH, Sherman JJ, Lentz G, LeResche L: Targeting temporomandibular disorder pain treatment to hormonal fluctuations: a randomized clinical trial. Pain 152:2074-2084, 2011

77. United Nations Department of Economic and Social Affairs: World Economics and Social Survey 2007. Development in an Aging World. New York, United Nations, 2007, pp 1-212

78. van der Windt D, Hay E, Jellema P, Main C: Psychosocial interventions for low back pain in primary care: lessons learned from recent trials. Spine 33:81-89, 2008

79. Vlaeyen J, Kole-Snijders A, Boeren R, van Eek H: Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 62:363-372, 1995

80. Vlaeyen JW, de Jong J, Geilen M, Heits P, van Breukelen G: The treatment of fear of movement/(re)injury in chronic low back pain: further evidence on the effectiveness of exposure in vivo. Clin J Pain 18:251-261, 2002

81. Von Korff M, Balderson BH, Saunders K, Miglioretti DL, Lin EH, Berry S, Moore JE, Turner JA: A trial of an activating intervention for chronic back pain in primary care and physical therapy settings. Pain 113:323-330, 2005

82. Ware J Jr, Kosinski M, Keller SD: A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 34:220-233, 1996

83. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden S, Hilibrand A, Goldberg H, Berven S, An H: Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 358:794-810, 2008

84. Weinstein JN, Tosteson TD, Lurie JD, Tosteson A, Blood E, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goodberg H, Berven S, An H: Surgical versus nonoperative treatment for lumbar spinal stenosis 4 year results of the spine patient outcomes research trial. Spine 35:1329-1338, 2010

85. Wideman TH, Adams H, Sullivan MJ: A prospective sequential analysis of the fear-avoidance model of pain. Pain 145:45-51, 2009

86. Williams DA, Cary M, Groner K, Chaplin W, Glazer L, Rodriguez A, Clauw D: Improving physical functional status in patients with fibromyalgia: a brief cognitive-behavioral intervention. J Rheumatol 29:1280-1286, 2002

87. Williams AC, McCracken LM: Cognitive-behavioral therapy for chronic pain: an overview with specific references to fear and avoidance. In: Asmundson GG, Vlaeyen JW, Crombez G (eds): Understanding and Treating Fear of Pain. London, Oxford University Press, 2004, pp 293-312

88. Williams P: Decision making in surgical treatment of chronic low back pain: the performance of prognostic tests to select patients for lumbar spinal fusion. Acta Orthop Suppl 84:1-35, 2013

89. Woby SR, Roach NK, Urmston M, Watson PJ: Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. Pain 117:137-144, 2005

90. Woods MP, Asmundson GJ: Evaluating the efficacy of graded in vivo exposure for the treatment of fear in patients with chronic back pain: a randomized controlled clinical trial. Pain 136:271-280, 2008

91. Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH: A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. J Orthop Sports Phys Ther 41:319-327, 2011

92. Zalon ML: Comparison of pain measures in surgical patients. J Nurs Meas 7:135-152, 1999

Original Article

Effect of the Use of Buzzy® during Phlebotomy on Pain and Individual Satisfaction in Blood Donors

Dilek Yılmaz, PhD,* **Yasemin Heper, MD,†**
and Leyla Gözler, BA†

Abstract:

Phlebotomy causes pain and discomfort to adults. The objective of this study was to investigate the effect of the use of Buzzy® on phlebotomy satisfaction and pain relating to the phlebotomy process in healthy adult blood donors voluntarily donating blood. This was a prospective, randomized, controlled experimental study. The research sample was made up of 90 healthy adult men. These individuals were randomly assigned to an experimental group (Buzzy group), a placebo control group, and a nonintervention control group. For the individuals in the experimental group, the ice wings of the Buzzy device, frozen solid in the refrigerator, were placed approximately 5 centimeters above the intervention site from 1 minute before the procedure until the end of the needle location process. When the device was operated, it applied vibration and cold to the site. For individuals in the placebo control group, the Buzzy device was also located approximately 5 centimeters above the intervention site from 1 minute before the procedure until the end of the needle location process, but with the ice wings at room temperature (unfrozen) and with the vibration switch remaining off. For the nonintervention control group, no intervention was implemented before the procedure. Immediately after entry to the vein, pain levels and levels of phlebotomy satisfaction were assessed in individuals in all groups. A statistically significant difference was determined between the mean pain and phlebotomy satisfaction scores of individuals in the experimental and control groups ($p < .05$). Results indicate that use of the Buzzy device was an effective method of reducing the pain of phlebotomy and increasing phlebotomy satisfaction in healthy adult male blood donors.

© 2017 by the American Society for Pain Management Nursing

*From the *Department of Nursing, School of Health; †Dr. Raşit Durusoy Blood Bank, Uludag University, Turkey.*

Address correspondence to Dilek Yılmaz, PhD, Department of Nursing, School of Health, Uludag University, Turkey. E-mail: dilekkara15@hotmail.com

Received November 27, 2016;
Revised January 25, 2017;
Accepted March 23, 2017.

This study was presented as an oral presentation on May 25-27, 2017, 4th Basic Nursing Care Conference in Mugla, Turkey.

1524-9042/$36.00
© 2017 by the American Society for Pain Management Nursing
http://dx.doi.org/10.1016/j.pmn.2017.03.005

INTRODUCTION

Phlebotomy is an invasive intervention that is widely practiced in health care (Aydin, Şahiner, & Çiftçi, 2016; Kiran, Kaur, & Marwaha, 2013; Whelan, Kunselman, Thomas, Moore, & Tamburro, 2014). It has been reported that individuals often experience stress and acute pain during phlebotomy (Ialongo & Bernardini, 2016; Whelan et al., 2014). In addition, it has been determined that hospital visits may be delayed because of fear of needles and symptoms such as pain (Pasero, 2003).

Control of the pain created by entry of the needle into the vein is achieved by both pharmacological and nonpharmacological methods (Schechter et al., 2007; Taddio et al., 2010). Hitherto, control of the pain of phlebotomy has been achieved by topical anesthetic creams (Eichenfield, Funk, Fallon-Friedlander, & Cunningham, 2002; Rogers & Ostrow, 2004) or vapocoolant sprays (Farion, Splinter, Newhook, Gaboury, & Splinter, 2008; Jimenez, Bradford, Seidel, Sousa, & Lynn, 2006). But it has been reported in studies that pharmacological methods used in pain management have limitations such as the possibility of side effects (Pershad, Steinberg, & Waters, 2008; Sethna et al., 2005; Zempsky et al., 2008), extra time and greater cost (Fein & Gorelick, 2006; Leahy et al., 2008; Sethna et al., 2005), and not being very suitable for use in health care environments where speed is necessary, such as in phlebotomy stations (Baxter, Cohen, McElvery, Lawson, & von Baeyer, 2011; MacLaren & Cohen, 2007; Inal & Kelleci, 2012). On the other hand, nonpharmacological approaches used in the control of pain include techniques such as listening to music, attracting the attention elsewhere, and blowing into a sphygmomanometer (Dutt-Gupta, Brown, & Mycama, 2007; Sinha, Tandon, & Singh, 2005). Although some data obtained from studies support each of these techniques, it is reported that there is no one single integrated method for optimal pain control (Inal & Kelleci, 2012). For all these reasons, in fast-tempo work environments, such as phlebotomy stations, there is a need for effective methods for the control of pain related to invasive procedures that are easy to use, do not entail high costs, and do not have the possibility of side effects.

Buzzy® (MMJ Labs, Atlanta, GA, USA) is a device designed to reduce the pain experienced during invasive procedures. It uses the combined application of external cold and vibration to help relieve pain and discomfort and can be used repeatedly on both children and adults (Home/Buzzy Helps, 2016) (Fig. 1). Studies have reported that the use of Buzzy gives positive results in reducing pain during entry to a vein (Baxter et al., 2011; Inal & Kelleci, 2012; Schreiber et al., 2016; Whelan et al., 2014).

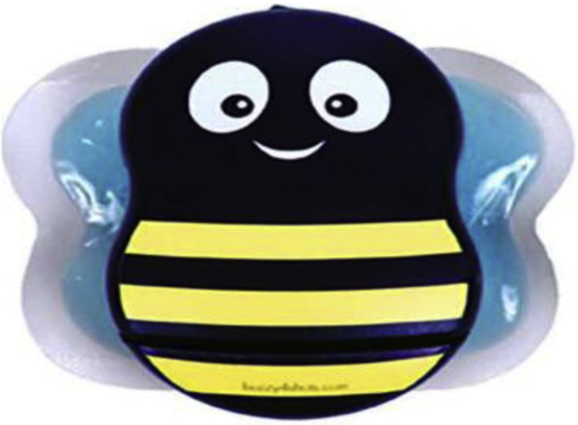


Figure 1. ■ Buzzy®.

In Turkey, nonpharmacological methods are not routinely used to reduce the pain of phlebotomy in voluntary blood donors. Studies in the literature examining the effect on pain of the use of Buzzy during phlebotomy have compared it either with a control group or with pharmacological methods. This study differs from others in that only adult male voluntary blood donors were included, individual satisfaction with the procedure was evaluated, and the procedure group was compared with control groups both with and without intervention. Therefore, this study will identify whether or not the Buzzy device has a placebo effect.

In this way, a need was felt for research based on the previously stated results to determine whether nonpharmacological methods can be effective and whether they can raise satisfaction in controlling the pain and discomfort of phlebotomy. It is thought that the results of this research will help health professionals to extend the use of a simple nonpharmacological method and will make a contribution to the literature.

The aim of this study was to investigate the effect of the use of Buzzy on phlebotomy satisfaction and the pain of phlebotomy in healthy adult men who were voluntarily donating blood.

MATERIAL AND METHODS

Research Design

This was a prospective randomized controlled experimental study.

Study Participants

The study was conducted between November 2015 and May 2016 at the blood center of a university hospital located in the Marmara region of Turkey. The research sample consisted of 90 healthy adult males who had come to the center to donate blood voluntarily. Only men were included in the study because most of those coming to the blood center where the research was conducted were men and because a gender factor in the perception of pain might have affected the results (Chung, Ng, & Wong, 2002; Dawson & List, 2009).

Criteria for inclusion in the study were as follows: not having heart or lung disease, not having any disorder that would affect the perception of pain, not having developed any complication that would hinder the procedure in the area in which it was to be performed, not having previously undergone phlebotomy, and participating voluntarily. The size of the research sample was determined statistically by power analysis. After a pilot study, a certain number of samples were collected and power analysis was performed on them using descriptive statistics. The sample size of each experimental group was calculated as 30 for 0.80 power and 0.05 type I error. The relevant calculation was made using the program PASS 13.0 (PASS, Kaysville, Utah, USA). Those for whom it was thought that entry to the vein might be unsuccessful at the first attempt and those whose vein was not easily seen were excluded. The flow diagram showing the steps followed in the study procedure is presented in Figure 2.

Enrolled individuals were allocated randomly into three groups: the experimental group (the Buzzy group), the placebo control group, and the nonintervention control group. A randomization sequence was created using R software 2008 (Windows, Vienna, Austria) with a 1:1 allocation using random block sizes of 3 and 6.

Ethical Considerations

The necessary legal permission to carry out the research was obtained from the Local Ethics Committee (Decision No. 2015-13/24). In addition, the individuals were given information concerning the research before the study was started. After information was given, a signed informed consent form was obtained from those who participated voluntarily in the study concerning their voluntary participation.

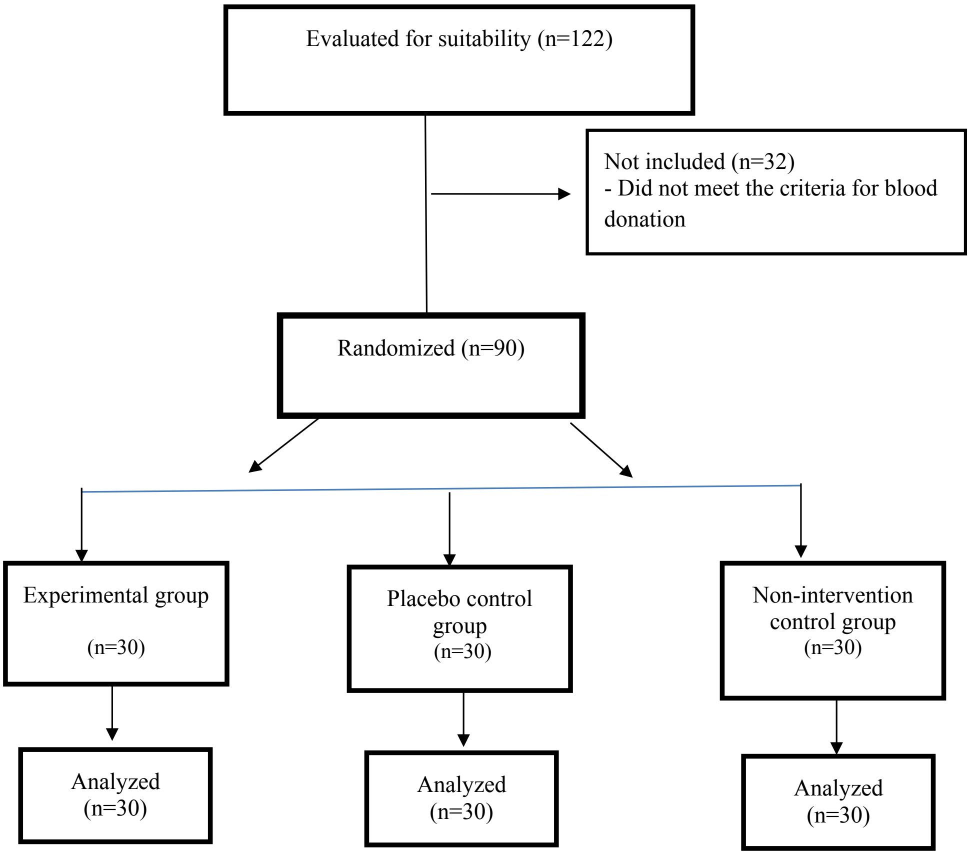

Evaluated for suitability (n=122)
Not included (n=32)
- Did not meet the criteria for blood donation
Randomized (n=90)
Experimental group (n=30)
Placebo control group (n=30)
Non-intervention control group (n=30)
Analyzed (n=30)
Analyzed (n=30)
Analyzed (n=30)

Figure 2. ■ Study flow chart.

Instruments

Individual Identification Form. An individual identification form collected information on individuals' age, height, and weight.

Visual Analog Scale. In evaluating the intensity of pain experienced by the individuals during the procedure, a 10-centimeter vertical visual analog scale was used. One end indicated no pain and the other end the worst possible pain (Kahl & Cleland, 2005). Pain intensity scores were assessed in millimeters.

Phlebotomy Satisfaction Evaluation Scale. In order to assess the level of satisfaction during phlebotomy, a scale developed by the researchers was used that consisted of a 100-millimeter vertical line with "I am very satisfied" at one end and "I am not at all satisfied" at the other.

Data Collection

After the participants' voluntary participation had been ensured, their identifying data were collected on the form, and then they were given information on the use of the visual analog scale. In the phlebotomy procedure, a standard needle and equipment were used on all participants. All phlebotomies were performed by the same researcher, and all were performed on the antecubital vein of the right arm. The standard phlebotomy protocol carried out on the individuals of the experimental and control groups is shown in Table 1.

Table 1.
Phlebotomy Protocol

The participant was placed in a supine position.
- Wearing gloves, the vein for the intervention was assessed by observation and palpation.
- An automatic tourniquet was attached 12 cm above the intervention site.
- The intervention site was cleaned with antiseptic solution with a single movement.
- The needle was held approximately 1 cm below the vein, which was to be entered at an angle of 30° to 45° to the skin.
- As the needle entered the vein, the entry angle was reduced to approximately 15° and the needle was advanced slowly in the vein.
- A check was made as to whether blood was entering the phlebotomy set.
- The needle position was fixed on the skin according to aseptic principles.
- When the phlebotomy process was completed, the tourniquet was released, the needle was removed aseptically, and the area was pressed with sterile gauze.

Protocol for Experimental and Control Groups

Experimental Group. For the individuals in the experimental group, the ice wings of the Buzzy device, frozen solid in the refrigerator, were placed approximately 5 centimeters above the intervention site from 1 minute before the procedure until the end of the needle location process. When the device was operated, it applied vibration and cold to the site. At the end of 1 minute, the device was removed and immediately afterward the vein entry procedure was implemented.

Placebo Control Group. For individuals in the placebo control group, the Buzzy device was also located approximately 5 centimeters above the intervention site from 1 minute before the procedure until the end of the needle location process, but with the ice wings at room temperature (unfrozen) and with the vibration switch remaining off. At the end of 1 minute, the device was removed and immediately afterward the vein entry procedure was implemented.

Nonintervention Control Group. For the nonintervention control group, no intervention was implemented before the procedure, and the standard vein entry procedure was used.

Immediately after the vein entry procedure had been completed, a researcher who was unaware of the method used assessed the pain levels of the participants in all groups with the visual analog scale and satisfaction with the phlebotomy procedure with the Phlebotomy Satisfaction Evaluation Scale. While the study was being conducted, the necessary steps were taken to prevent the groups from affecting each other: Participants did not see the methods applied to others, and they did not sit next to each other during the procedure.

Data Analysis

Analysis of the data collected in the study was performed with the IBM statistics program SPSS 22.0 (IBM Corp., Armonk, NY, USA). Numerical data were examined for normal distribution by the Shapiro-Wilk test. Distributions of identifying data were given as means and standard deviation. The Kruskal-Wallis test was used to compare more than two groups. The level of statistical significance was determined as $p < .05$.

RESULTS

Findings relating to the identifying characteristics of the groups included in the study are given in Table 2. Results of statistical analysis to determine the homogeneity of the groups indicated no statistically significant difference between the mean age and body mass index of the members of the three groups (Table 3).

Table 2.
Distribution of Identifying Characteristics of Groups

	Experimental Group	Placebo Control Group	Nonintervention Control Group	Test Value
Mean age (Years)	34.40 ± 1.7	36.26 ± 1.8	35.83 ± 1.7	K-W = 0.639 p = .726
Mean BMI	25.79 ± 0.6	25.77 ± 0.5	25.58 ± 0.5	K-W = 0.029 p = .986

BMI = body mass index; K-W = Kruskal-Wallis test.

Mean pain scores were found to be 20.93 ± 15.1 for the members of the experimental group, 35.40 ± 11.4 for the placebo control group, and 35.23 ± 19.3 for the nonintervention control group. Results of statistical analysis indicated a statistically significant difference between the mean pain scores of the members of the experimental and control groups (Table 3). The results of advanced analysis indicated a significant difference between the mean pain scores of the members of the experimental group and members of the placebo control and nonintervention control groups (p = .013 and p = .014 respectively) but no significant difference between the mean pain scores of members of the placebo control and nonintervention control groups (p = .999).

According to the findings of the study, the mean phlebotomy satisfaction score was 76.00 ± 23.7 for the experimental group, 61.90 ± 25.5 for the placebo control group, and 55.26 ± 34.8 for the nonintervention control group. A statistically significant difference was found between the mean scores for phlebotomy satisfaction of members of the experimental and control groups (Table 3). The results of advanced analysis indicated a significant difference between the mean phlebotomy satisfaction scores of members of the experimental group and members of the nonintervention control group (p = .016) but no significant difference between the mean phlebotomy satisfaction scores of the experimental group and the placebo control group or between the placebo control group and the nonintervention control group (p = .140 and p = .640, respectively).

DISCUSSION

The International Association for the Study of Pain describes pain as "an unpleasant sensory and emotional state and behavior that originates from any region of the body, depends on existing or possible tissue damage, or can be identified with this damage, and is affected by past experiences of the individual" (Abd El-Gawad & Elsayed, 2015; Aydin et al., 2016).

Individuals often experience pain related to the phlebotomy procedure (Ialongo & Bernardini, 2016; Whelan et al., 2014). The American Society for Pain Management Nursing recommends optimal pain control before and during painful interventions (Czarnecki et al., 2011). For this reason, pharmacological and nonpharmacological interventions are used in the optimal control of pain (Schechter et al., 2007; Taddio et al., 2010).

It was found as a result of this study that in adult male blood donors the mean scores for the pain caused by the intervention in the experimental group, to whom cold and vibration were applied during the phlebotomy procedure using the Buzzy device, were significantly lower than those of the control group. In a study by Baxter et al. (2011) it was found that the pain scores after the application to child patients of cold and vibration by means of Buzzy during intravenous

Table 3.
Distribution of Pain Score and Mean Satisfaction of Experimental and Control Groups

	Experimental Group	Placebo Control Group	Nonintervention Control Group	Test Value
Procedural pain score	20.93 ± 15.1	35.40 ± 11.4	35.23 ± 19.3	K-W = 10.883 p = .004
Mean satisfaction score	76.00 ± 23.7	61.90 ± 25.5	55.26 ± 34.8	K-W = 6.948 p = .031

K-W = Kruskal-Wallis test.

interventions was lower than a group that used a vapocoolant spray. Schreiber et al. (2016) reported that the application of cold and vibration by means of Buzzy reduced intervention pain in child patients (Schreiber et al., 2016). In studies of child patients by Inal and Kelleci (2012) and Whelan et al. (2014) relating to this topic, similar results were obtained. The findings of the present study support the results of these studies. However, these studies were all performed with groups of sick children. Our study was different in that only healthy adult men were included.

Different from these studies, Baxter et al. (2009) carried out a randomized controlled study investigating the effects of Buzzy and a vapocoolant spray on pain during vein entry in adult patients, and no significant difference was found between the mean pain scores of the groups using the two methods. The findings of this study are different from ours. This difference may stem from the use of the vapocoolant spray, which is a local anesthetic, in the control group, and may have reduced the perception of pain.

Baxter et al. (2009) stated that Buzzy involved three nonpharmacological mechanisms of pain relief. These were gate control theory, descending noxious inhibitory control, and distraction. According to gate control theory, as developed by Melzack and Wall in 1965, pain stimuli are carried by small-diameter fibers. Large-diameter fibers close the gate to stimuli carried by the small-diameter fibers. When the gate is open, the sense of pain and the resulting stimuli reach the level of consciousness, and pain is felt. If the gate is closed, it is asserted that the stimuli will not reach the level of consciousness and pain will not be felt (Baxter et al., 2011; Melzack & Wall, 1965). According to Nahra and Plaghki, "C fibers, which transmit slow pain and noxious temperature information, can also 'close' the fast pain gate. The mechanism seems to be mediated via intraneuron inhibition; noxious cold increases the pain threshold and decreases almost all $A_{\&}$ transmission when applied close to the nociception source" (Nahra & Plaghki, 2005). Finally, the technique of distraction depends on diverting the attention away from the pain, which may be helpful in pain management because, it is believed, an individual can only be conscious of one stimulus at a time (Demir, 2016).

In addition, it has been found in studies in the literature that the techniques of cold application (Aygün et al., 2013; Kiran et al., 2013), vibration application (Hollins, Roy, & Crane, 2003; Whelan et al., 2014) and directing the attention elsewhere (Abd El-Gawad & Elsayed, 2015; Aydin et al., 2016; Gupta et al., 2006; Usichenko, Pavlovic, Foellner, & Wendt, 2004) are effective in reducing the pain of invasive procedures according to the gate control theory. It was found in the present study that pain score averages were lower in the experimental group, where the techniques of cold, vibration, and distraction were used together by means of the Buzzy device. Seen from this viewpoint also, our study findings were found to support the literature.

In the present day, patient satisfaction is an important measure in evaluating the quality of service given (Uzun, 2001; Yıldız & Erdoğmuş, 2004). Reducing the feeling of pain experienced during phlebotomy also affects patient satisfaction. The findings of this study indicated that the mean phlebotomy satisfaction scores of members of the experimental group were higher than those of the members of the control groups. This result is thought to be because those in the experimental group felt less pain during the procedure than control group members.

Limitations

This study had a number of limitations. Principally, because the study was conducted with healthy men, its results cannot be generalized. Another important limitation is that the measurement of pain is participative. Lack of investigation of the effect of gender on pain is a further limitation.

CONCLUSION

It was found as a result of this study that the Buzzy device was an effective method of reducing the pain of phlebotomy and increasing phlebotomy satisfaction in healthy adult men.

Acknowledgments

The authors are grateful to all the participants in the study. This research is supported by Bursa Uludağ University Scientific Research Projects Management Unit (HDP(T)-2015-25).

REFERENCES

Abd El-Gawad, S. M., & Elsayed, L. A. (2015). Effect of interactive distraction versus cutaneous stimulation for venipuncture pain relief in school age children. *Journal of Nursing Education and Practice, 5*, 32–40.

Aydin, D., Şahiner, N. C., & Çiftçi, E. K. (2016). Comparison of the effectiveness of three different methods in decreasing pain during venipuncture in children: ball squeezing, balloon inflating and distraction cards. *Journal of Clinical Nursing, 25*(15-16), 2328–2335.

Aygün, H., Armağan, E., Özdemir, F., Köse, A., Selimoğlu, K., Köksal, Ö., & Aydin, S. A. (2013). Comparison of ice and lidocaine- prilocaine cream mixture in the reduction of pain during peripheral intravenous cannulation in emergency department patients. *The Journal of Academic Emergency Medicine, 12*, 27–29, (in Turkish).

Baxter, A. L., Cohen, L. L., McElvery, H. L., Lawson, M. L., & von Baeyer, C. L. (2011). An integration of vibration and cold relieves venipuncture pain in a pediatric emergency department. *Pediatric Emergency Care, 27*(12), 1151–1156.

Baxter, A. L., Leong, T., & Matthew, B. (2009). External thermomechanical stimulation versus vapocoolant for adult venipuncture pain: pilot data on a novel device. *The Clinical Journal of Pain, 25*(8), 705–710.

Chung, J. W. Y., Ng, W. M., & Wong, T. K. S. (2002). An experimental study on the use of manual pressure to reduce pain in intramuscular injections. *Journal of Clinical Nursing, 11*(4), 457–461.

Czarnecki, M. L., Turner, H., Collins, P. M., Doellman, D., Wrona, S., & Reynolds, J. (2011). Procedural pain management: a position statement with clinical practice recommendations. *Pain Management Nursing, 12*(2), 95–111.

Dawson, A., & List, T. (2009). Comparison of pain thresholds and pain tolerance levels between Middle Easterners and Swedes and between genders. *Journal of Oral Rehabilitation, 36*(4), 271–278.

Demir, Y. (2016). Nonpharmacological therapies in pain management. In G. Racz (Ed.), *Pain Management: Current Issues and Opinion*. Retrieved from. http://cdn.intechopen.com/pdfs/26152/InTechNon_pharmacological_therapies_in_pain_management.pdf. Accessed September 2016.

Dutt-Gupta, J., Brown, T., & Mycama, M. (2007). Effect of communication on pain during intravenous cannulation: a randomized control trial. *British Journal of Anesthesia, 99*(6), 871–879.

Eichenfield, L. F., Funk, A., Fallon-Friedlander, S., & Cunningham, B. B. (2002). A clinical study to evaluate the efficacy of ELA-Max (4% liposomal lidocaine) as compared with eutectic mixture of local anesthetics cream for pain reduction of venipuncture in children. *Pediatrics, 109*(6), 1093–1099.

Farion, K. J., Splinter, K. L., Newhook, K., Gaboury, I., & Splinter, W. M. (2008). The effect of vapocoolant spray on pain due to intravenous cannulation in children: a randomized controlled trial. *Canadian Medical Association Journal, 179*(1), 31–36.

Fein, J. A., & Gorelick, M. H. (2006). The decision to use topical anesthetic for intravenous insertion in the pediatric emergency department. *Academic Emergency Medicine, 13*(3), 264–268.

Gupta, D., Agarwal, A., Dhiraaj, S., Tandon, M., Kumar, M., & Singh, R. S. (2006). An evaluation of efficacy of balloon inflation on venous cannulation pain in children: a prospective, randomized, controlled study. *Anesthesia and Analgesia, 102*(5), 1372–1375.

Hollins, M., Roy, E. A., & Crane, S. A. (2003). Vibratory antinociception: effects of vibration amplitude and frequency. *The Journal of Pain, 4*(7), 381–391.

Home/Buzzy Helps. (2016). Available from https://buzzyhelps.com. Accessed September 2016.

Ialongo, C., & Bernardini, S. (2016). Phlebotomy, a bridge between laboratory and patient. *Biochemia Medica, 26*(1), 17–33.

Inal, S., & Kelleci, M. (2012). Relief of pain during blood specimen collection in pediatric patients. *MCN, the American Journal of Maternal Child Nursing, 37*(5), 339–345.

Jimenez, N., Bradford, H., Seidel, K. D., Sousa, M., & Lynn, A. M. (2006). A comparison of a needle-free injection system for local anesthesia versus EMLA for intravenous catheter insertion in the pediatric patient. *Anesthia and Analgesia, 102*(5), 411–414.

Kahl, C., & Cleland, J. A. (2005). Visual analogue scale, numeric pain rating scale and the McGill Pain Questionnaire: an overview of psychometric properties. *Physical Therapy Reviews, 10*(2), 123–128.

Kiran, N., Kaur, S., & Marwaha, R. K. (2013). Effect of ice pack application at the site prior to venipuncture on intensity of pain among children. *Nursing and Midwifery Research Journal, 9*, 160–167.

Leahy, S., Kennedy, R. M., Hesselgrave, J., Gurwitch, K., Barkey, M., & Millar, T. F. (2008). On the front lines: lessons learned in pediatric hospitals. *Pediatrics, 122*(Suppl 3), 161–170.

MacLaren, J. E., & Cohen, L. (2007). Interventions for paediatric procedure-related pain in primary care. *Paediatrics and Child Health, 12*(2), 111–116.

Melzack, R., & Wall, P. D. (1965). Pain mechanisms: a new theory. *Science, 150*, 971–979.

Nahra, H., & Plaghki, L. (2005). Innocuous skin cooling modulates perception and neurophysiological correlates of brief CO2 laser stimuli in humans. *European Journal of Pain, 9*(5), 521–530.

Pasero, C. (2003). Pain in the emergency department. *The American Journal of Nursing, 103*(1), 73–74.

Pershad, J., Steinberg, S. C., & Waters, T. M. (2008). Cost-effectiveness analysis of anesthetic agents during peripheral intravenous cannulation in the pediatric emergency department. *Archive of Pediatric and Adolescent Medicine, 162*(10), 952–961.

Rogers, T. L., & Ostrow, C. L. (2004). The use of EMLA cream to decrease venipuncture pain in children. *Journal of Pediatric Nursing, 19*(1), 33–39.

Schechter, N. L., Zempsky, W. T., Cohen, L. L., McGrath, P. J., McMurtry, C. M., & Bright, N. S. (2007). Pain reduction during pediatric immunizations: evidence-based review and recommendations. *Pediatrics, 119*(5), 1184–1198.

Schreiber, S., Cozzi, G., Rutigliano, R., Assandro, P., Tubaro, M., Wiel, L. C., Ronfani, L., & Barb, E. (2016). Analgesia by cooling vibration during venipuncture in children with cognitive impairment. *Acta Paediatrica, 105*(1), e12–e16.

Sethna, N. V., Verghese, S. T., Hannallah, R. S., Solodiuk, J. C., Zurakowski, D., & Berde, C. B. (2005). A randomized controlled trial to evaluate S-Caine Patch for reducing pain associated with vascular access in children. *Anesthesiology, 102*, 403–408.

Sinha, P. K., Tandon, A. A., & Singh, S. D. (2005). Evaluating the efficacy of Valsalva maneuver on venous cannulation pain: A prospective randomized study. *Anaesthesia and Analgesia, 101*(4), 1230–1232.

Taddio, A., Appleton, M., Bortolussi, R., Chambers, C., Dubey, V., Halperin, S., Hanrahan, A., Ipp, M., Lockett, D., MacDonald, N., Midmer, D., Mousmanis, P., Palda, V., Pielak, K., Riddell, R. P., Rieder, M., Scott, J., & Shah, V. (2010). Reducing the pain of childhood vaccination: an evidence-based clinical practice guideline. *Canadian Medical Association Journal, 182*(18), 43–55.

Usichenko, T. I., Pavlovic, D., Foellner, S., & Wendt, M. (2004). Reducing venipuncture pain by a cough trick: a randomized crossover volunteer study. *Anesthesia and Analgesia, 98*(2), 343–345.

Uzun, Ö. (2001). Patient satisfaction with nursing care at a university hospital in Turkey. *Journal of Nursing Care Quality, 16*(1), 24–33.

Whelan, H. M., Kunselman, A. R., Thomas, N. J., Moore, J., & Tamburro, R. F. (2014). The impact of a locally applied vibrating device on outpatient venipuncture in children. *Clinical Pediatrics, 53*(12), 1189–1195.

Yıldız, Z., & Erdoğmuş, Ş. (2004). Measuring patient satisfaction of the quality of health care: a study of hospitals in Turkey. *Journal of Medical Systems, 28*, 581–589.

Zempsky, W. T., Bean-Lijewski, J., Kauffman, R. E., Koh, J. L., Malviya, S. V., Rose, J. B., Richards, P. T., & Gennevois, D. J. (2008). Needle-free powder lidocaine delivery system provides rapid effective analgesia for venipuncture or cannulation pain in children: Randomized, double-blind comparison of venipuncture and venous cannulation pain after fast-onset needle-free powder lidocaine or placebo treatment trial. *Pediatrics, 121*(5), 979–987.

CE

1.5
ANCC
Contact
Hours

Impact of Music Therapy on Hospitalized Patients Post-Elective Orthopaedic Surgery

A Randomized Controlled Trial

Lisa M. Gallagher ▼ Vickie Gardner ▼ Debbie Bates ▼ Shelley Mason ▼ Jeanine Nemecek ▼
Jacquelyn Baker DiFiore ▼ James Bena ▼ Manshi Li ▼ Francois Bethoux

BACKGROUND: Music therapy (MT) research has demonstrated positive effects on fatigue, depressed mood, anxiety, and pain in perioperative care areas. However, there has been limited research on the effects of MT for surgical patients on orthopaedic units.

PURPOSE: The purpose of this study was to understand the impact of MT sessions on post-elective orthopaedic surgery patients' pain, mood, nausea, anxiety, use of narcotics and antiemetics, and length of stay.

METHODS: This was a randomized controlled study with an experimental arm (MT sessions) and a control arm (standard medical care). Patients received MT within 24 hours of admission to the unit, as well as every day of their stay. Same-day pre- and postdata were collected 30 minutes apart for both arms, including patient self-reported mood, pain, anxiety, and nausea. Use of medications and length of stay were gleaned from the electronic medical record.

RESULTS: Data were obtained for 163 patients, age 60.5 ± 11.1 years, 56% of whom were male. Joints targeted by surgeries were hips (54%), knees (42%), and shoulders (4%). There were significantly greater changes favoring the MT group on Day 1 (pain, anxiety, and mood), Day 2 (pain, anxiety, mood, and nausea), and Day 3 (pain, anxiety, and mood). Among participants with a pre–pain score of 2 or more on Day 1, a decrease of at least 2 points was noted in 36% of the MT group and 10% of the control group ($P <$ .001). Overall, 73% of MT patients versus 41% of control patients reported improved pain ($P <$ .001). No significant between-group differences in medications or length of stay were noted.

CONCLUSIONS: We observed greater same-day improvements of pain, emotional status, and nausea with MT sessions, compared to usual care, in patients hospitalized after elective orthopaedic surgeries. Effects on narcotic and antiemetic usage, as well as length of stay, were not observed. More research needs to be conducted to better understand the benefits of MT pre- and post-elective orthopaedic surgery.

Introduction

Total knee arthroplasties (TKAs) or total knee replacements (TKRs) are on the increase, with more than 600,000 surgeries completed annually in the United States (American Academy of Orthopaedic Surgeons, 2016). The number of hip arthroplasties (THA) has also increased, with more than 300,000 surgeries performed annually (Dotinga, 2015). These surgeries are generally performed for patients suffering from osteoarthritis, osteoarthrosis, osteonecrosis, or rheumatoid arthritis (Parker, 2011).

Patients recovering from TKAs, Birmingham hip resurfacing (BHR) or THA, and shoulder replacements or rotator cuff surgeries can be negatively impacted and their recovery complicated by experiencing severe pain, as well as a fear of anesthesia and discomfort (Antall & Kresevic, 2004; Eisenman & Cohen, 1995; Engwall & Duppils, 2009; Ignacio, Fai, Hui, Marie, & Goy, 2012;

Lisa M. Gallagher, MA, MT-BC, Research Program Manager, Arts and Medicine Institute, and Cleveland Clinic, Cleveland, OH.

Vickie Gardner, MSN, RN, Nurse Manager, Euclid Hospital, Cleveland, OH.

Debbie Bates, MMT, MT-BC, Music Therapy Manager, Arts and Medicine Institute, and Cleveland Clinic, Cleveland, OH.

Shelley Mason, CNP, RN, Certified Nurse Practitioner, Cleveland Clinic, Euclid Hospital, Cleveland, OH.

Jeanine Nemecek, MSN, RN, Director of Surgical Services, Euclid Hospital, Cleveland, OH.

Jacquelyn Baker DiFiore, RN, Cleveland Clinic, Euclid Hospital, Cleveland, OH; currently, Chief Nursing Officer, Ashtabula County Medical Center.

James Bena, MS, Lead Biostatistician, Cleveland Clinic, Quantitative Health Sciences, Cleveland, OH.

Manshi Li, MS, Statistical Programmer I, Cleveland Clinic, Quantitative Health Sciences, Cleveland, OH.

Francois Bethoux, MD, Cleveland Clinic Medical Director, Arts and Medicine Institute; Director, Mellen Center Rehabilitation Services, Neurological Institute.

The authors and planners do not declare any conflicts of interest, financial or otherwise.

DOI: 10.1097/NOR.0000000000000432

© 2018 by National Association of Orthopaedic Nurses

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

Lukas, 2004; Simcock et al., 2008). Attempting to decrease this pain can also help to decrease the patient's length of stay and recovery time while promoting patient satisfaction (Antall & Kresevic, 2004; Easter et al., 2010; Lin, 2011; Simcock et al., 2008). Utilizing a variety of nonpharmacologic techniques such as music, progressive muscle relaxation, massage, Reiki, art, imagery, meditation, relaxation therapy, and rhythmic breathing can be effective in addressing the physical and psychological aspects of pain and anxiety (Cepeda, Carr, Lau, & Alvarez, 2006; Engwall & Duppils, 2009; Evans, 2002; Horrigan, 2013; Lin, 2011; Parker, 2011; Pellino et al., 2005). These techniques are also beneficial in decreasing anxiety and pain, while increasing coping skills in patients post-TKA (Parker, 2011; Pellino et al., 2005).

Using music, guided imagery, and other complementary therapies may be helpful in not only decreasing pain but also decreasing the risk of sedation and confusion and other common side effects of pain medication, especially in elderly patients (Antall & Kresevic, 2004; Evans, 2002). They also may help to increase comfort and the ability to participate in physical therapy sooner, while decreasing the risk of complications, (Antall & Kresevic, 2004). Recorded music interventions have also been found to decrease respirations, heart rate, depression, anxiety, and the overall emotional burden in hospitalized patients (Phipps, Carroll, & Tsiantoulas, 2010). Using music in a therapeutic manner can be especially helpful in decreasing the anxiety and stress associated with being hospitalized, and it has been suggested that it is also a cost-effective, minimally invasive intervention for addressing pain, anxiety, and coping (Easter et al., 2010; Eckhouse et al., 2014; Gallagher, Lagman, Walsh, Davis, & Legrand, 2006; Lukas, 2004).

The use of music prior to various outpatient surgeries improved vital signs and decreased preoperative anxiety (Ni, Tsai, Lee, Kao, & Chen, 2011). Playing music during and after surgery has resulted in decreased pain, anxiety, and stress for patients, as well as the use of sedation and analgesia during surgery and narcotics after surgery (Eisenman & Cohen, 1995; Engwall & Duppils, 2009; Evans, 2002; Nilsson, Rawal, & Unosson, 2003; Simcock et al., 2008). The effect of music was positive when utilizing both general anesthesia and regional anesthesia (Eisenman & Cohen, 1995; Nilsson et al., 2003; Simcock et al., 2008). Music has also been found to be effective during and after surgery, as well as prior to ambulation in decreasing postsurgical pain (Good et al., 2001; Nilsson, Rawal, & Unosson, 2003).

Common music therapy (MT) interventions/goals to address pre- and postoperative anxiety and pain have included active music making and listening to live or recorded music (Bradt, Dileo, & Shim, 2013; MacDonald et al., 2003). Several studies have utilized patient-preferred music and found it to be effective, while other studies have utilized researcher-chosen music (Bradt et al., 2013; Eisenman & Cohen, 1995; Lin, 2011; MacDonald et al., 2003).

Simcock et al. (2008) asked patients scheduled for TKA or THA to choose three CDs that would be played throughout their surgeries. Results demonstrated decreased anxiety and pain when music was used during and after surgery, as well as decreased physiologic stress and stress hormones (Simcock et al., 2008). Music-focused relaxation has also been found to be helpful in decreasing the anxiety of patients receiving orthopaedic care (Eckhouse et al., 2014). Allred, Byers, and Sole (2010) investigated the use of music listening before and after the first ambulation post-TKA. They found statistically significant lower anxiety and pain over time and suggested that listening to music could help limit the adverse effects of opioid medications (Allred et al., 2010). Lin (2011) found that relaxation therapy was helpful in decreasing pain severity while promoting sleep and relaxation in patients who received total joint replacements. Listening to music before, during, and after orthopaedic surgery resulted in improved anxiety and pain management (Lukas, 2004).

While several studies have been published regarding the use of music, there are limited evidence-based studies of the utilization of MT with patients post-elective orthopaedic surgery. MT is the utilization of music interventions by a board-certified music therapist within a therapeutic relationship with a patient to accomplish individualized and specific goals (American Music Therapy Association, 2005). The main goal of this study was to understand the impact of MT sessions on patients' experiences post-elective orthopaedic surgery compared to usual care alone. It was hypothesized that MT would have a positive effect on patients' self-reported scores of pain, anxiety, mood, and nausea. It was also hypothesized that participation in MT would have a positive effect on length of stay and the use of narcotics and antiemetics.

Methods

Subjects

This study was approved, and a waiver of written informed consent was granted by the institutional review board at The Cleveland Clinic. All procedures followed were in accordance with the Helsinki Declaration of 1964 and its later amendments or comparable ethical standards (World Medical Association, 2016). Patients who were to receive elective orthopaedic surgery were recruited by the Pre-Admission Testing nurses (PATs) when they arrived at the hospital for their preadmission testing. Only those patients who were to be admitted to the orthopaedic unit at Euclid Hospital postsurgery were eligible for the study. Subjects also had to be at least 18 years old; cognitively able to consent to participate; and able to speak, read, and write English. Patients who ended up sharing a semiprivate room with another patient already participating in the study were excluded from participation. Nonelective orthopaedic surgical patients, those who did not have their testing done at Euclid Hospital, those whose surgeries were moved to another hospital, and patients on isolation or contact precautions were also excluded.

Informed consent was obtained by the PATs during preadmission testing, or by the Surgery Center nurse or a Research Assistant on the day of surgery (for those patients who were not ready to consent during preadmission testing). Patients who agreed to participate were randomly assigned to either the MT (experimental)

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

arm or the control arm of the study through the use of a computerized block randomization table. Block randomization was utilized to maintain balance between the two arms in the study, and it was based on the subject's data and time of consent. Allocation concealment was utilized, and the information was kept on a secure password protected drive. Only the PI on the study knew the order of randomization. Patients did not know to which group they had been assigned until after their surgeries, and study staff did not know until they received the schedule at the beginning of each week.

Once the patients reached the 27-bed unit, after surgery it was determined whether they were eligible to remain in the study. Patients participating in the study were assigned a private room whenever possible, or they were the only patients enrolled in the study in a semiprivate room. If the patient was assigned to the room of another patient who was already participating in the study, the newly arrived patient was excluded. If two patients from the MT arm were moved into a semiprivate room, the patient who was involved in the study the least amount of time was excluded unless the music therapist was able to see one patient while the other was out of the room and vice versa.

SESSIONS/PROCEDURES

Patients enrolled in the experimental arm of the study received an MT session every day of their hospitalization, with the first session occurring within the first 24 hours of admission to the orthopaedic unit or approved overflow unit. This was based on the time of the nurse's first note when admitting the patient to the unit. If the first session did not occur within this timeframe, the patient was removed from the study. If a patient declined to participate in an MT session, he or she was offered a session at another time that day. If a patient declined twice in 1 day, the music therapist returned the following day. This process continued throughout the patient's admission. The total number of MT sessions during each patient's hospitalization was recorded.

The MT sessions involved assessment and MT interventions. The board-certified music therapists collaborated with the patients to determine individualized goals for the session and also provided individualized interventions to address these goals. The patient had input regarding the types of interventions he or she wished to engage in during the session. MT methods used were receptive/listening, re-creative, composition, or improvization (Bruscia, 2014). These included such things as the patient choosing songs and then listening to live music performed by the music therapist; or the patient engaging in the session by playing or improvising on instruments, singing, discussing song lyrics, reminiscing/sharing memories, and/or participating in music-assisted relaxation techniques such as breathing, progressive muscle relaxation, and imagery. Multiple goals were addressed in some sessions, and multiple interventions could also be utilized within one session. MT sessions lasted approximately 30 minutes.

Patients in the control arm did not receive MT sessions; however, they were provided with usual care post-elective orthopaedic surgery. A research assistant or an investigator on the study, other than the music therapists who provided MT sessions, visited these patients daily and asked them to complete the mood, pain, nausea, and anxiety instruments. The first set of data was collected, and then the second set of data was collected 30 minutes later (the same length of time as a typical MT session).

DATA COLLECTED

Patient information included age, gender, race/ethnicity, diagnosis, type of surgery, comorbidities, and length of stay. Symptom severity measures were administered before and after MT interventions (pain, anxiety, nausea, and mood) by a research assistant or an investigator on the study, other than the music therapists who provided the MT sessions, via the use of an iPad. These individuals were not blinded to the intervention. The iPad was connected with the Cleveland Clinic secured network and the data were directly uploaded into a REDCap database. REDCap is a secure, web application designed to support data capture for research studies, providing user-friendly web-based case report forms, real-time data entry validation (e.g., for data types and range checks), audit trails, and a de-identified data export mechanism to common statistical packages such SPSS, SAS, Stata, and R/S-Plus (Harris et al., 2009). The system is protected behind a log-in and Secure Sockets Layer (SSL) encryption.

Four patient-reported variables were scored before and after the 30-minute time period, three on a 0-10 point numeric rating scale (pain, anxiety, and nausea), and mood on a 0-4 point scale based on the Rogers Happy/Sad Faces Assessment Tool (Rogers, 1981). For all of these scales, higher scores represent worse results.

Data collected from the inpatient pharmacy included for narcotic and antiemetic medications, name and dosage of medications used during hospitalization, the number of tablets and doses given, and whether or not the patient was on the medications prior to hospitalization. These data were matched with the data in the electronic medical record as to the date, time, and dosages of narcotics and antiemetics given throughout the patient's hospitalization. All medication-related data were entered into the REDCap database

STATISTICAL CONSIDERATIONS

A G*Power analysis for medium effect size indicated that 79 patients were needed in each arm to achieve the needed power. Allowing for attrition of approximately 20% in each arm, it was estimated that a sample of 200 would allow for sufficient power for statistical analysis. Therefore, 200 patients were recruited for the study.

STATISTICAL METHODS

Categorical factors were summarized using frequencies and percentages, while continuous measures were described using means and standard deviations for normally distributed variables and medians and quartiles for nonnormally distributed measures. For comparisons of patient characteristics, two-sample *t* tests, Pearson's chi-square tests, and Fisher's exact tests were

© 2018 by National Association of Orthopaedic Nurses

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

used as appropriate. For comparisons of groups on outcomes within day, similar tests were used. To compare groups on patient-level outcomes including medication use and length of stay, linear models were fit. To compare groups on changes in scores across days, linear mixed effect models were fit. An interaction between day and group was introduced, but was not significant in any of the models, and was removed. Mean differences between groups across days with 95% confidence intervals were calculated with as were mean differences. For the success measure of improvement or maintenance of a score of 0, logistic regression models with generalized estimating equations were fit. As above, an interaction between group and day was considered, but found to be non-significant and removed. Results from the logistic regression model are presented as predicted probabilities, and comparisons between groups are presented as odds ratios with 95% confidence intervals. Analyses were performed using SAS software (version 9.4; SAS, Cary, NC).

Results

PATIENT SAMPLE

A total of 292 patients were approached, 200 were enrolled in the study, 91 did not consent to participate, and one was ineligible. Reasons for exclusion are listed in Figure 1. Data were available for 164 patients. One patient who failed to complete any information for the study was excluded, leaving 163 patients: 79 in the control group and 84 in the experimental group (see Figure 1).

Table 1 shows descriptive summaries of patient demographic factors. No significant between-group differences at baseline were observed. For sake of ease, type of surgery was listed as knees, hips, or shoulders (see Table 1). However, there were different types of surgeries performed for each of these. For instance, knees included TKAs, TKRs, partial knee replacements, and conversion of unilateral to TKA. Hips included BHR, arthroscopy, hip revisions, total hip replacements, and a left revision of the femoral component-Birmingham. Shoulders included total shoulder arthroscopies and total shoulder replacements.

MUSIC THERAPY SESSIONS

Twenty-three different goals were addressed during the MT sessions, with multiple goals often addressed during the same session. Goals were individualized for each patient based on patients' needs, with the most frequently used goals addressing relaxation (27%), pain (22%), self/emotional expression (19%), anxiety (9%), support (4%), mood (4%), nausea (3%), coping (2%), reminiscence/memory sharing (2%), and spirituality (1%). All remaining goals were addressed in less than 1% of sessions. Eighteen different interventions were used during the sessions, with multiple interventions used in the majority.

The most frequent interventions utilized were music listening—live or recorded (30%); support/validation (24%); song choices (13%); instrument playing and/or instrumental improvization (7%); lyric and/or music discussion/analysis (7%), singing (6%); music-assisted relaxation—rhythmic breathing, progressive muscle relaxation and/or imagery (5%), and reminiscence/memory

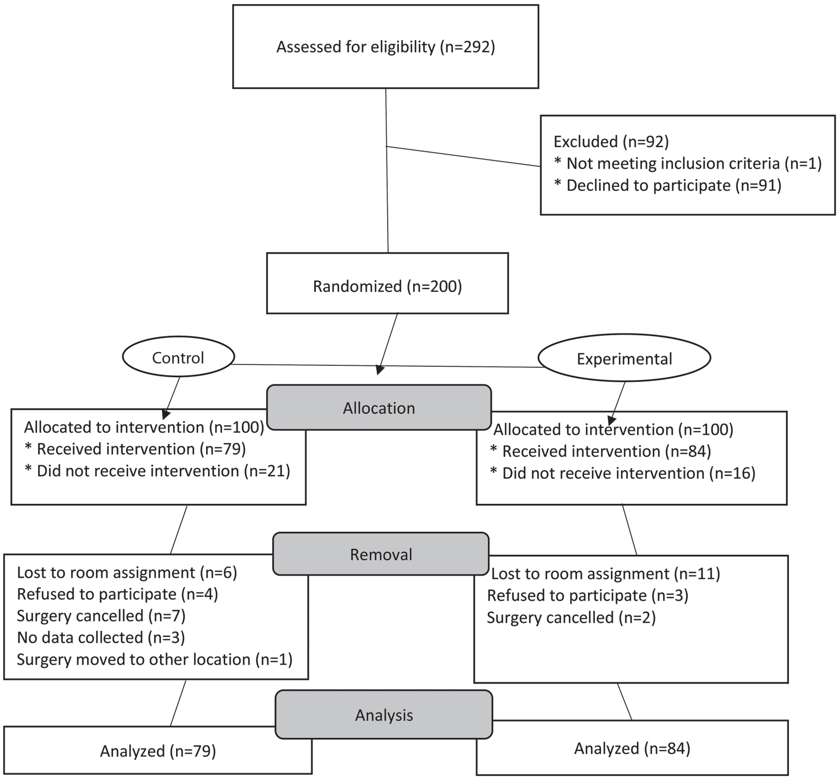

Assessed for eligibility (n=292)
Excluded (n=92)
* Not meeting inclusion criteria (n=1)
* Declined to participate (n=91)
Randomized (n=200)
Control
Experimental
Allocation
Allocated to intervention (n=100)
* Received intervention (n=79)
* Did not receive intervention (n=21)
Allocated to intervention (n=100)
* Received intervention (n=84)
* Did not receive intervention (n=16)
Removal
Lost to room assignment (n=6)
Refused to participate (n=4)
Surgery cancelled (n=7)
No data collected (n=3)
Surgery moved to other location (n=1)
Lost to room assignment (n=11)
Refused to participate (n=3)
Surgery cancelled (n=2)
Analysis
Analyzed (n=79)
Analyzed (n=84)

FIGURE 1. Enrollment flow diagram.

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

Table 1. Patient demographics

Factor	Total (*N* = 163)	Control Group (*n* = 79)	Experimental Group (*n* = 84)
Patient age	60.5 ± 11.1	59.9 ± 11.6	61.1 ± 10.6
Gender			
Male	92 (56.4)	44 (55.7)	48 (57.1)
Female	71 (43.6)	35 (44.3)	36 (42.9)
Race			
Missing	10 (6.1)	4 (5.1)	6 (7.1)
Hispanic	1 (0.61)	1 (1.3)	0 (0.0)
Non-Hispanic	152 (93.3)	74 (93.7)	78 (92.9)
Primary diagnosis			
Osteoarthritis	83 (50.9)	37 (46.8)	46 (54.8)
Osteoarthrosis	67 (41.1)	34 (43.0)	33 (39.3)
Avascular necrosis	2 (1.2)	2 (2.5)	0 (0.0)
Mechanical complication	4 (2.5)	2 (2.5)	2 (2.4)
Arthritis	7 (4.3)	4 (5.1)	3 (3.6)
Type of surgery			
Knee	69 (42.3)	34 (43.0)	35 (41.7)
Hip	88 (54.0)	41 (51.9)	47 (56.0)
Shoulder	6 (3.7)	4 (5.1)	2 (2.4)

Note. Values are presented as mean ± SD or *N* (column %).

sharing (4%). All other interventions were utilized in 1% of sessions or less. When looking at the interventions in terms of Bruscia's four methods, they break down as follows: receptive/listening (61%), recreative (12%), improvisation (1%), and composition (0%). In addition, 26% of interventions included, for example, verbal processing, therapeutic use of self, support and validation, and affirmation.

Following evaluation of length of stay, adequate data existed only through day 3, so later follow-up measurements were truncated at this point. Minimal clinically important differences (MCIDs) of at least 2 points on pain, nausea, and anxiety, and at least 1 in mood were calculated among patients with baseline levels that were at least as large as the MCID. Thus, for example, patients with a pain score of 1 pretherapy were not included in the MCID calculation because a 2-point decrease was not possible. Success was also evaluated using improvement from baseline or maintaining a score of 0 as a success.

OUTCOMES

Table 2 shows the distribution of follow-up days by group. By Day 3, about 60% of patients were lost to follow-up due to discharge. Table 3 shows summaries of patient-level outcomes. Summaries of total antiemetic use and length of stay were very similar across groups. For total narcotic use, although median levels were higher in the experimental group, quartiles were quite similar across groups.

Table 2. Follow-up Summary Statistics

Factor	Total (*N* = 163)	Control Group (*n* = 79)	Experimental Group (*n* = 84)
Day 1	163 (100)	79 (100)	84 (100)
Day 2	123 (75.5)	61 (77.2)	62 (73.8)
Day 3	70 (42.9)	37 (46.8)	33 (39.3)

Note. Values are presented as *N* (column %).

Tables 4A, 4B, and 4C show the outcomes by day. At Day 1, significantly better outcomes were seen in pain, anxiety, and mood among the experimental group compared with the control group. This was a consistent finding based on change scores, percentage of patients reaching MCID, or percentage of patients exhibiting improvement. At Day 2, all measures showed significantly greater changes in the experimental group, while at Day 3, all measures except nausea saw greater improvements in the experimental group.

Table 5 shows the comparisons of raw survey score changes pre- to post across all days. Those in the control group saw on average a 0.25 point improvement in their pain score, while those in the experimental group had their pain improve on average by more than 1.25 points (mean difference 1.03, $p < .001$). Anxiety and mood saw similar mean changes in terms of significance, while the change in nausea was smaller, yet still significant at the 0.05 level ($p = .044$). Analysis of the improvement score for each survey yielded similar findings. For pain, 41% of control patients reported an improvement in pain pre- to posttherapy, as compared with 73% of experimental patients. This corresponds to an odds ratio of 3.8 ($p < .001$). As above, significant differences were also seen in pain, anxiety, and nausea. The patient-level outcomes for medication and length of stay were also identified, but no significant differences between groups were observed.

Discussion

MT was found to consistently produce immediate improvement of pain and anxiety, and in some cases nausea, at a statistically significant level compared to usual care, during an inpatient stay after elective orthopaedic surgery. The efficacy of MT was demonstrated both over the entire intervention period and on each of the 3 days included in the analysis. Demonstrating significant changes in pain on Days 1 and 2 supports the findings of Ignacio et al. (2012), who also found significant changes in elective orthopaedic surgical patients on Days 1 and 2. They also found a decrease in anxiety on Day 2, which is consistent with our findings. Ignacio et al. (2012) also did not find any significant differences regarding the use of analgesic drugs. One difference here is that their patients received patient-controlled analgesia (PCA) after surgery, whereas our patients received either oral or injected medications, not PCA. The limitations, however, of Ignacio's study were the small sample size of 21

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

TABLE 3. **SUMMARY STATISTICS OF ADMISSION OUTCOMES**

Factor	Total (*N* = 163)	Control Group (*n* = 79)	Experimental Group (*n* = 84)
Total dose narcotics	140.0 [75.0, 225.0]	127.5 [75.0, 225.0]	150.0 [77.5, 225.0]
Total dose antiemetics	0.00 [0.00, 1.00]	0.00 [0.00, 1.00]	0.00 [0.00, 1.00]
Length of stay	2.0 [2.0, 3.0]	2.0 [2.0, 3.0]	2.0 [2.0, 3.0]

Note. Values are presented as median [P25, P75].

patients, as well as a lack of description of the music intervention. Allred et al. (2010) found that in patients having a TKA, a statistically significant decrease was noted in anxiety and pain as compared to the control group who quietly rested. This too is consistent with our findings. Neither of these studies evaluated nausea.

It was noted by Dotinga (2015) that between 2000 and 2010 the length of stay for hip replacements decreased from almost 5 days to slightly under 4 days. This compares with the average length of stay of 3 days in our study. This short duration of hospitalization could explain why we did not see any reduction in length of stay with MT.

Many of the published studies on the effects of music perioperatively involved small sample sizes (Antall & Kresevic, 2004; Eisenman & Cohen, 1995; Ignacio et al., 2012; Simcock et al., 2008); therefore, our finding on a much larger sample is a significant contribution to the literature. It is also important to note that many of the previous studies pre-, during, and/or postsurgery focused on only one intervention or utilized recorded music and/or recorded music-focused relaxation, instead of live MT interventions (Allred et al., 2010; Antall & Kresevic, 2004; Eckhouse et al., 2014; Eisenman & Cohen, 1995; Good et al., 2001; Lin, 2011; Lukas, 2004;

TABLE 4A. **SUMMARY STATISTICS, DAY 1 OUTCOMES**

Factor[a]	Total (*N* = 163)	Control Group (*n* = 79)	Experimental Group (*n* = 84)	*p*
Pretherapy pain	3.5 ± 2.3	3.3 ± 2.3	3.6 ± 2.3	.43[b]
Pretherapy anxiety	1.5 ± 2.0	1.5 ± 2.0	1.5 ± 2.0	.99[b]
Pretherapy nausea	0.69 ± 1.8	0.60 ± 1.6	0.77 ± 1.9	.55[b]
Pretherapy mood	1.2 ± 0.85	1.1 ± 0.92	1.2 ± 0.78	.53[b]
Posttherapy pain	2.9 ± 2.3	3.3 ± 2.2	2.6 ± 2.3	.065[b]
Posttherapy anxiety	0.71 ± 1.3	1.01 ± 1.6	0.43 ± 0.95	***.005***[b]
Posttherapy nausea	0.40 ± 1.2	0.40 ± 1.2	0.40 ± 1.2	.99[b]
Posttherapy mood	0.69 ± 0.84	1.04 ± 0.86	0.37 ± 0.67	<***.001***[b]
Change pain	0.54 ± 1.6	0.05 ± 1.7	1.00 ± 1.4	<***.001***[b]
Change anxiety	0.77 ± 1.4	0.46 ± 1.2	1.05 ± 1.6	***.009***[b]
Change nausea	0.29 ± 1.3	0.21 ± 1.2	0.37 ± 1.5	.43[b]
Change mood	0.49 ± 0.84	0.10 ± 0.57	0.86 ± 0.88	<***.001***[b]
MCID: Pain	30 (23.8)	6 (10.0)	24 (36.4)	<***.001***[c]
MCID: Anxiety	33 (55.0)	10 (33.3)	23 (76.7)	<***.001***[c]
MCID: Nausea	11 (47.8)	3 (30.0)	8 (61.5)	.21[d]
MCID: Mood	70 (55.1)	12 (21.4)	58 (81.7)	<***.001***[c]
Improvement: Pain	89 (54.9)	28 (35.9)	61 (72.6)	<***.001***[c]
Improvement: Anxiety	143 (88.3)	62 (79.5)	81 (96.4)	***.001***[d]
Improvement: Nausea	150 (93.2)	70 (89.7)	80 (96.4)	.12[c]
Improvement: Mood	103 (63.6)	32 (41.0)	71 (84.5)	<***.001***[c]

Note. Values presented as mean ± SD or *N* (column %). The values in bold italic represent values that are statistically significant. MCID = minimal clinically important difference.

[a] Data are not available for all subjects.

[b] ***t***-test.

[c] Pearson's chi-square test.

[d] Fisher's Exact test.

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

TABLE 4B. SUMMARY STATISTICS, DAY 2 OUTCOMES

Factor[a]	Total (*N* = 163)	Control Group (*n* = 79)	Experimental Group (*n* = 84)	*p*
Pretherapy pain	4.2 ± 2.4	4.0 ± 2.3	4.3 ± 2.5	.40[b]
Pretherapy anxiety	1.2 ± 1.9	1.2 ± 2.1	1.2 ± 1.8	.98[b]
Pretherapy nausea	0.82 ± 1.9	0.75 ± 1.7	0.89 ± 2.1	.70[b]
Pretherapy mood	1.2 ± 1.00	0.95 ± 0.85	1.4 ± 1.08	***.011***[b]
Posttherapy pain	3.1 ± 2.4	3.5 ± 2.4	2.8 ± 2.4	.14[b]
Posttherapy anxiety	0.84 ± 1.6	1.2 ± 1.8	0.52 ± 1.2	***.019***[b]
Posttherapy nausea	0.44 ± 1.5	0.72 ± 2.1	0.16 ± 0.61	***.046***[b]
Posttherapy mood	0.74 ± 0.84	0.95 ± 0.79	0.53 ± 0.84	***.006***[b]
Change pain	1.03 ± 1.6	0.52 ± 1.5	1.5 ± 1.6	<***.001***[b]
Change anxiety	0.39 ± 1.6	0.05 ± 1.6	0.73 ± 1.5	***.016***[b]
Change nausea	0.38 ± 1.5	0.03 ± 0.97	0.72 ± 1.8	***.010***[b]
Change mood	0.44 ± 0.77	0.00 ± 0.37	0.87 ± 0.82	<***.001***[b]
MCID: Pain	43 (41.0)	10 (19.6)	33 (61.1)	<***.001***[c]
MCID: Anxiety	17 (44.7)	5 (27.8)	12 (60.0)	.058[c]
MCID: Nausea	15 (68.2)	3 (30.0)	12 (100.0)	<***.001***[d]
MCID: Mood	45 (51.1)	4 (10.0)	41 (85.4)	<***.001***[d]
Improvement: Pain	77 (63.1)	31 (51.7)	46 (74.2)	***.010***[c]
Improvement: Anxiety	96 (78.7)	38 (63.3)	58 (93.5)	<***.001***[d]
Improvement: Nausea	109 (90.1)	50 (83.3)	59 (96.7)	***.016***[d]
Improvement: Mood	75 (61.5)	21 (35.0)	54 (87.1)	<***.001***[c]

Note. Values are presented as mean ± SD or *N* (column %). The values in bold italic represent values that are statistically significant.
MCID = minimal clinically important difference.
[a] Data are not available for all subjects.
[b] *t*-test.
[c] Pearson's chi-square test.
[d] Fisher's Exact test.

Ni et al., 2011; Ottaviani, Jean-Luc, Thomas & Pascal et al., 2012; Parker, 2011; Pellino et al., 2005; Simcock et al., 2008). Although the utilization of recorded music showed positive results, the use of live MT interventions also demonstrated positive results. We purposefully allowed the music therapists to use a variety of MT interventions to better reflect standard MT clinical practice. The sessions were therefore individualized, and promoted patient input, and this approach yielded significant results. These interventions also have the potential to affect more than pain, anxiety, and nausea. Although it was not a measure in our study, in some sessions patients expressed feeling depressed at the beginning of the session but stated that they had less depression at the end of the session. Therefore, MT has the potential to address a wide variety of patient needs.

One strength of this study was the use of a randomized, controlled, trial methodology. Another strength was the collection of pre- and post scores by a research assistant who was not involved with the treatment. The music therapists were blinded to the pre- and posttherapy scores. Although this makes the study methodologically stronger, it does not follow usual clinical practice. When conducting MT sessions in the clinical setting, the music therapist is usually not blind to the patient self-reports of symptom severity. This helps in identifying goals and implementing interventions. Being aware of postsession symptom severity may help adjust the treatment plan for future visits. The blinding of the music therapists may explain why pain was addressed as a goal in only 22% of sessions, anxiety in 9%, and mood in 4%.

Many of the patients expressed appreciation for the MT sessions. One patient, in particular, who was waiting for transport to come for her discharge, wanted MT one more time before she left. She told the music therapist that she attributed her ability to stop pain medication within 48 hours of surgery to holistic healing practices and the supportive MT sessions she had received. Even though some patients were dissatisfied with their group assignment, none withdrew from the study for this reason. No safety issues were identified during the study.

One limitation to the study is that the patients were not blinded to intervention as this was not possible. The research assistants who collected the pre- and postsession data also were not blinded to group assignment. This was also not feasible as, in the case of the experimental group, they had to let the music therapist know the patient was ready to be seen, and at the end of the

© 2018 by National Association of Orthopaedic Nurses
Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

Table 4c. Summary Statistics, Day 3 Outcomes

Factor[a]	Total (N = 163)	Control Group (n = 79)	Experimental Group (n = 84)	p
Pretherapy pain	3.4 ± 2.4	2.9 ± 2.4	4.0 ± 2.4	.060[b]
Pretherapy anxiety	1.6 ± 2.3	1.3 ± 2.1	1.9 ± 2.4	.24[b]
Pretherapy nausea	0.56 ± 1.3	0.64 ± 1.4	0.47 ± 1.2	.60[b]
Pretherapy mood	1.09 ± 0.93	1.06 ± 0.92	1.1 ± 0.94	.76[b]
Posttherapy pain	2.7 ± 2.3	2.8 ± 2.3	2.5 ± 2.4	.63[b]
Posttherapy anxiety	0.93 ± 1.8	1.3 ± 2.1	0.50 ± 1.2	.063[b]
Posttherapy nausea	0.250 ± 0.78	0.28 ± 0.66	0.22 ± 0.91	.76[b]
Posttherapy mood	0.65 ± 0.84	0.94 ± 0.95	0.31 ± 0.54	***.002***[b]
Change pain	0.74 ± 1.6	0.08 ± 1.1	1.5 ± 1.8	<***.001***[b]
Change anxiety	0.66 ± 1.7	-0.03 ± 0.91	1.4 ± 2.0	<***.001***[b]
Change nausea	0.31 ± 1.2	0.36 ± 1.4	0.36 ± 1.4	.70[b]
Change mood	0.44 ± 0.80	0.11 ± 0.62	0.81 ± 0.82	<***.001***[b]
MCID: Pain	15 (30.0)	3 (13.0)	12 (44.4)	***.029***[c]
MCID: Anxiety	13 (48.1)	1 (7.1)	12 (92.3)	<***.001***[c]
MCID: Nausea	5 (55.6)	3 (50.0)	2 (66.7)	.99[c]
MCID: Mood	25 (53.2)	6 (25.0)	19 (82.6)	<***.001***[c]
Improvement: Pain	36 (52.9)	14 (38.9)	22(68.8)	***.014***[d]
Improvement: Anxiety	54 (79.4)	25 (69.4)	29(90.6)	***.039***[c]
Improvement: Nausea	60 (88.2)	31 (86.1)	29(90.6)	.71[c]
Improvement: Mood	45 (66.2)	17 (47.2)	28(87.5)	<***.001***[c]

Note. Values are presented as mean ± SD or *N* (column %). The values in bold italic represent values that are statistically significant.
MCID = minimal clinically important difference.
[a] Data are not available for all subjects.
[b] *t* test.
[c] Fisher's Exact test.
[d] Pearson's chi-square test.

Table 5. Comparisons Survey Scores and Patient-Level Outcomes Overall

Cohort	Control: Mean (95% CI)	Experimental: Mean (95% CI)	Difference: Mean (95% CI)	p
Pain change	0.25 (−0.02, 0.52)	1.28 (1.02, 1.55)	1.03 (0.67, 1.40)	<***.001***[a]
Anxiety change	0.22 (−0.02, 0.46)	1.01 (0.77, 1.24)	0.79 (0.46, 1.11)	<***.001***[a]
Nausea change	0.18 (−0.03, 0.39)	0.47 (0.27, 0.68)	0.29 (0.01, 0.58)	***.044***[a]
Mood change	0.07 (−0.05, 0.19)	0.83 (0.71, 0.96)	0.76 (0.59, 0.94)	<***.001***[a]
Pain improvement	0.41 (0.34, 0.50)	0.73 (0.64, 0.80)	3.77 (2.27, 6.26)	<***.001***[b]
Anxiety improvement	0.71 (0.62, 0.78)	0.94 (0.90, 0.97)	6.69 (3.28, 13.63)	<***.001***[b]
Nausea improvement	0.86 (0.79, 0.91)	0.95 (0.89, 0.98)	3.17 (1.19, 8.48)	***.022***[b]
Mood improvement	0.41 (0.32, 0.50)	0.86 (0.79, 0.91)	9.13 (5.02, 16.62)	<***.001***[b]
Total dose of narcotics	175.04 (137.60, 212.48)	1.69.60 (133.30, 205.91)	−5.43 (−57.59, 46.72)	.84[c]
Total dose of antiemetics	0.68 (0.43, 0.94)	0.70 (0.45, 0.95)	0.02 (−0.34, 0.38)	.92[c]
Length of stay	2.62 (2.35, 2.89)	2.74 (2.48, 3.00)	0.12 (−0.26, 0.49)	.54[c]

Note. The values in bold italic represent values that are statistically significant.
[a] Linear mixed effect model result.
[b] Logistic regression with GEE result.
[c] Linear model result.

© 2018 by National Association of Orthopaedic Nurses

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

session the therapist had to let the research assistant know the session was over. In addition, data were collected on an inpatient population in a tertiary care center and is not necessarily generalizable to other settings.

Many with severe nausea refused to participate because of how they were feeling. This may be one reason why nausea was not a commonly addressed goal. It was difficult to tease out the narcotic and antiemetic usage, and compare the times they were given with the times of the MT session. This may explain why it was not possible to see any differences in the usage of these medications.

Although patients were appreciative of MT postsurgery, many stated that they wish they could have had it presurgery as they were experiencing anxiety waiting for the surgery to begin. Future research could test the effects of providing MT services before and after surgery. The retention of effect of MT sessions (particularly if the patients are taught ways to use music to address their symptoms at home) and the cost-effectiveness of MT interventions are other questions to address in future studies. Because of the increased number of knee and hip replacements done each year, it is important to find interventions that could help these patients deal with postsurgical symptoms and potentially hasten their recovery. Therefore, research regarding the impact, value, and efficacy of MT could prove to be beneficial to patients, to healthcare institutions, and to society.

Conclusion

This study demonstrated the efficacy of MT in improving pain, anxiety, mood, and nausea for patients following elective orthopaedic surgeries. To the best of our knowledge, this is one of the first studies to investigate the use of MT, as conducted by a board-certified music therapist, in this patient population. It is also unique in its use of a wide variety of interventions to address various goals, resulting in highly individualized sessions. Based on these findings, MT should be used more frequently post-elective orthopaedic surgeries to aid in managing symptoms such as pain and anxiety. Further research is needed to better understand the efficacy of MT before and after orthopaedic surgery in a variety of settings.

ACKNOWLEDGMENTS

The authors thank Maria Jukic for her support; Deborah Bucar, Mary Ann Donovan, Mary Ann Winch, and Joyce Smith, the Pre-Admission Testing Nurses; Stephanie Morris, Laura Higgins, Melissa Onabajo, Sarah Paczak Chappell, and Amy Foye, the music therapists who conducted the music therapy sessions for our patients; and Anastasia Pintur and Tammy Shella, the research assistants. This project was funded internally. No external funding was received.

REFERENCES

Allred, K. D., Byers, J. F., & Sole, M. L. (2010). The effect of music on postoperative pain and anxiety. *Pain Management Nursing*, *11*(1), 15–25. doi:10.1016/j.pmn.2008.12.002

American Academy of Orthopaedic Surgeons. (2016). *Beyond surgery day: The full impact of knee replacement*. Retrieved from http://www.anationinmotion.org/value/knee

American Music Therapy Association. (2005). *What is music therapy?* Retrieved from http://www.musictherapy.org/about/quotes/

Antall, G. F., & Kresevic, D. (2004). The use of guided imagery to manage pain in an elderly orthopaedic population. *Orthopaedic Nursing*, *23*(5), 335–340.

Bradt, J., Dileo, C., & Shim, M. (2013). Music interventions for preoperative anxiety (Review). *The Cochrane Library*, *6*(6), CD006908.

Bruscia, K. E. (2014). *Defining music therapy (3rd ed)*. University Park, IL: Barcelona.

Cepeda, M. S., Carr, D. B., Lau, J., & Alvarez, H. (2006). Music for pain relief. *Cochrane Database of Systematic Reviews*, *2*. doi:10.1002/14651858

Dotinga, R. (2015). Number of hip replacements has skyrocketed, U.S. report shows. HealthDay. Retrieved from https://consumer.healthday.com/senior-citizen-information-31/demographic-arthritis-news-37/number-of-hip-replacements-has-skyrocketed-u-s-report-shows-696419.html

Easter, B., DeBoer, L., Settlemyre, G., Starnes, C., Marlowe, V., & Tart, R. C. (2010). The impact of music on the PACU patient's perception of discomfort. *Journal of PeriAnesthesia Nursing*, *25*(2), 79–87. doi:10.1016/j.jopan.2010.01.017

Eckhouse, D. R., Hurd, M., Cotter-Schaufele, S., Sulo, S., Sokolowki, M., & Barbour, L. (2014). A randomized controlled trial to determine the effects of music and relaxation interventions on perceived anxiety in hospitalized patients receiving orthopaedic or cancer treatment. *Orthopaedic Nursing*, *33*(6), 342–351. doi:10.1097/NOR.0000000000000098

Eisenman, A., & Cohen, B. (1995). Music therapy for patients undergoing regional anesthesia. *AORN Journal*, *62*(6), 947–950.

Engwall, M., & Duppils, G. S. (2009). Music as a nursing intervention for postoperative pain: A systematic review. *Journal of PeriAnesthesia Nursing*, *24*(6), 370–383. doi:10.1016/j.jopan.2009.10.013

Evans, D. (2002). The effectiveness of music as an intervention for hospital patients: A systematic review. *Journal of Advanced Nursing*, *37*(1), 8–18.

Gallagher, L. M., Lagman, R., Walsh, D., Davis, M. P., & LeGrand, S. B. (2006). The clinical effects of music therapy in palliative medicine. *Support Care Cancer*, *14*, 859–866.

Good, M., Stanton-Hicks, M., Grass, J. A., Anderson, G. C., Lai, H. L., Roykulcharoen, V., & Adler, P. A. (2001). Relaxation and music to reduce postsurgical pain. *Journal of Advanced Nursing*, *33*(2), 208–215.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Information*, *42*(2), 377–381.

Horrigan, B. (2013). *Easing pain and improving patient satisfaction with integrative care*. Retrieved from http://www.hhnmag.com/articles/5871-easing-pain-and-improving-patient-satisfaction-with-integrative-care

Ignacio, J. J., Chan, M. F., Teo, S. H., Tsen, L. M., & Goy, R. (2012). Research in brief—The effect of music on pain, anxiety, and analgesic use on adults undergoing an orthopaedic surgery: A pilot study. *Singapore Nursing Journal*, *39*(4), 49–51.

Lin, P. -C. (2011). An evaluation of the effectiveness of relaxation therapy for patients receiving joint replacement surgery. *Journal of Clinical Nursing*, *21*, 601–608. doi:10.1111/j.1365-2702.2010.03406.x

© 2018 by National Association of Orthopaedic Nurses

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

Lukas, L. K. (2004). Orthopedic outpatients' perception of perioperative music listening as therapy. *Journal of Theory Construction & Testing*, *8*(1), 7–12.

MacDonald, R. A. R., Mitchell, L. A., Dillon, T., Serpell, M. G., Davies, J. B., & Ashley, E. A. (2003). An empirical investigation of the anxiolytic and pain reducing effects of music. *Psychology of Music*, *31*(2), 187–203.

Ni, C. H., Tsai, W. H., Lee, L. M., Kao, C. C., & Chen, Y. C. (2012). Minimising preoperative anxiety with music for day surgery patients—a randomized clinical trial. *Journal of Clinical Nursing*, *21*(5/6), 620–625.

Nilsson, U., Rawal, N., & Unosson, M. (2003). A comparison of intraoperative or postoperative exposure to music: A controlled trial of the effects on postoperative pain. *Anaesthesia*, *58*, 699–703.

Ottaviani, S., Jean-Luc, B., Thomas, B., & Pascal, R. (2012). Effect of music on anxiety and pain during joint lavage for knee osteoarthritis. *Clinical Rheumatology*, *31*, 531–534. doi:10.1007/s10067-011-1925-9

Parker, R. J. (2011). Evidence-based practice: Caring for a patient undergoing total knee arthroplasty. *Orthopaedic Nursing*, *30*(1), 4–8. doi:10/1097/NOR.0b013e3182057451

Pellino, T. A., Gordon, D. B., Engelke, Z. K., Busse, K. L., Colllins, M. A., Silver, C. E., & Norcross, N. J. (2005). Use of nonpharmacologic interventions for pain and anxiety after total hip and total knee arthroplasty. *Orthopedic Nursing*, *24*(3), 182–190.

Phipps, M. A., Carroll, D. L., & Tsiantoulas, A. (2010). Music as a therapeutic intervention on an inpatient neuroscience unit. *Complementary Therapies in Clinical Practice*, *16*, 138–142.

Rogers, A. (1981). In: Third World Congress on Pain of the International Association for the Study of Pain, Edinburgh, Scotland, September 4–11, 1981, Abstracts. *Pain Supplement*, *1*, S1–S319.

Simcock, X. C., Yoon, R. S., Chalmers, P., Geller, J. A., Kiernan, H. A., & Macaulay, W. (2008). Intraoperative music reduces perceived pain after total knee arthroplasty. *Journal of Knee Surgery*, *21*, 275–278.

World Medical Association. (2016). *WMA declaration of Helsinki: Ethical principles for medical research involving human subjects*. Retrieved from https://www.wma/net/en/30publications/10policies/b3

For additional continuing nursing education activities on orthopaedic nursing topics, go to nursingcenter.com/ce.

NATIONAL ASSOCIATION OF ORTHOPAEDIC NURSES

Orthopaedic Nursing

The International Leader in Practice and Education

Orthopaedic Nursing journal Column Opportunities

We are currently recruiting interested authors for the following 2 lead columnist positions:

Pathophysiology
Pharmacology

If you are interested in discussing one of these opportunities,
please e-mail Mary Rodts, Editor, at onjeditor@gmail.com
Please include a copy of your Curriculum Vitae

© 2018 by National Association of Orthopaedic Nurses

Copyright © 2018 by National Association of Orthopaedic Nurses. Unauthorized reproduction of this article is prohibited.

Original Article

The Mediating and Moderating Effect of Volunteering on Pain and Depression, Life Purpose, Well-Being, and Physical Activity

Elizabeth Salt, PhD, APRN,[*] Leslie J. Crofford, MD,[†] and Suzanne Segerstrom, PhD[‡]

Abstract:

To improve function and quality of life in patients with chronic pain, a prevalent and costly condition, an understanding of the relationships among well-being, physical activity, depression, and life purpose with pain is needed. Because of the role loss experienced by people with chronic pain, activities such as volunteering could have an important role in improving health and well-being. In one study, chronic pain patients who participated in volunteer activities reported both decreased pain and "a sense of purpose." The aim of this study is to test the relationships among pain and well-being, physical activity, depression, and life purpose and then to determine if volunteering activities mediated or moderated these relationships. This observational study was conducted in a large university setting in Kentucky and used a sample of 200 women older than age 50. We found that people with higher pain were more depressed and had lower life purpose and well-being. People who volunteered less had more pain, lower perceived life purpose, more depressive symptoms, and decreased physical activity. Volunteer activities did have a significant mediating effect on the relationship between pain and depression; approximately 9% of the relationship between pain and depression can be accounted for by volunteering. Moderation by volunteering was found between pain and life purpose. We identified important relationships among pain, volunteering, and health outcomes and found that volunteering has a role in improving depressive symptoms and life purpose in women with pain.

© 2017 by the American Society for Pain Management Nursing

*From the *University of Kentucky, College of Nursing, Lexington, Kentucky; †Vanderbilt University, Department of Medicine, Nashville, Tennessee; ‡University of Kentucky, Department of Psychology, Lexington, Kentucky.*

Address correspondence to Elizabeth Salt, PhD, APRN, University of Kentucky, College of Nursing, 315 College of Nursing Building, 751 Rose Street, Lexington, KY, 40536-0232. E-mail: egsalt0@uky.edu

Received August 24, 2016;
Revised April 5, 2017;
Accepted April 7, 2017.

This project was supported by National Institutes of Health grants: R01-AG046116, K02-033629, and UL1TR000117. This work was supported in part by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health.

1524-9042/$36.00
© 2017 by the American Society for Pain Management Nursing
http://dx.doi.org/10.1016/j.pmn.2017.04.004

INTRODUCTION

Chronic pain or persistent pain affects 37% of people residing in developing countries. The total (direct and indirect) U.S. costs attributed to this condition are estimated to be between $560 and $635 billion annually, and an estimated 3%-10% of the gross domestic product has been spent on this condition in European countries (Breivik, Eisenberg, & O'Brien, 2013; Phillips et al., 2016; Tsang et al., 2008). People with persistent pain have been found to have decreased well-being, physical activity, and life purpose and increased depression (Gureje, Von Korff, Simon, & Gater, 1998; Schleicher et al., 2005). In efforts to improve function and quality of life in this population, these factors are important. Despite the clinical importance of and vast resources used on a number of pharmacologic and nonpharmacologic strategies, the management of chronic pain remains a significant health problem (U.S. Department of Health and Human Services, 2016), suggesting a different approach to treatment is needed.

Volunteering is defined as providing a service without the intent of compensation (Klinedinst & Resnick, 2014). Health benefits of volunteering in older adults have been well-described, including (1) increased physical activity; (2) improved self-reported health, life satisfaction, and well-being; and (3) reduced depressive symptoms, pain, and mortality risk (Ahern & Hendryx, 2008; Ayalon, 2008; Cattan, Hogg, & Hardill, 2011; Choi, Stewart, & Dewey, 2013; Jenkinson et al., 2013; Klinedinst, Resnick, Yerges-Armstron, & Dorsey, 2015; Pillemer, Fuller-Rowell, Reid, & Wells, 2010; Veerasamy, Sambasivan, & Kumar, 2013). The health benefits of volunteering in people with pain conditions are not well-studied. Yet the loss of roles in people with chronic pain has been described (Harris, Morley, & Barton, 2003). In the only study identified, people with chronic pain who participated in volunteer activities reported significantly decreased pain. Participants in this study also reported "a sense of purpose" after volunteering (Arnstein, 2002).

It is possible that volunteer activities could mediate or moderate the relationships among pain and well-being, physical activity, depression, and life purpose. A mediational relationship, or an indirect effect, would imply that pain affects volunteer activities, which, in turn, affects life purpose, depression, physical activity, and depression (Hayes, 2013). A moderated relationship, or a contingent effect, would imply that volunteer activity influences the effect of pain on well-being, physical activity, depression, and life purpose (Hayes, 2013). Because quality of life and function are targeted clinical outcomes, a furthered understanding of these relationships could affect care of patients with chronic pain.

Therefore, the purpose of this study was to test the relationship between pain and well-being, physical activity, depression, and life purpose and then to determine if volunteering activities mediate or moderate these relationships in a sample of older women. The following hypotheses were tested:

1. There will be a significant relationship between pain and depression, physical activity, life purpose, and well-being.
2. Volunteer activities will have a significant mediating effect on the relationship between pain and well-being, physical activity, depression, and life purpose.
3. Volunteer activities will have a significant moderating effect on the relationship between pain and well-being, physical activity, depression, and life purpose.

METHODS

Sample

Because pain disproportionately affects older women (Johannes, Le, Zhou, Johnston, & Dworkin, 2010), the hypotheses were tested on a sample of 199 women older than age 50. One additional participant was not included in the present study because of incomplete daily diary measures for pain. Mean age of the sample was 61.9 years (standard deviation [SD] = 6.4 years) with a mean of 16.7 (SD = 2.3) years of education. The majority (99%) were Caucasian, with the rest African American (Table 1).

Procedure

Women were recruited using the Kentucky Women's Health Registry, a registry of more than 15,000 women residing in the state of Kentucky, to participate in an observational study titled "Daily Activity and Health in the Lives of Adult Women" (Kentucky Women's Health Registry, 2016). Women with pain were oversampled in this study, which addresses the study aims of investigating the effect of pain on well-being. More than half of the sample reported no pain (54%)

TABLE 1.
Demographics Table

Age	61.9 yr ± 6.4
Education	16.7 yr ± 2.3
Ethnicity	
Caucasian	197 (99%)
Other	2 (1%)

and the remaining 46% reported pain in one (27%) or more (19%) body locations in their registry survey.

Women aged 50-75 years who resided in a seven-county area around Lexington, Kentucky were invited to participate in this study. Those who had a physical condition that significantly affected mobility, body mass index >40, pacemaker, serious heart conditions (e.g., recent heart attack), serious medical conditions (e.g., autoimmune disease), serious mental disorders, or use of oral, inhaled, or injected corticosteroids (e.g., prednisone) in the 3 months before enrollment were excluded from participation. Baseline data were collected during an initial in-person appointment. Self-report and interviewer-delivered questionnaires were completed at baseline and after completing a week-long diary using REDCap computer software. Women were compensated for their time in participating in this study. All procedures were approved by the Medical Institutional Review Board at the University of Kentucky.

Measures

Pain. Pain severity, pain interference, and a pain composite score were calculated by averaging a 7-day self-report of pain using the Patient Reported Outcomes Measures Information System (PROMIS; Cella et al., 2010; Department of Health and Human Services, 2015). Pain severity was measured with one PROMIS item asking, "In the past 7 days, how intense was your pain at its worst?" (had no pain = 0, worst imaginable pain = 10; range: 0-10). The six PROMIS pain interference items (example item: "In the past 7 days, how much did pain interfere with your activities of daily living?"; not at all = 1, a little bit = 2, somewhat = 3, quite a bit = 4, very much = 5; range: 6-30) were averaged for the measurement of pain interference. The pain composite was an equally weighted combination of the severity and interference score after interference scores were transformed to have the same range as severity. The Cronbach's α was .96 for the pain composite measure in this study.

Well-Being. Well-being was measured with the 84-item Scales of Psychological Well-Being (SPWB; Ryff, 1989), which has six subscales measuring personal growth, purpose in life, environmental mastery, autonomy, personal relations with others, and self-acceptance. The SPWB uses a 6-point Likert-type scale (6 = strongly agree, 1 = strongly disagree) and a total score is calculated (range: 84-504); a higher score identifies people with higher well-being. The Cronbach's α in this sample was .95.

Life Purpose. Life purpose was measured using the Life Purpose subscale of the SPWB (e.g., "I have a sense of direction and purpose in life"; range: 14-84; Ryff, 1989). The Cronbach's α in this sample was .83.

Volunteering. Volunteer activities were measured by summing yes/no responses to two items ("Do volunteer work? Attend church or take part in church activities?") of the Community Healthy Activities Model Program for Seniors scale (CHAMPS; Stewart et al., 2001).

Physical Activity. Physical activity was measured by summing the frequency scores for aerobic activities on the CHAMPS (e.g., "Ride a bicycle or stationary cycle?"; range: 0-672 hours [4 weeks]; Stewart et al., 2001). The interclass correlation coefficient for the CHAMPS frequency score has been reported to be .58-.62 and the test-retest reliability at 2 weeks has been reported to be .62-.76. Medium correlations were reported between self-reported physical functioning and the CHAMPS score (Stewart et al., 2001).

Depression. Depression was measured by summing yes/no responses to the 30 items of the Geriatric Depression Scale (Yesavage et al., 1983; example item: "Are you hopeful about the future?"; range: 0-30) The Cronbach's α was .82 in this sample.

Data Analysis

Data were summarized using descriptive statistics including means and standard deviations or frequency distributions, as appropriate. Depression and pain were positively skewed. Therefore, analyses were conducted with both log-transformed and untransformed variables. For the purposes of moderation analysis, all predictor variables were mean centered (Aiken & West, 1991). For hypothesis 1, we investigated the relationship among pain (composite, severity, and interference) and volunteer activities, depression, well-being, life purpose, and physical activity using Pearson correlations in SPSS (Version 22.0; IBM Corp., Armonk, NY, USA). To test for a mediating effect of volunteer activities among pain and well-being, depression, physical activity, and life purpose (hypothesis 2), we used PROCESS macros provided by Hayes (2013). In these models $\hat{Y} = i_1 + cx$; $\hat{M} = i_2 + ax$; $\hat{Y} = i_2 + c'x + bM$, where x is the predictor (in this case, pain) and M is the mediator (in this case, volunteering). Hypothesis 3 (moderation) tested the main effects of pain and volunteering and their interaction on life purpose, physical activity, well-being, and depression. Because there were significant relationships among age and pain composite, pain interference, life purpose, well-being, and depression, we controlled for age in all mediation and moderation analyses. An α of .05 was set for all analyses.

RESULTS

Table 2 shows the means, standard deviations, and correlations among the study variables. Consistent with

Table 2.
Descriptive Statistics and Correlations among Variables (N = 199)

	Mean ± SD	Age	Pain Composite	Pain Severity	Pain Interference	Volunteering	Depression	Well-Being	Life Purpose	Physical Activity
Age	61.9 ± 6.4	—								
Pain composite	2.28 ± 2.5	−.141*	—							
Pain severity	1.52 ± 1.5	−.126	.963†	—						
Pain interference	.76 ± 1.17	−.145*	.940†	.813†	—					
Volunteering	1.17 ± .76	.135	−.130	−.137*	−.108	—				
Depression	3.37 ± 3.6	−.251†	.380†	.327†	.406†	−.247†	—			
Well–being	428.53 ± 41.2	.218†	−.253†	−.229†	−.256†	.138	−.743†	—		
Life purpose	71.19 ± 8.5	.163*	−.253†	−.233†	−.251†	.171*	−.662†	.866†	—	
Physical activity	22.6 ± 9.7	.069	−.009	−.027	.014	.150*	−.087	.140*	.110	—

*Correlation is significant at the .05 level (2-tailed).
†Correlation is significant at the .01 level (2-tailed).

hypothesis 1, people with increased pain and depression volunteered less and had lower reports of well-being and life purpose, and people who volunteered more had higher reports of life purpose. People who were more physically active volunteered more and reported higher well-being. Otherwise, physical activity did not have a significant relationship with pain, life purpose, and depression.

Hypothesis 2, in which volunteering mediates the relationships among pain and physical activity, depression, life purpose and well-being, was partially supported. Using transformed variables, the relationship between pain (composite and severity) and depression was mediated by volunteering (depression indirect effect = .03, .04, respectively; confidence interval [CI] = .0004-.0795, CI = .0005-.0995, respectively; Fig. 1). The relationships among pain (composite, severity, and interference) and well-being, physical activity, and life purpose were not mediated by volunteer activities using transformed or untransformed variables.

Hypothesis 3, in which volunteering moderates the relationships among pain and depression, life purpose, well-being, and physical activity, was also partially supported. Using transformed variables, volunteering did moderate the effect of pain severity and life purpose (interaction $\beta = 2.26$, $t = 2.19$, $p = .03$), as well as the effect of composite pain (interaction $\beta = 2.71$, $t = 2.08$, $p = .04$) (Fig. 2). People with higher pain composite had significantly lower life purpose than people with lower pain composite when they had volunteering scores ≤1.55. Similarly, people with higher pain severity had significantly lower life purpose than people with lower pain severity when they had a volunteering score of ≤1.46. This moderating effect was not found using the untransformed variables. Volunteer activities did not moderate the relationship between pain interference and life purpose. Volunteering did not moderate relationships among pain (composite, interference, or severity) and depression, physical activity, and well-being using the transformed or untransformed variables.

DISCUSSION

The purpose of this study was to test the relationships among pain and well-being, physical activity, depression, and life purpose and then to determine if volunteering activities mediate or moderate these relationships in a sample of older women. Hypothesis 1 was largely supported; people with higher pain were more depressed and people with higher pain and depression volunteered less and had lower perceived life purpose. People who volunteered

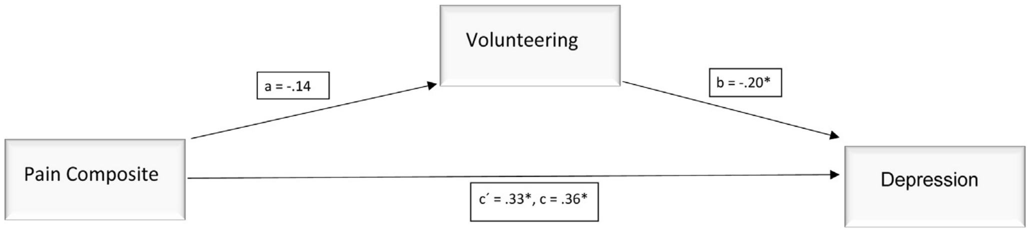


Figure 1. ■ Mediation model showing that the effect of pain composite on depression is mediated by volunteering. a = the direct path between pain composite and volunteering; b = the direct path between volunteering and depression; c' = the direct path between pain composite and depression; c = the total effect or the direct and indirect effect of pain composite and depression. *Significant at the .05 level.

more were more physically active. Hypothesis 2 was partially supported. Volunteer activities did have a significant mediating effect on the relationship between pain and depression. In regard to hypothesis 3, a moderating relationship between volunteering and pain was found for life purpose.

The positive relationships found between volunteering and indicators of psychological well-being and function and negative relationships between volunteering and psychological distress are supported by prior research (Arnstein, 2002; Gureje et al., 1998; Harris et al., 2003; Schleicher et al., 2005). The use of volunteering as a health care tool to improve pain and function in people with pain has not been studied. Thus, these findings are important in identifying a potential future direction for improved care of patients suffering from pain.

Hypothesis 2 was partially supported insofar as the relationship between pain and depression was mediated by volunteering. Approximately 9% of the relationship between pain and depression could be accounted for by volunteering (Fig. 1). These findings suggest that volunteer activities could be a potential target to improve depressive symptoms in people with pain. Because the mediating effect was only identified with transformed variables, it is likely that people with higher pain are not influencing this finding unduly; rather, differences in pain in the middle and

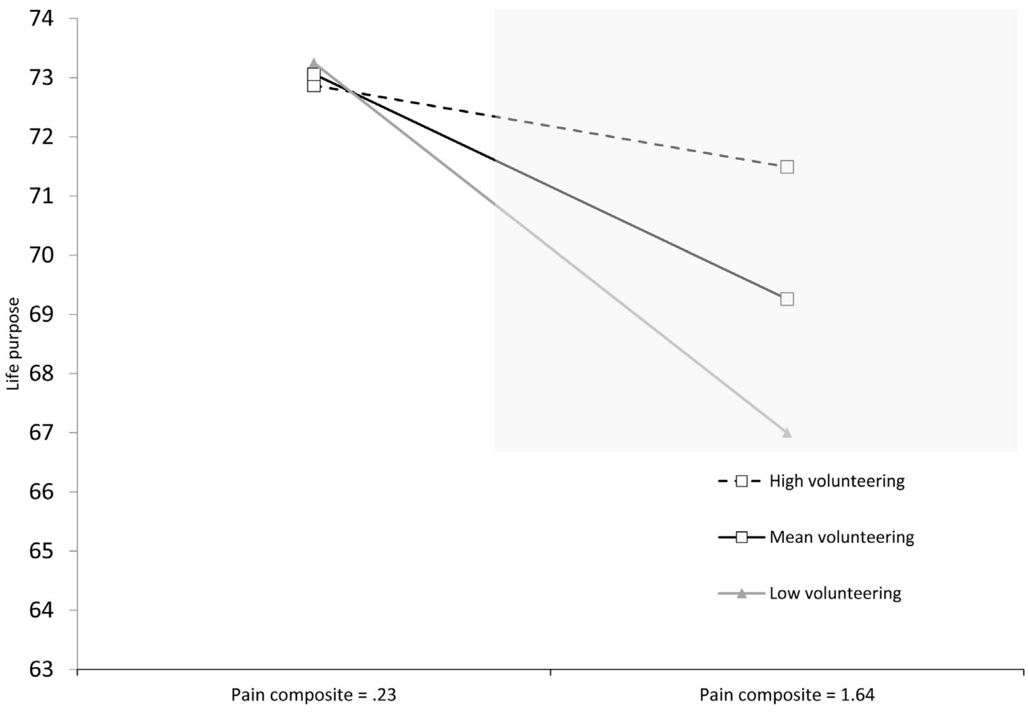


Figure 2. ■ Moderating effect of volunteering between pain composite and life purpose. Simple main effects of pain on life purpose are shown for high (+1 standard deviation [SD]), mean, and low (−1 SD) levels of volunteering. The effect of volunteering is statistically significant.

lower parts of the range are important. Thus, this finding is important for many people who experience pain at less than severe levels.

Our study findings also partially supported hypothesis 3. We found that volunteer activity moderated the relationship between pain and life purpose only. We found that women with pain who had a volunteer score of >1.55 had lower life purpose. These findings provide a further understanding the relationship between two constructs largely unexplored in pain research.

There are a number of limitations of this study. Because of the cross-sectional design, we were unable to establish temporal precedent or causality. Longitudinal and intervention studies are needed to establish temporal precedence (e.g., changes in pain precede changes in volunteering) and causality (e.g., experimental assignment to volunteering decreases the effect of pain on life purpose.) Second, the sample was demographically homogenous, which could affect the generalizability of study findings. Third, there are limitations related to our measures of physical activity and volunteer activities. Objective measures of physical activity with activity trackers and daily diaries of volunteer activities would likely be more accurate measures of these variables and should be considered for use in future studies.

Implications for Nursing Practice

These study findings have a number of clinical implications. First, these findings suggest that life purpose and volunteering, which are largely ignored in pain research, might have important associations with factors affecting health outcomes. Second, not only does volunteering have a significant relationship with the health indicators, it also play a mediating and moderating role, suggesting it might be a useful health care tool in this population.

CONCLUSION

This study described relationships among life purpose and volunteering and health outcomes such as well-being, pain, physical activity, and depression. It also described an important potential role of volunteering to improve depressive symptoms and increase life purpose in people with pain. Because people with higher life purpose have recently been found to have improved survival, this is clinically important (Steptoe, Deaton, & Stone, 2015).

REFERENCES

Ahern, M. M., & Hendryx, M. (2008). Community participation and the emergence of late-life depressive symptoms: Differences between women and men. *Journal of Women's Health, 17*(9), 1463–1470.

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.

Arnstein, P. (2002). From chronic pain patient to peer: Benefits and risks of volunteering. *Pain Management Nursing, 3*(3), 94–103.

Ayalon, L. (2008). Volunteering as a predictor of all-cause mortality: What aspects of volunteering really matter. *International Psychogeriatics, 20*(5), 1000–1013.

Breivik, H., Eisenberg, E., & O'Brien, T. (2013). The individual and societal burden of chronic pain in Europe: The case for strategic prioritization and action to improve knowledge and availability of appropriate care. *BMC Public Health, 13*, 1229.

Cattan, M., Hogg, E., & Hardill, I. (2011). Improving quality of life in ageing populations: Why can volunteering do? *Maturitias, 70*(4), 328–332.

Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., Amtmann, D., Bode, R., Buysse, D., Choi, S., Cook, K., Devellis, R., DeWalt, D., Fries, J. F., Gershon, R., Hahn, E. A., Lai, J. S., Pilkonis, P., Revicki, D., Rose, M., Weinfurt, K., & Hays, R. (2010). PROMIS Cooperative Group. The patient-reported outcomes measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *Journal of Clinical Epidemiology, 63*(11), 1179–1194.

Choi, K. S., Stewart, R., & Dewey, M. (2013). Participation in productive activities and depression among older Europeans: Survey of health, ageing and retirement. *Geriatric Psychiatry, 28*(11), 1157–1165.

Department of Health and Human Services. *PROMIS*. Retrieved from http://www.nihpromis.org/measures/availableinstruments. Accessed May 25, 2016

Gureje, O., Von Korff, M., Simon, G., & Gater, R. (1998). Persistent pain and well-being: A World health Organization study in primary care. *Journal of the American Medical Association, 280*(2), 147–151.

Harris, S., Morley, S., & Barton, S. B. (2003). Role loss and emotional adjustment in chronic pain. *Pain, 105*(1-2), 363–370.

Hayes, A. F. (2013). *Introduction to mediation, moderation and conditional process analysis: A regression-based approach*. New York: The Guilford Press.

Jenkinson, C. E., Dickens, A. P., Jones, K., Thompson-Coon, J., Taylor, R. S., Rogers, M., Bambra, C. L., Lang, I., & Richards, S. H. (2013). Is volunteering a public health intervention? A systematic review and meta-analysis of the health and survival of volunteers. *BMC Public Health, 13*, 773–783.

Johannes, C. B., Le, T. K., Zhou, X., Johnston, J. A., & Dworkin, R. H. (2010). The prevalence of chronic pain in U.S. adults. *Journal of Pain, 11*(11), 1230–1239.

Kentucky Women's Health Registry. Retrieved from https://obgyn.med.uky.edu/obgyn-kwhr. Accessed June 2016

Klinedinst, N., & Resnick, B. (2014). Volunteering and depressive symptoms among residents in a continuing care retirement community. *Journal of Gerontological Social Work, 57*(1), 52–71.

Klinedinst, N. J., Resnick, B., Yerges-Armstron, L. M., & Dorsey, S. G. (2015). The interplay of genetics, behavior, and pain with depressive symptoms in the elderly. *The Gerontologist, 55*(S1), S67–S77.

Phillips, K., Taylor, A., Mease, P. J., Simon, L. S., Conaghan, G., Choy, E. H., Singh, J. A., Strand, V., Gossec, L., Kaiser, U., de Wit, M., Ostelo, R., Maxwell, L., & Tugwell, P. S. (2016). Harmonizing pain outcomes measures: Results of the Pre-OMERACT meeting on partnerships for consensus on patient important pain outcomes domains between the Cochrane Musculoskeletal Group and OMERACT. *Journal of Rheumatology, 42*, 1943–1946.

Pillemer, K., Fuller-Rowell, T. E., Reid, M. C., & Wells, N. M. (2010). Environmental volunteering and health outcomes over a 20-year period. *The Gerontologist, 50*(5), 594–602.

Ryff, C. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality Social Psychology, 57*(6), 1069–1081.

Schleicher, H., Alonso, C., Shirtcliff, E. A., Muller, D., Loevinger, B., & Coe, C. L. (2005). In the face of pain: The relationship between psychological well-being and disability in women with fibromyalgia. *Psychotherapy & Psychosomatics, 74*(4), 231–239.

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health and ageing. *Lancet, 385*(9968), 640–648.

Stewart, A. L., Mills, K. M., King, A. C., Haskell, W. L., Gillis, D., & Ritter, P. L. (2001). CHAMPS physical activity questionnaire for older Adults: Outcomes for interventions. *Medicine and Science in Sports and Exercise, 33*(7), 1126–1141.

Tsang, A., Von Korff, M., Lee, S., Alonso, J., Karam, E., Angermeyer, M. C., Borges, G. L., Bromet, E. J., Demytteneare, K., de Girolamo, G., de Graaf, R., Gureje, O., Lepine, J. P., Haro, J. M., Levinson, D., Oakley Browne, M. A., Posada-Villa, J., Seedat, S., & Watanabe, M. (2008). Common chronic pain conditions in developed and developing countries: Gender and age differences and comorbidity with depression anxiety disorders. *Journal of Pain, 9*, 883–891.

U.S. Department of Health and Human Services, 2016. Pain Management. Retrieved from https://healthfinder.gov/FindServices/SearchContext.aspx?topic=624. Accessed May 2, 2016.

Veerasamy, C., Sambasivan, M., & Kumar, N. (2013). Individual skills based volunteerism and life satisfaction among healthcare volunteers in Malaysia: Role of employer encouragement, self-esteem and job performance, a cross-sectional study. *Volunteerism and Life Satisfaction, 8*(10), 1–10.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*(1), 37–49.

The LIFE Study: Lifestyle Interventions and Independence For Elders

Rationale, Design and Results

Jack M. Guralnik, M..D., Ph.D.

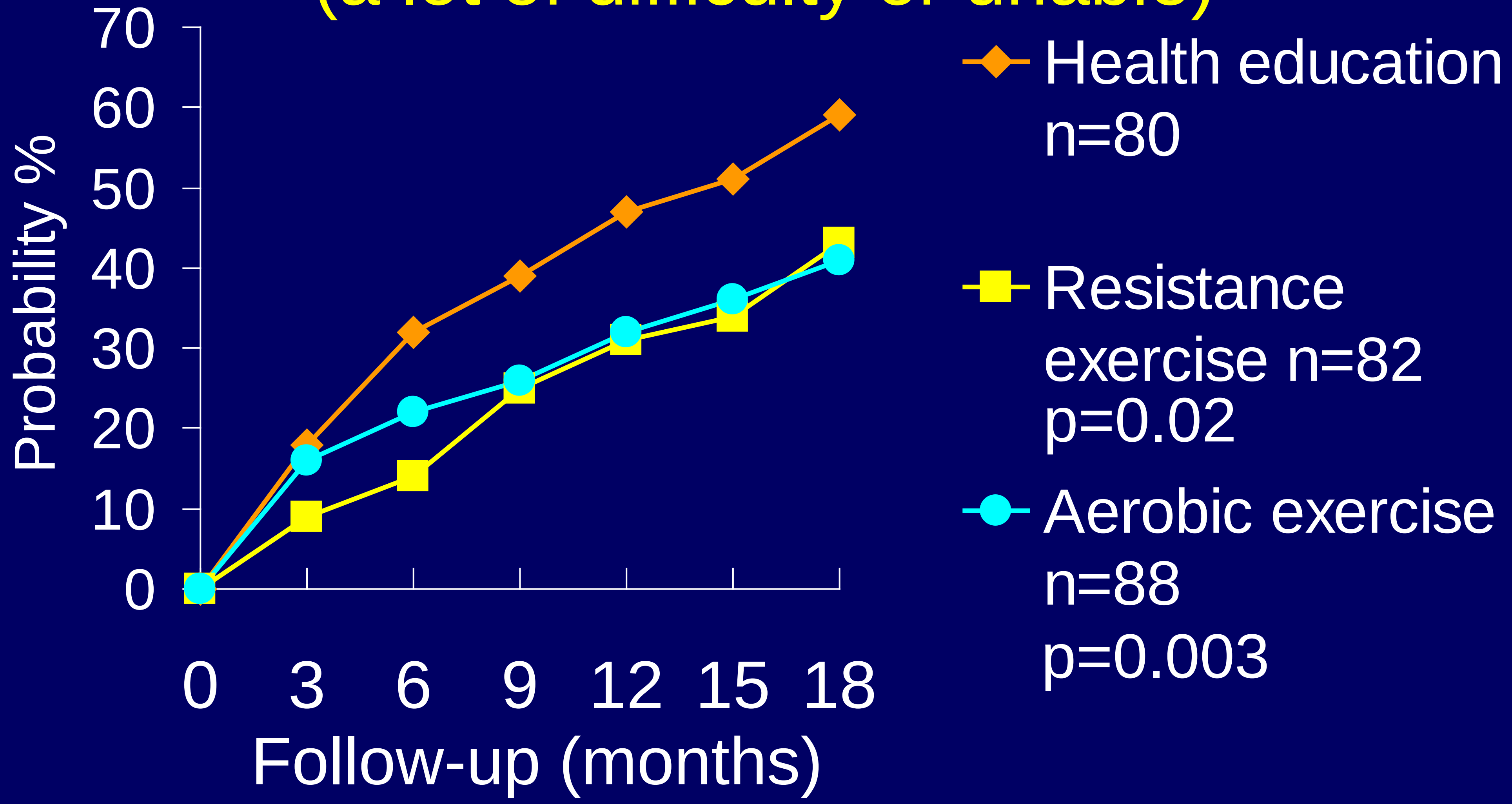
University of Maryland School of Medicine

The LIFE Study Journey

2000	First idea for a Phase 3 trial
2001-2003	Planning grant R21AG19353
2003-2009	LIFE Pilot U01AG022376
Sep 2009	LIFE funding U01AG022376
Feb 2010	Start randomization
Dec 2011	Randomization complete n=1635
Summer 2012	Release of baseline data
Dec 2013	Follow-up complete
May 2014	Publication of the main results
Summer 2014	Release of follow-up data

FAST - Incidence of ADL disability

(a lot of difficulty or unable)

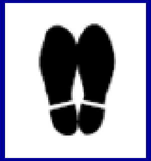


Penninx et al. *Arch. Int. Med. 2001*

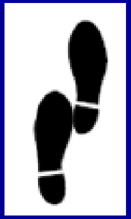
Short Physical Performance Battery

- Timed standing balance (up to 10 seconds)

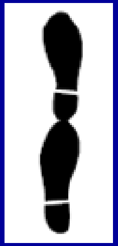
  Side-by-side stand



  Semi-tandem stand

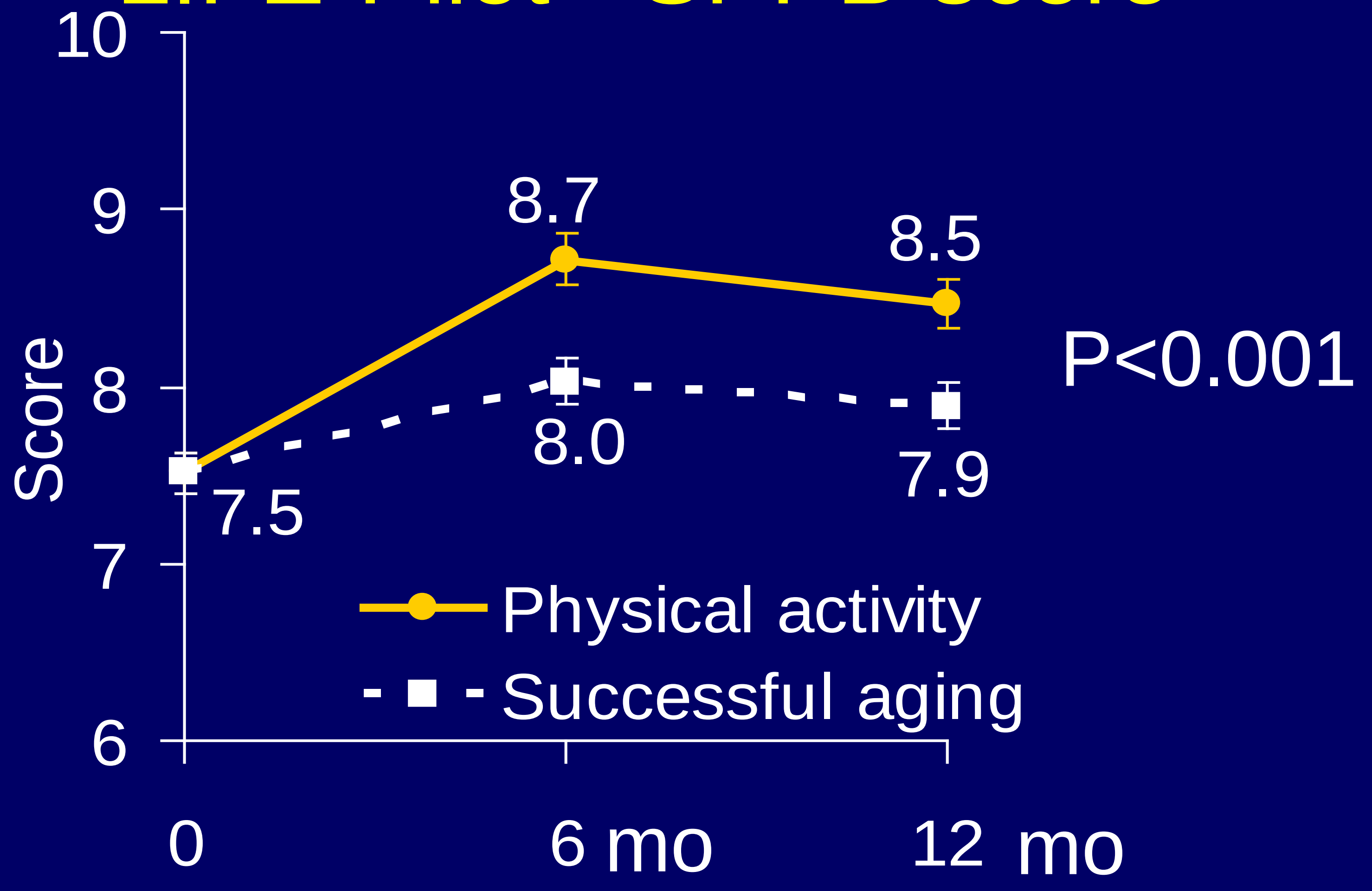


  Tandem stand



- Timed 4-meter walk
- Timed multiple (5) chair rises

LIFE-Pilot - SPPB score



Means estimated from repeated measures ANCOVA
adjusted for gender, field center and baseline values

Pahor et al *J Gerontol 2006;61:1157*

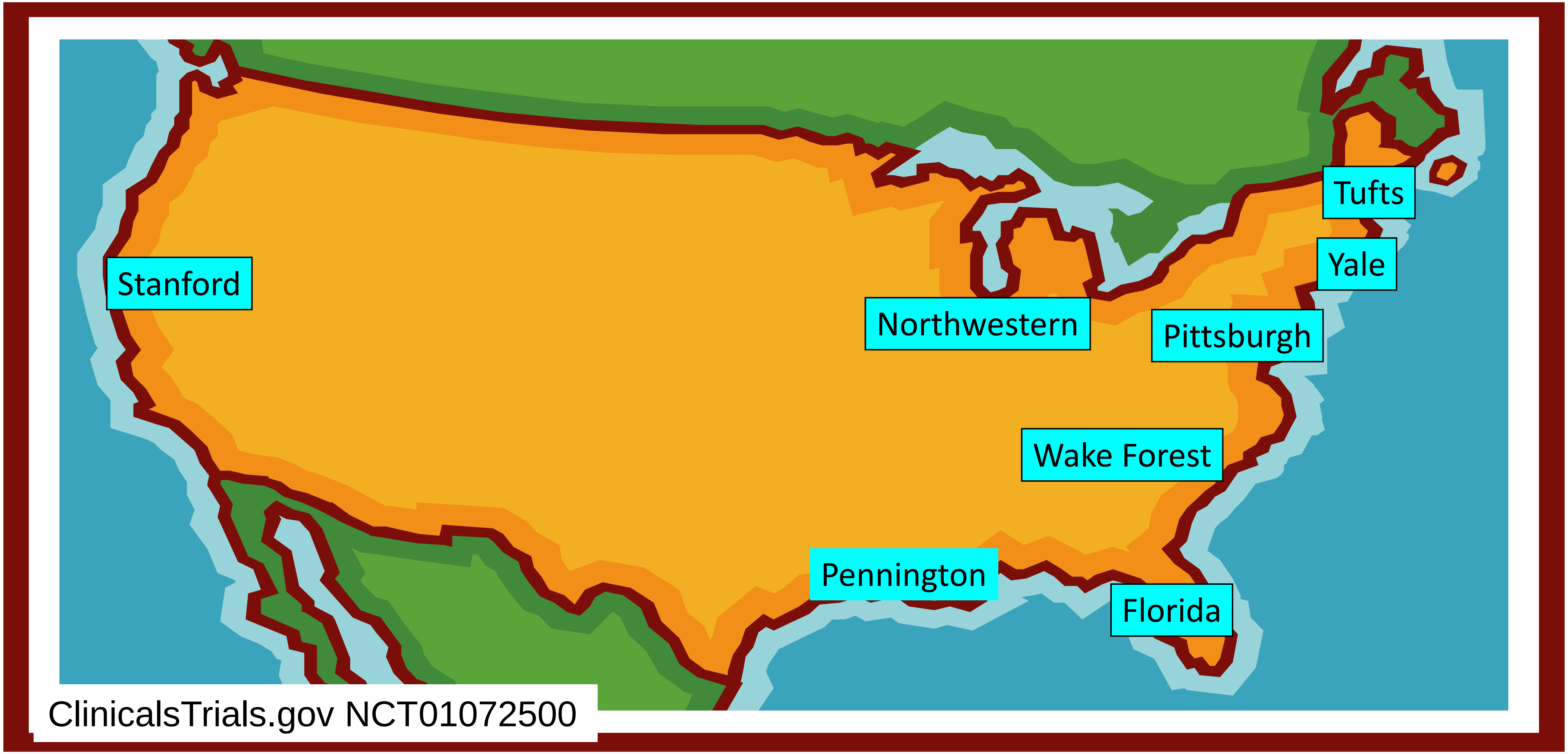
Lifestyle Interventions and Independence for Elders

The LIFE Study

Is a structured physical activity or a health education program more effective in reducing the risk of major mobility disability in older persons?

Fielding et al. J Gerontol 2011;66:1226
Pahor et al. JAMA 2014

- Multicenter, single-blinded, parallel randomized trial
- 8 field centers across the US
- Coordinating Center: University of Florida
- Data Management Quality Control: Wake Forest University
- February 2010 – December 2013


Stanford
Northwestern
Tufts
Yale
Pittsburgh
Wake Forest
Pennington
Florida

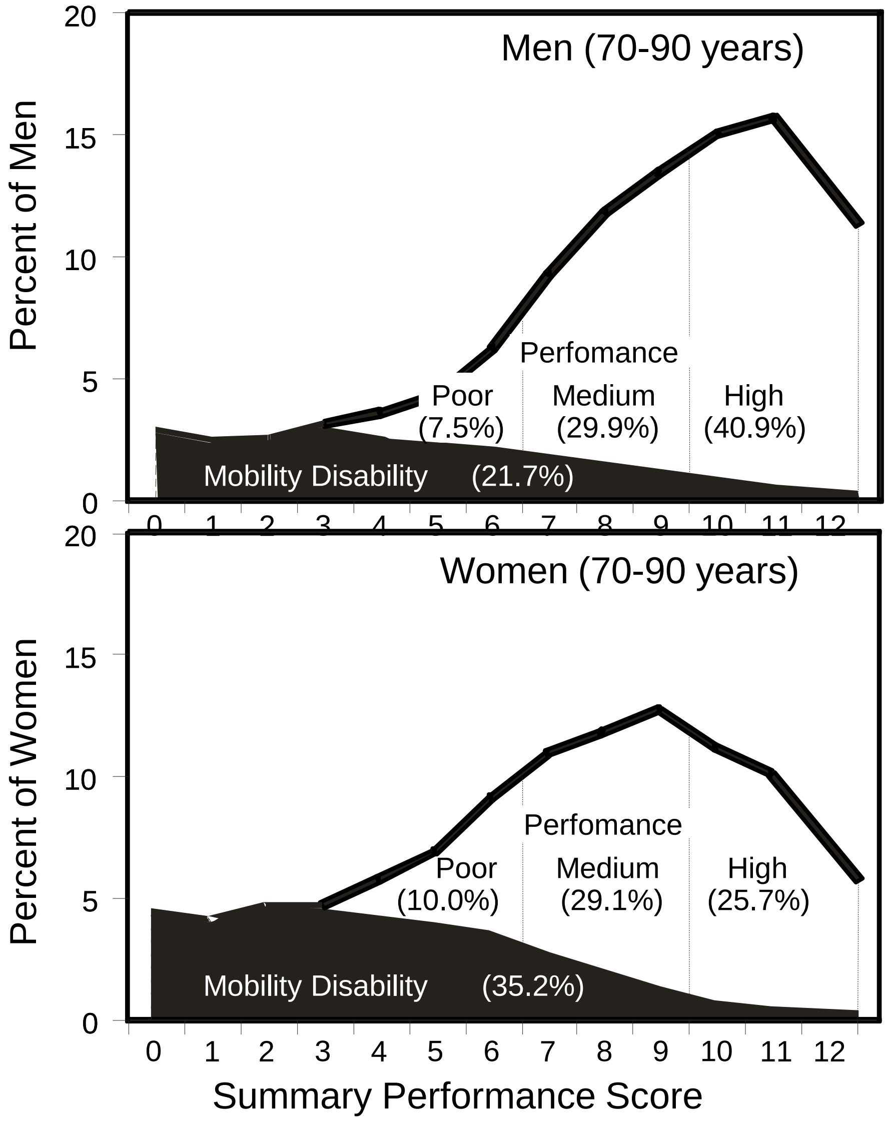
ClinicalsTrials.gov NCT01072500

LIFE Inclusion criteria

- Men and women 70-89 years
- Sedentary lifestyle (<20 min per week in structured PA, <150 min/week in moderate PA)
- Able to walk 400 m
- SPPB score $\leq$9 (45% <8)
- No major cognitive impairment
- Could safely participate in the intervention (medical history, physical exam and ECG)
- Gives informed consent, lives in the study area and does not plan to move

Fielding et al. J Gerontol 2011;66:1226
Pahor et al. JAMA 2014

Distribution of Performance in Those With & Without Mobility Disability



Ferrucci et al. J Am Geriatr Soc. 2000; 48:1102-10

Health Education also called Successful Aging intervention

- Health workshops relevant to older adults (e.g., healthful nutrition, how to effectively negotiate the health care system, how to travel safely, recommended screening, etc.)
- Short instructor-led program (5-10 min) of gentle stretching or flexibility exercises
- Frequency: weekly during the first 26 weeks, and then monthly (bi-monthly optional)

Fielding et al. J Gerontol 2011;66:1226
Pahor et al. JAMA 2014

Physical activity intervention

Center-based in a group setting + home

- Aerobic (walking)
- Strength (lower extremities)
- Balance
- Flexibility stretching
- Behavioral counseling (group and telephone)

Fielding et al. J Gerontol 2011;66:1226
Pahor et al. JAMA 2014

Frequency and duration

- Walking 3 to 6 days per week (2 times per week at the center) (minimum walking bout 10 minutes with goal of 30 minutes per bout)
- Strength training 3 times per week (10 minutes per session)

Fielding et al. J Gerontol 2011;66:1226
Pahor et al. JAMA 2014

Primary outcome:
Major Mobility Disability

Inability to walk 400 m at usual pace on a 20 m course - 10 laps (40 m per lap)

- Within 15 min without sitting
- Without help of a person or walker
- Use of a cane and stop for up to 1 min was acceptable

Fielding et al. J Gerontol 2011;66:1226
Pahor et al JAMA 2014

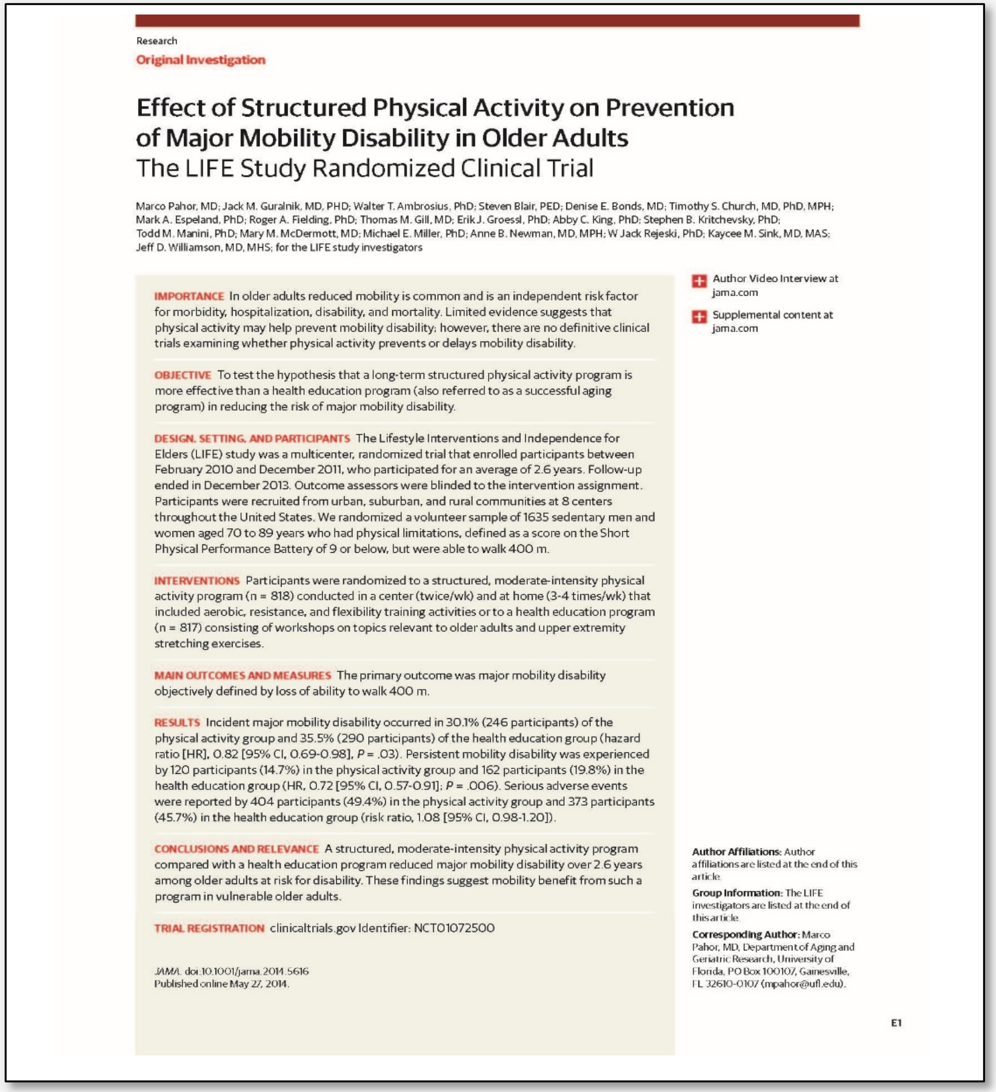
Conclusion

Among older adults at risk of disability, a structured moderate intensity physical activity program, compared with a health education program, reduces

- Major mobility disability by 18%
- Persistent mobility disability by 28%

Take Home Message

- Although highly prevalent and increasing in size, the older, more vulnerable population has been understudied and typically is not included in large randomized trials.
- LIFE demonstrates mobility benefit of a physical activity program, and provides a strong rationale for both health care providers and community health systems to promote physical activity for vulnerable ambulatory older persons



Research
Original Investigation

Effect of Structured Physical Activity on Prevention of Major Mobility Disability in Older Adults
The LIFE Study Randomized Clinical Trial

Marco Pahor, MD; Jack M. Guralnik, MD, PHD; Walter T. Ambrosius, PhD; Steven Blair, PED; Denise E. Bonds, MD; Timothy S. Church, MD, PhD, MPH; Mark A. Espeland, PhD; Roger A. Fielding, PhD; Thomas M. Gill, MD; Erik J. Groessl, PhD; Abby C. King, PhD; Stephen B. Kritchevsky, PhD; Todd M. Manini, PhD; Mary M. McDermott, MD; Michael E. Miller, PhD; Anne B. Newman, MD, MPH; W Jack Rejeski, PhD; Kaycee M. Sink, MD, MAS; Jeff D. Williamson, MD, MHS; for the LIFE study investigators

IMPORTANCE In older adults reduced mobility is common and is an independent risk factor for morbidity, hospitalization, disability, and mortality. Limited evidence suggests that physical activity may help prevent mobility disability; however, there are no definitive clinical trials examining whether physical activity prevents or delays mobility disability.

OBJECTIVE To test the hypothesis that a long-term structured physical activity program is more effective than a health education program (also referred to as a successful aging program) in reducing the risk of major mobility disability.

DESIGN, SETTING, AND PARTICIPANTS The Lifestyle Interventions and Independence for Elders (LIFE) study was a multicenter, randomized trial that enrolled participants between February 2010 and December 2011, who participated for an average of 2.6 years. Follow-up ended in December 2013. Outcome assessors were blinded to the intervention assignment. Participants were recruited from urban, suburban, and rural communities at 8 centers throughout the United States. We randomized a volunteer sample of 1635 sedentary men and women aged 70 to 89 years who had physical limitations, defined as a score on the Short Physical Performance Battery of 9 or below, but were able to walk 400 m.

INTERVENTIONS Participants were randomized to a structured, moderate-intensity physical activity program (n = 818) conducted in a center (twice/wk) and at home (3-4 times/wk) that included aerobic, resistance, and flexibility training activities or to a health education program (n = 817) consisting of workshops on topics relevant to older adults and upper extremity stretching exercises.

MAIN OUTCOMES AND MEASURES The primary outcome was major mobility disability objectively defined by loss of ability to walk 400 m.

RESULTS Incident major mobility disability occurred in 30.1% (246 participants) of the physical activity group and 35.5% (290 participants) of the health education group (hazard ratio [HR], 0.82 [95% CI, 0.69-0.98], $P = .03$). Persistent mobility disability was experienced by 120 participants (14.7%) in the physical activity group and 162 participants (19.8%) in the health education group (HR, 0.72 [95% CI, 0.57-0.91]; $P = .006$). Serious adverse events were reported by 404 participants (49.4%) in the physical activity group and 373 participants (45.7%) in the health education group (risk ratio, 1.08 [95% CI, 0.98-1.20]).

CONCLUSIONS AND RELEVANCE A structured, moderate-intensity physical activity program compared with a health education program reduced major mobility disability over 2.6 years among older adults at risk for disability. These findings suggest mobility benefit from such a program in vulnerable older adults.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01072500

JAMA. doi:10.1001/jama.2014.5616
Published online May 27, 2014.

+ Author Video Interview at jama.com

+ Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Group Information: The LIFE investigators are listed at the end of this article.

Corresponding Author: Marco Pahor, MD, Department of Aging and Geriatric Research, University of Florida, PO Box 100107, Gainesville, FL 32610-0107 (mpahor@ufl.edu).

The Journal of the American Medical Association

Pahor M, Guralnik JM, Ambrosius WT, et al.

Effect of Structured Physical Activity on Prevention of Major Mobility Disability in Older Adults: The LIFE Study Randomized Clinical Trial

Published online May 27, 2014

Available at jama.com and
on The JAMA Network Reader at
mobile.jamanetwork.com

The JAMA Network

Acknowledgment

- NIA U01AG22376
- NHLBI 3U01AG022376-05A2S

- NIA Intramural Research Program
- Claude D. Pepper Older Americans Independence Centers at the University of Florida (P30AG028740), Wake Forest University (P30 AG21332),Tufts University (1P30AG031679), University of Pittsburgh (P30AG024827), and Yale University (P30AG021342)
- NCRR/NCATS CTSA at Stanford University (UL1RR025744), University of Florida (U54RR025208), Yale University (UL1TR000142).
- Boston Rehabilitation Outcomes Center at Tufts University (1R24HD065688-01A1).

We are indebted to Evan C. Hadley, MD, Sergei Romashkan, MD, PhD from the NIA and Denise E. Bonds, MD from the NHLBI

Original Investigation

Effect of Structured Physical Activity on Prevention of Major Mobility Disability in Older Adults

The LIFE Study Randomized Clinical Trial

Marco Pahor, MD; Jack M. Guralnik, MD, PHD; Walter T. Ambrosius, PhD; Steven Blair, PED; Denise E. Bonds, MD; Timothy S. Church, MD, PhD, MPH; Mark A. Espeland, PhD; Roger A. Fielding, PhD; Thomas M. Gill, MD; Erik J. Groessl, PhD; Abby C. King, PhD; Stephen B. Kritchevsky, PhD; Todd M. Manini, PhD; Mary M. McDermott, MD; Michael E. Miller, PhD; Anne B. Newman, MD, MPH; W Jack Rejeski, PhD; Kaycee M. Sink, MD, MAS; Jeff D. Williamson, MD, MHS; for the LIFE study investigators

IMPORTANCE In older adults reduced mobility is common and is an independent risk factor for morbidity, hospitalization, disability, and mortality. Limited evidence suggests that physical activity may help prevent mobility disability; however, there are no definitive clinical trials examining whether physical activity prevents or delays mobility disability.

OBJECTIVE To test the hypothesis that a long-term structured physical activity program is more effective than a health education program (also referred to as a successful aging program) in reducing the risk of major mobility disability.

DESIGN, SETTING, AND PARTICIPANTS The Lifestyle Interventions and Independence for Elders (LIFE) study was a multicenter, randomized trial that enrolled participants between February 2010 and December 2011, who participated for an average of 2.6 years. Follow-up ended in December 2013. Outcome assessors were blinded to the intervention assignment. Participants were recruited from urban, suburban, and rural communities at 8 centers throughout the United States. We randomized a volunteer sample of 1635 sedentary men and women aged 70 to 89 years who had physical limitations, defined as a score on the Short Physical Performance Battery of 9 or below, but were able to walk 400 m.

INTERVENTIONS Participants were randomized to a structured, moderate-intensity physical activity program (n = 818) conducted in a center (twice/wk) and at home (3-4 times/wk) that included aerobic, resistance, and flexibility training activities or to a health education program (n = 817) consisting of workshops on topics relevant to older adults and upper extremity stretching exercises.

MAIN OUTCOMES AND MEASURES The primary outcome was major mobility disability objectively defined by loss of ability to walk 400 m.

RESULTS Incident major mobility disability occurred in 30.1% (246 participants) of the physical activity group and 35.5% (290 participants) of the health education group (hazard ratio [HR], 0.82 [95% CI, 0.69-0.98], *P* = .03). Persistent mobility disability was experienced by 120 participants (14.7%) in the physical activity group and 162 participants (19.8%) in the health education group (HR, 0.72 [95% CI, 0.57-0.91]; *P* = .006). Serious adverse events were reported by 404 participants (49.4%) in the physical activity group and 373 participants (45.7%) in the health education group (risk ratio, 1.08 [95% CI, 0.98-1.20]).

CONCLUSIONS AND RELEVANCE A structured, moderate-intensity physical activity program compared with a health education program reduced major mobility disability over 2.6 years among older adults at risk for disability. These findings suggest mobility benefit from such a program in vulnerable older adults.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01072500

JAMA. doi:10.1001/jama.2014.5616

Published online May 27, 2014.

Author Video Interview at jama.com

Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Group Information: The LIFE investigators are listed at the end of this article.

Corresponding Author: Marco Pahor, MD, Department of Aging and Geriatric Research, University of Florida, PO Box 100107, Gainesville, FL 32610-0107 (mpahor@ufl.edu).

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

The life expectancy of older Americans continues to increase, with persons 65 years or older representing the fastest growing segment of the US population.[1] Although prolongation of life remains an important public health goal, of even greater significance is the preservation of the capacity to live independently and to function well during late life.[2] Identification of proven interventions to prevent disability is an important public health challenge.[3]

Mobility—the ability to walk without assistance—is a critical characteristic for functioning independently.[4,5] Those who lose mobility have higher rates of morbidity, disability, and mortality[6-13] and yet are often excluded from clinical trials. Preserving the ability to walk 400 m, an excellent proxy for community ambulation, is central to maintaining a high quality of life and independence in the community.

To our knowledge, no trial has conclusively tested that physical activity can prevent or delay the onset of mobility disability over an extended follow-up. Therefore, we conducted the Lifestyle Interventions and Independence for Elders (LIFE) pilot study from 2004 to 2006 to plan for the phase 3 randomized trial.[14] As hypothesized, the LIFE pilot study (N = 424) showed significant improvements in walking speed and physical performance measures. The pilot was not powered for a disability end point, but showed a nonsignificant reduction in risk of major mobility disability in the physical activity group compared with the health education group (also referred to as the successful aging group). In the LIFE study, we hypothesized that a long-term structured physical activity program would reduce the risk of major mobility disability compared with a health education program.

Methods

Trial Design and Participants

The study protocol was approved by the institutional review boards at all participating sites. Written informed consent was obtained from all study participants. The trial was monitored by a data and safety monitoring board appointed by the National Institute on Aging. The LIFE study was a multicenter, single-blind, parallel randomized trial conducted at 8 centers across the United States (University of Florida, Gainesville and Jacksonville, Florida; Northwestern University, Chicago, Illinois; Pennington Biomedical Research Center, Baton Rouge, Louisiana; University of Pittsburgh, Pittsburgh, Pennsylvania; Stanford University, Stanford, California; Tufts University, Boston, Massachusetts; Wake Forest School of Medicine, Winston-Salem, North Carolina; and Yale University, New Haven, Connecticut) between February 2010 and December 2013. The Administrative Coordinating Center was located at the University of Florida and the Data Management, Analysis, and Quality Control Center at Wake Forest School of Medicine. The centers included rural, suburban, and urban communities.

Details of the methods were published previously.[15] Briefly, the eligibility criteria consisted of men and women aged 70 to 89 years who (1) were sedentary (reporting <20 min/wk of performing regular physical activity in the past month and <125 min/wk of moderate physical activity); (2) were at high risk for mobility disability based on lower extremity functional limitations measured by the Short Physical Performance Battery (SPPB)[16] with a score of 9 or lower out of 12 (45% of participants were targeted to have a score <8); (3) could walk 400 m in less than 15 minutes without sitting, leaning, or the help of another person or walker; (4) had no major cognitive impairment (measured by the Modified Mini-Mental State Examination[17] [3MSE] with a score of no more than 1.5 standard deviations below education- and race-specific norms); and (4) could safely participate in the intervention as determined by medical history, physical examination, and resting electrocardiography. Persons with 9 or more years of education who scored less than 80 (<76 if African American) and those with less than 9 years of education who scored less than 76 (<70 if African American or Spanish speaking) on the 3MSE were excluded.

Targeted mass mailings to the community was the primary recruitment strategy.[18]

Randomization

Participants were randomized to a physical activity group or to a health education program group (**Figure 1**) via a secure, web-based data management system using a permuted block algorithm (with random block lengths) stratified by field center and sex. Both groups received an initial individual 45-minute face-to-face introductory session by a health educator who described the intervention, communicated expectations, and answered questions.

Interventions

The physical activity intervention involved walking, with a goal of 150 min/wk, strength, flexibility, and balance training.[15] The intervention included attendance at 2 center-based visits per week and home-based activity 3 to 4 times per week for the duration of the study. A protocol was in place to restart the intervention for the participants who suspended the physical activity for medical reasons. The physical activity sessions were individualized and progressed toward a goal of 30 minutes of walking daily at moderate intensity, 10 minutes of primarily lower extremity strength training by means of ankle weights (2 sets of 10 repetitions), 10 minutes of balance training, and large muscle group flexibility exercises. The participants began with lighter intensity and gradually increased intensity over the first 2 to 3 weeks of the intervention. The Borg scale of self-perceived exertion,[19] which ranges from 6 to 20, was used to measure intensity of activity. Participants were asked to walk at an intensity of 13 (activity perception "somewhat hard"), and lower extremity strengthening exercises were performed at an intensity of 15 to 16.

The health education program focused on successful aging (termed *the successful aging group* in previous publications). The health education group attended weekly workshops of health education during the first 26 weeks, and then monthly sessions thereafter (bimonthly attendance was optional). Workshops included topics relevant to older adults, such as how to effectively negotiate the health care system, how to travel safely, preventive services and screenings rec-

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

ommended at different ages, where to go for reliable health information, nutrition, etc. The workshops did not include any physical activity topics. The program also included a 5- to 10-minute instructor-led program of gentle upper extremity stretching or flexibility exercises.

Measurements

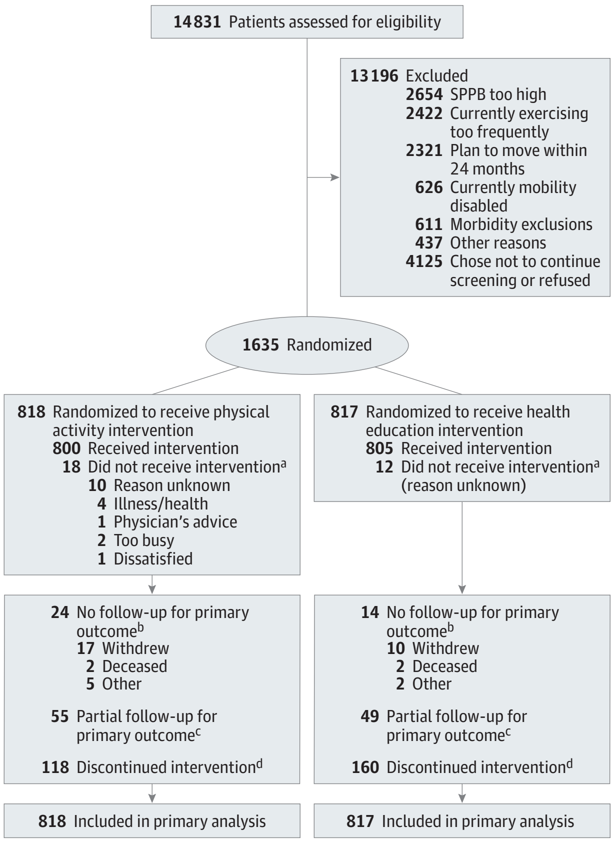
Participants were assessed every 6 months at clinic visits. Home, telephone, and proxy assessments were attempted if the participants could not come to the clinic. The assessment staff was blinded to the intervention and remained separate from the intervention team. Participants were asked not to disclose their assigned group and not to talk about their interventions during the assessment. Self-reported physical activity was ascertained by a separate set of unblinded assessors.

The main baseline assessments included self-reported demographic and contact information, medical and hospitalization history, medication inventory, electrocardiography, physical examination, Quality of Well-Being questionnaire,[20] health care utilization, physical activity assessed with the Community Healthy Activities Model Program for Seniors (CHAMPS) questionnaire,[21] and with accelerometry over 7-day periods (Actigraph Inc), cognitive testing, 400-m walk test,[22] the SPPB, body weight, blood pressure, and pulse rate. These measures were repeated during follow-up at varied intervals. Details of these measures and their frequency are described elsewhere.[15] The SPPB consisted of 4-m walk at usual pace, a timed repeated chair stand, and 3 increasingly difficult standing balance tests.[16,23] Each measure was assigned a categorical score ranging from 0 (inability to complete the test) to 4 (best performance). A summary score ranging from 0 (worst performers) to 12 (best performers) was calculated by summing the 3 component scores. Race and ethnicity were reported by the participants and were collected according to National Institutes of Health requirements. To minimize reporting bias, adverse events originating from the blinded assessments are presented.

Outcome Assessment

The primary outcome of major mobility disability was defined as the inability to complete a 400-m walk test within 15 minutes without sitting and without the help of another person or walker.[15] Use of a cane was acceptable. Participants were asked to walk 400 m at their usual pace, without overexerting, on a 20-m course for 10 laps (40 m/lap). Participants were allowed to stop for up to 1 minute for fatigue or related symptoms. When major mobility disability could not be objectively measured because of the inability of the participant to come to the clinic and absence of a suitable walking course at the participant's home, institution, or hospital, an alternative adjudication of the outcome was based on objective inability to walk 4 m in less than 10 seconds, or self-, proxy-, or medical record–reported inability to walk across a room. If participants met these alternative criteria, they would not be able to complete the 400-m walk within 15 minutes. Reports of death were tracked through regular surveillance. Two consecutive major mobility disability assessments or major mobility disability followed by death defined persistent mobility disability. Censoring was defined at the time of the last definitive assessment for major mobility disability.

Figure 1. Flow of Participants Through the Study


14831 Patients assessed for eligibility
13196 Excluded
2654 SPPB too high
2422 Currently exercising too frequently
2321 Plan to move within 24 months
626 Currently mobility disabled
611 Morbidity exclusions
437 Other reasons
4125 Chose not to continue screening or refused
1635 Randomized
818 Randomized to receive physical activity intervention
800 Received intervention
18 Did not receive intervention^a
10 Reason unknown
4 Illness/health
1 Physician's advice
2 Too busy
1 Dissatisfied
817 Randomized to receive health education intervention
805 Received intervention
12 Did not receive intervention^a (reason unknown)
24 No follow-up for primary outcome^b
17 Withdrew
2 Deceased
5 Other
55 Partial follow-up for primary outcome^c
118 Discontinued intervention^d
14 No follow-up for primary outcome^b
10 Withdrew
2 Deceased
2 Other
49 Partial follow-up for primary outcome^c
160 Discontinued intervention^d
818 Included in primary analysis
817 Included in primary analysis

SPPB indicates Short Physical Performance Battery.

[a] Participants who did not receive the allocated intervention (ie, attended no intervention sessions).

[b] For participants who did not have any major mobility disability assessments, we assigned 1 hour of follow-up time, because we knew that they were able to do the 400-m walk at baseline.

[c] Partial follow-up indicates participants who had censoring times prior to the last planned follow-up visit.

[d] Discontinuation of the intervention was operationalized as participants who did not attend at least 1 intervention session during their last 6 months of follow-up prior to the last planned follow-up visit date. Deaths and intervention withdrawals are included in these numbers. As an example, a participant may have discontinued the intervention in the initial 6 months of follow-up due to illness and then died prior to the 6-month assessment for the primary outcome. This participant would be reflected as missing the primary outcome due to death and also discontinuing the intervention.

At each contact, participants (or proxies, if the participant was not available) were questioned about outcomes and hospitalizations since the last visit. All records for hospitalizations were obtained and outcomes were reviewed and adjudicated independently by 2 experts who were blinded to the group randomization. If the 2 reviewers disagreed, the information was forwarded to the adjudication committee and a determination was made by consensus.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

Statistical Considerations

Power calculations for the primary outcome, time until the first postrandomization occurrence of major mobility disability, were based on a log-rank test with a 2-sided, .05 significance level. Based on the LIFE pilot study,[14] the annual incidence rate of major mobility disability in the health education group was assumed to increase from 18% in the first year to 21% after 2 years. We further assumed that recruitment would be uniform over 21 months, follow-up would average 31 months, and loss to follow-up would be 8% per year. Under these assumptions, randomization of 1600 participants provides 80% power to detect a 21% reduction, and 90% power to detect a 24% reduction in the hazard for major mobility disability in the physical activity participants. These effect-size targets were determined based on consistency with effects derived from observational research, the LIFE pilot experience, clinical relevance (around 20% reduction), and available funding resources.

Baseline characteristics were summarized by intervention group using mean and standard deviation, or percentages. Intervention adherence was calculated as the percentage of scheduled intervention sessions attended by participants. Self-reported minutes of activity and minutes spent in activity associated with more than 760 counts/min (by accelerometry)[24] were analyzed using mixed-effects analysis of covariance models for repeatedly measured outcomes with an unstructured parameterization for longitudinal covariance. Models contained the following terms: field center and sex (both used to stratify randomization), baseline value of the relevant physical activity measure, intervention, clinic visit, and intervention-by-visit interaction. Least squares means were obtained from these models and contrasts were used to estimate the average effects (95% CI) over the follow-up period. Risk ratios (95% CI) were calculated to determine the relative effect of the intervention on the proportion of participants reporting adverse effects. A test of equality of the risk ratios for hospitalization between baseline subgroups defined by SPPB levels (<8 vs ≥8) was performed using Poisson regression.

The effect of the intervention on the primary outcome (ie, time until the initial ascertainment of major mobility disability) was tested based on a 2-tailed significance of .05 using the intention-to-treat approach in which participants are grouped according to randomization assignment. To compare interventions, we used a likelihood ratio test from a Cox regression model, stratified by field center and sex. Failure time was measured from the time of randomization; follow-up was censored at the last successfully completed 400-m walk test. For participants who did not have any outcome assessments, we assigned 1 hour of follow-up time, because we knew that they completed the 400-m walk at baseline. An assessment for nonproportionality of hazards was made with the addition of the interaction between log (time) and intervention.[25] Interaction terms were entered into these Cox models and likelihood ratio tests were used to assess the consistency of the intervention effect across levels of baseline subgroups (ethnicity/race, sex, cardiovascular disease, diabetes, walking speed, and physical performance). The secondary end points were analyzed using the same approach as used for the primary outcome.

Sensitivity analyses were performed to investigate the effect of loss to follow-up on major mobility disability. These analyses used stabilized inverse probability weights that were a function of baseline covariates hypothesized to be predictive of loss-to-follow-up (ie, sex, race/ethnicity, age [≥80], history of diabetes, gait speed <0.8 m/s, low SPPB score [<8], 3MSE <90, clinical site, and living alone [yes/no]) and follow-up gait speed and SPPB scores to explore how the estimated hazard ratios and CIs may have been altered under these missing data assumptions. Statistical analyses were performed in SAS (SAS Institute), version 9.3, and R (Institute for Statistics and Mathematics).[26]

Results

Study Participants

From February 2010 to December 2011, we screened 14 831 participants; of these, 1635 were eligible and randomized (818 to the physical activity group and 817 to the health education group; Figure 1). Details regarding screening, recruitment yields, and baseline characteristics have been published.[18] Baseline characteristics were similar in the 2 groups (**Table 1**). The mean age was 78.9 years, 67.2% were women, 17.6% were African American, the average body mass index (calculated as weight in kilograms divided by height in meters squared) was 30.2, and the average SPPB score was 7.4. The mean follow-up for any contact (including telephone) was 2.6 years (median, 2.7 years; interquartile range [IQR], 2.3-3.1 years). The trial ended in December 2013, as planned in the study protocol.

Intervention Adherence

The physical activity group attended 63% of the scheduled sessions after excluding medical leave (SD, 27%; median [IQR], 71% [50%-83%]). A total of 479 participants (58.6%) went on medical leave at least once and 210 participants (25.7%) went more than once. The mean duration of medical leave was 135 days (SD, 203 days; median [IQR], 49 days [21-140]). Health education participants attended 73 of the scheduled sessions (SD, 25%; median [IQR], 82% [63%-90%]). Based on CHAMPS questionnaires, through the 24-month follow-up visit (the minimum planned intervention duration for all participants), the physical activity group maintained an average of 218 min/wk (95% CI, 210-227; average change from baseline, 138 minutes [95% CI, 129-146]) in walking and weight training activities, whereas the health education group maintained an average of 115 min/wk (95% CI, 106-123; average change from baseline, 34 minutes [95% CI, 24-42]; **Figure 2**). Thus, the physical activity intervention maintained a 104-minute difference (95% CI, 92-116; $P < .001$) in walking and weight training activities compared with the health education group during the initial 2 years in which all participants were followed up.

Based on accelerometry using a definition of more than 760 counts/min for moderate activity,[24] through follow-up,

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

on average, the physical activity group participated in 213 min/wk (95% CI, 205 to 221; average change from baseline, 15 minutes [95% CI, 7 to 23]) of moderate activity. The health education group maintained 173 min/wk (95% CI, 165 to 181; average change from baseline, −25 minutes [95% CI, −33 to −17]; Figure 2). Thus, the physical activity intervention maintained a 40-min/wk difference (95% CI, 29 to 52; *P* < .001) in moderate physical activity assessed with accelerometry, compared with the health education group during 2 years of follow-up.

Table 1. Baseline Characteristics of the Participants

	No. (%)[a]	
Characteristic	Physical Activity (n = 818)	Health Education (n = 817)
Age, mean (SD), y	78.7 (5.2)	79.1 (5.2)
Women	547 (66.9)	551 (67.4)
Ethnicity/race		
Hispanic	31 (3.8)	30 (3.7)
White	604 (73.8)	635 (77.7)
African American	163 (19.9)	125 (15.3)
SPPB score		
Mean (SD)	7.4 (1.6)	7.3 (1.6)
<8	353 (43.3)	378 (46.2)
400-m walking speed, mean (SD), m/s	0.83 (0.17)	0.82 (0.17)
BMI, mean (SD)	30.1 (5.7)	30.3 (6.2)
Walking/weight training activities, mean (SD), min/wk[b]	75.1 (125.6)	86.7 (134.5)
Median (IQR)	0 (0-105)	30 (0-105)
Accelerometry of moderate physical activity, mean (SD), min/wk[c]	193.7 (155.3)	202.1 (186.5)
Median (IQR)	161 (80-257) (n = 590)	153 (85-266) (n = 581)
3MSE score, 0-100 scale, mean (SD)	91.5 (5.5)	91.6 (5.3)
Conditions, No./total (%)		
Hypertension[b]	573/813 (70.5)	578/808 (71.5)
Diabetes[b]	199/815 (24.4)	216/813 (26.6)
Myocardial infarction[b]	60/815 (7.4)	69/812 (8.5)
Stroke[b]	57/814 (7.0)	52/814 (6.4)
Cancer[b]	178/814 (21.9)	192/815 (23.6)
Chronic pulmonary disease[b]	130/815 (16.0)	123/812 (15.2)

Abbreviations: 3MSE, Modified Mini-Mental State Examination; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); SPPB, Short Physical Performance Battery.

[a] Some values may slightly differ from those previously published[18] due to data updates.

[b] Self-reported.

[c] Moderate physical activity was defined for accelerometry based on the 760 counts/min cut point.[24]

Major Mobility Disability

Data for major mobility disability were obtained for 794 participants (97.1%) in the physical activity group and 803 participants (98.3%) in the health education group. Loss to follow-up was 4.0% annually. Major mobility disability was experienced by 246 participants (30.1%) in the physical activity group and 290 participants (35.5%) in the health education group (HR, 0.82 [95% CI, 0.69-0.98]; *P* = .03; **Figure 3**). Of the 246 and 290 physical activity and health education participants classified with major mobility disability, 42 participants (17%) of the physical activity group and 32 participants (11%) of the health education group resulted from alternative adjudications. The sensitivity analyses exploring the effect of loss to follow-up on conclusions altered the estimates of the HR and CI limits by less than 0.016 for all analyses (eAppendix in the Supplement). Persistent mobility disability was experienced by 120 participants (14.7%) in the physical activity group and 162 participants (19.8%) in the health education group (HR, 0.72 [95% CI, 0.57-0.91]; *P* = .006). Major mobility disability or death was experienced by 264 participants (32.3%) in the physical activity group and 309 participants (37.8%) in the health education group (HR, 0.82 [95% CI, 0.70-0.97]; *P* = .02).

In prespecified subgroup analyses, results for major mobility disability did not significantly differ when participants were categorized by ethnicity/race, sex, history of cardiovascular disease, history of diabetes, baseline walking speed, and baseline physical performance (**Figure 4**). The subgroup with lower physical function at baseline (SPPB <8), representing 44.7% of the study population and 71% of major mobility disability events (283 of 536 total events), received considerable benefit (HR, 0.75). In post-hoc analyses, the benefit of physical activity on major mobility disability was similar in participants with a 3MSE score of less than 90 and in those with a score of 90 or higher (Figure 4).

Safety

Serious adverse events were reported by 404 participants (49.4%) in the physical activity group and 373 participants (45.7%) in the health education group (risk ratio [RR], 1.08 [95% CI, 0.98-1.20], **Table 2**). For inpatient hospitalizations, 396 of 818 participants (48.4%) in the physical activity group and 360 of 817 participants (44.1%) in the health education group reported an event (RR, 1.10 [95% CI, 0.99-1.22]). The reasons for hospitalization were highly heterogeneous, most of them deemed unrelated to the intervention. Among those with SPPB score lower than 8, the RR was 1.04 (95% CI, 0.90-1.20); and among those with SPPB score of 8 or higher, the RR was 1.17 (95% CI, 1.00-1.36). The test of equality of RRs for hospitalization for physical activity vs health education between the 2 baseline SPPB subgroups was not significant (*P* = .44).

Discussion

The LIFE study showed that, over 2.6 years of follow-up, the physical activity intervention compared with the health education intervention significantly reduced major mobility disability (HR, 0.82; *P* = .03), persistent mobility disability (HR, 0.72; *P* = .006), and the combined outcome of major mobility disability or death (HR, 0.82; *P* = .02). The subgroup with lower physical function at baseline (SPPB <8),

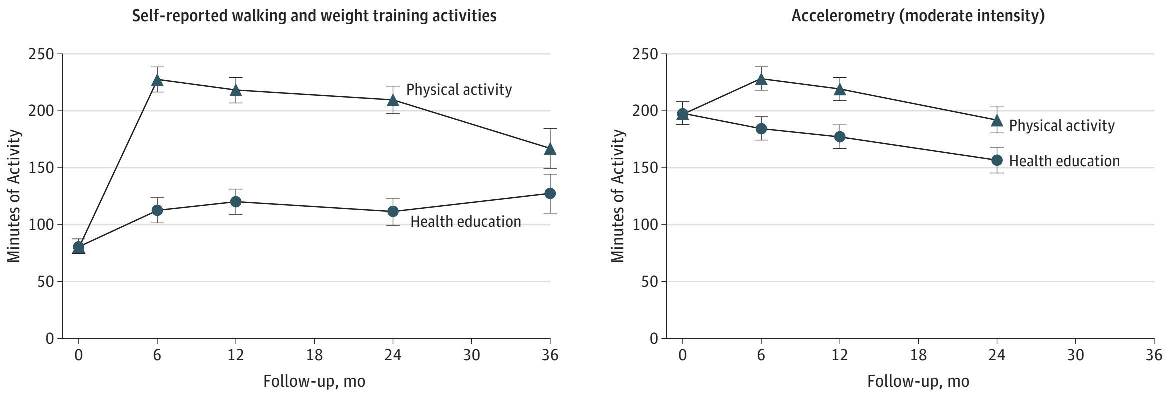
Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

representing 44.7% of the study population and 71% of major mobility disability events (283 of 536 total events), received considerable benefit (HR, 0.81). These results suggest the potential for structured physical activity as a feasible and effective intervention to reduce the burden of disability among vulnerable older persons, in spite of functional decline in late life. To our knowledge, the LIFE study is the largest and longest duration randomized trial of physical activity in older persons.

The LIFE study has important strengths, including the objectively measured primary outcome of major mobility disability that is a reliable,[22] well-validated, and important clinical and public health outcome in older people.[11] Participants at high risk for disability were recruited from 8 field centers spanning the United States, including urban, suburban, and rural settings, and included a high proportion of older adults from African American and Hispanic backgrounds. Although highly prevalent and increasing in size,

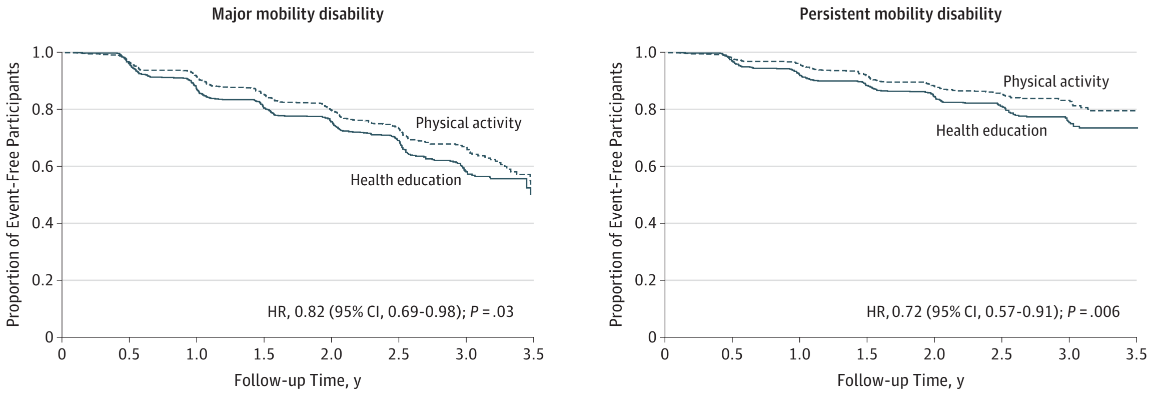
Figure 2. Self-reported and Accelerometry-Derived Physical Activity by Treatment Group



No. of participants	0	6	12	24	36	0	6	12	24
Physical activity	818	774	761	724	292	590	423	420	349
Health education	817	796	777	738	295	581	422	425	346

Plotted values represent least squares means (95% CI) from a mixed-effects model adjusting for clinical site and sex (both used to stratify randomization) and the baseline self-reported walking/weight training activities or accelerometry counts. In addition to the above-mentioned terms, the model contained a term for the intervention group, follow-up clinic visit, and the intervention × visit interaction. All participants had expected follow-up through 24 months and approximately 47% of randomized participants had expected visits at 36 months. Accelerometry data were not collected at the 36-month visit. Baseline values represent the overall mean of both groups combined (a standard practice when using the analysis of covariance method): this is the assumed value for both groups when obtaining least squares means at follow-up using mixed-effects analysis of covariance. The baseline, prerandomization value is reflected by follow-up time 0. The *P* value is less than .001 for the average intervention effect.

Figure 3. Effect of a Moderate Physical Activity Intervention on the Onset of Major Mobility Disability and Persistent Mobility Disability



	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5
No. at risk																
Physical activity	818	758	706	646	559	378	182	11	818	761	726	673	579	393	188	12
Health education	817	765	680	617	540	358	162	13	817	762	707	655	567	371	178	10
Events																
Physical activity	0	29	67	115	155	197	224	246	0	18	32	64	88	104	113	120
Health education	0	33	105	155	190	232	277	286	0	25	64	91	118	138	158	162

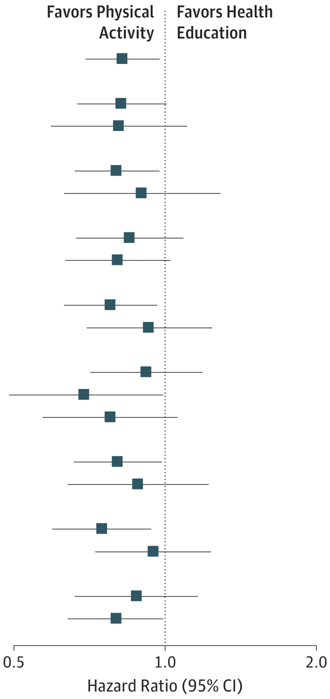
HR indicates hazard ratio. The graph for major mobility disability was truncated at 3.5 years and the health education group had 4 additional failures between 3.5 and 3.6 years of follow-up. Number of events represents cumulative events and adjusted HRs and *P* values are from proportional hazards regression models defined in the Methods section.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

Figure 4. Hazard Ratio of Major Mobility Disability for Physical Activity vs Health Education According to Subgroups

Subgroup	Physical Activity: Events, No.	Physical Activity: Total Participants	Health Education: Events, No.	Health Education: Total Participants	Hazard Ratio (95% CI)	Interaction *P* Value
Overall	246	818	290	817	0.82 (0.69-0.98)	
Sex						
Women	171	547	204	551	0.82 (0.67-1.01)	.95
Men	75	271	86	266	0.81 (0.59-1.11)	
Ethnicity/race						
Non-Hispanic white	182	604	234	635	0.80 (0.66-0.98)	.58
Other	64	211	56	180	0.90 (0.63-1.29)	
Age, y						
70-79	123	477	138	455	0.85 (0.67-1.09)	.76
≥80	123	341	152	362	0.81 (0.63-1.03)	
History of CVD						
No CVD	155	582	187	563	0.78 (0.63-0.97)	.34
CVD	91	236	103	254	0.93 (0.70-1.24)	
History of diabetes						
None	114	406	126	414	0.92 (0.71-1.19)	.41
Impaired fasting glucose	59	192	68	165	0.69 (0.49-0.99)	
Diabetes	73	220	96	238	0.78 (0.57-1.06)	
Gait speed						
<0.8 m/s	173	485	210	508	0.81 (0.66-0.99)	.63
≥0.8 m/s	73	333	80	309	0.88 (0.64-1.22)	
SPPB						
<8	135	353	177	378	0.75 (0.60-0.94)	.19
8 or 9	111	465	113	439	0.95 (0.73-1.23)	
3MSE (post hoc)						
<90	95	261	108	261	0.88 (0.66-1.16)	.58
≥90	151	557	182	556	0.80 (0.64-0.99)	



3MSE indicates Modified Mini-Mental State Examination; CVD, cardiovascular disease; SPPB, Short Physical Performance Battery. *P* values were obtained from likelihood ratios tests of the interaction terms added to the Cox regression model.

the older, more vulnerable population has been understudied and typically is not included in large randomized trials. Retention throughout the follow-up was excellent. The adherence rates to the physical activity intervention were similar or higher than those achieved in other much shorter studies involving older adults.[27-29] The physical activity program was likely successful in part because of the adherence and lifestyle motivation procedures.[30] The participants were reimbursed for their transportation costs, which added to the cost of the intervention, but likely contributed to the high levels of attendance. According to initial cost data collected in the LIFE study, the physical activity intervention cost, including transportation, was approximately $4900 per participant over the 2.6 years of average participation ($1815/year). The physical activity intervention was designed to be simple for widespread implementation in a variety of communities and settings, because it does not require any special equipment.

The LIFE study has limitations. We could not ascertain whether participants who were excluded because of their high level of physical function or severe cognitive deficits would also benefit from physical activity. The participants were recruited from the community, but may have been self-referred, so they may not be fully representative of all people in the community. The average follow-up duration of 2.6 years was relatively short vs the estimated average 9-year life expectancy of the LIFE cohort.[31] Ideally, it would be useful to assess the effect of the intervention on the quality of the remaining years of life. The study, which was powered based on assumptions of 21% to 24% risk reduction, achieved an HR of 0.82 and an absolute risk difference of 5.4%. Although the effect size was slightly lower than planned, we believe that it is clinically relevant given the major health effect of mobility disability and the lack of proven interventions to avert mobility disability in vulnerable older populations. In addition, persistent mobility disability was significantly reduced by a larger degree in the physical activity group (HR, 0.72), indicating that physical activity not only prevents the onset of major mobility disability, but also favors improved recovery in those who lose mobility.

Based on observational cohorts,[32] we expected a lower hospitalization rate in the physical activity group. In the LIFE study, physical activity did not decrease the hospitalizations rate. We found a higher rate of hospitalizations in the physical activity group that did not reach statistical significance. The hospitalizations comprised a range of heterogeneous diagnoses mostly deemed unrelated to the intervention. Our finding may have several explanations. First, physical activity may unmask symptoms resulting in earlier detection of underlying medical conditions. For example, sedentary older persons with subclinical left ventricular dysfunction may observe heart failure symptoms when they start moderate physical activity. Second, the physical activity group's more frequent contact and testing of vital signs at each intervention session may have led to a higher rate of recognition of health events. Third, the stress of exercise in

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

Table 2. All Deaths and Number of Participants Reporting Adverse Events at Blinded Assessments

	Physical Activity Group (n = 818)		Health Education Group (n = 817)		
Event Type	Participants, No. (%)	Events, No.	Participants, No. (%)	Events, No.	Risk Ratio (95% CI)[a]
Serious adverse events					
All serious adverse events	404 (49.4)	879	373 (45.7)	774	1.08 (0.98-1.20)
Death	48 (5.9)	48	42 (5.1)	42	1.14 (0.76-1.71)
Life-threatening event	11 (1.3)	11	8 (1.0)	8	1.37 (0.56-3.40)
Persistent disability/incapacity	33 (4.0)	51	26 (3.2)	45	1.27 (0.77-2.10)
All inpatient hospitalizations	396 (48.4)	777	360 (44.1)	681	1.10 (0.99-1.22)
Any other serious events	7 (0.9)	8	8 (1.0)	10	0.87 (0.32-2.40)
Most frequent hospitalization diagnoses					
Infection	74 (.9.0)	95	57 (7.0)	68	1.30 (0.93-1.81)
Surgical procedure	68 (8.3)	76	73 (8.9)	84	0.93 (0.68-1.28)
Fall, syncope, dizziness, vertigo	54 (6.4)	58	53 (6.5)	62	1.02 (0.71-1.49)
Atrial fibrillation/flutter	24 (2.9)	28	20 (2.4)	23	1.20 (0.67-2.15)
Heart failure	18 (2.2)	21	14 (1.7)	20	1.28 (0.64-2.56)
Stroke/TIA/intracranial hemorrhage	29 (3.5)	33	28 (3.4)	34	1.03 (0.62-1.72)
MI/chest pain/acute coronary syndrome	33 (4.0)	42	25 (3.1)	27	1.32 (0.79-2.20)
Fracture	27 (3.3)	29	26 (3.2)	27	1.04 (0.61-1.76)
Neoplasm	17 (2.1)	17	17 (2.1)	20	1.00 (0.51-1.94)
Arthritis/back, neck, or bone pain	30 (3.7)	31	33 (4.0)	35	0.91 (0.56-1.47)
Symptoms resulting in at least 1 wk of restricted activity[b]					
All cases	198 (24.2)	253	198 (24.2)	249	1.00 (0.84-1.19)
Fall	47 (5.7)	53	71 (8.7)	81	0.66 (0.46-0.94)
Fatigue	38 (4.6)	46	41 (5.0)	45	0.93 (0.60-1.42)
Muscle or joint aching	32 (3.9)	37	40 (4.9)	43	0.80 (0.51-1.26)
Back pain	36 (4.4)	41	33 (4.0)	35	1.09 (0.69-1.73)
Muscle or joint stiffness	26 (3.2)	30	33 (4.0)	35	0.79 (0.48-1.30)
Foot pain	17 (2.1)	17	18 (2.2)	18	0.94 (0.49-1.82)
Dizziness	18 (2.2)	19	14 (1.7)	15	1.28 (0.64-2.56)
Shortness of breath	15 (1.8)	16	20 (2.4)	22	0.75 (0.39-1.45)
Fainting	16 (2.0)	18	10 (1.2)	11	1.60 (0.73-3.50)
Abnormal heart rhythm	9 (1.1)	9	8 (1.0)	8	1.12 (0.44-2.90)
Other symptom	84 (10.3)	96	71 (8.7)	75	1.18 (0.87-1.60)

Abbreviations: MI, myocardial infarction; TIA, transient ischemic attack.

[a] Risk ratio compares the proportion of participants reporting any events in the physical activity group vs the health education group, with asymptomatic 95% CIs.

[b] Symptoms resulting in at least 1 week of restricted activity may also lead to serious adverse events. Thus, events reported in this section of Table 2 may also be reflected as serious adverse events or hospitalizations.

the context of lowered homeostatic reserve in vulnerable participants[33] may have led to a higher risk of adverse events. However, our data do not support this explanation. The hospitalization results were not significantly different among those with SPPB score less than 8, and those with a score 8 or 9. Finally, there may be no causal association between physical activity and hospitalizations.

Physical activity did not decrease the death rate. We found a higher rate of mortality in the physical activity group that did not reach statistical significance, and which was compatible with benefit or harm of physical activity (Table 2). Given the small number of events the data regarding mortality are inconclusive. Further studies are needed to assess the effects of physical activity on mortality and hospitalizations in vulnerable older adults.

Conclusions

A structured moderate-intensity physical activity program compared with a health education program reduced major mobility disability over 2.6 years among older adults at risk of disability. These findings suggest mobility benefit from such a program in vulnerable older adults.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

ARTICLE INFORMATION

Published Online: May 27, 2014. doi:10.1001/jama.2014.5616.

Author Affiliations: Department of Aging and Geriatric Research, University of Florida, Gainesville (Pahor, Guralnik, Manini); Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore (Guralnik); Department of Internal Medicine, Wake Forest University and School of Medicine, Winston-Salem, North Carolina (Ambrosius, Espeland, Kritchevsky, Miller, Rejeski, Sink, Williamson); Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia (Blair); Division of Cardiac Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Bonds); Department of Preventative Medicine, Pennington Biomedical Research Center, Louisiana (Church); Jean Mayer United States Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts (Fielding); Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut (Gill); Veterans Affairs San Diego Healthcare System, San Diego, California (Groessl); Department of Family and Preventive Medicine, University of California, San Diego, San Diego (Groessl); Department of Health Research and Policy and Department of Medicine, Stanford University, School of Medicine, Stanford, California (King); Department of Medicine and Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (McDermott); Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Newman).

Author Contributions: Dr Miller 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: Pahor, Guralnik, Blair, Church, Espeland, Fielding, Gill, Groessl, King, Kritchevsky, Manini, Miller, Newman, Rejeski, Sink, Williamson.
Acquisition, analysis, or interpretation of data: Pahor, Guralnik, Ambrosius, Bonds, Church, Espeland, Fielding, Gill, Manini, McDermott, Miller, Newman, Sink, Williamson.
Drafting of the manuscript: Pahor, Espeland, Fielding, Manini, McDermott, Miller, Rejeski.
Critical revision of the manuscript for important intellectual content: Pahor, Guralnik, Ambrosius, Blair, Bonds, Church, Espeland, Fielding, Gill, Groessl, King, Kritchevsky, Manini, McDermott, Miller, Newman, Sink.
Statistical analysis: Ambrosius, Espeland, Miller.
Obtained funding: Pahor, Guralnik, Church, Fielding, Gill, Groessl, King, Manini, McDermott, Miller, Newman, Rejeski, Williamson.
Administrative, technical, or material support: Pahor, Guralnik, Bonds, Fielding, Gill, King, Manini, Miller, Newman, Williamson.
Study supervision: Pahor, Guralnik, Fielding, Gill, King, Kritchevsky, Manini, McDermott, Rejeski, Williamson.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Blair reports receiving grant funding from BodyMedia, the Coca-Cola Company, and Technogym. Dr Fielding reports receiving grant funding from the Dairy Research Institute, Department of Defense, Nestle, Regeneron Pharmaceuticals, and Unilever; consulting for Dairy Management, Eli Lilly, Essentient, Merck, Nestle, and Regeneron; and serving on the board for Aging in Motion, Ammonett, Cytokinetics, Myosyntax, Nestle, and Segterra. Dr McDermott reports being a medical editor for Informed Medical Decisions Foundation. Dr Sink reports receiving grant funding from the Alzheimer's Association and the Donald W. Reynolds Foundation; and being a site principal investigator for a multicenter trial of medical food for patients with Alzheimer disease. No other disclosures were reported.

Funding/Support: The Lifestyle Interventions and Independence for Elders Study is funded by cooperative agreement UO1AG22376 from the National Institutes of Health (NIH) and National Institute on Aging; supplement 3U01AG022376-05A2S from the National Heart, Lung, and Blood Institute; and was sponsored in part by the Intramural Research Program. The research is partially supported by the Claude D. Pepper Older Americans Independence Centers at the University of Florida (1 P30 AG028740), Wake Forest University (1 P30 AG21332),Tufts University (1P30AG031679), University of Pittsburgh (P30 AG024827), and Yale University (P30AG021342) and the NIH/NCRR CTSA at Stanford University (UL1 RR025744), at University of Florida (U54RR025208) and at Yale University (UL1 TR000142). Tufts University is also supported by the Boston Rehabilitation Outcomes Center (1R24HD065688-01A1). LIFE investigators are also partially supported by the following: Dr Thomas Gill (Yale University) is the recipient of an Academic Leadership Award (K07AG3587) from the National Institute on Aging. Dr Carlos Fragoso (Spirometry Reading Center, Yale University) is the recipient of a Career Development Award from the Department of Veterans Affairs. Dr Roger Fielding (Tufts University) is partially supported by the US Department of Agriculture, under agreement 58-1950-0-014.

Role of the Sponsors: The NIH sponsor was a voting member (1 of 12 votes) of the LIFE Steering Committee, which approved the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, and approval of the manuscript; and decision to submit the manuscript for publication.

Research Investigators for the LIFE Study: Administrative Coordinating Center, University of Florida, Gainesville, FL: Marco Pahor, MD—Principal Investigator of the LIFE Study; Jack M. Guralnik, MD, PhD—Coinvestigator of the LIFE Study (University of Maryland School of Medicine, Baltimore, MD); Christiaan Leeuwenburgh, PhD; Connie Caudle; Lauren Crump, MPH; Latonia Holmes; Jocelyn Lee, PhD; Ching-ju Lu, MPH. ***Data Management, Analysis and Quality Control Center***: Wake Forest University, Winston Salem, NC: Michael E. Miller, PhD—DMAQC Principal Investigator; Mark A. Espeland, PhD—DMAQC Coinvestigator; Walter T. Ambrosius, PhD; William Applegate, MD; Daniel P. Beavers, PhD, MS; Robert P. Byington, PhD, MPH, FAHA; Delilah Cook, CCRP; Curt D. Furberg, MD, PhD; Lea N. Harvin, BS; Leora Henkin, MPH, Med; John Hepler, MA; Fang-Chi Hsu, PhD; Laura Lovato, MS; Wesley Roberson, BSBA; Julia Rushing, BSPH, MStat; Scott Rushing, BS; Cynthia L. Stowe, MPM; Michael P. Walkup, MS; Don Hire, BS; W. Jack Rejeski, PhD; Jeffrey A. Katula, PhD, MA; Peter H. Brubaker, PhD; Shannon L. Mihalko, PhD; Janine M. Jennings, PhD; National Institutes of Health, Bethesda, MD; Evan C. Hadley, MD (National Institute on Aging); Sergei Romashkan, MD, PhD (National Institute on Aging); Kushang V. Patel, PhD (National Institute on Aging); National Heart, Lung and Blood Institute, Bethesda, MD; Denise Bonds, MD, MPH. ***Field Centers***: Northwestern University, Chicago, IL: Mary M. McDermott, MD—Field Center Principal Investigator; Bonnie Spring, PhD—Field Center Coinvestigator; Joshua Hauser, MD—Field Center Coinvestigator; Diana Kerwin, MD—Field Center Coinvestigator; Kathryn Domanchuk, BS; Rex Graff, MS; Alvito Rego, MA. Pennington Biomedical Research Center, Baton Rouge, LA: Timothy S. Church, MD, PhD, MPH—Field Center Principal Investigator; Steven N. Blair, PED (University of South Carolina); Valerie H. Myers, PhD; Ron Monce, PA-C; Nathan E. Britt, NP; Melissa Nauta Harris, BS; Ami Parks McGucken, MPA, BS; Ruben Rodarte, MBA, MS, BS; Heidi K. Millet, MPA, BS; Catrine Tudor-Locke, PhD, FACSM; Ben P. Butitta, BS; Sheletta G. Donatto, MS, RD, LDN, CDE; Shannon H. Cocreham, BS; Stanford University, Palo Alto, CA; Abby C. King, PhD—Field Center Principal Investigator; Cynthia M. Castro, PhD; William L. Haskell, PhD; Randall S. Stafford, MD, PhD; Leslie A. Pruitt, PhD; Kathy Berra, MSN, NP-C, FAAN; Veronica Yank, MD; Tufts University, Boston, MA; Roger A. Fielding, PhD—Field Center Principal Investigator; Miriam E. Nelson, PhD—Field Center Coinvestigator; Sara C. Folta, PhD—Field Center Coinvestigator; Edward M. Phillips, MD; Christine K. Liu, MD; Erica C. McDavitt, MS; Kieran F. Reid, PhD, MPH; Won S. Kim, BS; Vince E. Beard, BS; University of Florida, Gainesville, FL; Todd M. Manini, PhD—Field Center Principal Investigator; Marco Pahor, MD—Field Center Coinvestigator; Stephen D. Anton, PhD; Susan Nayfield, MD; Thomas W. Buford, PhD; Michael Marsiske, PhD; Bhanuprasad D. Sandesara, MD; Jeffrey D. Knaggs, BS; Megan S. Lorow, BS; William C. Marena, MT, CCRC; Irina Korytov, MD; Holly L. Morris, MSN, RN, CCRC (Brooks Rehabilitation Clinical Research Center, Jacksonville, FL); Margo Fitch, PT (Brooks Rehabilitation Clinical Research Center, Jacksonville, FL); Floris F. Singletary, MS, CCC-SLP (Brooks Rehabilitation Clinical Research Center, Jacksonville, FL); Jackie Causer, BSH, RN (Brooks Rehabilitation Clinical Research Center, Jacksonville, FL); Katie A. Radcliff, MA (Brooks Rehabilitation Clinical Research Center, Jacksonville, FL); University of Pittsburgh, Pittsburgh, PA; Anne B. Newman, MD, MPH—Field Center Principal Investigator; Stephanie A. Studenski, MD, MPH—Field Center Coinvestigator; Bret H. Goodpaster, PhD; Nancy W. Glynn, PhD; Oscar Lopez, MD; Neelesh K. Nadkarni, MD, PhD; Kathy Williams, RN, BSEd, MHSA; Mark A. Newman, PhD; George Grove, MS; Janet T. Bonk, MPH, RN; Jennifer Rush, MPH; Piera Kost, BA (deceased); Diane G. Ives, MPH; Wake Forest University, Winston Salem, NC; Stephen B. Kritchevsky, PhD—Field Center Principal Investigator; Anthony P. Marsh, PhD—Field Center Coinvestigator; Tina E. Brinkley, PhD; Jamehl S. Demons, MD; Kaycee M. Sink, MD, MAS; Kimberly Kennedy, BA, CCRC; Rachel Shertzer-Skinner, MA, CCRC; Abbie Wrights,

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

MS; Rose Fries, RN, CCRC; Deborah Barr, MA, RHEd, CHES; Yale University, New Haven, CT; Thomas M. Gill, MD—Field Center Principal Investigator; Robert S. Axtell, PhD, FACSM—Field Center Coinvestigator (Southern Connecticut State University, Exercise Science Department); Susan S. Kashaf, MD, MPH (VA Connecticut Healthcare System); Nathalie de Rekeneire, MD, MS; Joanne M. McGloin, MDiv, MS, MBA; Karen C. Wu, RN; Denise M. Shepard, RN, MBA; Barbara Fennelly, MA, RN; Lynne P. Iannone, MS, CCRP; Raeleen Mautner, PhD; Theresa Sweeney Barnett, MS, APRN; Sean N. Halpin, MA; Matthew J. Brennan, MA; Julie A. Bugaj, MS; Maria A. Zenoni, MS; Bridget M. Mignosa, AS. Cognition Coordinating Center, Wake Forest University, Winston Salem, NC; Jeff Williamson, MD, MHS—Center Principal Investigator; Kaycee M Sink, MD, MAS—Center Coinvestigator; Hugh C. Hendrie, MB, ChB, DSc (Indiana University); Stephen R. Rapp, PhD; Joe Verghese, MB, BS (Albert Einstein College of Medicine of Yeshiva University); Nancy Woolard; Mark Espeland, PhD; Janine Jennings, PhD; Electrocardiogram Reading Center, University of Florida, Gainesville, FL; Carl J. Pepine MD, MACC; Mario Ariet, PhD; Eileen Handberg, PhD, ARNP; Daniel Deluca, BS; James Hill, MD, MS, FACC; Anita Szady, MD. Spirometry Reading Center, Yale University, New Haven, CT; Geoffrey L. Chupp, MD; Gail M. Flynn, RCP, CRFT; Thomas M. Gill, MD; John L. Hankinson, PhD (Hankinson Consulting, Inc.); Carlos A. Vaz Fragoso, MD; Cost Effectiveness Analysis Center; Erik J. Groessl, PhD (University of California, San Diego and VA San Diego Healthcare System); Robert M. Kaplan, PhD (Office of Behavioral and Social Sciences Research, National Institutes of Health).

Disclaimer: Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the US Department of Agriculture. Dr McDermott, senior editor at *JAMA*, did not participate in the evaluation of this article or the decision to publish the study.

Additional Contributions: We thank Evan C. Hadley, MD, and Sergei Romashkan, MD, PhD, from the National Institute on Aging (Bethesda, Maryland), for their substantial intellectual contribution to the development and implementation of the LIFE Study. Dr Hadley and Dr Romashkan are federal employees fully paid by the NIH. They did not receive any additional compensation from the study.

REFERENCES

1. Werner CA. The older population. http://www.census.gov/prod/cen2010/briefs/c2010br-09.pdf. Accessed May 7, 2014.

2. Katz S, Branch LG, Branson MH, Papsidero JA, Beck JC, Greer DS. Active life expectancy. *N Engl J Med*. 1983;309(20):1218-1224.

3. Branch LG, Guralnik JM, Foley DJ, et al. Active life expectancy for 10 000 Caucasian men and women in 3 communities. *J Gerontol*. 1991;46(4):M145-M150.

4. Lonergan ET, Krevans JR. A national agenda for research on aging. *N Engl J Med*. 1991;324(25):1825-1828.

5. Guralnik JM, LaCroix AZ, Abbott RD, et al. Maintaining mobility in late life. *Am J Epidemiol*. 1993;137(8):845-857.

6. Branch LG, Jette AM. A prospective study of long-term care institutionalization among the aged. *Am J Public Health*. 1982;72(12):1373-1379.

7. Corti MC, Guralnik JM, Salive ME, Sorkin JD. Serum albumin level and physical disability as predictors of mortality in older persons. *JAMA*. 1994;272(13):1036-1042.

8. Khokhar SR, Stern Y, Bell K, et al. Persistent mobility deficit in the absence of deficits in activities of daily living. *J Am Geriatr Soc*. 2001;49(11):1539-1543.

9. Hirvensalo M, Rantanen T, Heikkinen E. Mobility difficulties and physical activity as predictors of mortality and loss of independence in the community-living older population. *J Am Geriatr Soc*. 2000;48(5):493-498.

10. Lampinen P, Heikkinen E. Reduced mobility and physical activity as predictors of depressive symptoms among community-dwelling older adults. *Aging Clin Exp Res*. 2003;15(3):205-211.

11. Newman AB, Simonsick EM, Naydeck BL, et al. Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. *JAMA*. 2006;295(17):2018-2026.

12. Shumway-Cook A, Patla A, Stewart A, Ferrucci L, Ciol MA, Guralnik JM. Environmental components of mobility disability in community-living older persons. *J Am Geriatr Soc*. 2003;51(3):393-398.

13. Shumway-Cook A, Patla AE, Stewart A, Ferrucci L, Ciol MA, Guralnik JM. Environmental demands associated with community mobility in older adults with and without mobility disabilities. *Phys Ther*. 2002;82(7):670-681.

14. Pahor M, Blair SN, Espeland M, et al. Effects of a physical activity intervention on measures of physical performance: results of the lifestyle interventions and independence for Elders Pilot (LIFE-P) study. *J Gerontol A Biol Sci Med Sci*. 2006;61(11):1157-1165.

15. Fielding RA, Rejeski WJ, Blair S, et al. The Lifestyle Interventions and Independence for Elders Study. *J Gerontol A Biol Sci Med Sci*. 2011;66(11):1226-1237.

16. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med*. 1995;332(9):556-561.

17. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry*. 1987;48(8):314-318.

18. Marsh AP, Lovato LC, Glynn NW, et al. Lifestyle interventions and independence for elders study. *J Gerontol A Biol Sci Med Sci*. 2013;68(12):1549-1558.

19. Borg G. *Perceived Exertion and Pain Scales.* Champaign, IL: Human Kinetics; 1988.

20. Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults. *Med Care*. 1998;36(9):1349-1360.

21. Stewart AL, Verboncoeur CJ, McLellan BY, et al. Physical activity outcomes of CHAMPS II. *J Gerontol A Biol Sci Med Sci*. 2001;56(8):M465-M470.

22. Rolland YM, Cesari M, Miller ME, Penninx BWJH, Atkinson HH, Pahor M. Reliability of the 400-m usual-pace walk test as an assessment of mobility limitation in older adults. *J Am Geriatr Soc*. 2004;52(6):972-976.

23. Guralnik JM, Ferrucci L, Pieper CF, et al. Lower extremity function and subsequent disability. *J Gerontol A Biol Sci Med Sci*. 2000;55(4):M221-M231.

24. Matthew CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc*. 2005;37(11)(suppl):S512-S522.

25. Therneau TM, Grambsch PM. *Modeling Survival Data, Extending the Cox Model.* New York, NY: Springer Science; 2000.

26. Van der Wal WM, Geskus RB. An R package for inverse probability weighting. J Stat Softw. 2011;43(13):1-23. http://www.jstatsoft.org/v43/i13/paper. Accessed May 7, 2014.

27. Messier SP, Loeser RF, Miller GD, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis. *Arthritis Rheum*. 2004;50(5):1501-1510.

28. Ettinger WH Jr, Burns R, Messier SP, et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis: the Fitness Arthritis and Seniors Trial (FAST). *JAMA*. 1997;277(1):25-31.

29. Berry MJ, Rejeski WJ, Adair NE, Ettinger WH Jr, Zaccaro DJ, Sevick MA. A randomized, controlled trial comparing long-term and short-term exercise in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil*. 2003;23(1):60-68.

30. Rejeski WJ, Axtell R, Fielding R, et al. Promoting physical activity for elders with compromised function. *Clin Interv Aging*. 2013;8:1119-1131.

31. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA*. 2011;305(1):50-58.

32. Physical Activity Guidelines Advisory Committee. *Physical Activity Guidelines Advisory Committee Report*. Washington, DC: US Dept of Health and Human Services; 2008.

33. Szanton SL, Allen JK, Seplaki CL, Bandeen-Roche K, Fried LP. Allostatic load and frailty in the Women's Health and Aging studies. *Biol Res Nurs*. 2009;10(3):248-256.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/27/2014

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

Whey protein, amino acids, and vitamin D supplementation with physical activity increases fat-free mass and strength, functionality, and quality of life and decreases inflammation in sarcopenic elderly[1,2]

Mariangela Rondanelli,[3] Catherine Klersy,[6] Gilles Terracol,[7] Jacopo Talluri,[8] Roberto Maugeri,[7] Davide Guido,[4] Milena A Faliva,[3] Bruno S Solerte,[5] Marisa Fioravanti,[5] Henry Lukaski,[9] and Simone Perna[3]*

[3]Department of Public Health and Experimental and Forensic Medicine, Section of Human Nutrition and Dietetics, Azienda Human Service of Pavia, [4]Medical and Genomic Statistics Unit, Department of Brain and Behavioral Sciences, and [5]Department of Internal Medicine, University of Pavia, Pavia, Italy; [6]Biometry and Clinical Epidemiology Service, San Matteo Hospital, Pavia, Italy; [7]Research and Clinical Investigation Department, Company Medical Dietetics, Savigliano, Italy; [8]Research and Clinical Investigation Department, Akern, Pontassieve, Italy; and [9]Department of Kinesiology and Public Health Education, Hyslop Sports Center, University of North Dakota, Grand Forks, ND

ABSTRACT

Background: Interventions to attenuate the adverse effects of age-related loss of skeletal muscle and function include increased physical activity and nutritional supplementation.

Objective: This study tested the hypothesis that nutritional supplementation with whey protein (22 g), essential amino acids (10.9 g, including 4 g leucine), and vitamin D [2.5 μg (100 IU)] concurrent with regular, controlled physical activity would increase fat-free mass, strength, physical function, and quality of life, and reduce the risk of malnutrition in sarcopenic elderly persons.

Design: A total of 130 sarcopenic elderly people (53 men and 77 women; mean age: 80.3 y) participated in a 12-wk randomized, double-blind, placebo-controlled supplementation trial. All participants concurrently took part in a controlled physical activity program. We examined body composition with dual-energy X-ray absorptiometry, muscle strength with a handgrip dynamometer, and blood biochemical indexes of nutritional and health status, and evaluated global nutritional status, physical function, and quality of life before and after the 12 wk of intervention.

Results: Compared with physical activity and placebo, supplementation plus physical activity increased fat-free mass (1.7-kg gain, $P < 0.001$), relative skeletal muscle mass ($P = 0.009$), android distribution of fat ($P = 0.021$), handgrip strength ($P = 0.001$), standardized summary scores for physical components ($P = 0.030$), activities of daily living ($P = 0.001$), mini nutritional assessment ($P = 0.003$), and insulin-like growth factor I ($P = 0.002$), and lowered C-reactive protein ($P = 0.038$).

Conclusion: Supplementation with whey protein, essential amino acids, and vitamin D, in conjunction with age-appropriate exercise, not only boosts fat-free mass and strength but also enhances other aspects that contribute to well-being in sarcopenic elderly. This trial was registered at clinicaltrials.gov as NCT02402608. *Am J Clin Nutr* 2016;103:830–40.

Keywords: amino acids, dietary supplement, elderly, insulin-like growth factor I, fat-free mass, relative skeletal muscle mass, sarcopenia, vitamin D, whey protein

INTRODUCTION

Human aging involves changes in body structure and function. Older adults experience a progressive, generalized loss of skeletal muscle and a decrease in physical function, with an inherent risk of disability, poor quality of life, and death (1). Rosenberg (2) proposed the term "sarcopenia" to describe this age-related depletion of skeletal muscle mass and loss of strength.

The etiology and mechanisms of sarcopenia are complex and multifactorial (3). Primary sarcopenia is a consequence of the aging process (e.g., reduced neurological function, altered muscle fiber type distribution, and increased protein turnover). Secondary sarcopenia, however, is linked with inactivity (e.g., bed rest or a sedentary lifestyle) or chronic disease (e.g., organ failure, malignancy, inflammation, or endocrine disease). There is growing evidence that nutritional factors (e.g., an inadequate intake of protein, energy, and certain micronutrients; malabsorption; and drug-induced anorexia) contribute to secondary sarcopenia (4).

Interventions for sarcopenia include nutrition, because nutrition can have a positive impact on protein anabolism. Increasing the quantity (e.g., in excess of the recommended dietary intake) (5) and quality (e.g., essential amino acids, specifically leucine) of dietary protein stimulates muscle protein synthesis in the elderly (6). Increased intake of vitamin D stimulates gene expression and boosts muscle protein synthesis, facilitates neuromuscular function (7, 8), and enhances strength and balance (9, 10). It also reduces the inflammation that is associated with decreased muscle strength in the elderly (11). Because older adults risk having a low intake of high-quality protein, as well as

---

[1] Supported by the Italian Ministry of the University and Research to the University of Pavia Department of Public Health.

[2] The funders had no role in the conception of the study or the writing of the manuscript.

*To whom correspondence should be addressed. E-mail: mariangela.rondanelli@unipv.it.

Received April 26, 2015. Accepted for publication January 4, 2016.

First published online February 10, 2016; doi: 10.3945/ajcn.115.113357.

vitamin D insufficiency, strategies to boost intake are recommended to attenuate the loss of muscle and its adverse effects in these people (12, 13).

Physical activity can also slow the loss of muscle mass and improve function. Strength training with resistance exercise universally strengthens muscles (14). The addition of dynamic exercise to resistance training also contributes substantial benefits to physical function (15).

The complex etiology of sarcopenia calls for integrated interventions in a practical approach (16). We therefore designed a randomized, double-blind, placebo-controlled supplementation trial that combined whey protein, essential amino acids, and vitamin D with regular physical activity for all participants. We set out to test whether, compared with placebo, supplementation would increase fat-free mass (FFM)[10] (primary outcome) while improving strength, nutritional status, inflammation, and measures of quality of life and physical function (secondary outcomes).

METHODS

Participants

The study was approved by the institutional review board at the University of Pavia and was conducted after approval from the ethics committee of the Department of Internal Medicine and Medical Therapy at the University of Pavia. Participants gave their written consent to this study (NCT02402608).

We evaluated elderly men and women admitted to the geriatric physical medicine and rehabilitation division at the Santa Margherita Hospital, Azienda Human Service of Pavia in Pavia, Italy. Before participation, each person had complete medical screening, including vital signs, blood tests, urine tests, and a 12-lead electrocardiogram. Anyone with evidence of heart disease, kidney or liver disease, or any other disease that might influence the results of the study was excluded. Data were gathered from the end of January 2013 to the end of June 2014. Eligible persons were aged ≥65 y and had an appendicular skeletal FFM divided by height squared that was 2 SD below the mean for young adults (17), hence, relative muscle mass <7.26 kg/m^2 for men and <5.5 kg/m^2 for women. They had to have no acute illness or severe liver, heart, or kidney dysfunction, and body weight had to have been stable for 6 mo. Anyone with altered glycometabolic control, thyroid disorders, other endocrinopathies, or cancers, and any patients treated with steroids and heparin or who had total walking incapacity were excluded. The participants selected had to have similar physical ability, assessed with the activities of daily living (ADL) score, and normal cognitive function or only mild cognitive disturbance as defined by a Mini-Mental State Examination >20 (18).

Body composition, nutritional status, and food intake

Body composition (FFM, fat mass, and gynoid and android fat distribution) was measured by dual-energy X-ray absorptiometry (DXA) with the use of a Lunar Prodigy DXA (GE Medical Systems). The in vivo CVs were 0.89% and 0.48% for whole-body fat (fat mass) and FFM, respectively. The relative skeletal muscle mass (RSMM) was taken as the sum of the fat-free soft tissue mass of arms and legs (19).

Body weight was measured to the nearest 0.1 kg on a precision scale with the participants wearing light clothing, without shoes, with the use of a standardized technique (20). Waist measurements were taken at the midpoint between the lowest rib and the top of the hip bone (iliac crest), with the use of a standardized technique (20).

We assessed the hydration of these elderly adults with bioelectrical impedance, because changes in fluid status affect the soft tissue composition estimated by DXA (21, 22). Whole-body resistance and reactance were measured with the patient lying supine on a nonconductive surface with the use of a phase-sensitive, single-frequency impedance plethysmograph [400-μA, 50-kHz alternating current (BIA-101; RJL/Akern Systems)]. Adhesive surface electrodes were placed on the right hand and foot, and measurements were taken according to the guidelines of the NIH Technology Assessment Conference Statement (23).

Resistance and reactance were standardized by the standing height of each individual (i.e., resistance divided by height and reactance divided by height), expressed in ohms/m and plotted on the resistance-reactance graph (24). Bioelectrical impedance vector analysis (BIVA) expresses tissue hydration status and body cell mass solely while considering the impedance vector relative to a population of healthy individuals (24); this was a valid method for detecting changes in hydration (classified as under-, normal or overhydration) and body fluid volume changes (25). Sex-specific bivariate reference intervals were available for the Italian healthy population as 50%, 75%, and 95% tolerance ellipses on the resistance-reactance graph.

A mini nutritional assessment (MNA) was done for all participants (26). The MNA uses simple measurements and a brief questionnaire involving an anthropometric assessment (weight, height, and weight loss), a general assessment (lifestyle, medication, and mobility), and a dietary assessment (number of meals, food and fluid intake, self-assessment of autonomy of eating, and self-perception of health and nutrition). Patients ate 3 meals daily.

Dietary schedule

Food intake was based on a balanced diet (with standard caloric and macro- and micronutrient content) provided by the hospital kitchen, which consisted of a repeating 4-wk rotating menu, so the diet remained similar throughout the study.

A trained dietitian used a calibrated dietetic spring scale to weigh all foods served and returned for 3 consecutive days at the beginning and end of the study. Nurses who served any foods to the participants between meals recorded the amount eaten, in household measurements. A computer program (DR3 v3.1.0; Sintesi Informatica) was used to calculate the energy and the macronutrient content of food consumed.

Handgrip

The JAMAR Hand Dynamometer (Jamar 5030J1; Sammons Preston Rolyan; accuracy 0.6 N) was used to assess muscle function with the use of a standardized procedure (27).

---

[10] Abbreviations used: ADL, activities of daily living; CRP, C-reactive protein; DXA, dual-energy X-ray absorptiometry; FFM, fat-free mass; GH, growth hormone; IGF-I, insulin-like growth factor I; MNA, mini nutritional assessment; PCS, physical component summary; PRT, progressive resistance training; RSMM, relative skeletal muscle mass; SF-36, Short-Form 36-Item Health Survey.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

Biochemical analyses

Fasting venous blood samples were drawn with the participants seated. Blood was collected and handled under strictly standardized conditions. Blood samples were collected into vacuum tubes without anticoagulant, left for 1 h at room temperature, and then centrifuged for 15 min at 1500 × *g* at 20°C. The serum was then transferred into plastic tubes, rapidly frozen, and stored at –80°C until analysis (<1 mo later). Whole blood (with the use of EDTA as an anticoagulant) was used for hematologic variables. Clinical chemistry markers were detected on the Roche Cobas Integra 400 plus analyzer (Roche Diagnostics), with specially designed commercial kits provided by the manufacturer. Cobas Integra 400 is a random, continuous-access, sample-selective analyzer that provides absorbance photometry for measuring enzymes and substrates, turbidimetry for specific proteins, and ion-selective electrode potentiometry for serum electrolytes. Serum total and LDL cholesterol, triglycerides, HDL cholesterol, total proteins, total bilirubin, iron, glucose, uric acid, creatinine, and liver enzymes such as alanine transaminase, aspartate transaminase, and γ-glutamyltransferase were measured by enzymatic-colorimetric methods. C-reactive protein (CRP) was determined by a nephelometric high-sensitivity CRP (Dade Behring).

Erythrocyte, white blood cell, and platelet counts; hemoglobin concentrations; mean cell volumes; and mean cell hemoglobin concentrations were measured with the use of a Coulter automated cell counter (MAX-M; Beckman Coulter). Serum albumin was analyzed with the use of a nephelometric method (Behring Nephelometric Analyzer II, Behring Diagnostics), with a 2% CV.

Serum samples for insulin-like growth factor I (IGF-I) assay were collected at admission and after the 12 wk of treatment; samples were pretreated to release IGF-I from binding proteins and then assayed with quality controls. Serum IGF-I concentrations were measured with the use of a solid-phase quantitative ELISA kit (R&D Systems); the minimum detectable dose of IGF-I was 0.026 ng/mL. Intra- and interassay CVs were 4% and 7.9%, respectively.

Health-related quality of life

The participants were tested with the Short-Form 36-Item Health Survey (SF-36) (28) to assess their quality of life. This questionnaire is a valid generic measure that is used for rating health-related quality of life in several research fields because of its validity, high internal consistency, and high test–retest reliability. The SF-36 scales were summarized in 2 dimensions. The first 5 make up the "physical health" dimension, and the last 5 the "mental health" dimension. The vitality and general health scales are parts of both dimensions. Thus, each dimension includes 3 specific and 2 overlapping scales. The standardized summary scores for physical and mental components were calculated and used separately as outcome measures. The quality-of-life SF-36 was administered before and after the treatment period.

Function

The participants' ability to care for themselves was assessed with the Katz Index of Independence in Activities of Daily Living (29).

Intervention

Physical activity

A comprehensive physical fitness and muscle mass enhancement training program of moderate intensity was provided for all participants (30). The exercise intervention was supervised by trained personnel and consisted of 20-min exercise sessions daily, 5 times/wk for 12 wk. Each session consisted of a 5-min warm-up, 5 min of strengthening exercises, 5 min of balance and gait training, and 5 min of cool-down. The strengthening exercises were done in a progressive sequence from seated to standing positions (31). For each type of exercise, participants were instructed to repeat the movements ≤8 times. Intensity was maintained at ∼12–14 on the Borg Rate of Perceived Exertion scale (32). The principal investigator, with the exercise instructor and assistant trainers, assessed each individual's ability to increase intensity.

For the chair exercise, repetitions of toe raises, heel raises, knee lifts, knee extensions, and others were done while seated on a chair. Hip flexions, lateral leg raises, and repetitions of other exercises were done while standing upright behind the chair, holding the back of the chair for stability.

For the ankle-weight exercise, to strengthen the legs, a fixed weight was placed on the ankle while participants did strengthening exercises. Weights of 0.50, 0.75, 1.00, and 1.50 kg were prepared and used in accordance with each participant's strength as the resistance progressively increased. The exercises with the use of these ankle weights included seated knee flexion and extension and standing knee flexion and extension.

In the exercises with the use of a resistance band, resistance bands were used to strengthen the upper and lower body. Lower-body exercises included leg extension and hip flexion. Upper-body exercises included double-arm pull-downs and biceps curls.

For balance and gait training, exercises included standing on one leg, multidirectional weight shifts, a tandem stand, and a tandem walk. Participants practiced proper gait mechanics that focused on maintaining stability during walking and increasing stride length, toe elevation of the forward limb, heel elevation of the rear limb, frequency of stepping, and heel–floor angle. Exercises included raising the toes (dorsiflexion) during the forward swing of the leg, kicking off the floor with the ball of the foot, walking with directional changes, and gait pattern variations. In spring and summer, these exercises were done outdoors.

Dietary supplement

The intervention treatment included an oral essential amino acid, whey protein, and vitamin D mixture (**Tables 1** and **2**). The control group was given a placebo that consisted of an isocaloric amount of maltodextrin with the same flavor and appearance as the intervention product. Subjects were randomly assigned to receive one portion containing the dietary supplement or placebo (32 g) orally 1 time/d at 1200 with meals for 12 wk.

Participants were assigned to a treatment according to a coded (AB) block randomization table prepared by an independent statistician. Investigators were blinded to the randomization table, the code assignments, and the procedure. As people were enrolled they were assigned a progressive number. A research dietitian, blinded to the randomization schedule provided by the statistician, distributed the supplements to participants each day. Supplements were in powder form and packed in numerically

The American Journal of Clinical Nutrition

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

TABLE 1
Nutritional content of the dietary supplement[1]

	Energy value		% RDA[2] per 32-g dose
	Per 100 g	Per 32 g	
Kilojoules	1466	469	
Kilocalories	351	112	
Nutrients, g			
Whey protein	68.9	22	
Lipids	1.1	0.4	
SFAs	0.2	0.0	
Total carbohydrates	14.8	4.7	
Simple carbohydrates	2.6	0.8	
Complex carbohydrates	3.9	1.2	
Polyols	8.3	2.7	
Fiber	6.9	2.2	
Fructo-oligosaccharides	3.2	1.0	
Minerals, mg			
Calcium	25.8	8.3	1
Phosphorus	76.3	24.4	3
Sodium	917.4	293.6	
Magnesium	140.7	45.0	12
Iron	0.8	0.3	2
Vitamins, μg (IU)			
D_3 (cholecalciferol)	7.8 (312)	2.5 (100)	50

[1] SAI Nutrition.

[2] RDA, Recommended Dietary Allowance.

coded packages. Instructions on each bottle included the amount of water to be added; the water and contents of the bottle were then mixed and stirred for 60 s until the product was ready for consumption. Participants were instructed to eat their normal amounts of food in addition to the dietary supplement. All supplements were provided by SDM, Savigliano, Italy.

Safety was judged based on the absence of serious side effects with the supplement, i.e., gastrointestinal symptoms such as nausea and diarrhea. Every day, after administering the supplement, the dietitian asked about any unwanted side effects. No participant refused to take the supplement, and no side effects were reported.

Statistical analysis

Study design

This was a randomized, controlled, double-blind, parallel-group superiority clinical trial to compare the efficacy of whey protein, essential amino acid, and vitamin D supplementation or placebo in improving FFM or strength in sarcopenic elderly people in a hospital and rehabilitation division. The primary endpoint of the study was comparison of the increase in FFM after supplementation in the 2 groups. Secondary endpoints included the comparison of anthropometric characteristics (RSMM, fat mass, gynoid and android fat, and waist circumference), muscle strength (handgrip), quality of life [SF-36 mental component summary and physical component summary (PCS)], hormonal status (IGF-I), inflammation (CRP), and ADL. Finally, we assessed the correlations between several biomarkers, independently of treatment assignment, as exploratory endpoints.

To understand the links between variables better, we investigated the correlations between measures of primary endpoint muscle mass (FFM and RSMM), strength (handgrip), inflammation (CRP), and quality of life (SF-36 mental component summary and PCS), then between muscle mass (FFM and RSMM), strength (handgrip), function (ADL), and hormonal status (IGF-I). Finally, we analyzed the correlations between primary endpoint muscular mass (FFM and RSMM) and strength (handgrip).

Sample size

We based our sample size calculation on the findings of Borsheim et al. (6), and considered an expected mean ± SD increase of 1.1 ± 1.2 kg in the supplement group, and 0.5 ± 1.2 kg in the placebo group, with a power of 80% and an α level (2-tailed) of 5%, as well as 10% attrition. This gave a sample size of 140 patients (70/group).

Random assignment and masking

A random-blocks 1:1 random assignment list was prepared by a statistician. The treatment assignment sequence was masked to the investigator with the use of opaque envelopes. Blindness was maintained by providing the patients with undistinguishable products.

Statistical analysis

We used Stata 13. A 2-sided P value < 0.05 was considered to be significant. Continuous variables were summarized by treatment groups as means ± SDs or medians (25th, 75th percentiles), and categorical variables were summarized as counts and percentages. To compare changes in FFM between groups, a general linear regression model was fitted with FFM as the dependent variable, and treatment, time, and the interaction of treatment with time were used as independent variables. Huber–White robust SEs were computed with subject as the cluster variable to account for within-patient correlations of measurements and clustering for patients. Similarly, changes within groups also were analyzed by fitting to each group the same model with time as the sole independent variable. Mean

TABLE 2
Aminogram of the essential amino acids (g) in the dietary supplement

Amino acid	Value
Essential amino acids/g of product ready for use (32 g)	
L-Ile	1.0
L-Leu	4.0
L-Lys	1.5
L-Thr	1.1
L-Trp	0.3
L-Val	1.0
Nonessential amino acids/g of product ready for use (32 g)	
DL-Met	0.6
L-Cys	0.4
L-Phe	0.5
L-Tyr	0.5
Asp	1.8
Ser	0.8
Glu	5.2
Pro	1.0
Gly	0.3
Ala	0.8
Arg	0.8

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

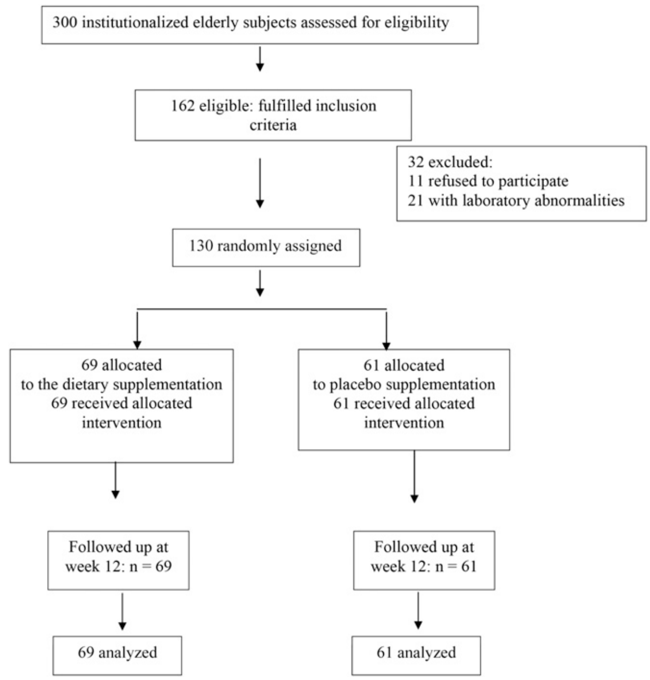

300 institutionalized elderly subjects assessed for eligibility
162 eligible: fulfilled inclusion criteria
32 excluded:
11 refused to participate
21 with laboratory abnormalities
130 randomly assigned
69 allocated to the dietary supplementation
69 received allocated intervention
61 allocated to placebo supplementation
61 received allocated intervention
Followed up at week 12: n = 69
Followed up at week 12: n = 61
69 analyzed
61 analyzed

FIGURE 1 Flow diagram of trial supplementation with dietary supplement compared with placebo in sarcopenic elderly people. The diagram indicates the number of individuals analyzed for the main outcome (effect on fat-free mass).

differences (changes) and 95% CIs were computed. Secondary endpoints were analyzed similarly.

For tertiary endpoints, partial correlations (and 95% CIs) were computed by fitting linear regression models while adjusting for treatment and time. Robust SEs were computed because of heteroscedasticity. All model assumptions were verified graphically (residuals compared with fitted plot) and were satisfied. Finally, a Fisher's exact test was used for comparison of the proportion of patients who became nonsarcopenic.

We used a Hotelling's T^2 (33) test on the calculated BIVA distribution with the use of pretreatment and compared the vectors to the population vector of healthy Italian adults to detect possible imbalances in the hydration levels of the study participants. Impedance individual vectors were plotted with BIVA software (25).

RESULTS

In all, 162 people were enrolled and 130 were randomly assigned; 32 were excluded because of refusal to participate (n = 11) or laboratory abnormalities (n = 21) (**Figure 1**). Participants were recruited from January 2013 through June 2014. Their main details are shown in **Table 3**. The placebo and supplement groups were similar on all counts.

The FFM response was significantly different between the 2 groups, with a mean difference of 1.7 kg (95% CI: 0.9, 2.5; P-interaction for treatment × time < 0.001). Fat-free mass increased significantly in the supplemented group (1.4 kg, $P < 0.001$), with no noteworthy change in the placebo group (−0.3 kg) (**Table 4**). Among secondary endpoints, responses were also significantly different in the 2 groups for treatment × time, as follows: RSMM, P-interaction = 0.009; android distribution of fat, P-interaction = 0.021; handgrip, P-interaction = 0.001; PCS, P-interaction = 0.030; ADL, P-interaction = 0.001; MNA, P-interaction = 0.003; IGF-I, P-interaction = 0.002; and CRP, P-interaction = 0.038. Specifically, RSMM, handgrip, PCS, ADL, MNA, and IGF-I increased significantly in the treatment group, whereas this was not the case in the placebo group. PCS slightly increased in the treatment group (P = 0.06), but not in the placebo group. Conversely, percentage android distribution of fat significantly decreased in the treatment group, but not in the placebo group, whereas CRP slightly decreased in the treatment group and slightly increased in the placebo group, although not significantly in either case (Table 4). The substantial improvements in RSMM and muscle strength in the supplementation group improved the classification of 68% of the elderly people in that group from sarcopenic to nonsarcopenic, but none of the participants in the placebo group showed improvement (Fisher's exact test $P < 0.001$).

No treatment effects were seen for waist circumference, fat mass, or gynoid percentage distribution of fat (Table 4), although gynoid fat percentage decreased significantly in both the supplemented and the placebo groups. Routine blood test results for clinical chemistry did not differ with respect to changes over time (data not shown). Changes in nutritional intake in treated and control patients did not differ between groups either. The dietary intake of both groups (not including the supplementation or placebo) is shown in **Table 5**.

TABLE 3
Baseline characteristics of the study participants[1]

Characteristic	Dietary supplement group (n = 69)	Placebo group (n = 61)
Age, y	80.77 ± 6.29	80.21 ± 8.54
Male	29 (42)	24 (39)
Smoker	3 (4)	5 (8)
Level of schooling, y	7 (3–11)	5 (2–9)
Fat-free mass, g	39,895 ± 8132	38,714 ± 8371
Fat mass, g	17,813 ± 6780	19,210 ± 9182
Gynoid, %	35.79 ± 9.67	37.67 ± 10.60
Android, %	34.21 ± 10.76	34.26 ± 12.85
RSMM, kg/m^2	6.60 ± 1.19	6.36 ± 1.32
MNA score	17.84 ± 3.07	17.84 ± 3.57
Weight, kg	59.47 ± 11.16	59.39 ± 13.51
BMI, kg/m^2	23.85 ± 3.63	23.93 ± 4.60
Wrist circumference, cm	16.29 ± 1.75	16.08 ± 1.42
Arm circumference, cm	25.22 ± 3.36	25.02 ± 3.80
Calf circumference, cm	30.43 ± 3.13	29.95 ± 4.55
Waist circumference, cm	88.95 ± 9.74	89.01 ± 10.15
MMSE score	21.78 ± 3.70	20.5 ± 4.93
ADL score	3.97 ± 1.19	4.03 ± 1.08
SF-36 MCS score	46.65 ± 10.7	44.0 ± 9.7
SF-36 PCS score	34.1 ± 10.2	37.1 ± 11.0
Proteins, g/dL	6.67 ± 0.55	6.56 ± 0.61
Albumin, g/dL	3.76 ± 0.54	3.6 ± 0.55
Creatinine, mg/dL	0.95 ± 0.7	0.91 ± 0.38
CRP, mg/L	0.30 (0.14–1.23)	0.33 (0.16–1.03)
IGF-I, ng/mL	80.6 ± 33.8	82.7 ± 38.8
Handgrip, kg	16.63 ± 4.99	19.62 ± 6.01

[1] Data are means ± SDs, medians (25–75th percentiles), or n (%). ADL, activities of daily living; CRP, C-reactive protein; IGF-I, insulin-like growth factor I; MCS, mental component summary; MMSE, Mini-Mental State Examination; MNA, mini nutritional assessment; PCS, physical component summary; RSMM, relative skeletal muscle mass; SF-36, Short-Form 36-Item Health Survey.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

TABLE 4
Effects of supplementation compared with placebo in exercise-trained elderly people[1]

	Dietary supplement group (n = 69)		Placebo group (n = 61)		Treatment effect	
Variable	Mean change (95% CI)	Intragroup P^2	Mean change (95% CI)	Intragroup P^2	Mean difference (95% CI)	P^3
Fat-free mass,[4] g	1382 (847, 1918)	<0.001	−312 (−930, 307)	0.316	1695 (892, 2498)	<0.001
Fat mass, g	−345 (−747, 57.18)	0.092	−484 (−1049, 81.74)	0.092	−114 (−786, 559)	0.689
Gynoid, %	−1.39 (−2.22, −0.56)	0.001	−0.92 (−1.83, −0.02)	0.046	0.54 (−0.67, 1.75)	0.451
Android, %	−2.03 (−2.99, −1.06)	0.001	−0.26 (−1.43, 0.92)	0.66	1.80 (0.30, 3.29)	0.021
RSMM, kg/m^2	0.21 (0.07, 0.35)	0.004	−0.06 (−0.21, 0.90)	0.42	0.27 (0.07, 0.47)	0.009
MNA score	1.76 (1.23, 2.28)	<0.001	0.24 (−0.63, 1.11)	0.585	1.52 (0.51, 2.52)	0.003
Weight, kg	1.12 (0.37, 1.87)	0.004	−0.89 (−1.62, −0.15)	0.019	2.00 (0.97, 3.04)	<0.001
BMI, kg/m^2	0.42 (0.11, 0.72)	0.008	−0.42 (−0.70, −0.14)	0.004	0.84 (0.43, −1.25)	<0.001
Waist circumference, cm	4.93 (−0.86, 10.72)	0.094	2.27 (−1.72, 6.25)	0.259	2.67 (−4.29, 9.62)	0.449
ADL score	0.54 (0.39, 0.68)	<0.001	−0.61 (−0.79, −0.42)	<0.001	1.14 (0.91, 1.38)	<0.001
SF-36 MCS score	4.50 (2.68, 6.32)	<0.001	2.48 (0.21, 4.75)	0.033	2.02 (−0.85, 4.89)	0.166
SF-36 PCS score	1.32 (−0.05, 2.68)	0.059	−0.77 (−2.10, 0.58)	0.249	2.09 (0.21, 3.97)	0.030
CRP, mg/dL	−0.19 (−0.57, 0.19)	0.329	0.44 (−0.02, 0.90)	0.061	0.63 (0.04, 1.22)	0.038
IGF-I, ng/mL	20.7 (11.0, 30.4)	<0.001	1.8 (−4.2, 7.8)	0.541	19.7 (7.1, 32.3)	0.002
Handgrip, kg	3.20 (2.23, 4.18)	<0.001	−0.47 (−1.07, 0.12)	0.117	3.68 (2.55, 4.81)	<0.001

[1] ADL, activities of daily living; CRP, C-reactive protein; IGF-I, insulin-like growth factor I; MCS, mental component summary; MNA, mini nutritional assessment; PCS, physical component summary; RSMM, relative skeletal muscle mass; SF-36, Short-Form 36-Item Health Survey.

[2] From within-treatment regression model—test for main effect of time within each treatment arm.

[3] From between-treatment regression model—test for treatment × time interaction.

[4] Primary endpoint. Regression model for repeated measures.

The reference bivariate tolerance ellipses (50%, 75%, and 95% of the distribution of the values for the general Italian population) for elderly men were used for the qualitative and semiquantitative assessment of body composition and hydration in each individual. The 95% CI ellipses for the mean vectors of the treated group before and after supplementation were drawn to compare these groups. SDs between the mean vectors were found with Hotelling's T^2 test for vector analysis, which is a multivariate extension of Student's test for unpaired data in comparisons of mean vectors from 2 groups. Two mean vectors have a significantly different ($P < 0.05$) position in the resistance-reactance graph if their 95% CI ellipses are separated according to Hotelling's T^2 test. Overlapping ellipses are not significantly different ($P > 0.05$). Ellipses were plotted with BIVA software (25). Mean group vectors before and after treatments were within the reference sex–specific 50% tolerance ellipse. Thus, hydration was classified as normal. Hotelling's T^2 test indicated a nonsignificant (T^2= 1.7756, $P > 0.05$) difference between the study group vector and the reference population 50% CI ellipse, confirming that the hydration of FFM was normal, so the DXA results were not influenced by altered soft-tissue hydration (21, 22).

Among the exploratory endpoints, a weak but statistically significant correlation was found between handgrip strength and RSMM (R = 23%; 95% CI: 11%, 34%; P = 0.0014) (**Figure 2**), between handgrip strength and FFM (R = 27%; 95% CI: 16%, 38%; P = 0.003) (**Figure 3**), and between IGF-I and FFM (R = 15%; 95% CI: 2%, 29%; P = 0.041) (**Figure 4**), while adjusting for treatment and time.

The dietary supplement was well tolerated, and there were no serious adverse events. Compliance was 100%.

DISCUSSION

This study found a significant beneficial effect of supplementation with whey protein, essential amino acids, and vitamin D compared with placebo in elderly sarcopenic adults

TABLE 5
Nutritional intake of supplemented and control participants at beginning of study and after 12 wk[1]

	Placebo group (n = 69)		Dietary supplement group (n = 61)			
Daily nutritional intake	Baseline	12 wk	Baseline	12 wk	P^2	P^3
Energy, kcal/d	1622 ± 350	1615 ± 273	1600 ± 215	1573 ± 339	NS	NS
Proteins, g/d	59 ± 8	60 ± 9	54 ± 12	55 ± 11	NS	NS
Fat, g/d	54 ± 12	55 ± 11	52 ± 9	53 ± 14	NS	NS
Carbohydrates, g/d	225 ± 4	220 ± 5	214 ± 3	212 ± 4	NS	NS
Vitamin D, IU/d	299 ± 79	298 ± 87	301 ± 92	296 ± 89	NS	NS

[1] Data are means ± SDs with the use of the Carnovale E Marletta L food composition tables, Italian National Institute of Nutrition, Rome, 1997.

[2] From within-treatment regression model—test for main effect of time within each treatment arm.

[3] From between-treatment regression model—test for treatment × time interaction.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

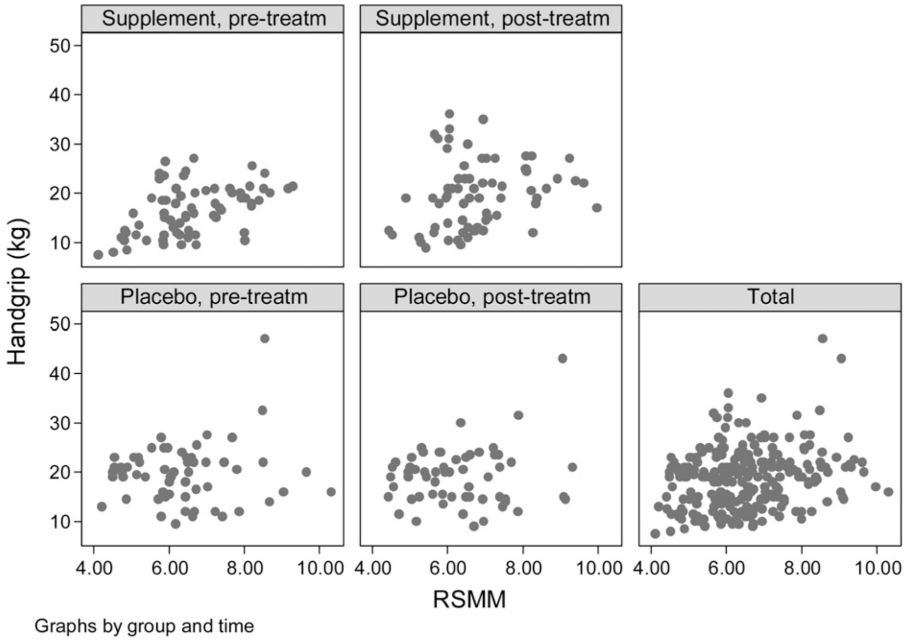


FIGURE 2 Correlation between handgrip strength and RSMM ($R_{\text{total}} = 0.23$; 95% CI: 0.11, 0.34; $P = 0.0014$) for each treatment arm, before and after treatment, and overall. After dropping 2 outliers for handgrip (>40), the results were the same. *R* is computed as the partial correlation, adjusted for treatment and time, from the repeated-measures model; thus, each individual is represented twice in the "total" graph. RSMM, relative skeletal muscle mass; treatm, treatment.

participating in controlled resistance training, with a gain of 1.7 kg in FFM. Supplementation significantly improved RSMM and muscle strength and, in fact, 68% of sarcopenic people became nonsarcopenic (19).

Key strengths of the study included the comprehensive assessment of the main causal factors of sarcopenia in a well-defined elderly population. Nutritional supplementation, independently of increased physical activity, also improved some factors that contribute to sarcopenia. Supplementation attenuated the inflammatory state, as seen by the significant drops in CRP concentrations, and enhanced the anabolic growth hormone (GH) IGF-I hormone axis, with significant increases in IGF-I concentrations and a reduction in the indexes of malnutrition assessed with the MNA. Dietary supplementation also boosted various

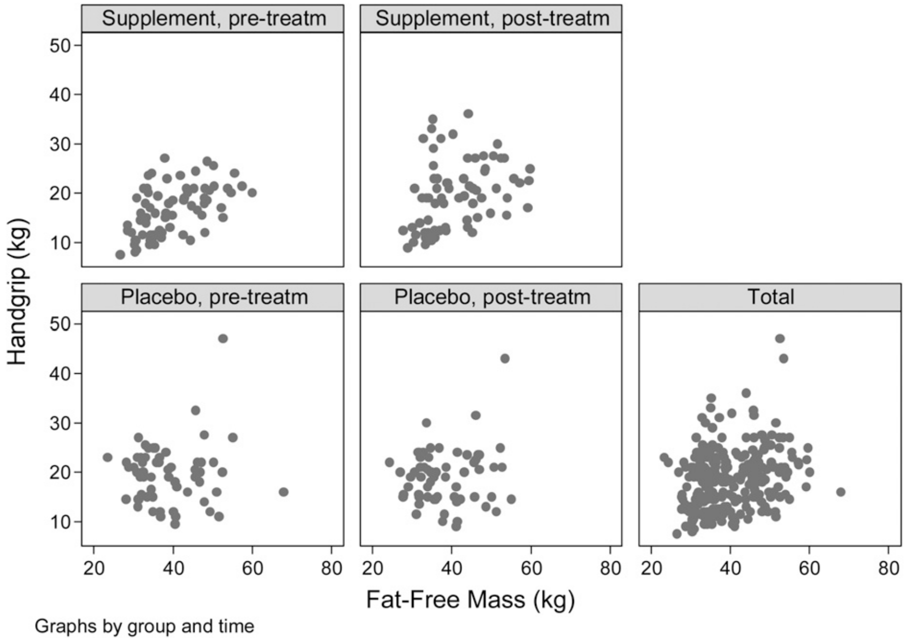


FIGURE 3 Correlation between handgrip strength and fat-free mass ($R_{\text{total}} = 0.27$; 95% CI: 0.16, 0.38; $P = 0.003$) for each treatment arm, before and after treatment, and overall. After dropping 2 outliers for handgrip (>40), the results were the same. *R* is computed as the partial correlation, adjusted for treatment and time, from the repeated-measures model; thus, each individual is represented twice in the "total" graph. treatm, treatment.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

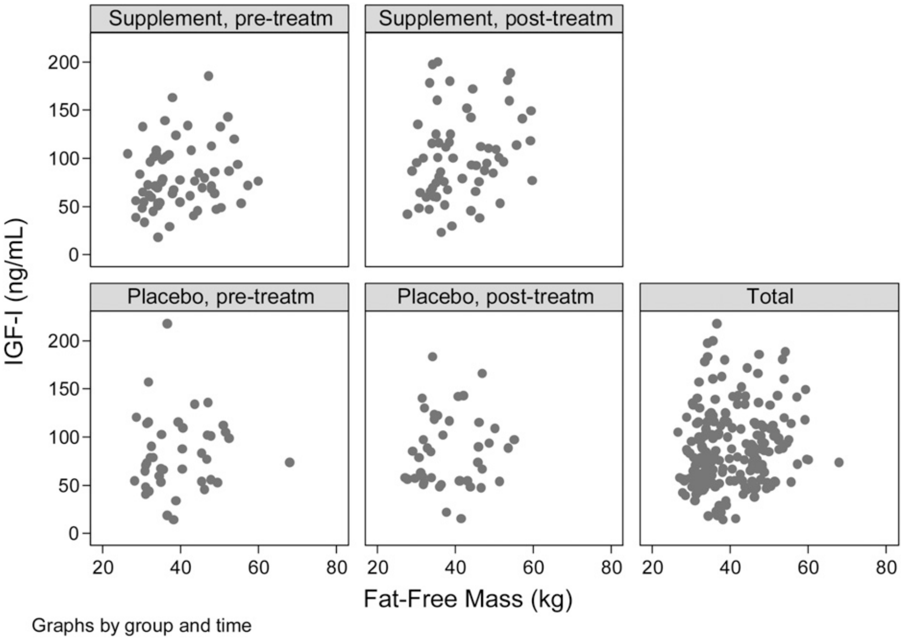


FIGURE 4 Correlation between IGF-I and fat-free mass (R_{total} =0.15; 95% CI: 0.2, 0.29; P = 0.041) for each treatment arm, before and after treatment, and overall. R is computed as the partial correlation, adjusted for treatment and time, from the repeated-measures model; thus, each individual is represented twice in the "total" graph. IGF-I, insulin-like growth factor I; treatm, treatment.

measures of function in the participants, assessed by the ADL, and their quality of life, particularly its physical component, assessed by the physical component of the SF-36.

Both groups followed a physical activity plan, but beneficial results were seen only in the supplemented group, indicating that physical activity is important, but not sufficient to achieve a significant result. However, the physical activity was gentle and nonintensive and this might have explained the lack of increase in FFM in the placebo group. These results are in agreement with the study by Raguso et al. (34), which showed that leisure-time physical activity does not seem to prevent loss of muscle mass.

We decided on 12 wk of resistance training on the basis of previous research showing that substantial muscle hypertrophy can occur within this period (34, 35) and that substantial diet-related differences in muscle hypertrophy responses may also arise within this time frame (34, 36).

Both aerobic and resistance-type exercise training have been shown to improve the rate of decline in muscle mass and strength with age (37). Progressive resistance training (PRT) is the most commonly used resistance therapy in older people. A Cochrane review of 121 randomized controlled trials of PRT in older people showed that doing PRT 2–3 times/wk improved physical function, gait speed, timed get-up-and-go, climbing stairs, and balance, and, more importantly, had a significant effect on muscle strength, especially in the high-intensity training groups (38). Even in very old nursing home residents, PRT achieved substantial improvements in muscle fiber cross-sectional area (3–9%), muscle strength (100%), and physical performance such as gait speed and stair climbing (38, 39).

The majority of studies show that resistance exercise training must be carried out at high intensity to achieve substantial improvements in muscle strength, but, for sarcopenic people, high-intensity resistance training may not be realistic or practical. The elderly people recruited for this study would not have been able to maintain high-intensity resistance exercise training. We therefore selected an age-appropriate, tolerable, and sustainable exercise program that consisted of resistance and aerobic exercises. Compliance was 100%. These age-related resistance and aerobic exercises in older people increased their strength by 38% and resulted in significant reductions in CRP (40).

Reducing inflammation is one mechanism that can improve age-related muscle loss through either direct catabolic effects or indirect mechanisms (through higher GH and IGF-I concentrations, less anorexia, etc.) (41)

Protein supplementation combined with physical exercise, particularly resistance training, has yielded mixed results on body composition, muscle hypertrophy, strength, and physical function in the elderly (6, 39, 42–47), even though most studies have focused on healthy older adults, with limited data from trials on sarcopenic individuals, in which nutritional but not specifically protein and amino acid supplementation was a focus (39). Moreover, in these studies, the doses of protein supplemented varied between 7.4 and 15 g/serving. These differences make it difficult to compare studies.

Another key finding is the positive effect of nutritional supplementation on IGF-I concentrations. IGF-1 contributes to improving muscle function by increasing production of muscle satellite cells and stimulating production of muscle contractile proteins. The age- related decline in GH concentrations, combined with lower IGF-I concentrations contributes to the development of sarcopenia (48). IGF-I is perhaps the most important mediator of muscle growth and repair (49), possibly through the use of protein kinase B–mechanistic target of rapamycin–p70 ribosomal protein S6 kinase signaling.

The composition and timing of the supplement are novel aspects of this study. To counteract protein catabolism, the elderly must increase the anabolic stimulus, consuming 30 g protein/meal (50–52). The combination of whey protein and essential

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES LIBRARY on April 26, 2017

amino acids providing Leu is important (53, 54). Whey protein increases postprandial plasma amino acid availability, further stimulating muscle protein synthesis (53, 55–58), more than casein (59, 60). Whey contains a high concentration of Leu, which stimulates skeletal muscle protein synthesis (61). Thus, whey protein and essential amino acids that contain Leu are recommended interventions for sarcopenia (62, 63), and the effect of nutritional therapies for sarcopenia can be enhanced by a comprehensive approach (64).

A low-caloric dietary supplement can be taken with a meal without problems of gastric emptying; the supplement can even be taken by overweight or obese sarcopenic individuals (sarcopenic obesity), because being sarcopenic does not necessarily mean being underweight (65).

The present findings are in agreement with previous reports of improvements in muscle strength with exercise and whey protein supplementation (20 g/d) in frail elderly people after ∼12 wk of supplementation, as in this study (6) and in an acute situation (66, 67). Verreijen et al. (68) reported that a high whey protein, Leu, and vitamin D–enriched supplement similar to the supplement used in this study—except for the vitamin D content—compared with isocaloric control preserved appendicular muscle mass in obese older adults during a hypocaloric diet and resistance exercise program (3 times/wk) for 13 wk and might therefore reduce the risk of sarcopenia. However, Verreijen et al. (68) found no beneficial effect of supplementation on muscle strength or function—which contrasts with the findings of the present study. Whether differences in the physical activity interventions explain the lack of functional improvement remains to be established.

The effect of nutritional therapies for sarcopenia can be enhanced by a comprehensive approach (64). This is why the supplement we used also contained vitamin D. A recent meta-analysis (69) of the results of 30 randomized, double-blind, placebo-controlled clinical trials indicated that daily supplementation with ≥400 IU of vitamin D_3 increased skeletal muscle strength on average by 17%. The intervention supplement we used contained 2.5 μg (100 IU) vitamin D. We selected this dose of vitamin D for the supplement, whereas the participants receiving the control diet consumed 120 g halibut 2 times/wk, canned tuna 2 times/wk, cod 2 times/wk, and 2 eggs/wk, which provided them with a mean of 300 IU of vitamin D daily; they also participated in balance and gait training outdoors. Our result is in line with the meta-analysis by Beaudart et al. (69), because the increase in muscle strength averaged 21%. This gain in muscle strength has been suggested as the mechanism behind a reduction in falls of 23–53%, in addition to a reduction in fractures in older nursing or residential home residents given vitamin D (70–72).

A potential implication of our findings is that patients with sarcopenia should consider the use of specific supplements combined with appropriate physical activity to attenuate loss or increase skeletal muscle mass.

A limitation of this study was that we did not assay blood vitamin D concentrations. Vitamin D status and its relation to physical training with and without supplementation are important questions that await investigation. Another important limitation was that we were not able to assess the effects of vitamin D supplementation separately from essential amino acid supplementation, although this type of experimental design would require many more participants than were available and a new sample size calculation.

In conclusion, aging causes the loss of many of the anabolic signals and an increase in catabolic signals to muscle that are present in young adulthood, but this study suggests that whey protein, essential amino acid, and vitamin D supplementation, together with gentle physical activity, can produce changes in catabolic mediators, lowering inflammatory markers such as CRP, and improving anabolic markers such as IGF-I. This shift results in a significant increase in FFM (+1.7 kg) and muscle strength, proving effective in the treatment of sarcopenia, with improvements in physical function and quality of life.

The authors' responsibilities were as follows—MR, CK, GT, JT, and RM: designed the study; MAF and SP: conducted the study; CK and DG: analyzed the data; BSS and MF: conducted the blood tests; MR and CK: wrote the manuscript; MR: had primary responsibility for the design and final content of the article; HL: revised the manuscript; and all authors: read and approved the final manuscript. None of the authors reported a conflict of interest related to the study.

REFERENCES

1. Morley JE, Baumgartner RN, Roubenoff R, Mayer J, Nair KS. Sarcopenia. J Lab Clin Med 2001;137:231–43.
2. Rosenberg IH. Sarcopenia: origins and clinical relevance. J Nutr 1997;127:990S–1S.
3. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, Martin FC, Michel JP, Rolland Y, Schneider SM, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. European Working Group on Sarcopenia in Older People. Age Ageing 2010;39:412–23.
4. Cerri AP, Bellelli G, Mazzone A, Pittella F, Landi F, Zambon A, Annoni G. Sarcopenia and malnutrition in acutely ill hospitalized elderly: prevalence and outcomes. Clin Nutr 2015;34:745–51.
5. Houston DK, Nicklas BJ, Ding J, Harris TB, Tylavsky FA, Newman AB, Lee JS, Sahyoun NR, Visser M, Kritchevsky SB, et al. Dietary protein intake is associated with lean mass change in older, community-dwelling adults: the Health, Aging, and Body Composition (Health ABC) Study. Am J Clin Nutr 2008;87:150–5.
6. Børsheim E, Bui QU, Tissier S, Kobayashi H, Ferrando AA, Wolfe RR. Effect of amino acid supplementation on muscle mass, strength and physical function in elderly. Clin Nutr 2008;27:189–95.
7. Wicherts IS, van Schoor NM, Boeke AJ, Visser M, Deeg DJ, Smit J, Knol DL, Lips P. Vitamin D status predicts physical performance and its decline in older persons. J Clin Endocrinol Metab 2007;92:2058–65.
8. Ceglia L, Harris SS. Vitamin D and its role in skeletal muscle. Calcif Tissue Int 2013;92:151–62.
9. Latham NK, Anderson CS, Reid IR. Effects of vitamin D supplementation on strength, physical performance, and falls in older persons: a systematic review. J Am Geriatr Soc 2003;51:1219–26.
10. Muir SW, Montero-Odasso M. Effect of vitamin D supplementation on muscle strength, gait and balance in older adults: a systematic review and meta-analysis. J Am Geriatr Soc 2011;59:2291–300.
11. Schaap LA, Pluijm SM, Deeg DJ, Visser M. Inflammatory markers and loss of muscle mass (sarcopenia) and strength. Am J Med 2006;119:526.e9–17.
12. Waters DL, Baumgartner RN, Garry PJ, Vellas B. Advantages of dietary, exercise-related, and therapeutic interventions to prevent and treat sarcopenia in adult patients: an update. Clin Interv Aging 2010;5:259–70.
13. Candow DG, Forbes SC, Little JP, Cornish SM, Pinkoski C, Chilibeck PD. Effect of nutritional interventions and resistance exercise on aging muscle mass and strength. Biogerontology 2012;13:345–58.
14. Latham NK, Bennett DA, Stretton CM, Anderson CS. Systematic review of progressive resistance strength training in older adults. J Gerontol A Biol Sci Med Sci 2004;59:48–61.
15. Liu CJ, Latham NK. Progressive resistance strength training for improving physical function in older adults. Cochrane Database Syst Rev 2009;CD002759.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES LIBRARY on April 26, 2017

16. Cesari M, Vellas B. Sarcopenia: a novel clinical condition or still a matter for research? J Am Med Dir Assoc 2012;13:766–7.
17. Baumgartner RN, Koehler K, Gallagher D, Romero L, Heymsfield SB, Ross RR, Garry PJ, Lindeman RD. Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol 1998;147:755–63.
18. 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.
19. Janssen I, Baumgartner RN, Ross R, Rosenberg IH, Roubenoff R. Skeletal muscle cutpoints associated with elevated physical disability risk in older men and women. Am J Epidemiol 2004;159:413–21.
20. Frisancho AR. New standards of weight and body composition by frame size and height for assessment of nutritional status of adults and the elderly. Am J Clin Nutr 1984;40:808–19.
21. St-Onge MP, Wang Z, Horlick M, Wang J, Heymsfield SB. Dual-energy X-ray absorptiometry lean soft tissue hydration: independent contributions of intra- and extracellular water. Am J Physiol Endocrinol Metab 2004;287:E842–7.
22. Moon JR, Stout JR, Smith-Ryan AE, Kendall KL, Fukuda DH, Cramer JT, Moon SE. Tracking fat-free mass changes in elderly men and women using single-frequency bioimpedance and dual-energy X-ray absorptiometry: a four-compartment model comparison. Eur J Clin Nutr 2013;67:S40–6.
23. Piccoli A, Rossi B, Pillon L, Bucciante G. A new method for monitoring body fluid variation by bioimpedance analysis: the RXc graph. Kidney Int 1994;46:534–9.
24. Lukaski HC. Requirements for clinical use of bioelectrical impedance analysis (BIA). Ann N Y Acad Sci 1999;873:72–6.
25. Piccoli A, Piazza P, Noventa D, Pillon L, Zaccaria M. A new method for monitoring hydration at high altitude by bioimpedance analysis. Med Sci Sports Exerc 1996;28:1517–22.
26. Guigoz Y, Vellas B, Garry PJ. Assessing the nutritional status of the elderly: the Mini Nutritional Assessment as part of the geriatric evaluation. Nutr Rev 1996;54:S59–65.
27. Spijkerman DC, Snijders CJ, Stijnen T, Lankhorst GJ. Standardization of grip strength measurements. Effects on repeatability and peak force. Scand J Rehabil Med 1991;23:203–6.
28. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: manual and interpretation guide. Boston: The Health Institute; 1993.
29. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10:20–30.
30. Kim HK, Suzuki T, Saito K, Yoshida H, Kobayashi H, Kato H, Katayama M. Effects of exercise and amino acid supplementation on body composition and physical function in community-dwelling elderly Japanese sarcopenic women: a randomized controlled trial. J Am Geriatr Soc 2012;60:16–23.
31. American College of Sports Medicine. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. Med Sci Sports Exerc 2009;41:687–708.
32. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14:377–81.
33. Hotelling H. The generalization of Student's ratio. Ann Math Stat 1931;2:360–78.
34. Raguso CA, Kyle U, Kossovsky MP, Roynette C, Paoloni-Giacobino A, Hans D, Genton L, Pichard C. A 3-year longitudinal study on body composition changes in the elderly: role of physical exercise. Clin Nutr 2006;25:573–80.
35. Charette SL, McEvoy L, Pyka G, Snow-Harter C, Guido D, Wiswell RA, Marcus R. Muscle hypertrophy response to resistance training in older women. J Appl Physiol 1991;70:1912–6.
36. Meredith CN, Frontera WR, O'Reilly KP, Evans WJ. Body composition in elderly men: effect of dietary modification during strength training. J Am Geriatr Soc 1992;40:155–62.
37. Frankel JE, Bean JF, Frontera WR. Exercise in the elderly: research and clinical practice. Clin Geriatr Med 2006;22:239–56.
38. Crocker T, Forster A, Young J, Brown L, Ozer S, Smith J, Green J, Hardy J, Burns E, Glidewell E, et al. Physical rehabilitation for older people in long-term care. Cochrane Database Syst Rev 2013;2:CD004294. DOI: 10.1002/14651858.CD004294.
39. Fiatarone MA, O'Neill EF, Ryan ND, Clements KM, Solares GR, Nelson ME, Roberts SB, Kehayias JJ, Lipsitz LA, Evans WJ. Exercise training and nutritional supplementation for physical frailty in very elderly people. N Engl J Med 1994;330:1769–75.
40. Stewart LK, Flynn MG, Campbell WW, Craig BA, Robinson JP, Timmerman KL, McFarlin BK, Coen PM, Talbert E. The influence of exercise training on inflammatory cytokines and C-reactive protein. Med Sci Sports Exerc 2007;39:1714–9.
41. Roubenoff R. Catabolism of aging: is it an inflammatory process? Curr Opin Clin Nutr Metab Care 2003;6:295–9.
42. Kukuljan S, Nowson CA, Sanders K, Daly RM. Effects of resistance exercise and fortified milk on skeletal muscle mass, muscle size, and functional performance in middle-aged and older men: an 18-mo randomized controlled trial. J Appl Physiol 2009;107:1864–73.
43. Iglay HB, Thyfault JP, Apolzan JW, Campbell WW. Resistance training and dietary protein: effects on glucose tolerance and contents of skeletal muscle insulin signaling proteins in older persons. Am J Clin Nutr 2007;85:1005–13.
44. Rosendahl E, Lindelöf N, Littbrand H, Yifter-Lindgren E, Lundin-Olsson L, Håglin L, Gustafson Y, Nyberg L. High-intensity functional exercise program and protein-enriched energy supplement for older persons dependent in activities of daily living: a randomised controlled trial. Aust J Physiother 2006;52:105–13.
45. Holm L, Olesen JL, Matsumoto K, Doi T, Mizuno M, Alsted TJ, Mackey AL, Schwarz P, Kjaer M. Protein-containing nutrient supplementation following strength training enhances the effect on muscle mass, strength, and bone formation in postmenopausal women. J Appl Physiol 2008;105:274–81.
46. Bunout B, Barrera G, de la Maza P, Avendano M, Gattas V, Petermann M, Hirsh S. Effects of nutritional supplementation and resistance training on muscle strength in free living elders. Results of one year follow. J Nutr Health Aging 2004;8:68–75.
47. Haub MD, Wells AM, Tarnopolsky MA, Campbell WW. Effect of protein source on resistive-training-induced changes in body composition and muscle size in older men. Am J Clin Nutr 2002;76:511–7.
48. Ferrucci L, Penninx BWJH, Volpato S, Harris TB, Bandeen-Roche K, Balfour J, Leveille SG, Fried LP, Md JM. Change in muscle strength explains accelerated decline of physical function in older women with high interleukin-6 serum levels. J Am Geriatr Soc 2002;50:1947–54.
49. Philippou A, Maridaki M, Halapas A, Koutsilieris M. The role of the insulin-like growth factor 1 (IGF-1) in skeletal muscle physiology. In Vivo 2007;21:45–54.
50. Solerte SB, Gazzaruso C, Bonacasa R, Rondanelli M, Zamboni M, Basso C, Locatelli E, Schifino N, Giustina A, Fioravanti M. Nutritional supplements with oral amino acid mixtures increases whole-body lean mass and insulin sensitivity in elderly subjects with sarcopenia. Am J Cardiol 2008;101:69E–77E.
51. Dillon EL, Sheffield-Moore M, Paddon-Jones D, Gilkison C, Sanford AP, Casperson SL, Jiang J, Chinkes DL, Urban RJ. Amino acid supplementation increases lean body mass, basal muscle protein synthesis, and insulin-like growth factor-I expression in older women. J Clin Endocrinol Metab 2009;94:1630–7.
52. Paddon-Jones D, Rasmussen BB. Dietary protein recommendations and the prevention of sarcopenia. Curr Opin Clin Nutr Metab Care 2009;12:86–90.
53. Boirie Y, Dangin M, Gachon P, Vasson MP, Maubois JL, Beaufrere B. Slow and fast dietary proteins differently modulate postprandial protein accretion. Proc Natl Acad Sci USA 1997;94:14930–5.
54. Koopman R, Crombach N, Gijsen AP, Walrand S, Fauquant J, Kies AK, Lemosquet S, Saris WH, Boirie Y, van Loon LJ. Ingestion of a protein hydrolysate is accompanied by an accelerated in vivo digestion and absorption rate when compared with its intact protein. Am J Clin Nutr 2009;90:106–15.
55. Dangin M, Boirie Y, Garcia-Rodenas C, Gachon P, Fauquant J, Callier P, Ballèvre O, Beaufrère B. The digestion rate of protein is an independent regulating factor of postprandial protein retention. Am J Physiol Endocrinol Metab 2001;280:E340–8.
56. Bohé J, Low A, Wolfe RR, Rennie MJ. Human muscle protein synthesis is modulated by extracellular, not intramuscular amino acid availability: a dose-response study. J Physiol 2003;552:315–24.
57. Tang JE, Moore DR, Kujbida GW, Tarnopolsky MA, Phillips SM. Ingestion of whey hydrolysate, casein, or soy protein isolate: effects on mixed muscle protein synthesis at rest and following resistance exercise in young men. J Appl Physiol 2009;107:987–92.
58. Pennings B, Boirie Y, Senden JM, Gijsen AP, Kuipers H, van Loon LJ. Whey protein stimulates postprandial muscle protein accretion more effectively than do casein and casein hydrolysate in older men. Am J Clin Nutr 2011;93:997–1005.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

59. Dangin M, Boirie Y, Guillet C, Beaufrere B. Influence of the protein digestion rate on protein turnover in young and elderly subjects. J Nutr 2002;132:3228S–33S.
60. Dangin M, Guillet C, Garcia-Rodenas C, Gachon P, Bouteloup-Demange C, Reiffers-Magnani K, Fauquant J, Ballèvre O, Beaufrère B. The rate of protein digestion affects protein gain differently during aging in humans. J Physiol 2003;549:635–44.
61. Volpi E, Ferrando AA, Yeckel CW, Tipton KD, Wolfe RR. Exogenous amino acids stimulate net muscle protein synthesis in the elderly. J Clin Invest 1998;101:2000–7.
62. Bauer J, Biolo G, Cederholm T, Cesari M, Cruz-Jentoft AJ, Morley JE, Phillips S, Sieber C, Stehle P, Teta D, et al. Evidence-based recommendations for optimal dietary protein intake in older people: a position paper from the PROT-AGE Study Group. J Am Med Dir Assoc 2013;14:542–59.
63. Rondanelli M, Faliva M, Monteferrario F, Peroni G, Repaci E, Allieri F, Perna S. Novel insights on nutrient management of sarcopenia in elderly. Biomed Res Int 2015;2015:524948.
64. Wakabayashi H, Sakuma K. Comprehensive approach to sarcopenia treatment. Curr Clin Pharmacol 2014;9:171–80.
65. Zamboni M, Mazzali G, Fantin F, Rossi A, Di Francesco V. Sarcopenic obesity: a new category of obesity in the elderly. Nutr Metab Cardiovasc Dis 2008;18:388–95.
66. Tieland M, Dirks ML, van der Zwaluw N, Verdijk LB, van de Rest O, de Groot LC, van Loon LJ. Protein supplementation increases muscle mass gain during prolonged resistance-type exercise training in frail elderly people: a randomized, double-blind, placebo-controlledtrial. J Am Med Dir Assoc 2012;13:713–9.
67. Drummond MJ, Dreyer HC, Pennings B, Fry CS, Dhanani S, Dillon EL, Sheffield-Moore M, Volpi E, Rasmussen BB. Skeletal muscle protein anabolic response to resistance exercise and essential amino acids is delayed withaging. J Appl Physiol 1985; 2008:1452–61.
68. Verreijen AM, Verlaan S, Engberink MF, Swinkels S, de Vogel-van den Bosch J, Weijs PJM. A high whey protein–, leucine-, and vitamin D–enriched supplement preserves muscle mass during intentional weight loss in obese older adults: a double-blind randomized controlled trial. Am J Clin Nutr 2015;101:279–86.
69. Beaudart C, Buckinx F, Rabenda V, Gillain S, Cavalier E, Slomian J, Petermans J, Reginster JY, Bruyère O. The effects of vitamin D on skeletal muscle strength, muscle mass, and muscle power: a systematic review and meta-analysis of randomized controlled trials. J Clin Endocrinol Metab 2014;99:4336–45.
70. Bischoff HA, Stahelin HB, Dick W, Akos R, Knecht M, Salis C, Nebiker M, Theiler R, Pfeifer M, Begerow B, et al. Effects of vitamin D and calcium supplementation on falls: a randomized controlled trial. J Bone Miner Res 2003;18:343–51.
71. Flicker L, MacInnis RJ, Stein MS, Scherer SC, Mead KE, Nowson CA, Thomas J, Lowndes C, Hopper JL, Wark JD. Should older people in residential care receive vitamin D to prevent falls? Results of a randomized trial. J Am Geriatr Soc 2005;53:1881–8.
72. Chapuy MC, Arlot ME, Duboeuf F, Brun J, Crouzet B, Arnaud S, Delmas PD, Meunier PJ. Vitamin D3 and calcium to prevent hip fractures in the elderly women. N Engl J Med 1992;327:1637–42.

Downloaded from ajcn.nutrition.org at UNIVERSITY OF MARYLAND HEALTH SCIENCES AND HUMAN SERVICES LIBRARY on April 26, 2017

ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle (2017)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.12170

Pre-operative oral nutritional supplementation with dietary advice versus dietary advice alone in weight-losing patients with colorectal cancer: single-blind randomized controlled trial

Sorrel T. Burden[1,2,6]*, Debra J. Gibson[1,6], Simon Lal[2,4,6], James Hill[3,4,6], Mark Pilling[1], Mattias Soop[2,4,6], Aswatha Ramesh[5,6] & Chris Todd[1,6]

[1]*School of Health Sciences, The University of Manchester, Manchester, UK;* [2]*Salford Royal NHS Foundation Trust, Salford, UK;* [3]*Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK;* [4]*School of Medicine, The University of Manchester, Manchester, UK;* [5]*University Hospital South Manchester NHS Foundation Trust, Wythenshawe, UK;* [6]*Manchester Academic Health Sciences Centre (MAHSC), Manchester, UK*

Abstract

Background Pre-operative weight loss has been consistently associated with increased post-operative morbidity. The study aims to determine if pre-operative oral nutritional supplements (ONSs) with dietary advice reduce post-operative complications.

Methods Single-blinded randomized controlled trial. People with colorectal cancer scheduled for surgery with pre-operative weight loss >1 kg/3–6 months were randomized by using stratified blocks (1:1 ratio) in six hospitals (1 November 2013–28 February 2015). Intervention group was given 250 mL/day ONS (10.1 KJ and 0.096 g protein per mL) and dietary advice. Control group received dietary advice alone. Oral nutritional supplements were administered from diagnosis to the day preceding surgery. Research team was masked to group allocation. Primary outcome was patients with one or more surgical site infection (SSI) or chest infection; secondary outcomes included percentage weight loss, total complications, and body composition measurements. Intention-to-treat analysis was performed with both unadjusted and adjusted analyses. A sample size of 88 was required.

Results Of 101 participants, (55 ONS, 46 controls) 97 had surgery. In intention-to-treat analysis, there were 21/45 (47%) patients with an infection—either an SSI or chest infection in the control group vs. 17/55 (30%) in the ONS group. The odds ratio of a patient incurring either an SSI or chest infection was 0.532 ($P = 0.135$ confidence interval 0.232 to 1.218) in the unadjusted analysis and when adjusted for random differences at baseline (age, gender, percentage weight loss, and cancer staging) was 0.341 ($P = 0.031$, confidence interval 0.128 to 0.909). Pre-operative percentage weight loss at the first time point after randomization was 4.1% [interquartile range (IQR) 1.7–7.0] in ONS group vs. 6.7% (IQR 2.6–10.8) in controls (Mann–Whitney U $P = 0.021$) and post-operatively was 7.4% (IQR 4.3–10.0) in ONS group vs. 10.2% (IQR 5.1–18.5) in controls ($P = 0.016$).

Conclusions Compared with dietary advice alone, ONS resulted in patients having fewer infections and less weight loss following surgery for colorectal cancer. We have demonstrated that pre-operative oral nutritional supplementation can improve clinical outcome in weight losing patients with colorectal cancer.

Keywords Colorectal; Supplements; Weight loss; Infections; RCT; Cancer

Received: 18 August 2016; Accepted: 2 November 2016

**Correspondence to:* Dr Sorrel Burden, School of Health Sciences, Jean McFarlane Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK. *Email* Sorrel.burden@manchester.ac.uk

© 2017 The Authors. Journal of Cachexia, Sarcopenia and Muscle published by John Wiley & Sons Ltd on behalf of the Society of Sarcopenia, Cachexia and Wasting Disorders
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Introduction

Internationally, in terms of annual incidence, colorectal cancer is the third most common cancer in men and the second in women.[1] Surgery, combined with either neo-adjuvant or adjuvant chemotherapy and/or radiotherapy in selected patients, is the mainstay of curative treatment for colon and rectal malignancies.[2] Malnutrition and weight loss have long been associated with an increased post-operative morbidity and mortality.[3–5] However, despite this, neither pre-operative nutritional assessment nor nutritional screening is commonly practiced or integrated into the care pathways for patients with colorectal cancer.[2]

In pre-operative patients with colorectal cancer, the prevalence of malnutrition has been reported as 36.4% by using subjective global assessment, whilst clinically significant or severe weight loss has been reported in 39%[6] and any weight loss in 47% of people with colorectal cancer.[7] The combination of a low muscle mass and impaired physical function referred to as sarcopenia[8] has been identified in 12% of people with pre-operative colorectal cancer by using computed tomography to measure muscle mass and is associated with older age, lower body mass index, and increased rate of post-operative complications.[9] A review of the literature on sarcopenia in abdominal malignancy concluded that sarcopenia is predictive of poorer clinical outcomes, increased morbidity, and increased hospital length of stay (LoS).[10] Within an 'enhanced recovery after surgery' (ERAS) programme for colorectal cancer, patients who were malnourished were found to be at increased risk of post-operative morbidity, delayed recovery of gastrointestinal function, and prolonged hospital stay.[11] Thus, there is considerable evidence that pre-operative malnutrition and sarcopenia are important prognostic factors for post-operative complications.[12] It is therefore reasonable to test whether or not nutritional interventions can decrease post-operative morbidity to improve clinical endpoints. Enhanced recovery after surgery programmes is now widespread, and patients with colorectal cancer are now cared for in their homes for an increased length of time both pre-operatively and post-operatively because admission is often on the day of surgery and post-surgical time to discharge has been drastically reduced.[13] Decreased hospital LoS with ERAS programmes places more emphasis on pre-operative and post-operative supportive interventions that can be delivered in individuals' own homes.

Pre-operative nutritional interventions in gastrointestinal surgery have been evaluated in a Cochrane review, which includes a number of studies on immune-enhancing nutrition, oral nutritional supplements (ONSs), and parenteral nutrition.[14] Studies have demonstrated a reduction in post-operative complications with the use of pre-operative immune-enhancing nutrition.[15] However, many of these studies include well-nourished patients, excluding those on neo-adjuvant anticancer therapy, and were conducted prior to the implementation of ERAS programmes.[14] Studies that have looked at nutritional supplements in all participants undergoing colorectal surgery have demonstrated mixed results.[16,17]

All trials of ONS compared with either standard care or dietary advice in those with colorectal cancer included a mixture of well-nourished and malnourished participants.[18] In one trial in people with colorectal cancer, a subgroup analysis of those who had lost weight demonstrated a reduction in surgical site infections (SSIs) in the ONS group compared with the controls.[17] There is a paucity of evidence on ONS in people who have lost weight diagnosed with colorectal cancer during the pre-operative period, although it has been repeatedly demonstrated in the evidence base that there is an association between a poor pre-operative nutritional status and poor clinical outcomes.

The aim of this study was to determine if pre-operative ONS with dietary advice, compared with dietary advice only, can reduce post-operative infections in people prior to surgical resection for colorectal cancer who have previously lost weight.

Materials and Methods

In this multi-centre, single-blinded, randomized controlled trial, we studied people with colorectal cancer who had lost weight pre-operatively to determine the effectiveness of oral nutritional supplementation. The protocol was amended after commencement to incorporate the comprehensive complication index (CCI),[19] an extension to the widely used grading system of post-operative complications.[20] The protocol (version 6 August 2013) is available at www.manchester.ac.uk/research/sorrel.burden/publications.

Study participants

The sample was recruited from colorectal surgical clinics in six hospitals in the northwest of England from 1 November 2013 to 28 February 2015. All hospitals had an ERAS protocol in colorectal surgery in place.

Participants were recruited at colorectal clinics. Data collection took place in the participant's residence for baseline and pre-operative time points and either on a hospital ward or participant's residence for the post-operative visit. At baseline, the participants' characteristics were recorded along with nutritional status measurements. Baseline visits occurred when the participants agreed to be in the study within a couple of days of surgical teams, informing the participants that they were suitable for an operation. Participants were included in the study if they had a primary colorectal tumour, were over 18 years old, listed for radical surgery, had capacity for informed consent, and reported unintentional weight loss over the previous 3–6 months ($>$1 kg). This weight loss was based on a subgroup analysis from a previous trial that demonstrated in all participants with colorectal cancer who had lost weight a significant reduction in wound infections in the group receiving ONS compared with controls.[17] Participants were excluded if they were pregnant or had a pacemaker precluding the use of bioelectrical impedance analysis (BIA), already on a similar nutritional supplement, or had insulin dependent diabetes.

Ethical approval and trial registration

This study has been approved by the National Research Ethics Service Committee Northwest (12/NW/0208) and been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All study participants provided written informed consent prior to inclusion. The trial was pre-registered ISRCTN: NCT24668100.

Randomization

The participants were randomly allocated on a 1:1 ratio by using blocks of two ensuring equal numbers in each group. Allocation was stratified according to tumour site (rectal vs. colon) and surgical approaches (open vs. laparoscopic). Four lists of random numbers were produced by a statistician, and an independent researcher set up the randomization procedure for each of the strata. Sequentially numbered opaque sealed envelopes were used, which allowed block randomization sequence allocation to be implemented and ensured sequence allocation concealment. The participants were randomized to either dietary advice alone (control) or ONS and dietary advice (intervention). Identification and recruitment of participants were undertaken by National Institute Health Research cancer nurses or colorectal specialist nurses at each site. Randomization was undertaken after participants consented prior to the baseline measurements.

Intervention

The intervention comprised oral supplementation (Fortisip Compact®, 10.1 KJ, 0.096 g of protein per mL, Nutricia UK) at a dose of 250 mL daily. Supplements were started at the point of allocation of an operation date. A minimum of 5 days pre-operative treatment was given. A sealed box containing sufficient supplements for 7 days was left with the participants randomized to the intervention for each week prior to the planned operation date. A mixture of vanilla and strawberry flavours were provided. Dietary advice was given to all participants in the form of a leaflet (see Supplementary Material), which was left with the participant and which formed the basis of a structured discussion with the research assistant (DG), a nutritionist, at the baseline visit. The dietary advice given aimed to increase energy and protein intake through dietary means by increasing the amount of high fat, sugar, and protein-rich foods in the diet. The leaflet also recommended the use of dietary supplements high in energy and protein that could be purchased from high street retailers. All participants were advised of potential side effects of the ONS and advised to discontinue the intervention if adverse effects were experienced.

Controls

Following randomization, procedures for control group participants were identical to those for the intervention group. This group received the dietary leaflet and discussion with the nutritionist, and for the purposes of blinding, control participants were given sealed cardboard boxes of identical weight and appearance as the ONS group at the time of group allocation. These boxes contained bottled water in 125 mL bottles. Thus, the same quantity of either bottled water or ONS (14 bottles) was in each cardboard box. Similar to the intervention group, further supplies of water in sealed boxes were delivered as required up to 24 h prior to admission date for surgery. The research team was blind to the intervention, but the participants were not.

Adherence

All participants were asked to keep a diary of the drinks they consumed for each week prior to surgery after recruitment. They were asked how much and how frequently they consumed the drinks. This diary was then given back to the researcher in a sealed envelope and was not opened until the unblinding at the end of the trial.

Data collection

The participants were seen three times by the research assistant (DG) (i) at baseline, (ii) 24–48 h pre-operatively, and (iii) 5–7 days post operatively. Nutritional status measurements were recorded at each time point, and clinical outcomes were recorded from Day 1 up to 30 days post operatively.

Outcomes

The primary outcome was patients with one or more chest or SSI defined by using the US Centre for Disease Control definitions.[21] Secondary outcome measures included post-operative complications recorded prospectively from the participants' medical records using a standard classification.[20]

This classification was used to determine the CCI.[19] Hospital LoS was recorded from the day of surgery until date of discharge. The participants received a telephone interview 30 days after surgery, and any self-reported potential complications were recorded and followed up in the participants' medical records, although primary care records were not accessed due to permission restrictions.

Nutritional status measurements were recorded as secondary outcomes. Weight was measured to the nearest 0.1 kg on calibrated scales (Seca 875 flat scale) with shoes removed. Height was measured by using a portable stadiometer to the nearest 0.1 cm (Harpenden pocket stadiometer Practical Metrology, Sussex, UK) with participants looking straight ahead and shoes removed. Height and weight were required to calculate baseline body mass index, percentage weight loss, fat-free mass index (FFMI), and fat mass index (FMI) so are not reported as standalone measures. To calculate percentage weight loss, the participants were asked to recall their previous weight (3–6 months ago); this was used with actual weight recorded at all time points to determine percentage weight loss. Percentage weight loss is included as it is part of the criteria that is used to define malnutrition[22] and has universally been used as a prognostic variable for predicting post-operative outcome.[6] Handgrip strength was measured by using the non-dominant hand (Takei 5001 Grip Dynamometer Analogue). Three measurements were taken and the mean recorded. Bioelectrical impedance analysis was measured (Bodystat 1500 machine, Isle of Man, Bodystat Ltd) with participant being adequately hydrated and in a supine position. Fat-free mass and fat mass measured by BIA were standardized to FFMI and FMI by dividing by height squared.[23]

The nutritional screening and assessment tools used were (i) patient-generated subjective global assessment (PG-SGA)[24] and (ii) the malnutrition universal screening tool,[25] which were completed during the baseline visit to the participant's home. Only PG-SGA was recorded at subsequent visits. Dietary intake was assessed at each time point by using 24 h semi-structured dietary recall method which was self-reported. For the dietary recall, the participants were asked semi-structured questions on all food and drink eaten in the previous day, including quantities of food consumed by using household measures, volumes of fluids, cooking methods, and ingredients in any recipes. For the dietary recalls, food portions were converted to grams and analysed by using Microdiet (Version 2, Downlee Systems Limited, UK) to estimate total energy and protein intakes. The energy and protein content of the ONS consumed for the corresponding day from the diaries was added to the nutrient content assessed by 24 h recall for the pre-operative time point to estimate daily total energy and protein intakes in the intervention group. Quality of life data, anthropometric measurements, and data on adherence to nutritional aspects of ERAS protocols are available in the Supplementary Material.

Power and sample size

We calculated that 88 patients needed to complete the trial (44 in each arm) to meet the number required to detect a difference in infective complications (chest and SSI defined by CDC definitions[21]) using 80% power, alpha = 0.05 (based on p1 = 0.26 and p2 = 0.54 on one-sided significance using χ^2 test of equal proportions).[17] We allowed for 12% dropout or non-completion and recruited 101 patients.

Data analysis

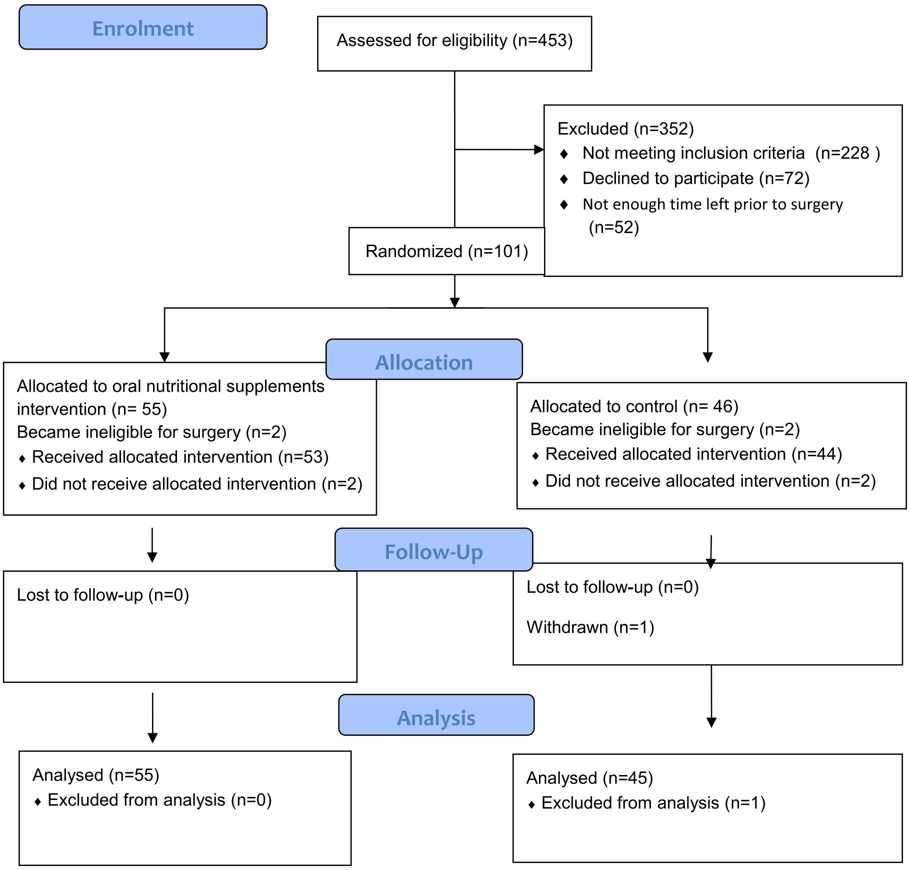
Means and standard deviations are used to describe normally distributed interval data, and for non-normally distributed data, median and interquartile ranges (IQRs) are displayed. Categorical data are displayed by using numbers and percentages. Confidence intervals (CIs) were calculated for primary and secondary outcomes where appropriate. To determine if there were any effects from random differences at baseline from prognostic variables, a logistic regression model was used. The dependent variable was infections, and the covariants were age, gender, baseline percentage weight loss, and cancer staging. Adjusted and unadjusted models were performed. Differences between groups were determined by using independent Student's *t*-tests for normally distributed interval data, and for skewed data, Mann–Whitney U-tests were performed. For dichotomous variables, χ^2 or Fisher's exact tests were used. For nominal data, Kruskal–Wallis test was used to determine the differences between groups. Data were analysed by using SPSS version 22.[26] There were no interim stopping guidelines for this trial.

Results

CONSORT diagram

We recruited and randomly allocated 101 patients, of whom 96 completed the trial. The CONSORT flow diagram of the participants through the trial is shown in Figure 1. One participant withdrew consent from the control group prior to the pre-operative visit. Participant characteristics and baseline measurements are shown in Table 1. Overall, there were more participants in the intervention group. At the point of recruiting participants, if it was undecided if surgery was open or laparoscopic, the default used was open-surgery stratum for randomization. The two arms of the trial were well matched with similar proportions of participants within each stratum, site of cancer, and type of operation (laparoscopic or open). Data were complete for the

Figure 1 CONSORT flow diagram.


Enrolment
Assessed for eligibility (n=453)
Excluded (n=352)
Not meeting inclusion criteria (n=228)
Declined to participate (n=72)
Not enough time left prior to surgery (n=52)
Randomized (n=101)
Allocation
Allocated to oral nutritional supplements intervention (n= 55)
Became ineligible for surgery (n=2)
Received allocated intervention (n=53)
Did not receive allocated intervention (n=2)
Allocated to control (n= 46)
Became ineligible for surgery (n=2)
Received allocated intervention (n=44)
Did not receive allocated intervention (n=2)
Follow-Up
Lost to follow-up (n=0)
Lost to follow-up (n=0)
Withdrawn (n=1)
Analysis
Analysed (n=55)
Excluded from analysis (n=0)
Analysed (n=45)
Excluded from analysis (n=1)

majority of baseline assessments, and missing data were recorded (Table 1). The trial was discontinued when recruitment numbers were sufficient to meet the power equations for both arms.

Primary outcome

The data on the primary outcome of infectious complications on an intention-to-treat analysis are shown in Table 2. The odds of a patient having a chest or an SSI in the ONS group compared with the control group was 0.532 ($P = 0.135$, CI 0.232 to 1.218) with an unadjusted analysis. However, when adjusted for random effects at baseline, the odds of a patient in the ONS group compared with the controls of having an infection (either chest or surgical site) was 0.341 ($P = 0.031$ CI 0.128 to 0.909). Each type of infection was evaluated and shown in Table 3. A significant difference was demonstrated between the ONS group and controls for SSI (odds ratio 0.41, CI 0.16 to 1.00, χ^2 $P = 0.044$) with a lower rate of SSI in the intervention arm compared with the control (20 vs. 38%). However, there was no difference for chest infections (Fisher's exact $P = 0.359$).

Secondary outcomes

On intention to treat analysis for total complications, there were no significant differences demonstrated. A total of 48 participants had a complication: 23/55 (42%) in the intervention group and 25/45 (56%) in the control group (χ^2 $P = 0.114$). For complications graded I or II, there were 13/55 (24%) in the intervention group and 11/45 (24%) in the control group, and for grades III to IVb, there were 9 (16%) and 10 (22%) complications in the ONS and control groups, respectively. A total of five people died, one in the ONS and four in the control ($P > 0.05$). The median CCI score was 29.6 (IQR 20.9–47.3) and 29.6 (IQR 20.9–43.3) for the intervention and the control groups, respectively (Mann–Whitney U $P = 0.984$). Hospital LoS was recorded for 92 participants, and in the ONS group, the median LoS was 7 days (IQR 4.0–10.5) and in the controls group also 7 days (IQR 4.0–10.0; Mann–Whitney U $P = 0.630$). Multivariate analysis for complications and LoS are included in the Supplementary Material but showed no difference between groups with regard to infections, although unsurprisingly, there was a significant effect from disease staging (Table S13).

Table 1 Participants' characteristics and baseline clinical details

	Control Dietary advice only	Intervention ONS and dietary advice
Randomization	n = 46	*n* = 55
Age mean (SD)	68.9 (11.49)	70.5 (11.66)
Gender *n* (%)		
Male	32 (70)	35 (64)
Female	14 (30)	20 (36)
Occupation *n* (%)		
Professional	13 (28)	19 (35)
Skilled	16 (35)	19 (35)
Unskilled	14 (30)	12 (22)
Unemployed	0 (0)	2 (4)
Missing	3 (7)	3 (5)
Site of surgery *n* (%)		
Colon	29 (63)	35 (64)
Rectum	17 (37)	20 (36)
Type of surgery *n* (%)		
Laparoscopic	30 (65)	37 (67)
Open	16 (35)	18 (33)
Smoking status *n* (%)		
Never	12 (26)	27 (49)
Ex smoker	21 (45)	18 (33)
Current	9 (20)	10 (18)
Missing	4 (9)	0 (0)
Body mass index		
Mean (SD)	25.5 (4.54)	25.9 (4.8)
Missing 0		
Percentage weight loss		
Median (IQR)	6.8 (3.4–12.1)	4.90 (2.2–8.8)
Missing 0		
MUST *n* (%)		
0	13 (28)	28 (51)
1	16 (35)	13 (24)
2	14 (30)	9 (16)
3	1 (2)	1 (2)
4	0 (0)	4 (7)
Missing 0	2 (4)	
Handgrip strength		
mean (SD)	24.9 (9.3)	25.0 (10.6)
Missing	4	0
Patient-generated SGA		
median (IQR)	9 (4–12)	6 (4–10)
Missing 0		
Cancer staging *n* (%)		
I	1 (2)	2 (4)
II	4 (9)	9 (16)
III	26 (56)	26 (47)
IV	12 (26)	11 (20)
Villous adenoma	1 (2)	1 (2)
Missing	1 (2)	4 (7)
Did not have surgery	2 (4)	2 (3)
Anaesthetic risk score *n* (%)		
Normal health	5 (11)	2 (4)
Mild systemic disease	19 (41)	24 (44)
Severe systemic disease	8 (17)	7 (13)
Missing data	14 (30)	22 (40)
Neo-adjuvant treatment		
Radiotherapy short course	11 (23)	15 (27)
Chemotherapy short course	6 (13)	7 (18)
Chemotherapy long course	2 (4)	3 (5)
Number of comorbidities *n* (%)		
0	6 (13)	15 (27)
1	7 (15)	8 (15)
2	12 (26)	9 (16)
3	4 (9)	9 (16)
More than 4	8 (17)	8(15)
Missing	9 (20)	6 (11)

IQR, interquartile range; MUST, malnutrition universal screening tool; ONS, oral nutritional supplement; SGA, subjective global assessment.

Nutritional status and body composition

Nutritional status measurements are shown in Table 4. There was a significant difference in the percentage of weight lost between groups both pre-operatively (intervention 4.1% IQR 1.7–7.0 vs. controls 6.7% IQR 2.6–10, $P = 0.021$) and post-operatively (intervention 7.4% IQR 4.3–10 vs. controls 10.2% IQR 5.1–18.5, $P = 0.016$). However, there was no significant difference observed for the other measures of nutritional status recorded nor handgrip strength or PG-SGA ($P > 0.05$ in all cases). The changes in measurements among baseline, pre-operative, and post-operative time points for BIA are shown in Table 5, and descriptive statistics for BIA with between-group comparisons are in the supplementary material (Table S12). There were no differences in BIA between groups at the pre-operative or post-operative time points. However, for the mean difference between the intervention and the control from baseline to pre-operative time points for FFMI (ONS group -0.345 kg/m^2, IQR -3.241 to 0.160 vs. control group 0.100 kg/m^2, -0.520 to 0.3150 Mann–Whitney U $P = 0.008$), there was a significant difference but not for FMI (ONS group 0.105 kg/m^2, IQR -0.170 to 0.315 vs. control group -0.150 kg/m^2 IQR -0.310 to 0.135 $P = 0.083$).

Dietary intake

Data for self-report energy and protein intake by using 24-h recall are shown in Table 6.

Provision and adherence of ONS

Oral nutritional supplements were provided to the participants for a median of 8 days (IQR 5–15). Of the 53 participants in the intervention arm, 39 (74%) returned a diary detailing their adherence to the ONS. For participants who returned a diary, 29 (74%) participants managed all the supplements (two cartons daily), 2 (5%) participants reported that they managed one and a half cartons, 3 (8%) participants reported that they managed one carton, 3 (8%) managed half a carton, and 2 (5%) participants reported that they did not consume any of the ONS. Intolerance of ONS was reported by seven participants who did not manage to follow the ONS regimen; these included nausea reported by four participants, abdominal discomfort reported by three participants, and diarrhoea reported by two (two participants reported more than one intolerance symptoms). Thus, seven participants reported that they did not tolerate the supplements due to unpalatability.

Blood loss and duration of operation were included as variables at baseline but were missing in the majority of instances in medical records, so therefore are not reported.

Table 2 Logistic regression showing adjusted and unadjusted analyses for primary outcome (patients with one or more infections either a chest or surgical site) as dependent variable and independent variables age, gender, cancer staging, baseline percentage weight loss, and treatment group

	Unadjusted odds ratio	*P*-value	95% CI	Adjusted odds ratio	*P*-value	95% CI
Weight loss (%)	0.952	0.182	0.885 to 1.024	0.922	0.059	0.848 to 1.003
Age (years)	0.998	0.928	0.964 to 1.034	0.998	0.920	0.956 to 1.041
Gender	1.500	0.366	0.623 to 3.613	0.976	0.963	0.347 to 2.747
TNM staging						
Stage 1	4.250	0.341	0.216 to 83.517	8.903	0.239	0.234 to 338.3
Stage 2	8.500	0.112	0.609 to 118.637	11.113	0.079	0.760 to 162.5
Stage 3	2.361	0.275	0.505 to 11.049	3.034	0.185	0.588 to 15.662
Stage 4	4.250	0.020	1.252 to14.427	5.223	0.013	1.420 to 19.220
Treatment group (ONS/control)	0.532	0.135	0.232 to 1.218	0.341	0.031	0.128 to 0.909

CI, confidence interval; ONS, oral nutritional supplement; TNM, staging, tumour, nodal, metastases. Hosmer and Lemeshow test $P = 0.792$ for adjusted analysis.

Table 3 Intention to treat analysis for number of participants with chest, surgical site, or urinary tract infections

	Control		Intervention		
	n = 45(%)	95% CI	*n* = 55(%)	95% CI	*P*-value
Surgical site infection	17 (38)	25.1 to 52.4	11 (20)	11.6 to 32.4	[a]0.044
Chest infection	3 (7)	2.3 to 17.9	5 (9)	3.9 to 19.6	[b]0.359
Urinary tract infection	6 (13)	6.3 to 26.2	4 (7)	2.9 to 17.3	[a]0.315

CI, confidence interval.

[a] χ^2.

[b] Fisher's exact test.

Table 4 Nutritional status measurements and screening and assessment tools in control and intervention groups

	24–48 h pre-operative				5–7 days post-operative			
	n	Control	ONS	*P*-value	*n*	Control	ONS	*P*-value
Handgrip mean (SD)	70	25.0 (8.52)	25.2 (10.07)	0.723	70	23.2 (7.85)	24.9 (9.89)	[a]0.394
Percentage weight loss median (IQR)	73	6.7 (2.6–10.8)	4.1 (1.7–7.0)	0.021	79	10.2 (5.1–18.5)	7.4 (4.3–10.0)	[b]0.016
PG-SGA score median (IQR)	69	6.5 (3.0–9.7)	4.0 (2.0–9.0)	0.215	72	12.0 (8.0–15.0)	10.0 (6.5–13.5)	[b]0.062

IQR, interquartile range; ONS, oral nutritional supplement; PG-SGA, patient-generated subjective global assessment.

[a] Independent Student's *t*-tests.

[b] Mann–Whitney U-test.

Table 5 Changes in bioelectrical impedance analysis between baseline and measurements at pre-operative and post-operative time points

	Difference between baseline and pre-operative measurements (*n* = 69)			Difference between baseline and post-operative measurements (*n* = 64)		
	Controls	ONS	*P*-value	controls	ONS	*P*-value
Fat-free mass index kg/m^2 median (IQR)	0.100 (−0.520, 0.3150)	−0.345 (−3.241, 0.160)	0.008[a]	0.720 (−0.190, 1.865)	0.080 (−1.010, 0.960)	0.100
Fat mass index kg/m^2 median (IQR)	−0.150 (−0.310, 0.135)	0.105 (−0.170, 0.315)	0.083[a]	0.200 (−0.220, 0.885)	0.410 (−0.100, 1.090)	0.242

IQR, interquartile range; ONS, oral nutritional supplement.

[a] Mann–Whitney U-test.

Table 6 Dietary intake at each time point for energy and protein intakes, including additional nutrition from oral nutritional supplements at pre-operative time point

	Energy (KJ) Median (IQR)			Protein (g) Median (IQR)		
Time point *n* = participants	Control	ONS	*P*-value	Control	ONS	*P*-value
Baseline *n* = 93	6085(4743–7493)	6407 (4233–8193)	0.760	68 (48–83)	57 (41–76)	0.271
Pre-operative *n* = 70	6350 (4714–6350)	8120 (6490–9831)	0.001	63 (49–78)	79 (67–97)	0.018
Post-operative *n* = 89	4499 (3218–6416)	5302 (3973–7173)	0.282	46 (31–70)	60 (43–70)	0.181

IQR, interquartile range; ONS, oral nutritional supplement.
Mann–Whitney U-tests.

Discussion

This is the first single-blind randomized controlled trial of pre-operative ONS based on standard nutritional recommendations[27] in people with colorectal cancer who have lost weight. We demonstrate on intention-to-treat analysis significantly fewer infections in the ONS and dietary advice group compared with controls who received dietary advice alone. Prior to surgery, the participants who received ONS with dietary advice lost significantly less weight, and this difference was maintained post-operatively. Percentage weight loss has long been regarded as a prognostic indicator for post-operative morbidity,[28] and this study demonstrates that a nutritional intervention is clearly linked to fewer post-operative infectious complications. The primary outcome was participants who incurred one or more infections (chest or surgical site) post operative.

It is notable in the analysis of BIA data that the participants who lost weight had less of a difference in muscle lost between baseline and pre-operative time points. This suggests that when people with colorectal cancer lose weight, they are losing fat-free mass. Fat-free mass is commonly used as a surrogate marker for skeletal muscle mass.[29] Reduction in skeletal muscle mass is linked to sarcopenia, which is known to be associated with reduced function and increased frailty in older people.[28] This is important as sarcopenia has been consistently linked to a poorer post-operative outcome and is also known to impact negatively on other cancer therapies.[30] Preventing weight loss and subsequent skeletal muscle mass would therefore seem to be a logical therapeutic strategy, especially given that it can be achieved by using a relatively inexpensive nutritional intervention that is easy to administer in the community pre-operatively.

The compliance rates for such an intervention in this trial are more than acceptable with two-thirds of participants managing more than 75% of the recommended dose. Most participants who were randomized to the ONS were able to consume some of the drink, and we recorded a 71% adherence rate. This supports the argument that weight loss in colorectal cancer during the perioperative period is preventable, and the benefits of increasing nutritional intake pre-operatively are sustained throughout the perioperative period.

Oral nutritional supplementation was effective at increasing the nutritional intake of the participants randomized to the intervention in relation to energy and protein. In this study, we did not look at the micronutrient profile of the participants' intake, although this may be important if subclinical levels or deficiencies of vitamins and minerals are present prior to surgery. The ONS used in this study contained a full profile of vitamins and minerals in amounts proportionate to the macronutrient composition. So, it is unclear if it were either macronutrients in ONS that are having a positive effect or indeed the mixture of nutrients within the substrate. Micronutrients have been shown to be important in the perioperative period,[31] and there are some vitamins and minerals which may be deficient particularly in older people.[32] It is also of note that ONSs supplemented with immune-enhancing agents administered pre-operatively have resulted in a reduction in infections in patients with colorectal cancer.[33]

People with colorectal cancer lose weight due to symptom load, psychological distress, and adjuvant treatment effects.[34] Weight loss in the perioperative period influences people's lives during recovery and post-operative rehabilitation.[34] Other researchers have highlighted that nutritional information is a priority for individuals and their families.[35] In a recent study of body composition, only 5% of people with colorectal cancer were found to be cachexic based on measurements of fat-free mass pre-operatively, suggesting that weight loss is a treatable consequence of colorectal cancer.[36] In this study, the participants recruited had lost more than 1 kg of weight. This was based on a subgroup analysis of a previous trial, which showed that all weight-losing participants who were randomized to receive ONS compared with controls had significantly fewer wound infections.

This is the first single-blind trial that we are aware of with ONS based on standard nutritional requirements in people with colorectal cancer. Improvements to decrease bias in future would be to double blind the trial so both participants and researchers are unaware of the allocation. Whilst this design was considered for this trial, it was not possible to obtain a non-active drink packaged as a placebo. The initial power equations indicated that there should be a 28% difference in chest and SSIs between groups, although the actual difference was 21%. This is most likely due to evolving practice in surgery-improving perioperative management as there has been the introduction of ERAS programmes and also an increase in laparoscopic surgery, resulting in fewer complications overall.[37,38] Nutrition is a supportive therapy in surgical oncology, and studies evaluating supportive therapies are subject to confounding factors that include technological developments and procedural alterations to the primary therapy as well as patient variables. This trial was designed to determine effectiveness and therefore recruited participants at the time point that was appropriate for each individual within the care pathway. This meant that the participants were given ONS for different lengths of time pre-operatively. Future trials could standardize the length of time participants received ONS but also account for other confounders such as neo-adjuvant therapies by stratification because these factors also influence pre-operative management. Interestingly, the regression analysis showed that disease staging had a significant effect on infections and also on patients with one or more complications. This would therefore be a factor to consider in future trials.

The limitations of the study include reliance on participant recall to determine weight loss over 3–6 months which may be subject to bias. However, there is no reason to believe that this potential source of bias was differentially distributed across groups. There were also random differences in baseline variables which included differences in percentage weight loss and staging of disease, albeit these were adjusted for in the analysis. Also, individuals were required to fill in diaries to assess adherence, again potentially subject to reporting bias. An alternate method of determining adherence is to collect the empty cartons, or it may be preferable to develop a medication event monitoring system, although such systems are costly and themselves have disadvantages.[39] Dietary assessment was undertaken by 24 h recall which only allows dietary intake to be assessed over a short period. However, a reasonable level of validity has been demonstrated with trained nutritionists and with the assessment of macronutrients.[40] Technological developments surrounding the use of smartphone applications and user-friendly databases offer more options in future trials to directly record dietary intake for nutritional analyses.

We aimed to collect data pre-operatively and post-operatively, although the participants were reluctant in some instances to see the researcher immediately pre-operatively, possibly due to anxiety levels about impending surgery. Likewise, some were reluctant post-operatively as they were still recovering from surgery. This led to some missing data at these time points. Future research may need to consider dropout rates in the region of 20–25%. Alternative outcomes which are part of routine care may be used to determine nutritional status or body composition.[36] This could include the use of computed tomography for body composition and using food diaries undertaken as part of ERAS programmes.

The findings of this RCT are encouraging in that we demonstrate a reduction of post-operative infections by pre-operative nutritional supplementation. This supports the use of ONS in people who have lost weight prior to surgery for colorectal cancer and concurs with other research promoting pre-operative optimization or prehabilitation.[41] Nutritional optimization of people with colorectal cancer should start with nutritional screening by using a validated screening tool[25] then implementing nutrition intervention into pre-operative assessment protocols. The opportunities for nutritional interventions should be recognized for patients having neo-adjuvant treatments as their pre-operative duration is longer than for patients proceeding directly to surgery.

On the basis of this trial, it seems likely that ONS increases energy and protein intakes, which results in less perioperative weight loss and preservation of skeletal muscle mass, resulting in a positive effect on clinical outcome in people who were otherwise losing weight prior to surgery for colorectal cancer.

Acknowledgements

The authors would like to thank the participants for taking part and the oncology research nurse (J Allsop) and colorectal specialist nurses (D Hitchen, M Parker, H Ashby, and D West) for assistance with recruiting participants. The authors would like to thank Nutricia UK for the provision of supplements for the trial.

All authors contributed to the study design. Data collection was undertaken by SB and DG, and analysis was undertaken by SB supported by MP. All authors contributed to interpretation of results and approved the final manuscript. All authors approved the final manuscript.

ClinicalTrials.gov NCT24668100.

The trial was pre- registered ISRCTN: NCT24668100. This study has been approved by the National Research Ethics Service Committee Northwest (12/NW/0208) and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All study participants provided written informed consent prior to inclusion. The authors certify that they comply with the ethical guidelines for Publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2015.[42]

This study was funded by Macmillan Cancer Support and British Dietetics Association.

Online supplementary material

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Conflict of interest

Dr Sorrel Burden received a travel grant received to attend a scientific meeting from Nutricia UK in 2013.

All other authors have no conflict of interests to declare.

Reference

1. Ferlay J SI, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0,Cancer Incidence and Mortality Worldwide. In: [Internet]. ICN, ed. Available from: http://globocan.iarc.fr,: accessed 27/11/2015 2013.
2. National Institute for Clinical Excellence. Colorectal cancer: the diagnosis and management of colorectal cancer. NICE

Guideline 131 guidance.nice.org.uk/cg131: Accessed 22 October 2014; 2011.

3. Garth AK, Newsome CM, Simmance N, Crowe TC. Nutritional status, nutrition practices and post-operative complications in patients with gastrointestinal cancer. *J Hum Nutr Diet* 2010;**23**:393–401.
4. Hardy SE, Dubin JA, Holford TR, Gill TM. Transitions between states of disability and independence among older persons. *Am J Epidemiol* 2005;**161**:575–584.
5. Schwegler I, von Holzen A, Gutzwiller JP, Schlumpf R, Muhlebach S, Stanga Z. Nutritional risk is a clinical predictor of postoperative mortality and morbidity in surgery for colorectal cancer. *Br J Surg* 2010;**97**:92–97.
6. Barbosa LR, Lacerda-Filho A, Barbosa LC. Immediate preoperative nutritional status of patients with colorectal cancer: a warning. *Arq Gastroenterol*; **51**:331–336.
7. Burden ST, Hill J, Shaffer JL, Todd C. Nutritional status of preoperative colorectal cancer patients. *Journal of human nutrition and dietetics : the official journal of the British Dietetic Association* 2010;**23**:402–407.
8. Muscaritoli M, Anker SD, Argil J, Aversa Z, Bauer JM, Biolo G, et al. Consensus definition of sarcopenia, cachexia and pre-cachexia: joint document elaborated by special interest groups (SIG) “cachexia–anorexia in chronic wasting diseases” and “nutrition in geriatrics”. *Clin Nutr* 2010;**29**:154–159.
9. Huang DD, Wang SL, Zhuang CL, Zheng BS, Lu JX, Chen FF, et al. Sarcopenia, as defined by low muscle mass, strength and physical performance, predicts complications after surgery for colorectal cancer. *Colorectal Dis* 2015;**17**:O256–O264.
10. Gibson DJ, Burden ST, Strauss BJ, Todd C, Lal S. The role of computed tomography in evaluating body composition and the influence of reduced muscle mass on clinical outcome in abdominal malignancy: a systematic review. *Eur J Clin Nutr* 2015;**10**:1079–1086.
11. Lohsiriwat V. The influence of preoperative nutritional status on the outcomes of an enhanced recovery after surgery (ERAS) programme for colorectal cancer surgery. *Tech Coloproctol* 2014;**18**:1075–1080.
12. Reisinger KW, van Vugt JL, Tegels JJ, Snijders C, Hulsewe KW, Hoofwijk AG, et al. Functional compromise reflected by sarcopenia, frailty, and nutritional depletion predicts adverse postoperative outcome after colorectal cancer surgery. *Ann Surg*; **261**:345–352.
13. Nygren J, Soop M, Thorell A, Hausel J, Ljungqvist O. An enhanced-recovery protocol improves outcome after colorectal resection already during the first year: a single-center experience in 168 consecutive patients. *Journal of the colon and rectum* 2009;**52**:978–985.
14. Burden S, Todd C, Hill J, Lal S. Pre-operative nutrition support in patients undergoing gastrointestinal surgery. *Cochrane Database Syst Rev* 2012;**11**:CD008879.
15. Gianotti L, Braga M, Nespoli L, Radaelli G, Beneduce A, Di Carlo V. A randomized controlled trial of preoperative oral supplementation with a specialized diet in patients with gastrointestinal cancer. *Gastroenterology* 2002;**122**:1763–1767.
16. Smedley F, Bowling T, James M, Stokes E, Goodger C, O’Connor O, et al. Randomized clinical trial of the effects of preoperative and postoperative oral nutritional supplements on clinical course and cost of care. *Br J Surg* 2004;**91**:983–990.
17. Burden ST, Hill J, Shaffer JL, Campbell M, Todd C. An unblinded randomised controlled trial of preoperative oral supplements in colorectal cancer patients. *J Hum Nutr Diet* 2011;**24**:441–448.
18. Burden S, Todd C, Hill J, Lal S. Preoperative nutrition in patients undergoing gastrointestinal surgery. *Cochrane Database Syst Rev* 2012;**11**:CD008879.
19. Slankamenac K, Graf R, Barkun J, Puhan MA, Clavien P-A. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Ann Surg* 2013;**258**:1–7.
20. Dindo D, Demartines N, Clavien PA. Classification of surgical complications a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;**240**:205–213.
21. Horan TC, Andrus M, Dudeck MA. CDC/NHSN surveillance definition of health care-associated infection and criteria for specific types of infections in the acute care setting. *Am J Infect Control* 2008;**36**:309–332.
22. National Institute for Health and Clinical Excellence. *Clinical Guideline 32 Nutrition support in Adults: Oral Nutrition Support, Enteral Tube Feeding and Parenteral Nutrition*. London: National Institute for Clinical Excellence; 2006.
23. VanItallie TB, Yang MU, Heymsfield SB, Funk RC, Boileau RA. Height-normalized indices of the body’s fat-free mass and fat mass: potentially useful indicators of nutritional status. *Am J Clin Nutr* 1990;**52**:953–959.
24. Bauer JCS, Ferguson M. Use of the scored patient-generated subjective global assessment (PG-SGA) as a nutrition assessment. *Eur J Clin Nutr* 2002;**56**:779–785.
25. Elia M, Chairman of MAG. The MUST Report: Nutritional screening of adults Redditch Worcs: BAPEN 2003.
26. IBM.Corp. SPSS Statistics for Windows. New York Armonk Released 2011.
27. The Commission of the European Communities March 1999 on dietary foods for special medical purposes. Commission Directive 1999 21 EC of 25 Official Journal of the European Communities 1999.
28. Dasgupta M, Rolfson DB, Stolee P, Borrie MJ, Speechley M. Frailty is associated with postoperative complications in older adults with medical problems. *Arch Gerontol Geriatr* 2009;**48**:78–83.
29. Schutz Y, Kyle UUG, Pichard C. Fat-free mass index and fat mass index percentiles in Caucasians aged 18–98 y. *International Journal of Obesity & Related Metabolic Disorders* 2002;**26**:953–960.
30. Baracos V, Kazemi-Bajestani SMR. Clinical outcomes related to muscle mass in humans with cancer and catabolic illnesses. *Int J Biochem Cell Biol* 2013;**45**:2302–2308.
31. Sriram K, Perumal K, Alemzadeh G, Osei A, Voronov G. The relationship between immediate preoperative serum 25-hydroxyvitamin D3 levels and cardiac function, dysglycemia, length of stay, and 30-d readmissions in cardiac surgery patients. *Nutrition* 2015;**31**:820–826.
32. ter Borg S, Verlaan S, Hemsworth J, Mijnarends DM, Schols JMGA, Luiking YC, et al. Micronutrient intakes and potential inadequacies of community-dwelling older adults: a systematic review. *Br J Nutr* 2015;**113**:1195–1206.
33. Moya P, Miranda E, Soriano-Irigaray L, Arroyo A, Aguilar M-d-M, Bellón M, et al. Perioperative immunonutrition in normo-nourished patients undergoing laparoscopic colorectal resection. *Surg Endosc* 2016;1–8.
34. Burden ST, Stamataki Z, Hill J, Molasiotis A, Todd C. An exploration of food and the lived experience of individuals after treatment for colorectal cancer using a phenomenological approach. *J Hum Nutr Diet* 2016;**29**:137–145.
35. Short V, Atkinson C, Ness AR, Thomas S, Burden S, Sutton E. Patient experiences of perioperative nutrition within an enhanced recovery after surgery programme for colorectal surgery. *Colorectal Dis* 2016;**18**:74–80.
36. Gibson DJ, Lal S, Strauss B, Todd C, Pilling M, Burden S. Computed tomography compared with standard clinical measurements to assess body composition, facilitating the identification of sarcopenia and cachexia in colorectal malignancy. *J Cachexia Sarcopenia Muscle* 2015;**6**:69.
37. Rawlinson A, Kang P, Evans J, Khanna A. A systematic review of enhanced recovery protocols in colorectal surgery. *The Annals of The Royal College of Surgeons of England* 2011;**93**:583–588.
38. The Colon cancer Laparoscopic or Open Resection Study Group. Laparoscopic surgery versus open surgery for colon cancer: short-term outcomes of a randomised trial. *Lancet Oncol* 2005;**6**:477–484.
39. Farmer KC. Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. *Clin Ther* 1999;**21**:1074–1090.
40. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, et al. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br J Nutr* 1994;**72**:619–643.
41. Carli F, Scheede-Bergdahl C. Prehabilitation to enhance perioperative care. *Anesthesiol Clin* 2015;**33**:17–33.
42. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2015. *J Cachexia Sarcopenia Muscle* 2015;**6**:315–316.

Session One Activity

Reviewing an abstract: Abstract 1

Please review the abstract provided and try to answer the following questions. Be as specific as possible.

1. What is the population of interest?

2. What is the intervention?

3. What is the comparator?

4. What are the outcomes of interest?

5. What questions do you have about why the study was designed this way?

Session One Activity

Reviewing an abstract: Abstract 2

Please review the abstract provided and try to answer the following questions. Be as specific as possible.

1. What is the population of interest?
2. What is the intervention?
3. What is the comparator?
4. What are the outcomes of interest?
5. What questions do you have about why the study was designed this way?

Session One Activity

Identifying and analyzing clinical trials

Using clinicaltrials.org, identify a clinical trial--ideally one you are not yet familiar with. Then answer as many of the following questions as you can:

1. What is the question the research is trying to answer?

2. What are the inclusion criteria? What are the exclusion criteria?

3. Why do you think these particular exclusion and inclusion criteria were selected?

4. How long will this trial take? Why do you think it will take that amount of time (and not longer or shorter)?

5. From a patient and caretaker perspective, what questions do you have about this trial?

Alliance
FOR AGING RESEARCH

Catalyzing Innovation
for Healthy Aging

Session Two

Recent Trends in Incorporating Patient and Family Caregiver Perspectives

Sue Peschin, MHS
President and CEO
Alliance for Aging Research

“Experience is the mother of science.”

Anonymous

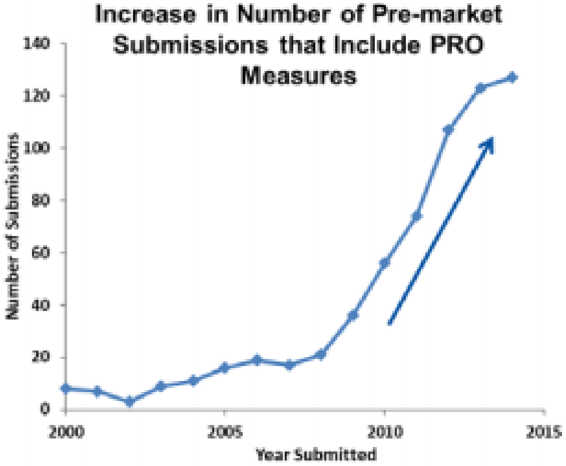
Intro quote to new documentary, “Will I Be Next?”

Patient-Centered Outcomes Research (PCOR) in Practice

Patient-centered practices in medical research and healthcare are on the rise!

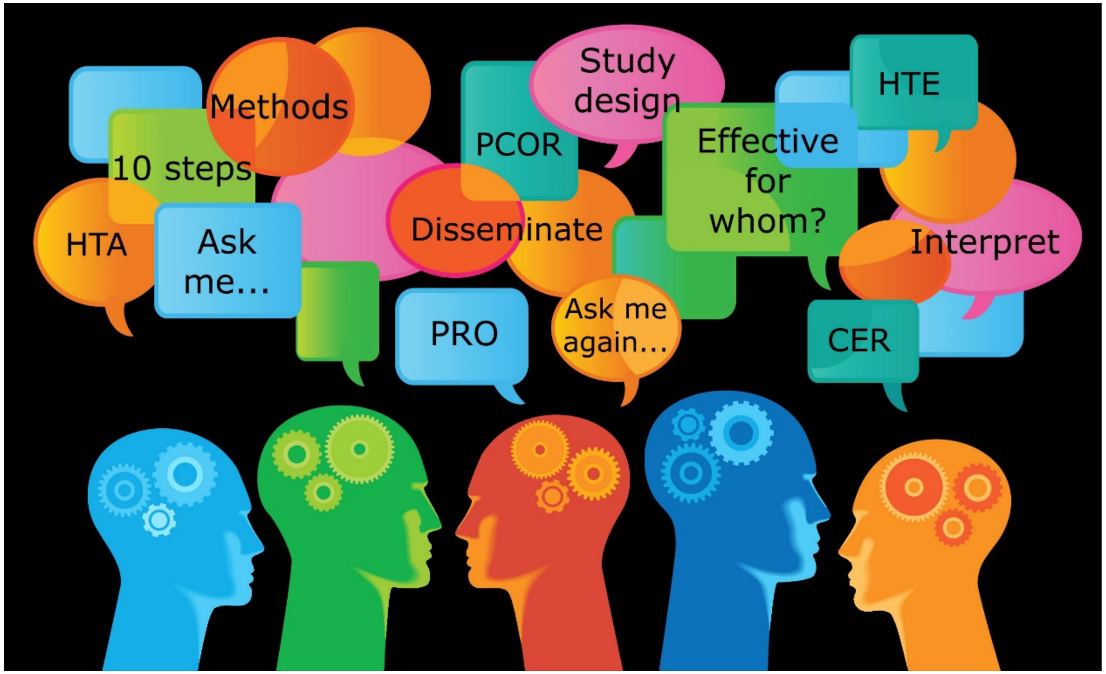
- The mindset of healthcare stakeholders is changing
- Momentum is building to incorporate patient preferences into the biomedical research & development system

SWEET, SWEET VICTORY!
WAY TO GO
memeshappen.com


Increase in Number of Pre-market Submissions that Include PRO Measures

PCOR in Practice

Congress, federal agencies, and private industry have started to better incorporate the patient perspective


Methods
10 steps
HTA
Ask me…
PCOR
Study design
Disseminate
PRO
Ask me again…
Effective for whom?
HTE
Interpret
CER

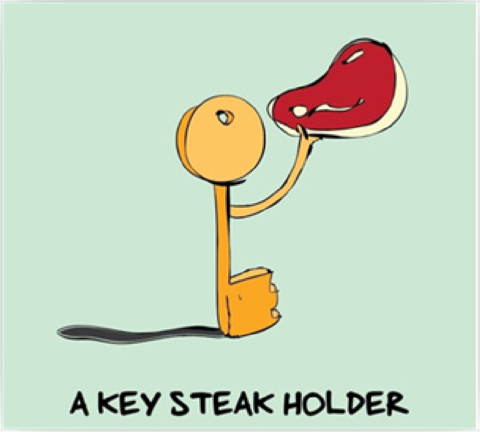
Congress & PCORI

Congress recognized that traditional health research hadn’t been able to answer questions or clarify choices that patients and their caregivers face

Their solution:

in 2010, Congress authorized the creation of the **Patient-Centered Outcomes Research Institute—PCORI—**as part of healthcare reform

Congress & PCORI
What PCORI Does



PCORI strives to improve the quality and relevance of evidence available to help patients, caregivers, clinicians, employers, insurers, and policy makers make better-informed health decisions.

PCORI has two main focus areas:

- Comparative Effectiveness Research
- Patient and Family Caregiver Engagement Projects, such as the Senior Patient and Family Caregiver Network!

Congress & PCORI
PCORI's goals

1. Increase information available to support health decisions
2. Speed the implementation and use of PCOR
3. Influence clinical and healthcare research funded by others

Congress & 21st Century Cures

The 21st Century Cures Act strives to accelerate both medical product development and market release

The law builds on FDA's ongoing work to incorporate the patient perspective

Catalyzing Innovation for Healthy Aging

Federal Agencies: FDA

Dr. Janet Woodcock, *Director of the Center for Drug Evaluation and Research, U.S. Food and Drug Administration (FDA),* on the new era of patient empowerment:

“It turns out that what is really bothering the patient and what is really bothering the doctor can be radically different things….patients are true experts in their disease.”

“It's clear you have to start with an understanding of the impact of the disease on the people who have it, and what they value most in terms of alleviation before you set up a measurement and go forward with truly patient-focused drug development.”

Federal Agencies: FDA

Patient-Focused Drug Development Initiative (PFDD)

Part of FDA’s commitments under PDUFA V, enacted in 2012, they convened 24 meetings on specific disease areas in FY 2013-17, including neuropathic pain and sarcopenia.

Federal Agencies: FDA

Patient-Focused Drug Development Initiative (PFDD)

The PFDD meetings:

- Support FDA staff in conducting the risk benefit-risk assessment;
- Assist in advising industry on their drug development programs; and
- Support drug development more broadly through identifying specific areas of unmet need

PATIENT FOCUSED DRUG DEVELOPMENT
PUBLIC MEETING

Federal Agencies: Centers for Medicare & Medicaid Services (CMS)

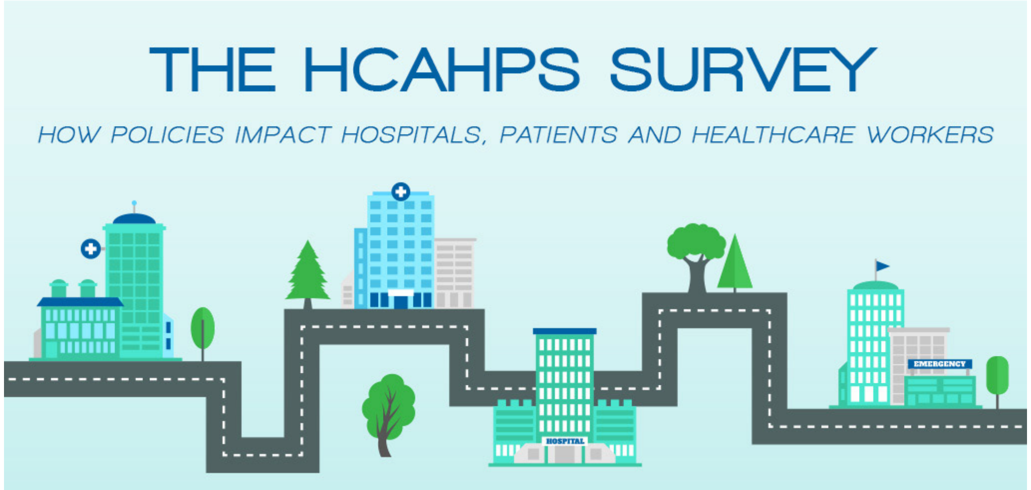
Consumer Assessment of Healthcare Providers and Systems (CAHPS)

- Patient experience surveys focus on how patients experienced or perceived key aspects of their care
- NOT satisfaction surveys
- Types of care provider surveyed include:
  1. health care providers and plans,
  2. hospitals,
  3. home health care agencies,
  4. doctors,
  5. health and drug plans

YOUR
VOICE
MATTERS!

Federal Agencies: CMS

One CAHPS Survey, the **Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Survey**, is the first national, standardized, publicly reported survey of patients' perspectives of their hospital care and experience


THE HCAHPS SURVEY
HOW POLICIES IMPACT HOSPITALS, PATIENTS AND HEALTHCARE WORKERS
HOSPITAL
EMERGENCY

Federal Agencies: CMS

Other CMS CAHPS Surveys:

- Home Health CAHPS
- HCBS CAHPS Survey
- Fee-for-Service CAHPS
- Medicare Advantage and Prescription Drug Plan CAHPS
- In-Center Hemodialysis CAHPS
- Nationwide Adult Medicaid CAHPS
- Hospice
- Outpatient and Ambulatory Surgery
- CAHPS® Survey for Accountable Care Organizations Participating in Medicare Initiatives
- CAHPS
- CAHPS for PQRS
- CAHPS for MIPS

IT'S SURVEY TIME

Federal Agencies: CMS

Ways to Get Involved:

1. Measures that focus on a particular condition, population, or specialty—identify your interest areas
2. Reach out to the professional society or advocacy group to become a subject matter expert
3. Join standing committees at specialty associations, broader professional associations, and the National Quality Forum (NQF)

GET INVOLVED

Federal Agencies: CMS

Ways to Get Involved:

4. Attend open meetings at NQF (see qualityforum.org for schedules) and/or give comments
5. Submit comments on measures currently in development or participate in the public comment period
6. Become a member of a Technical Expert Panel (TEP)

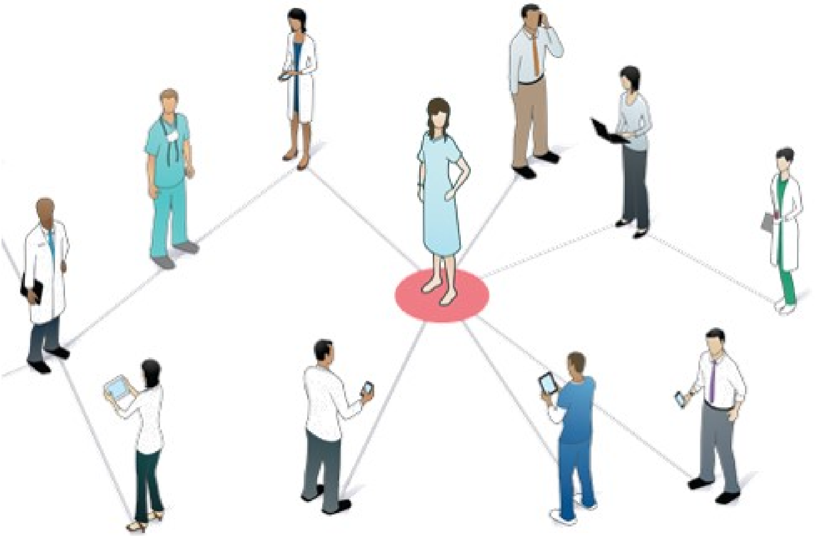
For more information, see the CMS Events Calendar:
https://go.cms.gov/2L7B0CZ

Private Industry Efforts

Establishing the relationship of value to patient outcomes

- How best to engage from the beginning?
- Advocacy and market research groups—no longer only commercial aspects of launch and post-launch activities, such as disease awareness and education

The patient perspective on value is especially important now, as focus on cost is central in political discussion



“Opportunities don't often come along. So, when they do, you have to grab them.”

Audrey Hepburn

Catalyzing Innovation for Healthy Aging

38 YEARS
HELP HOPE EDUCATION
American Chronic Pain Association

Involvement in PCORI

What is needed

- Understanding
- Time commitment
- Expectations
- Follow up

pcori
PATIENT-CENTERED OUTCOMES RESEARCH INSTITUTE

38 YEARS
HELP HOPE EDUCATION
American Chronic Pain Association

Understanding

- What is the topic being investigated
  - Pain
  - Cancer
  - Diabetes
- Who all is involved
  - Patients Stakeholders
  - Clinicians
  - Researchers
- What is your role
  - Patient Partners & Families (have real life experience)
  - Advisor

Time Commitment

- Depending on the study team and topic
- Review documents
- Time committed to study varies
  - Focus groups
  - Peer groups
  - One or two face-to face meetings each year of study
  - Study can be 3 to 5 years
- Calls
  - Weekly calls
  - Monthly calls
  - Quarterly calls

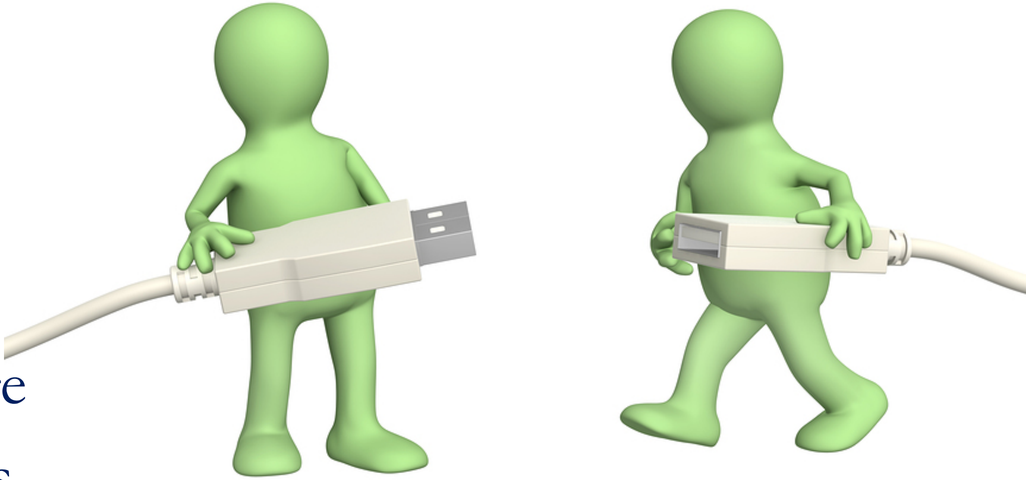
38 YEARS
HELP HOPE EDUCATION
American Chronic Pain Association

Expectations

- It is what you make it
- Share your experience
- Have a positive impact on having the patient/caregiver voice heard
- The patient voice is valuable in all levels of care and research
- Create change in practice

Follow Up

- It is important that the project you are involved in ensures that the outcomes
  - Published
  - Easily accessible
  - Creates positive change
  - Connects with patients needs



Follow Up

Involving Patients as Equal Partners

Kaiser Permanente, Portland, OR

https://youtu.be/8WtNjU2RF7Q

Gail Hunt

Founder, National Alliance for Caregiving

Member, Patient-Centered Outcomes Research Institute Board of Governors

How We Engage Patients and Other Stakeholders

Engagement Awards

Research Partners

Advisory Panels

Merit and Peer Review

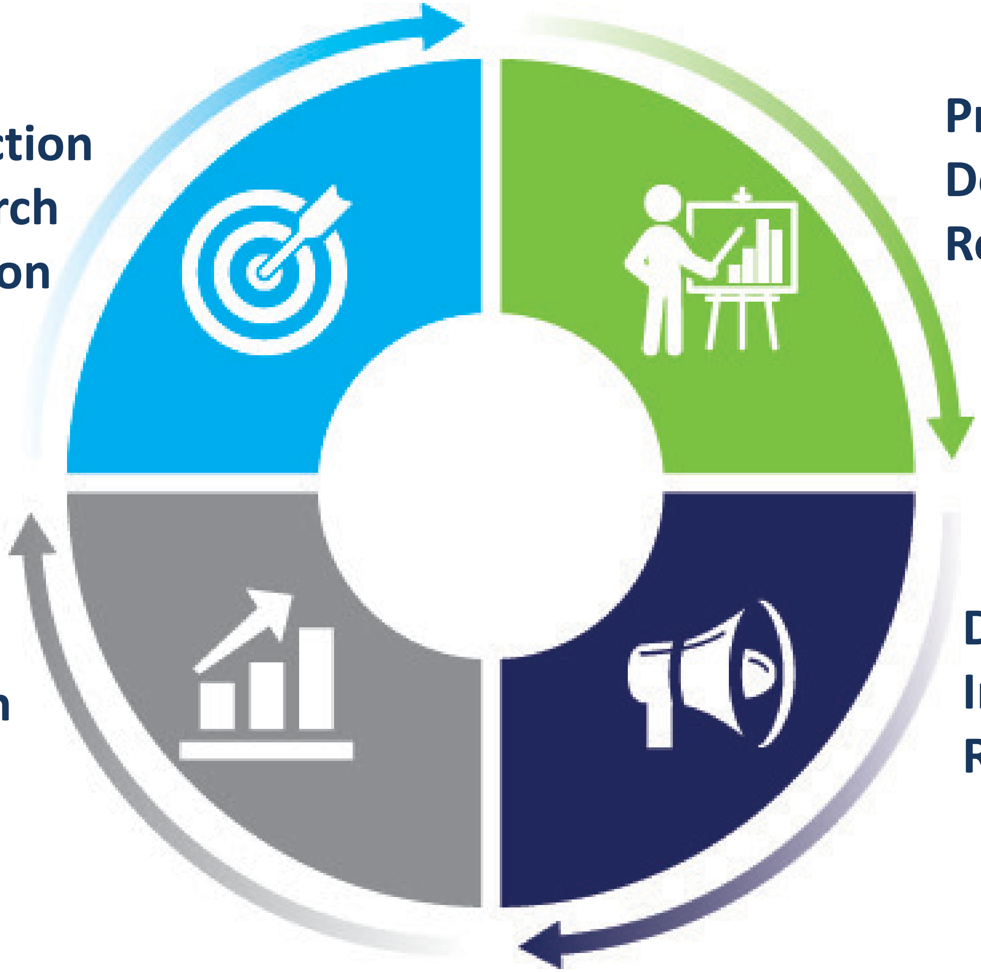
Webinars and Workshops

Ambassadors

Specific Opportunities for Patients/Advocacy Groups

- **Eugene Washington PCORI Engagement Awards**. Support projects to build a community of patients and other stakeholders who can participate in CER/PCOR and help disseminate the results of PCORI-funded studies.
- **PCORI Ambassadors**. Individuals and organizations who can share PCORI's vision and mission with their communities, participate as full partners in research, and help ensure the sharing and use of information generated by PCORI-funded projects.
- **Merit Review/Peer Review**. Patients and caregivers are integral to our process for assessing proposals for PCORI research funding and the results of our funded projects, to ensure patient-centeredness, meaningful engagement, and will yield evidence likely to help patients make better-informed healthcare decisions.
- **Advisory Panels.** Our Advisory Panel on Patient Engagement ensures we are patient-centered in everything we do, while patient/caregiver representatives are central to the four other panels that advise us on specific areas of science.

Engagement as a Path to Useful, High-Quality Research



Topic Selection and Research Prioritization

Proposal Review; Design and Conduct of Research

Dissemination and Implementation of Results

Evaluation

PCORnet, the National Patient-Centered Clinical Research Network

Developed with PCORI funding to:

- Improve the nation's capacity to conduct clinical research faster, more efficiently and less expensively, with greater power
- Establish a large, highly representative, national patient-centered clinical research network with a focus on conducting randomized and observational comparative studies
- Support a learning US healthcare system, which would allow for large-scale research to be conducted with greater accuracy and efficiency within real-world care-delivery systems

pcornet® The National Patient-Centered Clinical Research Network

PCORI, PCORnet, and the People-Centered Research Foundation

- PCORI’s funding of PCORnet will be ramping down as part of a long-planned effort to ensure network sustainability.
- PCRF is a nonprofit charitable organization created by PCORnet leaders and initially funded by PCORI to oversee the network in the future. It will serve as PCORnet’s program office and coordinate funding to ensure long-term sustainability.
- Under PCRF’s direction, PCORnet will evolve to ensure that all parts of the network are committed and able to conduct highly efficient, networked, and patient-centered studies, supported by a range of public and private research funders.

How PCORI-Funded Projects Work with Patient Groups

The Healthy eHeart Alliance Patient-Powered Research Network

- Cardiovascular disease research network within PCORnet.
- Patient groups involved include American Heart Association, Mended Hearts, StopAfib.org, and Sudden Arrhythmia Death Syndromes Foundation.
- Awardee under Partnership to Conduct Research within PCORnet initiative.

Comparative Effectiveness of Health System-based versus Community-based Dementia Care

- Seeks to improve care management for people with dementia and their caregivers. Stakeholder advisors include representatives of a number of Alzheimer’s disease organizations and support groups.

Optimizing Care for Patients with Dementia: A comparison of two non-pharmacological treatment approaches

- Seeks to improve quality of life for nursing home residents with dementia through alternative approaches to reduce disruptive behaviors. Stakeholder advisors include Alzheimer's Assn of Northern CA/NV and resident caregiver.

FDA and Patient/Caregiver Engagement

Debbe McCall, MBA

Patient PI and elected Chair, Health eHeart Alliance

SGE/Patient Representative, FDA Cardiovascular AdComm

Moderator, Atrial Fibrillation Support Forum on Facebook

@DebbeMcCall on Twitter and LinkedIn

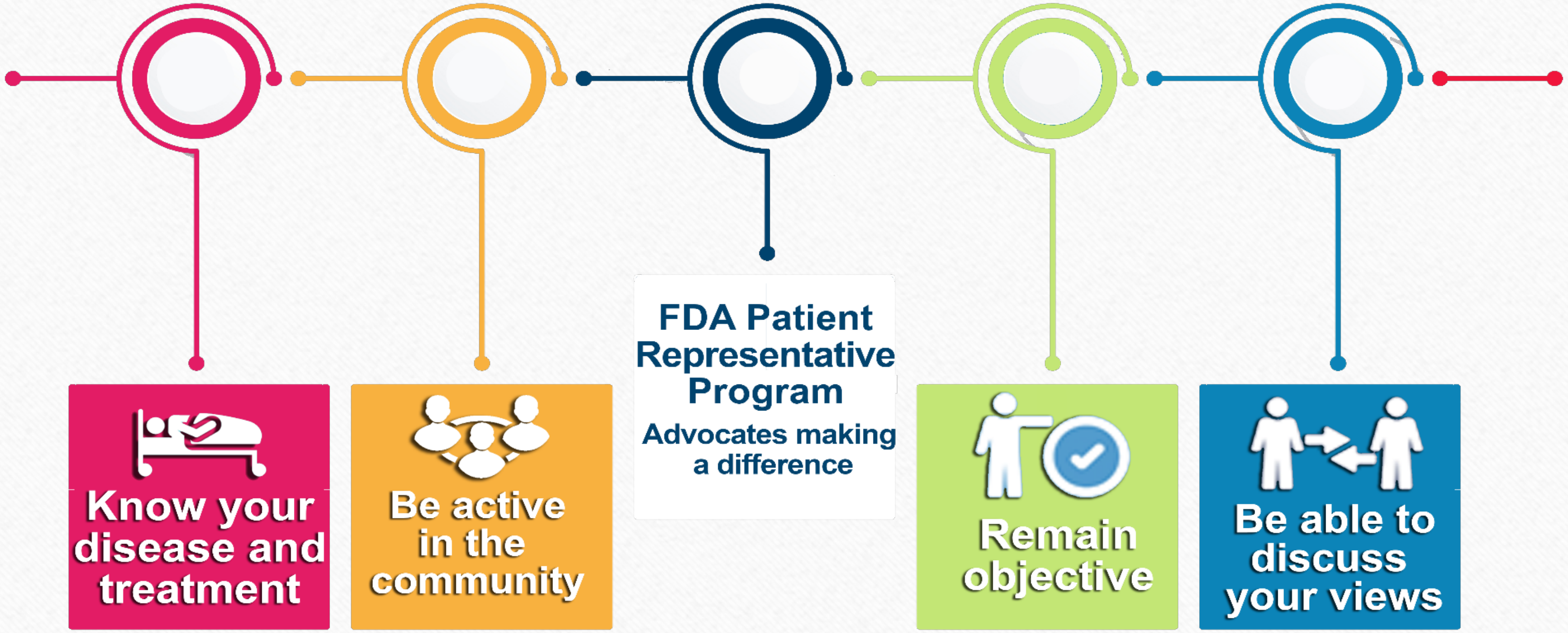
#TwitterMed

Disclaimer

I am a special government employee (SGE) for the FDA. I do not represent the FDA. Any and all information is based on my experience. All errors are mine.

Be curious.
Volunteer

FDA Patient Rep Program

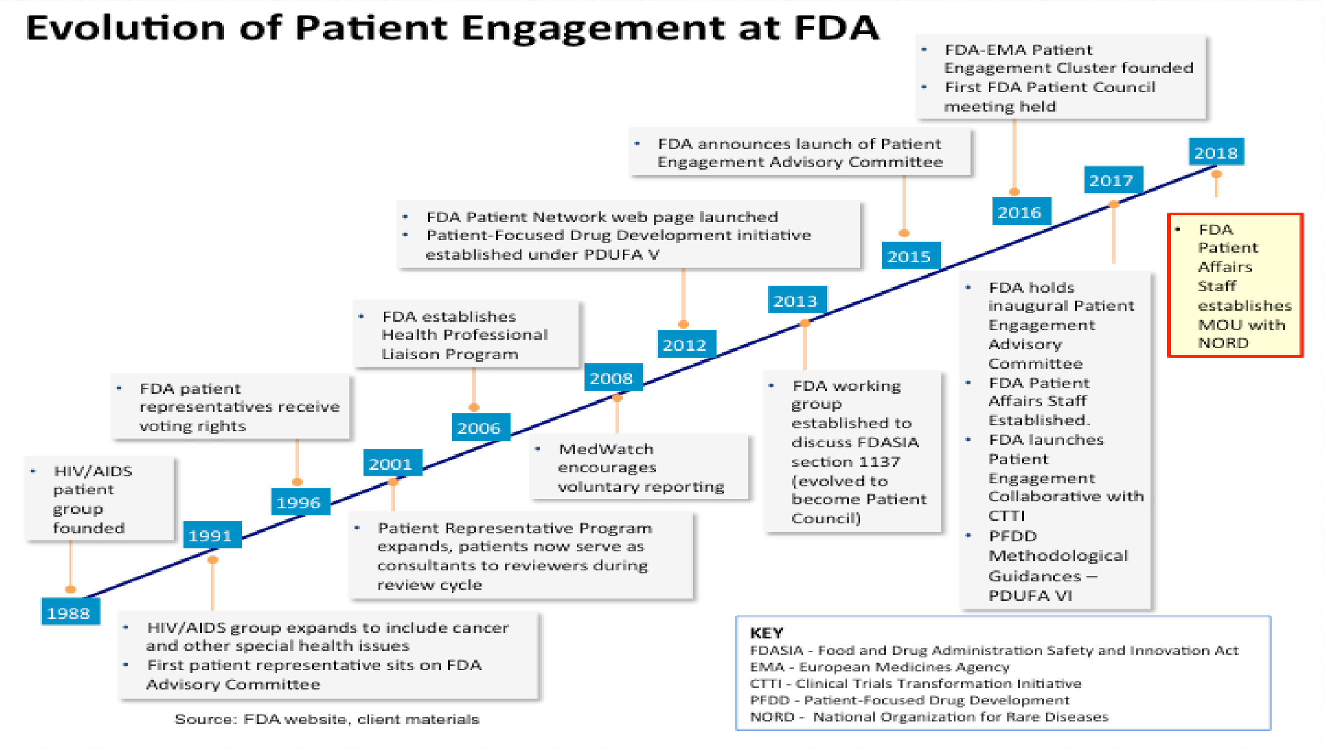


FDA Patient Representative Program

Advocates making a difference

- Know your disease and treatment
- Be active in the community
- Remain objective
- Be able to discuss your views

History of FDA Patient Representatives



Evolution of Patient Engagement at FDA

1988
- HIV/AIDS patient group founded

1991
- HIV/AIDS group expands to include cancer and other special health issues
- First patient representative sits on FDA Advisory Committee

1996
- FDA patient representatives receive voting rights

2001
- Patient Representative Program expands, patients now serve as consultants to reviewers during review cycle

2006
- FDA establishes Health Professional Liaison Program

2008
- MedWatch encourages voluntary reporting

2012
- FDA Patient Network web page launched
- Patient-Focused Drug Development initiative established under PDUFA V

2013
- FDA working group established to discuss FDASIA section 1137 (evolved to become Patient Council)

2015
- FDA announces launch of Patient Engagement Advisory Committee

2016
- FDA-EMA Patient Engagement Cluster founded
- First FDA Patient Council meeting held

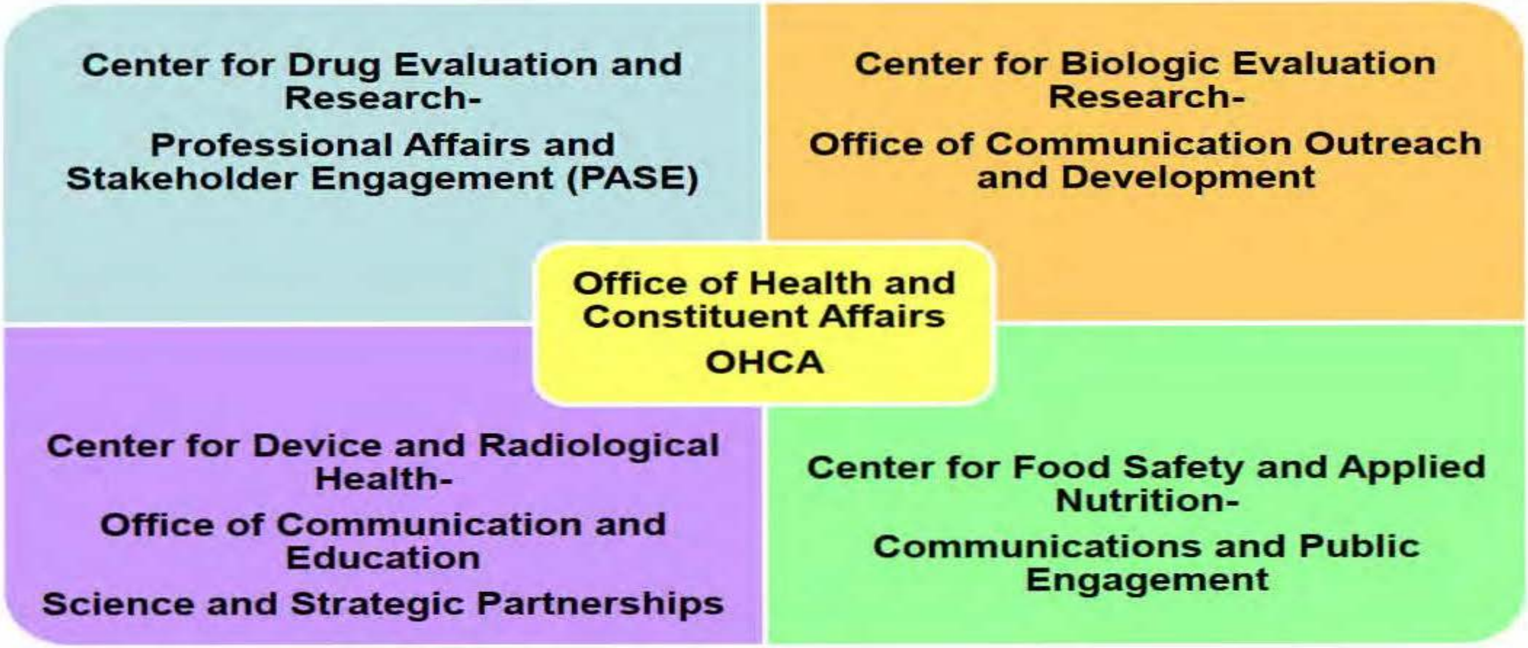
2017
- FDA holds inaugural Patient Engagement Advisory Committee
- FDA Patient Affairs Staff Established.
- FDA launches Patient Engagement Collaborative with CTTI
- PFDD Methodological Guidances – PDUFA VI

2018
- FDA Patient Affairs Staff establishes MOU with NORD

KEY
FDASIA - Food and Drug Administration Safety and Innovation Act
EMA - European Medicines Agency
CTTI - Clinical Trials Transformation Initiative
PFDD - Patient-Focused Drug Development
NORD - National Organization for Rare Diseases

Source: FDA website, client materials

FDA structure



Center for Drug Evaluation and Research-
Professional Affairs and Stakeholder Engagement (PASE)

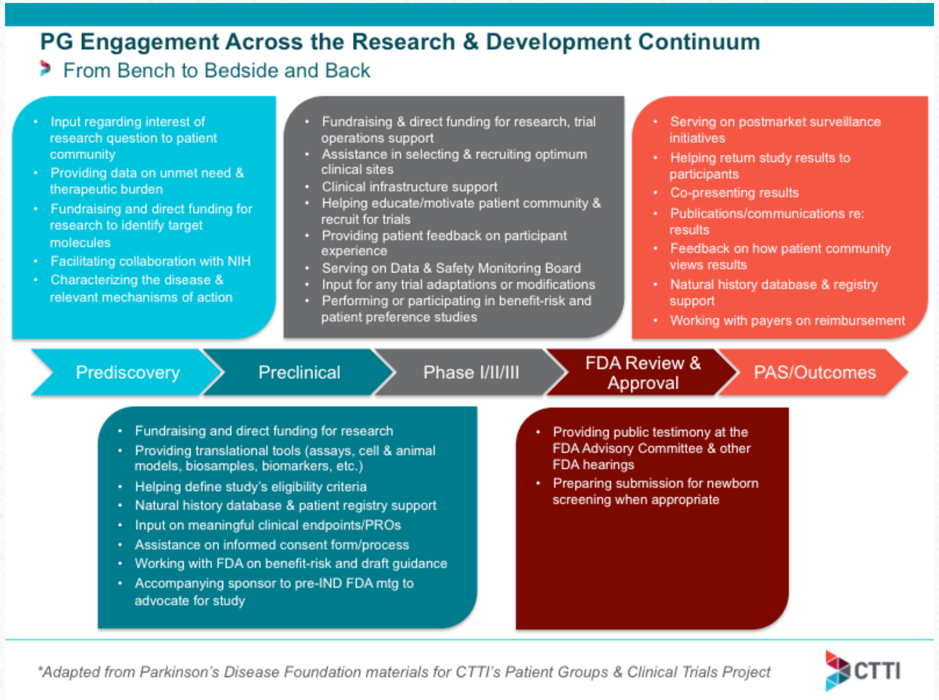
Center for Biologic Evaluation Research-
Office of Communication Outreach and Development

Office of Health and Constituent Affairs
OHCA

Center for Device and Radiological Health-
Office of Communication and Education
Science and Strategic Partnerships

Center for Food Safety and Applied Nutrition-
Communications and Public Engagement

Patient Engagement Across the R&D Continuum



PG Engagement Across the Research & Development Continuum

From Bench to Bedside and Back

- Input regarding interest of research question to patient community
- Providing data on unmet need & therapeutic burden
- Fundraising and direct funding for research to identify target molecules
- Facilitating collaboration with NIH
- Characterizing the disease & relevant mechanisms of action

- Fundraising & direct funding for research, trial operations support
- Assistance in selecting & recruiting optimum clinical sites
- Clinical infrastructure support
- Helping educate/motivate patient community & recruit for trials
- Providing patient feedback on participant experience
- Serving on Data & Safety Monitoring Board
- Input for any trial adaptations or modifications
- Performing or participating in benefit-risk and patient preference studies

- Serving on postmarket surveillance initiatives
- Helping return study results to participants
- Co-presenting results
- Publications/communications re: results
- Feedback on how patient community views results
- Natural history database & registry support
- Working with payers on reimbursement

Prediscovery → Preclinical → Phase I/II/III → FDA Review & Approval → PAS/Outcomes

- Fundraising and direct funding for research
- Providing translational tools (assays, cell & animal models, biosamples, biomarkers, etc.)
- Helping define study's eligibility criteria
- Natural history database & patient registry support
- Input on meaningful clinical endpoints/PROs
- Assistance on informed consent form/process
- Working with FDA on benefit-risk and draft guidance
- Accompanying sponsor to pre-IND FDA mtg to advocate for study

- Providing public testimony at the FDA Advisory Committee & other FDA hearings
- Preparing submission for newborn screening when appropriate

**Adapted from Parkinson's Disease Foundation materials for CTTI's Patient Groups & Clinical Trials Project*

CTTI

FDA Patient Network
Outgrowth of the Patient Program

Broadening opportunities for patient engagement

- Patient Representative
- Website
- Bi-weekly email newsletter
- Webinars
- In-person meetings and training
- Involvement in think tanks
- Submit written comments to the federal register
- Attend a public advisory meeting (DC area)
- Attend a policy meeting (DC area)
- MedWatch

U.S. Food and Drug Administration
Protecting and Promoting Your Health
For Patients
PATIENT NETWORK NEWS
from THE OFFICE OF HEALTH AND CONSTITUENT AFFAIRS
Get the Patient Network Newsletter
Navigate the For Patients Section

Why Does the FDA have Advisory Committees (AdComm)

The FDA, uses advisory committees to:

- obtain advice from experts who work outside of the government.
- work towards an open and transparent government.
- encourages patients, healthcare providers and other interested people to share their views during the open public hearing or by submitting comments to the open docket.

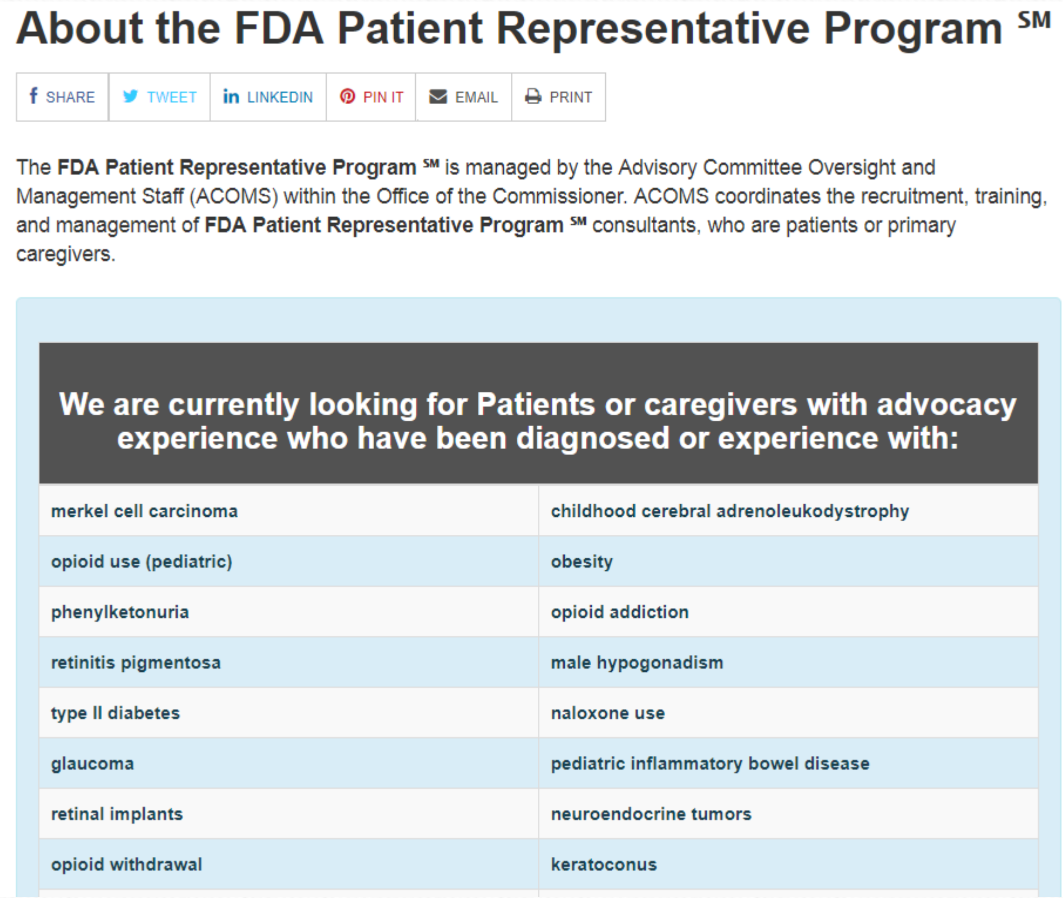
The primary role of the FDA advisory committee is to:

- provide independent expert advice to the Agency in its evaluation of these regulated products.
- help the agency move toward making sound decisions based upon reasonable application of sound scientific principles.

Patient Representatives in the FDA

- >200 Patient Reps
- Representing >125 diseases or conditions
- On >40 different advisory committees
- On average, there are 60 AdComms per year

How to Apply to be a Patient Representative

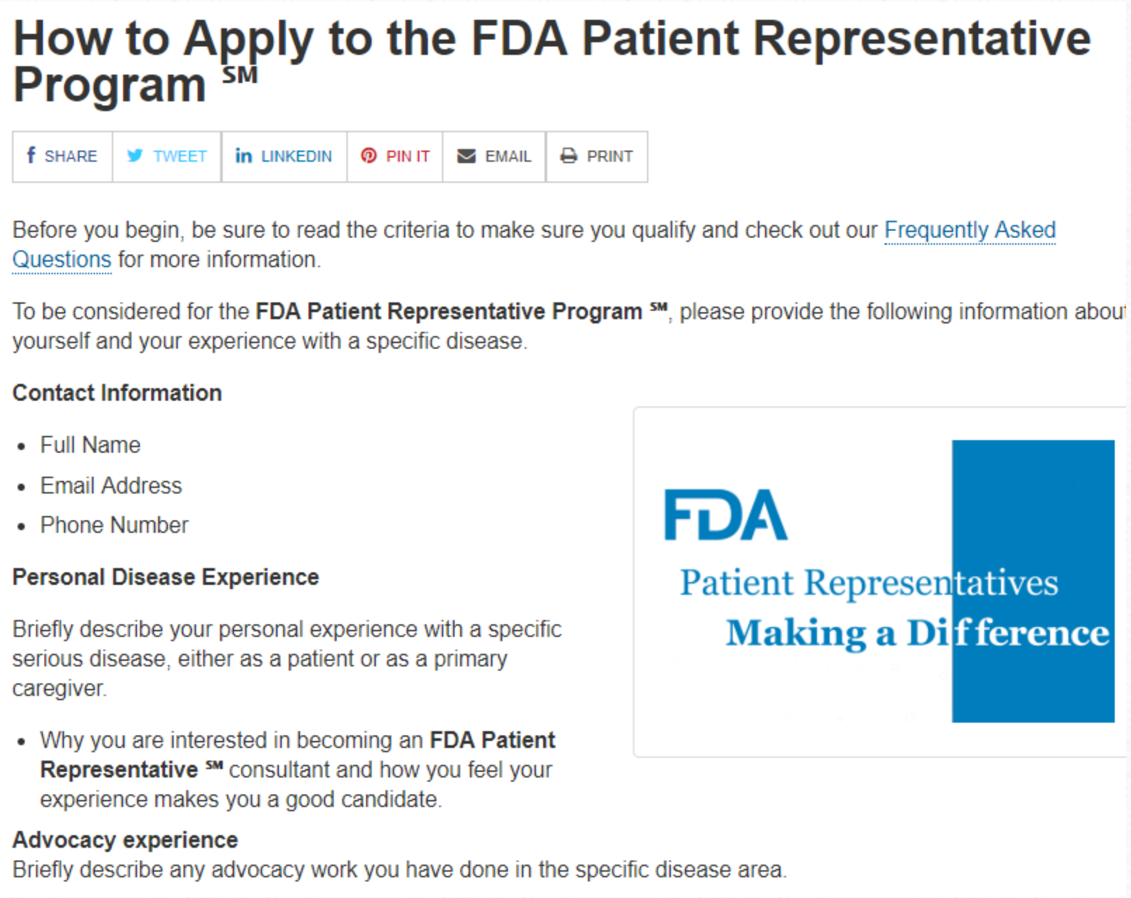


About the FDA Patient Representative Program SM

The **FDA Patient Representative Program** SM is managed by the Advisory Committee Oversight and Management Staff (ACOMS) within the Office of the Commissioner. ACOMS coordinates the recruitment, training, and management of **FDA Patient Representative Program** SM consultants, who are patients or primary caregivers.

We are currently looking for Patients or caregivers with advocacy experience who have been diagnosed or experience with:

merkel cell carcinoma	childhood cerebral adrenoleukodystrophy
opioid use (pediatric)	obesity
phenylketonuria	opioid addiction
retinitis pigmentosa	male hypogonadism
type II diabetes	naloxone use
glaucoma	pediatric inflammatory bowel disease
retinal implants	neuroendocrine tumors
opioid withdrawal	keratoconus



How to Apply to the FDA Patient Representative Program SM

Before you begin, be sure to read the criteria to make sure you qualify and check out our Frequently Asked Questions for more information.

To be considered for the **FDA Patient Representative Program** SM, please provide the following information about yourself and your experience with a specific disease.

Contact Information

- Full Name
- Email Address
- Phone Number

Personal Disease Experience

Briefly describe your personal experience with a specific serious disease, either as a patient or as a primary caregiver.

- Why you are interested in becoming an **FDA Patient Representative** SM consultant and how you feel your experience makes you a good candidate.

Advocacy experience

Briefly describe any advocacy work you have done in the specific disease area.

Being on an FDA AdComm

- You get training
- You get an assigned FDA officer
- Conflict of Interest (COI)
- Study materials from sponsor & the FDA – CONFIDENTIAL!
- Structured day
- Sequestered during the day
- You are representing not only yourself
- It doesn't happen without you

Be curious.

Volunteer

Handy dandy links

- About the Patient Rep program - https://www.fda.gov/ForPatients/PatientEngagement/ucm505721.htm
- Federal Register - https://www.fda.gov/RegulatoryInformation/Dockets/FR/default.htm
- Calendar of FDA Sponsored Public meetings - https://www.fda.gov/forpatients/default.htm
- Clinical Trials – What Patients Need to Know - https://www.fda.gov/ForPatients/ClinicalTrials/default.htm
- Listen to Webinars with FDA Experts - https://www.fda.gov/ForPatients/About/ucm410054.htm
- How Drugs or Devices are Developed or Approved - https://www.fda.gov/ForPatients/Approvals/Drugs/default.htm
- MedWatch - https://www.fda.gov/Safety/MedWatch/default.htm

Alliance for Aging Research
Research Advocacy Action Plan
June 2018

Name ____________________________________

We hope that you will you take the next step and seek out opportunities in your community to engage in research. Working with your colleagues at your table, please identify a goal, and what it will take to accomplish this goal. If this primary goal does not work out, do you have a second choice of activity in mind?

FIRST GOAL

1. **What do you want to accomplish?**

2. **Who has information or other resources that can help you accomplish this goal? This could be people you are already know, or do not know.**

3. **What is the first step to making this happen?**

4. **What barriers might you face? If this happens, how will you respond?**

5. **What can the Alliance for Aging Research do to support your efforts?**

SECOND GOAL

1. **What do you want to accomplish?**

2. **Who has information or other resources that can help you accomplish this goal? This could be people you are already know, or do not know.**

3. **What is the first step to making this happen?**

4. **What barriers might you face? If this happens, how will you respond?**

5. **What can the Alliance for Aging Research do to support your efforts?**

Senior Patient & Family Caregiver Network
Advocate Training
June 11-13, 2018

List of Attendees

Presenters & Staff

First Name	Last Name	Affiliation	Email	State
Sara	Collina, JD	Blueberry Hill Strategies	shcollina@gmail.com	DC
Penney	Cowan	American Chronic Pain Association	pcowan@theacpa.org	CA
Sarah	DiGiovine	Alliance for Aging Research	sdigiovine@agingresearch.org	DC
Jack	Guralnik, MD, PhD	University of Maryland School of Medicine	jguralnik@epi.umaryland.edu	MD
Lia	Hotchkiss	Patient-Centered Outcomes Research Institute	lhotchkiss@pcori.org	DC
Gail	Hunt	National Alliance for Caregiving	gailhunt@caregiving.org	DC
Alan	Jacobson, MD	Loma Linda VA Medical Center	akjacobson@linkline.com	CA
Mary	Lyons, MSN	Advanced Practice Registered Nurse/Clinical Nurse Specialist	lyons1528@gmail.com	IL
Beth	Mathews-Bradshaw	Leidos	beth.anne.mathews-bradshaw@leidos.com	DC
Debbe	McCall, MBA	Health eHeart Alliance Steering Committee	debbe.mccall@gmail.com	CA
George	Perry, PhD	The University of Texas at San Antonio	George.Perry@utsa.edu	TX
Susan	Peschin, MHS	Alliance for Aging Research	speschin@agingresearch.org	DC
Geri	Taylor	2017 SP&FCN Trainee	gera23dine@gmail.com	NY
James	Taylor	2017 SP&FCN Trainee	jmilburnta@aol.com	NY

Industry Representatives

First Name	Last Name	Affiliation	Email	State
Pamela	Anderson, MBA	Johnson & Johnson	pander22@its.jnj.com	NJ
Stephani	Shivers	LiveWell	sshivers@livewell.org	CT

Advocates

First Name	Last Name	Email	State
Charlotte	Cummings	cmc100east@gmail.com	IL
Holly	Hench	hollyhench@comcast.net	MA
Bob	Hoshaw		IL
Janet	Hoshaw	jhoshaw05@gmail.com	IL
Patricia	Hubert	patricia.hubert@shu.edu	NJ
Kevin	Kauffman	kevin@kauffman.net	PA
Jill	Lindsay	Jill.lindsay@va.gov	VA
Shon	Lowe	shongirls043@gmail.com	IL
Nicholas	Mentus	nicholasmentus@hotmail.com	FL
Terrie	Montgomery	terriemontgomery736@yahoo.com	IL
Tom	Norris	tomn482171@aol.com	CA
Nan	Rand	nanrand@gmail.com	ME
Ellen	Ryske	ejr1234567@gmail.com	IL
Joan	Sanborn	joanandsharky@gmail.com	NJ
Barbara	Santay	basantay@aol.com	IL
Damian	Santay	damiansantay@gmail.com	IL
Yvonne	Selden	yselden@theacpa.org	CA
Michele	Straube	mstraube@mindspring.com	UT
Jennifer	Ventre	jenven1018@gmail.com	NJ
Jeffrey	Wexler	jeffawex@gmail.com	NY

Alliance FOR AGING RESEARCH

Catalyzing Innovation
for Healthy Aging

Senior Patient & Family Caregiver Network

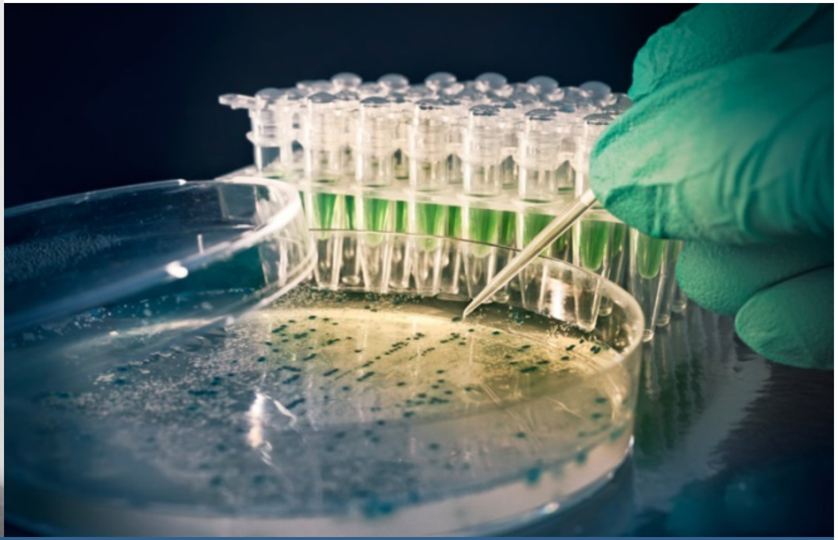
Webinar One

Patient-Centered Outcomes Research

Sue Peschin, President and CEO, Alliance for Aging Research
Sara Collina, Curriculum Developer, LEIDOS

Participants will:

- ✓ Understand what to expect at the upcoming Training
- ✓ Learn what *Research Advocacy* is and why it matters
- ✓ Explore the key elements of *Patient-Centered Outcomes Research*



WHO WE ARE

The Alliance for Aging Research is the leading non-profit organization dedicated to accelerating the pace of scientific discoveries and their application in order to vastly improve the universal human experience of aging and health.

WWW.AGINGRESEARCH.ORG

Catalyzing Innovation for Healthy Aging

Senior Patient and Family Caregiver Network

PATIENT-CENTERED OUTCOMES RESEARCH

2 years of funding received from PCORI to launch the first ever Senior Patient & Family Caregiver Network

The Challenges of AFib Research
Alan Jacobson, MD

Senior Patient and Family Caregiver Network

- Organized by the Alliance for Aging Research
- Funded by the Patient-Centered Outcomes Research Institute (PCORI)
  - Comparative Effectiveness Research
  - Patient and Family Caregiver Engagement Projects, such as the Senior Patient and Family Caregiver Network!
- Older adults with Alzheimer's disease, atrial fibrillation (AFib), chronic pain/disability, and/or sarcopenia; and, family caregivers
- Different levels of knowledge and experience—let's all learn from each other

Senior Patient and Family Caregiver Network

- Participate in two webinars in May 2018 prior to the workshop—this is the first one!
- Watch the 6-minute Alliance for Aging Research video, *Pay it Forward: Volunteering for a Clinical Trial,* between now and the second webinar—we will send everyone the link
- Webinar #2 on Understanding Clinical Trials: Wednesday, May 30, 2018 at 1:00 pm Eastern
- Complete the *Progress for Patients Video Training*. A link to the training will be sent to you via e-mail after the next webinar

Catalyzing Innovation for Healthy Aging

Senior Patient and Family Caregiver Network

- Participate in the in-person workshop from Monday evening June 11 to Wednesday mid-day on June 13 at the Marriott O'Hare in Chicago, IL
    - Dress is casual—wear what is comfortable
    - Bring a sweater in case it's cold in the room
- Participate in a post-workshop interview to refine the curriculum in July 2018
- Provide feedback on a revised curriculum (July-September 2018)
- Stay in touch with online network

How will you use this training?

- Volunteer opportunities to provide input into the medical research process from the patient/caregiver perspective at the national or local level

- Participants will receive:
    - Covered travel, lodging, and a stipend of $400 for full participation
    - A Certificate of Completion for participating in the training

Research Advocacy

- What is *research advocacy*?
- What is the purpose of *research advocacy training*?

Patient-Centered Outcomes Research

- What is medical (or health care) research?
- What is outcomes research?
- What makes outcomes research patient-centered?

Patient-Centered Outcomes Research

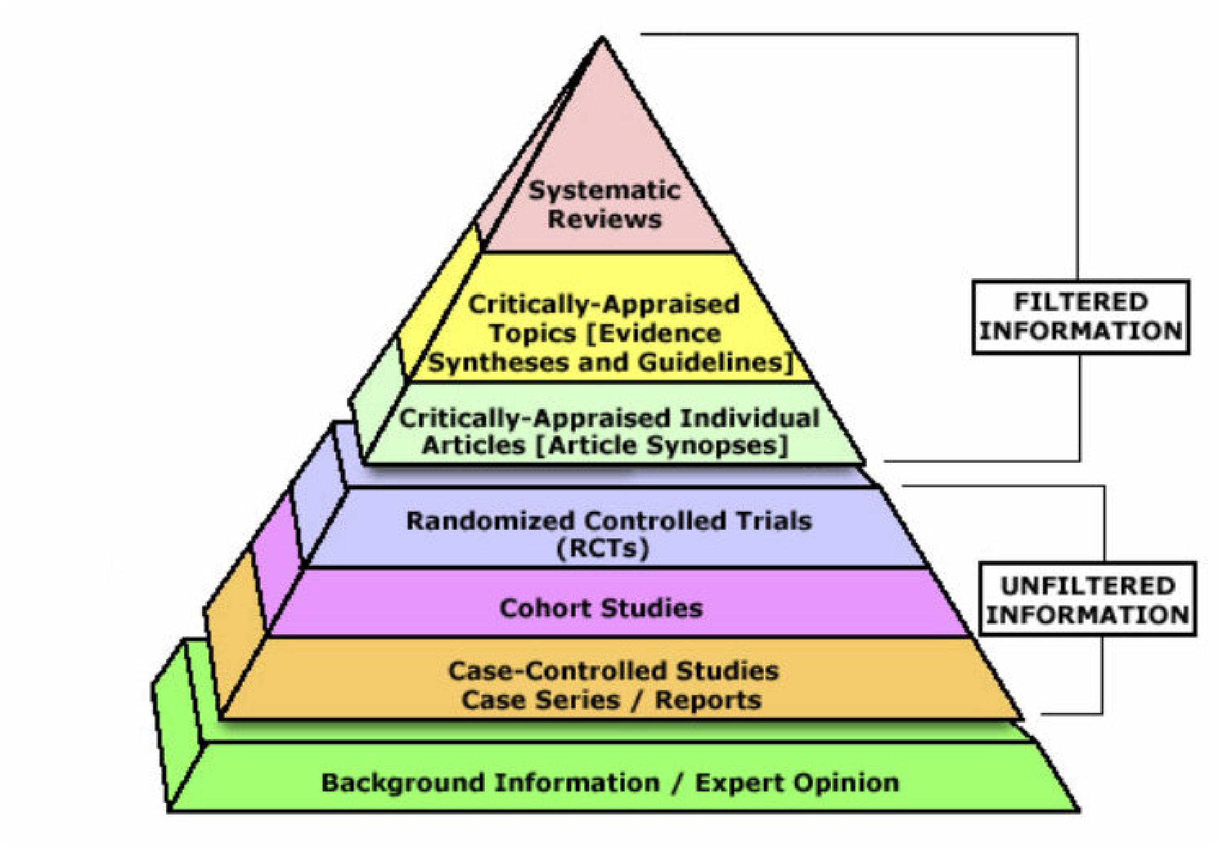
What is medical (or health care) research?

The key ingredient for…

Wait, aren't ALL health care decisions based on at least some evidence?

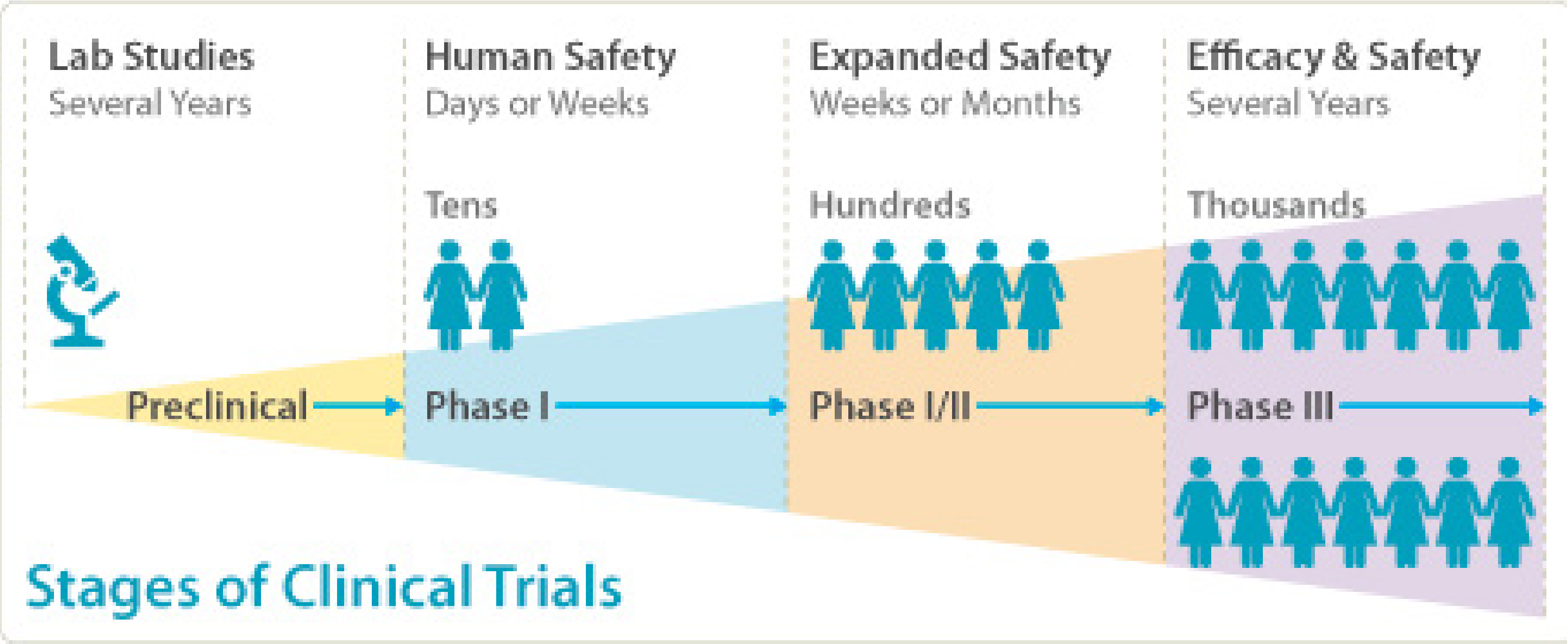
Evidenced-Based Health Care

Kinds of Research



Systematic Reviews
Critically-Appraised Topics [Evidence Syntheses and Guidelines]
Critically-Appraised Individual Articles [Article Synopses]
Randomized Controlled Trials (RCTs)
Cohort Studies
Case-Controlled Studies Case Series / Reports
Background Information / Expert Opinion
FILTERED INFORMATION
UNFILTERED INFORMATION

Clinical Trials


Lab Studies
Several Years
Human Safety
Days or Weeks
Expanded Safety
Weeks or Months
Efficacy & Safety
Several Years
Tens
Hundreds
Thousands
Preclinical
Phase I
Phase I/II
Phase III
Stages of Clinical Trials

Patient-Centered Outcomes Research

What is outcomes research? Focus is on end result.

Prospective studies: researchers follow participants into the future to record when and how they developed a particular outcome

Retrospective studies: researchers jump back in time to look at records of patients and follow their histories to determine when, why, and how they developed a particular outcome

Patient-Centered Outcomes Research

There are two types of data used to measure outcomes:

Data that is **quantitative**
Can be expressed as a number

Data that is **qualitative**
Cannot be expressed as a number

Who Funds Medical Research?

- Federal, State, and Local Governments
- Universities and Colleges
- Foundations
- Medical Research Organizations
- Diseased-Focused Organizations
- Industry (Pharmaceutical, Biotechnology, etc.)

Who Funds Medical Research?

Foundations
Medical Research Organizations
Diseased-Focused Organizations
Universities and Colleges
State and Local Governments

Federal Government

Industry

Patient-Centered Outcomes Research

Comparative Effectiveness Research (CER)

The direct comparison of two or more treatments
to determine what works best for which patients.

Patient-Centered Outcomes Research (PCOR)

A kind of comparative effectiveness research that
specifically answers patient-centered questions.

What is Patient-Centered Outcomes Research?

Given my personal characteristics, conditions, and preferences, what should I expect will happen to me?

What are my options, and what are the potential benefits and harms of those options?

How can clinicians and the care delivery systems they work in help me make the best decisions about my health and health care?

What is Patient Centered Outcomes Research?

Some Terms to Know

Efficacy Trials

Could it work in ideal settings?

Effectiveness or Pragmatic Trials

Does it work in the real world?

What is Patient-Centered Outcomes Research?

A Few More Terms to Know

Patient **Reported** Outcomes

Any report about a patient's health condition that comes directly from the patient, without interpretation by a clinician or anyone else.

Patient **Engagement**

Including patients in the research process itself.

Patient-**Centered** Outcomes

A health result (event or nonevent) that actually matters to patients.

Patient Centered Outcomes Research Institute (PCORI)

What makes PCOR different?

Patients can participate in PLANNING the research.

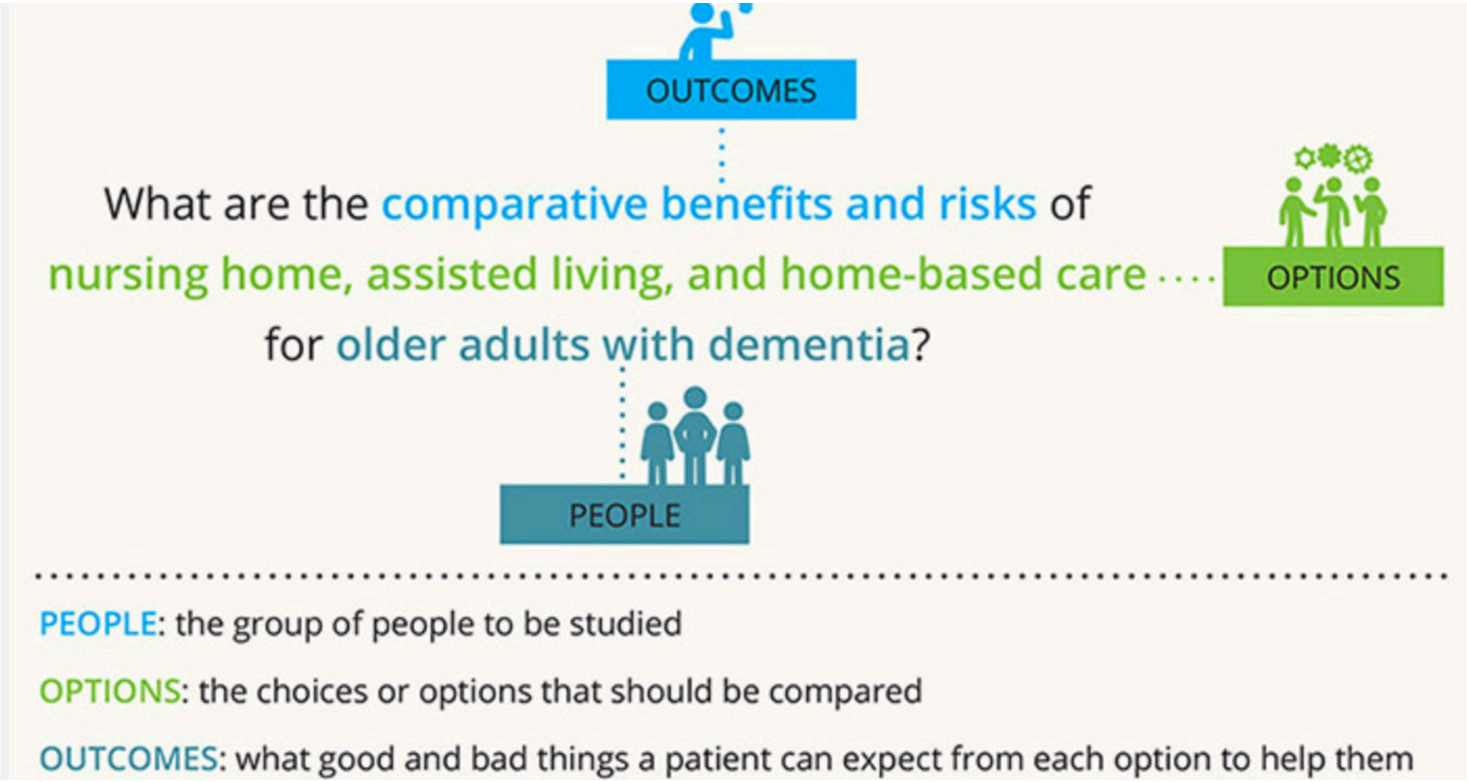
Patients can participate in CONDUCTING the research.

Patients can participate in DISSEMINATING the research.

Translating our concerns into research questions

THE PEOPLE	Who are the people that should be studied? This is the population of interest.
THE OPTIONS **INTERVENTION and COMPARATOR**	What options should be compared? These are the decisions the research is intended to inform.
THE OUTCOMES	How can people make informed choices between options? These are the factors that people will consider when making a decision between/among options.

Translating our concerns into research questions



OUTCOMES

What are the comparative benefits and risks of nursing home, assisted living, and home-based care for older adults with dementia?

OPTIONS

PEOPLE

PEOPLE: the group of people to be studied

OPTIONS: the choices or options that should be compared

OUTCOMES: what good and bad things a patient can expect from each option to help them

Did we succeed?

- ✓ Understand what to expect at the upcoming Training
- ✓ Learn what *Research Advocacy* is and why it matters
- ✓ Explore the key elements of *Patient-Centered Outcomes Research*

If you have questions after the webinar, please email Sue Peschin at speschin@agingresearch.org, or call me at 202-688-1246

Thank you!

Alliance
FOR AGING RESEARCH

Catalyzing Innovation
for Healthy Aging

Webinar Two

Clinical Trial Research

How do clinical trials work?

**Jack Guralnik, M.D., Ph.D.
and Alan Jacobson, M.D.**

Participants will be able to:

- ✓ Explain key elements of clinical trial design
- ✓ Extract key information from a scientific abstract

Randomized Controlled Trials

Jack M. Guralnik, M.D., Ph.D.
Department of Epidemiology and Public Health
University of Maryland School of Medicine

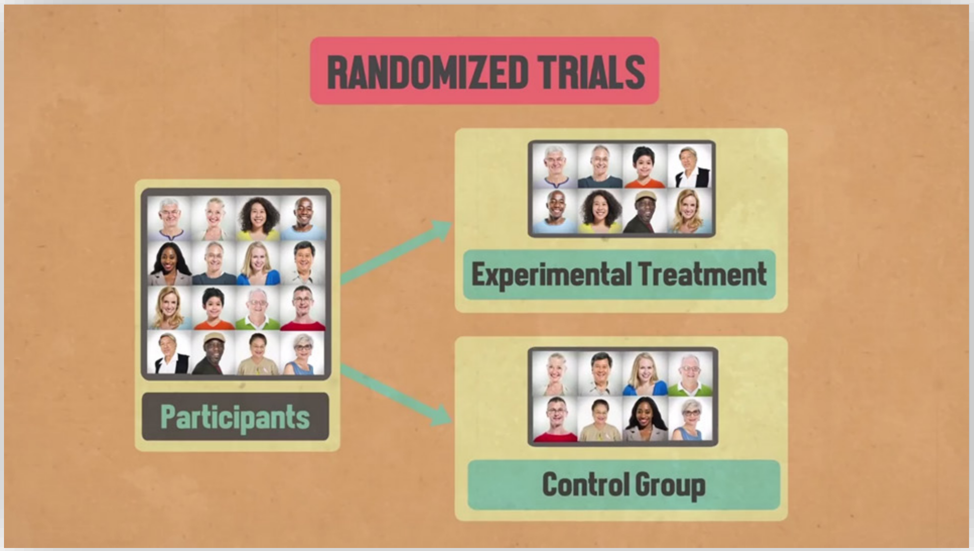
Background

- The randomized trial is considered the ideal design for evaluating both the effectiveness and the side effects of new forms of intervention[1]
- The randomized controlled trial is at present the unchallenged source of the highest standard of evidence used to guide clinical decision making[2]

1. Gordis L. Epidemiology. 4th edition. Philadelphia: Saunders (Elsevier); 2009.
2. Lavori PW, Kelsey J. Introduction and overview. Epidemiol Rev. 2002;24:1-3.

Randomized Controlled Trials

- Treated and untreated participants are followed over time to determine whether they experience the outcome
- Assignment to treatment or non-treatment is by randomization


RANDOMIZED TRIALS
Participants
Experimental Treatment
Control Group

Randomization

- Process by which all participants have equal probability of being assigned to the treated group or the untreated group
- Removes the potential for conscious or unconscious bias in the allocation of subjects to the treatment groups

Timing of RCTs

- Must have preliminary evidence of treatment's efficacy and safety
- Must know enough about treatment to know which outcomes to assess
- Before treatment becomes part of standard medical practice

Timing is Everything

Equipoise

- A state of genuine uncertainty about the benefits or harms that may result from different exposures or interventions. A state of equipoise is an indication for a randomized controlled trial, because there are no ethical concerns about one regimen being better for a particular patient.

Porta M (ed.). A dictionary of epidemiology. 5th ed. New York: Oxford University Press; 2008

Avoidance of Bias in RCTs

- Generation of truly random allocation sequence
- Concealment of allocation sequence
- Blinded outcome assessment

Blinding of Outcome Assessment

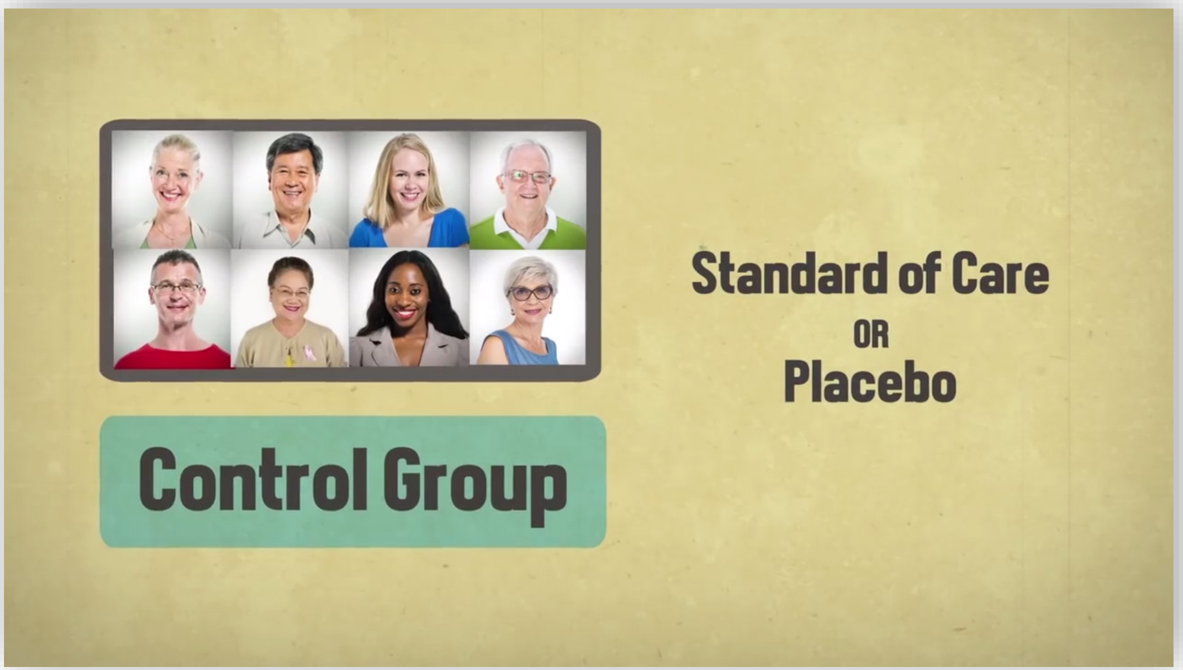
- Knowledge of participant’s group allocation could bias outcome assessment
- Blinding:
    - Participants
    - Research staff who are assessing the outcome
    - Health care professionals caring for the patient
    - Data analysts

Blinding of Outcome Assessment

- Blinding may not always be possible
  - Effectiveness of an exercise intervention in patients after myocardial infarction
- Side effects may affect ability to maintain blinding
  - Nausea, hair loss

Treatment of Controls

- No treatment
- Placebo
- Standard treatment


Control Group
Standard of Care
OR
Placebo

Placebo Effect

- Placebo:
  - A treatment that appears identical to the study treatment but that lacks the active component(s)
- Placebo effect:
  - Apparently beneficial effect of a treatment resulting solely from administration of the treatment

Purpose of Placebo Group

- To maintain blinding
- To strengthen bond between participant and study
- To control for placebo effect

Intention-to-Treat Approach

- Study participants who do not adhere to treatment protocol or who switch groups are analyzed according to original group assignment
- Answers the question, “How does the treatment work in the people to whom it is targeted?”
- Simulates the “real world”

Generalizability

- Study population
  - Systematic differences between study and target populations (eligibility criteria)
  - Volunteerism
- Trial conditions
  - Difference between trial conditions and "real world" conditions

Ethical Issues

- Is it ethical to randomize people
  - to receive the experimental treatment?
  - to not receive the experimental treatment?
- Is the sample size too small?
- Is the sample size too big?
- Informed consent
- Interim analyses, stopping rules

Strengths of RCTs

- Study design with the greatest ability to provide valid results
- Randomization prevents bias that may occur when allocating participants to groups
- Randomization usually results in groups that are comparable to each other in regard to known and unknown confounding variables

Limitations of RCTs

- Only useful for studying potentially beneficial factors
- Potential participants may be reluctant to agree to randomization
- Generalizability
- Timing/equipoise
- Expense

Alliance
FOR AGING RESEARCH

Five Concepts that Really Matter in Clinical Trial Design with Examples from the EAFT Trial

Alan Jacobson, M.D.
Loma Linda Veterans
Administration Medical Center

European Atrial Fibrillation Trial

Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke, EAFT (European Atrial Fibrillation Trial) Study Group, *The Lancet*, Volume 342, Issue 8882, 1993, Pages 1255-1262

European Atrial Fibrillation Trial - Abstract

Several studies have established the value of anticoagulation for primary prevention of thromboembolic events in patients with non-rheumatic atrial fibrillation (NRAF). However, in patients with a recent transient ischaemic attack (TIA) or minor ischaemic stroke the preventive benefit of anticoagulation or aspirin remains unclear. Physicians in 108 centres from 13 countries collaborated to study this question.

European Atrial Fibrillation Trial - Abstract

1007 NRAF patients with a recent TIA or minor ischaemic stroke were randomised to open anticoagulation or double-blind treatment with either 300 mg aspirin per day or placebo (group 1, 669). Patients with contraindications to anticoagulation were randomised to receive aspirin or placebo (group 2, 338). The measure of outcome was death from vascular disease, any stroke, myocardial infarction, or systemic embolism.

European Atrial Fibrillation Trial - Abstract

During mean follow-up of 2·3 years, the annual rate of outcome events was 8% in patients assigned to anticoagulants vs 17% in placebo-treated patients in group 1 (hazard ratio [HR] 0·53; 95% confidence interval [CI] 0·36-0·79). The risk of stroke alone was reduced from 12% to 4% per year (HR 0·34; 95% CI 0·20-0·57). Among all patients assigned to aspirin (groups 1 and 2), the annual incidence of outcome events was 15%, against 19% in those on placebo (HR 0·83; 95% CI 0·65-1·05). Anticoagulation was significantly more effective than aspirin (HR 0·60; 95% CI 0·41-0·87). The incidence of major bleeding events was low, both on anticoagulation (2·8% per year) and on aspirin (0·9% per year). No intracranial bleeds were identified in patients assigned to anticoagulation.

European Atrial Fibrillation Trial - Abstract

We conclude that anticoagulation is effective in reducing the risk of recurrent vascular events in NRAF patients with a recent TIA or minor ischaemic stroke. In absolute terms: 90 vascular events (mainly strokes) are prevented if 1000 patients are treated with anticoagulation for one year. Aspirin is a safe, though less effective, alternative when anticoagulation is contraindicated; it prevents 40 vascular events each year for every 1000 treated patients.

1. Bias/Randomization

Bias/Randomization

- Observational (Real World) vs. Experimental / Interventional
- Superiority vs. Noninferiority vs. Equivalence
- Prospective vs. Retrospective
- Randomization
- Blinding

2. Protocol/Reproducibility

Protocol/Reproducibility

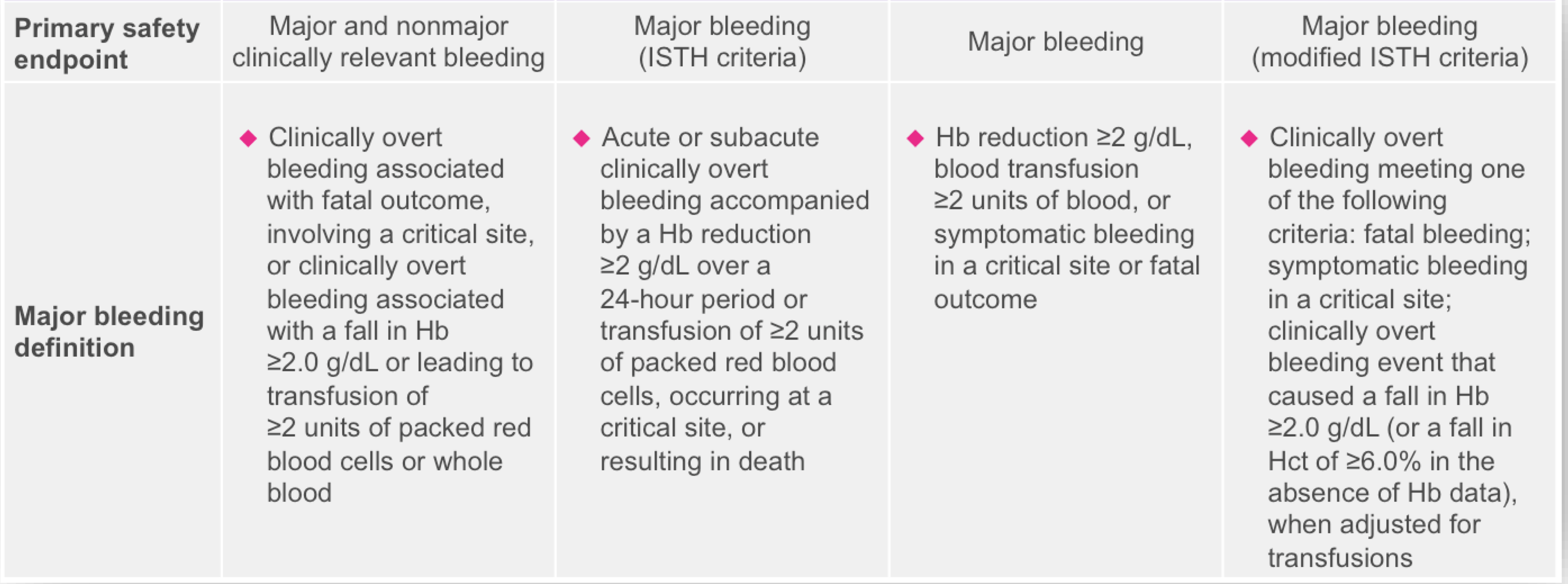
Alliance FOR AGING RESEARCH

- Eligible population
  - Selection Criteria – change selection criteria, will get a different answer to the same question
  - Eligible pool – is the recruitment population a subset of all affected patients
    - Patients in a particular country
    - Patients in a particular health care system
- Baseline Characteristics
  - Age
  - Gender
  - Disease States
- Confounders
  - Differences in laboratory testing in different countries (Anticoagulation, Troponin)

AGE
BACKGROUND
RACE
YOU
GENDER

3. Endpoints/Outcomes

Endpoints/Outcomes



Primary safety endpoint	Major and nonmajor clinically relevant bleeding	Major bleeding (ISTH criteria)	Major bleeding	Major bleeding (modified ISTH criteria)
Major bleeding definition	◆ Clinically overt bleeding associated with fatal outcome, involving a critical site, or clinically overt bleeding associated with a fall in Hb ≥2.0 g/dL or leading to transfusion of ≥2 units of packed red blood cells or whole blood	◆ Acute or subacute clinically overt bleeding accompanied by a Hb reduction ≥2 g/dL over a 24-hour period or transfusion of ≥2 units of packed red blood cells, occurring at a critical site, or resulting in death	◆ Hb reduction ≥2 g/dL, blood transfusion ≥2 units of blood, or symptomatic bleeding in a critical site or fatal outcome	◆ Clinically overt bleeding meeting one of the following criteria: fatal bleeding; symptomatic bleeding in a critical site; clinically overt bleeding event that caused a fall in Hb ≥2.0 g/dL (or a fall in Hct of ≥6.0% in the absence of Hb data), when adjusted for transfusions

Outcomes

- Definitions
- Relative Risk reduction (RRR)
- Absolute Risk reduction (ARR)
- Number needed to treat
- Scientists are just as capable of “spin” as politicians:
  - In some cases, will use RRR to overemphasize benefit
    - 25% reduction in events (from 12% to 9%)
  - And then use ARR to minimize risk
    - Only a 1% increase in bleeding (from 1% to 2%)

4. Selection Criteria (Inclusion/Exclusion)

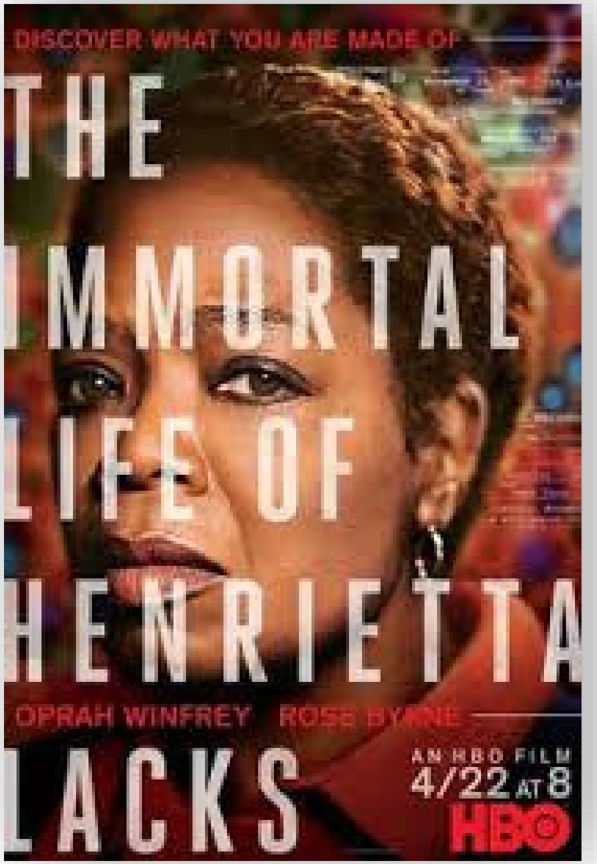
Selection Criteria

- How Specific?
    - Differential Diagnosis for Stroke vs. TIA
    - High blood pressure – treated, untreated, for how long
    - Smoking – total pack years exposure, how long since quit
    - Congestive heart failure – multiple subcategories and severity
    - Diabetes – great variations in level of control

5. Informed Consent

Informed Consent

- Tuskegee Airmen
- Declaration of Helsinki
- Common Rule
- HIPPA
- Third World vaccine trials


DISCOVER WHAT YOU ARE MADE OF
THE IMMORTAL LIFE OF HENRIETTA LACKS
OPRAH WINFREY ROSE BYRNE
AN HBO FILM
4/22 AT 8
HBO

SENIOR PATIENT AND FAMILY CAREGIVER NETWORK
ADVOCATE TRAINING
JUNE 11-13, 2018
Glossary

Absolute Risk Reduction (ARR): Absolute risk of a disease is the risk of developing the disease over a time period. Absolute Risk Reduction (ARR) is the change in the risk of an outcome in relation to a comparison treatment or activity.

Anticoagulant: Medicines that help prevent blood clots

Accuracy: The closeness of agreement between a data value and the true value.

Association: A connection or relationship between things.

Adverse Event (AE): An undesirable experience associated with the use of a medical product in a patient. An event is considered a Serious Adverse Event (SAE) when the patient outcome is death; life-threatening hospitalization; disability; congenital anomaly/birth defect; or rapid intervention is required to prevent permanent impairment.

Bias: A systematic error in sampling or testing that encourages one outcome or answer over others.

Biologic: A therapeutic agent derived from living things.

Biologics License Application (BLA): A form submitted to the Food and Drug Administration (FDA) after a Phase III trial that requests permission to label and market a biological product.

Blinding: The process of keeping secret the assignment of participants to study groups from researchers, participants, or both. This is done to minimize bias.

Causation: When changes in one variable directly cause changes in the other. In the clinical trial context, cause and effect can only be effectively studied through randomization. An association between two items does not necessarily mean that one caused the other.

Clinical Trial Phases: Clinical trials are conducted in a series of steps, called phases, to answer a separate research question.

- **Phase I:** Researchers test a new drug or treatment in a small group of people for the first time to evaluate its safety, determine a safe dosage range, and identify side effects.
- **Phase II:** The drug or treatment is given to a larger group of people to see if it is effective and to further evaluate its safety.
- **Phase III:** The drug or treatment is given to large groups of people to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow the drug or treatment to be used safely.
- **Phase IV:** Studies are done after the drug or treatment has been marketed to gather information on the drug's effect in various populations and any side effects associated with long-term use.

Cohort: A group of individuals who share a common exposure, experience, or characteristic. For example, a study may choose to follow a group, or cohort, of individuals who were exposed to contaminated water.

Collection Methods: The process of gathering and measuring information on variables of interest in an established, systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes.

Community-based participatory research (CBPR): A research approach that engages community partners in each stage of the process. CBPR differs from patient-centered outcomes research (PCOR) in that it is always steeped in community engagement, nurtures partnerships to realize shared outcomes over the long term, and often occurs outside of the clinical setting. PCOR can use a CBPR approach.

Comparative effectiveness research: Research focusing on building and evaluating evidence that assesses the benefits and risks of two or more methods that are designed to address the prevention, diagnosis, treatment, or monitoring of a clinical condition, or to improve health care delivery.

Control group: A group in an experimental study that serves as a comparison group. The experimental treatment, procedure, or program is not given to those in the control group; instead, this group receives either the usual available care, or an alternative such as a placebo.

Demographics: Personal information collected about an individual such as name, country of origin, birth date, race/ethnicity, occupation, education level, and income level.

Descriptive Research: A study in which information is collected without changing the environment (that is, nothing is manipulated).

Drug: A substance recognized by an official pharmacopeia or formulary intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease.

Efficacy: The performance of an intervention under ideal and controlled circumstances.

Effectiveness: The performance of an intervention under "real-world" conditions.

Endpoint: A direct measure of something substantial such as improved survival, improvement in systems or functional capacity, or decrease in the chance of developing a disease complication.

Equipoise: Genuine uncertainty as to the balance of benefits and harms that may result from two or more interventions; this genuine uncertainty makes randomization in clinical trials ethical.

Equivalence Trails: Aim to show the new drug/treatment is no better and no worse than a standard treatment.

Exclusion Criteria: Factors that are used to exclude people from participating in a clinical trial.

Experimental Research: A research design that uses manipulation and controlled testing to understand causality.

Generalizable: Extending research results or patterns found in a sample population to the wider population (which the sample represents).

Hypothesis: A prediction or explanation about future data based on previously collected data.

Inclusion Criteria: Factors that allow someone to participate in a clinical trial.

Informed Consent: The continuous process of ensuring that participation in research is voluntary. The process includes informing participants about the purpose of the research and the risks involved in participating.

Institutional Review Board (IRB): An independent group that reviews, approves, and monitors research plans and conduct to ensure that the safety and interests of research participants are protected.

Intention to Treat (ITT): A comparison of the treatment groups that includes all patients as originally allocated after randomization. ITT ignores noncompliance, protocol deviations, withdrawal, and anything that happens after randomization. This is the recommended method in superiority trials to avoid bias.

Intervention: A treatment or action taken to prevent or treat disease, or improve health in other ways.

Investigational New Drug Application (IND): A form submitted to the Food and Drug Administration (FDA) requesting permission to study a drug in humans for the first time. In limited circumstances, an IND Exemption can be requested.

Investigator's Brochure: A summary of the clinical and nonclinical data of an investigational product (IP).

Ischemic: Describes restriction in blood supply to tissues

Mean/Medium: The **mean** is the "average," the sum of all the numbers divided by the number of numbers. The **median** is the "middle" value in the list of numbers.

Meta-Analysis: A scientific, statistical method for combining data from several studies to gain more precise evidence of a treatment's effects.

Myocardial infarction: Heart attack

New Drug Application (NDA): A form submitted to the Food and Drug Administration (FDA) after a Phase III trial that requests permission to label and market a drug.

Noninferiority Trials: Aim to show that a new drug/treatment is no worse than standard treatment.

Number Needed to Treat: The average number of patients who need to be treated to prevent one additional bad outcome.

Observational Research: Studies that observe and measure variables of interest without assigning treatments to the subjects. Data can be collected prospectively (defining the question first) or retrospectively (answering a question using historical data).

On-Treatment Analysis: Also called per-protocol analysis, this is a comparison of treatment groups that includes only patients who adhered perfectly to the clinical trial instructions (completed the treatment)

Patient engagement: The inclusion of patients in the research process, from topic selection through study design and conduct, to dissemination of findings.

Patient-reported outcomes (PRO): A health outcome directly reported by the patient who experiences it.

Placebo: An inactive drug that may be used in research.

Placebo Effect: A beneficial effect that cannot be attributed to the properties of the placebo itself, and must therefore be due to the patient's belief in that treatment. When an inactive drug or treatment worsens symptoms this is called a **Nocebo Effect.**

Principal Investigator (PI): The lead researcher responsible for all aspects of a research study.

Pragmatic Trials: A kind of research that take place in a real-world environment, as opposed to a research setting. Pragmatic trials tend to exclude fewer people, and minimize the burden on trial participants so that the patient experience of those enrolled in the study is similar to the experience of patients who are not enrolled in the study.

Protocol: A detailed plan developed by a research team that must be followed when carrying out the study.

Randomized Controlled Clinical Trial: A study design that randomly assigns participants to receive one of two (or more) approaches to treatment. Randomization helps to minimize bias.

Relative Risk Reduction (RRR): Relative risk compares the risk in two different groups of people. Relative Risk Reduction (RRR) is the *ratio* of the probability of an event occurring in an exposed group to the probability of the event occurring in a comparison (non-exposed) group.

Reliability: The degree to which the result of a measurement, calculation, or specification can be depended on to be accurate.

Reproducibility: The ability of another researcher or group to accurately reproduce the results of a research study, using either the same or very similar data.

Risk-Benefit Analysis: A comparison of the risks and inconveniences on individuals with the anticipated benefit(s) of the study. The anticipated benefits of a trial must outweigh the potential risks.

Sample Size: The number of people who are enrolled in a study, often expressed as “n.” n=250 means 250 people were enrolled.

Standards of Care: A process that a clinician should follow to diagnose and treat a certain type of patient, illness, or clinical circumstance.

Superiority Trials: Aim to show that one treatment/drug is superior to another than a standard treatment.

Temporal Association: Two or more events that occur around the same time but may be unrelated, chance occurrences.

Variables: An attribute or property of a person, event, or object that is known to vary in a given study.