

Mind-Body Medicine and Immune System Outcomes: A Systematic Review

Helané Wahbeh^{1,2,*}, Ashley Haywood², Karen Kaufman², Noelle Harling² and Heather Zwickey²

¹Oregon Health & Science University, Portland, Oregon and ²Helfgott Research Institute, National College of Natural Medicine, Portland

Abstract: *Objectives:* The objectives of this systematic review were to: 1) characterize mind-body medicine studies that assessed immune outcomes, 2) evaluate the quality of mind-body medicine studies measuring immune system effects, and 3) systematically evaluate the evidence for mind-body interventions effect on immune system outcomes.

Data Sources: Data sources included MEDLINE, PsycINFO, CINAHL, Alt HealthWatch, Allied and Complementary Medicine Database, Cochrane Library, Database of Abstracts of Reviews of Effects, Health Technology Assessment Database.

Review Methods: Selection criteria included publications in any language, all study designs and participant types, mind-body interventions, and any immune outcome. Quantitative and qualitative data were extracted and quality assessments were made independently by 2 reviewers. Meta-analysis was not possible due to study heterogeneity. The Natural Standard evidence-based validated grading rationale™ was employed to evaluate the evidence.

Results: 111 studies with 4,777 subjects were reviewed, 81 of which were RCT's. The three largest intervention type categories were Relaxation Training (n=25), Cognitive Based Stress Management (n=22), and Hypnosis (n=21). Half the studies were conducted with healthy subjects (n=51). HIV (n=18), cancer (n=13) and allergies (n=7) were the most prominent conditions examined in the studies comprising of non-healthy subjects. Natural killer cell and CD4 T lymphocyte measures were the most commonly studied outcomes.

Conclusions: Most categories had limited or inconclusive evidence. Relaxation training had the strongest scientific evidence for affecting immune outcomes. Immunoglobulin A had the strongest scientific evidence for positive effects from mind-body medicine. Issues for mind-body medicine studies with immune outcomes are discussed and recommendations are made to help improve future clinical trials.

INTRODUCTION

An increasing number of people in the United States are using complementary and alternative medicine with mind-body medicine being the most commonly used form [1]. Mind-body medicine focuses on the relationships between the brain, mind, body, and behavior, and their effect on health and disease. According to the National Center for Complementary and Alternative Medicine, it encompasses a large group of therapies such as hypnosis, meditation, yoga, biofeedback, tai chi, and visual imagery [2]. Positive benefits of mind-body medicine are observed in numerous conditions [3] including headaches [4], coronary artery disease [5], chronic pain [6], mood, quality-of-life, and coping improvement. These therapies have also been shown to ameliorate disease and treatment-related symptoms, such as chemotherapy-induced nausea, vomiting, and pain in patients with cancer [7]. Mind-body modalities are commonly incorporated into treatment plans due to the low physical and emotional risk, the relatively low cost, and their ability to enable patients to take a more active role in their treatment.

The key premise of mind-body medicine is that a person's mental state influences their physical health. The exact

mechanisms underlying the health-promoting effects are unknown. However, evidence exists supporting the brain and central nervous system's influence on immune function and thus potentially on immune outcomes [8]. The study of these interactions, psychoneuroimmunology, has been a growing field since its introduction by Robert Ader and Nicholas Cohen in 1975 [9]. Despite increased research studies and reviews, there have been limited studies examining mind-body medicine and immune outcomes and thus, this topic is the focus of this review.

Psychoneuroimmunology researchers are posed with the challenging problem of selecting appropriate immune outcome for their studies, with a multitude of available immune outcomes to select for any given study. Assessing all aspects of the immune system in a single study is usually not feasible [10]. Thus, most investigators attempt to measure multiple immune measures relevant to the research question. It is unclear as of yet, which if any immune outcomes are most sensitive to mind-body medicine effects in general, or if they are only study-specific. It may also be that mind-body medicine's influence on immune outcomes is dependent upon the health and/or disease states of the participants. Physicians often observe immune system improvements in health on a clinical level. However, being able to demonstrate these improvements through rigorous research methods is challenging.

*Address correspondence to this author at the Oregon Health & Science University, 3181 SW Sam Jackson Park Rd. CR120, Portland, 97239, Oregon; Tel: 503-494-3528; Fax: 503-494-9520; E-mail: wahbeh@ohsu.edu

The research questions for the systematic review were, “What is the state of research literature in mind-body medicine and immune measures?” and “What is the evidence for mind-body therapies affecting immune outcomes?” Additionally, the authors hoped to gain insight into the most effective mind-body medicine and most sensitive immune outcomes for future trials. It was anticipated however, that generalizations about these immune outcomes may not actually be able to be made due to heterogeneity of the field. The study objectives were to: 1) characterize mind-body medicine studies that assessed immune outcomes, 2) evaluate the quality of these studies, and 3) systematically evaluate the evidence for mind-body interventions effect on immune system outcomes.

METHODS

Literature Search and Retrieval

Comprehensive searches were conducted by a research librarian using MEDLINE® (1950-10/25/2007), PsycINFO® (1967-10/25/2007), CINAHL® (1982-2/13/2007), Alt Health-Watch (1984-10/26/2007), AMED (ca. 1980 to 10/26//2007), Cochrane Library: CENTRAL (10/25/2007), Cochrane Database of Systematic Reviews (10/25/2007), Database of Abstracts of Reviews of Effects (10/25/2007), and Health Technology Assessment Database (10/30/2007). Search terms included immune system terms and mind-body medicine terms dependent on the search strategy required for each database (a comprehensive description of the search strategy is available from the first author). The search parameters for subject type and modalities were kept broad to maximize recall because it was anticipated that the inclusion criteria “immune outcomes” would be limiting. Reference lists of relevant studies were also reviewed for eligible papers.

Selection of Studies

Inclusion criteria included: 1) any published scientific literature regardless of peer review or paper type in any language; 2) all participant types; 3) modalities comprising mind-body interventions (hypnosis, imagery, meditation, mental healing, mind-body relations, all relaxation techniques, biofeedback, cognitive-behavioral therapies, group support, autogenic training, spirituality, and prayer, and emotional disclosure); 4) a sample size greater or equal to five; 5) study designs including randomized controlled trials (RCT's), non-randomized controlled trials (NRCT's), prospective and retrospective observational studies with controls, case-control studies, and uncontrolled pre-post studies; and 5) studies examining any measurable immune outcome. Because the study focused specifically on mind-body effects, interventions that incorporated body movement as the primary therapy (i.e. yoga, qi-gong, tai-chi) were excluded as exercise is known to influence immune function [11-13].

The primary author reviewed titles and abstracts according to study inclusion and exclusion criteria. A second reader reviewed and confirmed included and excluded studies. The full-text of studies meeting criteria, and those with insufficient information to determine eligibility from the abstract were retrieved.

Assessment of Methodological Quality

Numerous authors have noted difficulty in finding appropriate quality assessment tools [14-16]. After reviewing numerous instruments, a Quality Assessment Tool modeled after the “Aid to the Evaluation of Therapeutic Studies” developed by Reisch *et al.*, [17] and modified as recommended by Deeks [18] was used to rate the quality of the studies. This instrument was chosen because it included all relevant constructs (blinding, randomization, adequate reporting, etc.), provided a quantitative score, and adjusted for study design. The instrument adjusts for study design by removing questions from the total score for NRCT and uncontrolled trials about randomization, comparisons between groups, and blinding. The result is an adjusted score on a scale of 0-100, 100 being a higher quality study. The items for the assessment tool are listed in Table 4.

Two reviewers assessed the methodological quality of studies using the above instrument independently during data extraction with a third reviewer designated to resolve disagreements through consensus. Descriptive statistics of the quality scores were evaluated. Scores were analyzed between studies before and after the year 2001 due to the release of another meta-analysis, which made recommendations for improving study design and reporting for this field [19].

Additionally, an immunologist (HZ) reviewed the immune outcomes used for each study to determine appropriateness of outcomes. As the foundation of immunology knowledge has grown, certain outcomes have been found to be *in vitro* artifact (T suppressor cells, etc.), normal levels have been defined (cytokines), and many more receptors have been identified. Each outcome was assessed on a case by case basis for appropriate utilization depending on subject type, length of intervention, and assay method. For example, visualization had positive evidence for neutrophil adherence, but was rated as inappropriate by the immunologist (HZ). Neutrophils *in vivo* are attached to vessel walls with adhesion molecules. When measuring in serum, neutrophil concentrations are low and include activated neutrophils. *In vitro*, neutrophil adherence refers to the adhesion to the plastic labware, does not include adhesion molecules, and is irrelevant to health and immune function.

Data Collection

Uniformly trained research staff from the Helfgott Research Institute (Portland, Oregon) collected study data using a pre-tested data extraction form. A single reviewer extracted data and another independent reviewer verified the accuracy and completeness of the data extraction. Any discrepancies were resolved by consensus. The following data were collected: study design, number of treatment arms, setting, participant type, primary health condition (including acute or chronic), inclusion and exclusion criteria, number of subjects, mean age, study population (i.e., students, women), intervention type, length of each session, frequency of treatment, length of treatment period, total exposure time, home practice details, group or individual practice, type of outcome, sample type (i.e. blood, saliva), when outcomes were measured, means or mean differences when available, and *p*-values for each measure. If there were multiple time-points of measurement, values were taken from the time-points immediately before and after the intervention. For example,

if outcomes were measured after a 4 week intervention and also 6 weeks later, only the results at 4 weeks were included. The same data were also extracted on any and all control groups. All study data were managed with Microsoft Excel™ and an Access™ relational database (Microsoft Corporation, Redmond, Washington). Statistical tests were conducted in SPSS 16.0 (SPSS, Inc, Chicago, Illinois).

Study Classification

Twelve intervention study categories were developed. Brief descriptions of the categories are listed in Table 1. A full description of these therapies is beyond the scope of this paper and has already been discussed in the following review article [6]. Also, rather than collapse these categories into larger ones such as Relaxation or Psychotherapy-like, the categories were kept more specific to assess the finer distinctions between modalities. Many studies included more than one mind-body modality making the classification unclear. In these cases, studies were categorized according to the reported purpose of the study.

We considered reducing the selection criteria and/or focusing the review on just a specific outcome or disease state. However, the variability between studies was such that grouping them into smaller categories resulted in groups of one or two studies. This undermined our purpose of coming to broader generalizations about the outcomes. Thus, we

decided to continue with the broader selection and conducted a qualitative rather than quantitative review.

Evidence Grading

Due to study heterogeneity, a meta-analysis was not possible. The heterogeneity resulted from differences in subject type, intervention and implementation variation, and immune outcome type. The Natural Standard evidence-based validated grading rationale™ was used to provide a general understanding of the available evidence to guide future research, rather than attempt to definitively evaluate whether mind-body modalities affect immune outcomes. The Natural Standard evidence-based validated grading rationale™ is an objective grading criteria derived from validated instruments for evaluating study quality, including the 5-point scale developed by Jadad *et al.*, in which a score below 4 is considered to indicate lesser quality methodologically [20] [<http://www.nlm.nih.gov/medlineplus/druginfo/natural/grading.html>] (Table 2).

Resultant grades reflect the level of available scientific evidence in support of the effects of a given therapy for a specific indication. For this study, grades were assessed with only higher quality papers defined as a quality assessment score greater than 72, which was the median quality assessment score of all the studies. Gradings were then repeated with all studies, regardless of quality score, to determine if

Table 1. Description, Number of Studies, and Subjects in Intervention Categories

Category	Description	n	# of Studies
Relaxation	A therapy that promotes muscular and mental relaxation thought to increase parasympathetic activation and decrease sympathetic activation resulting in a slower heart rate, lower blood pressure, slower breath rate, and reducing muscle tension ²⁰⁻²² .	1070	25
CBSM	Training and education in cognitive restructuring, assertiveness skills behavior change strategies, and stress response with training in one or many of the following: progressive muscle relaxation, autogenic training, meditation, guided imagery and breathing exercises, stress management techniques ²³ .	1361	22
Hypnosis	Attention and focused concentration with a relative suspension of peripheral awareness ²⁴ documented by fMRI ^{25, 26} and EEG ²⁷ studies.	618	21
Visualization	Patient's imagination used to visualize a specific health outcome ^{28, 29} .	315	10
CBT	Psychotherapy based on the idea that thoughts cause emotions and behaviors and aims to change the way a person thinks in order to improve emotion and behaviors ³⁰ .	375	9
Disclosure	Verbal or written expression of emotional experience ³¹ .	303	7
Support	Group intervention that stated support was the major component ³² .	435	6
MBSR	Structured group program that employs instruction and practice in mindfulness meditation, education and discussions, and intensive home practice ³³ .	134	3
Biofeedback	Measures physiological markers like heart rate, breathing rate, electromyography, electroencephalography, or electrodermal activity and displays the results back to the patient to aid in self-modulation ³⁴ .	47	2
Humor	Induction of laughter in the patient through various mediums ³⁵ .	73	2
Meditation	Self-observation of mental activity, attentional focus training, and cultivating an attitude that highlights process rather than content ³⁶ .	20	2
Music	Listening to music or singing ³⁷ .	26	2

* CBSM-cognitive based stress management; CBT-cognitive based therapy; MBSR-mindfulness-based stress reduction.

Table 2. Natural Standard Evidence-Based Validated Grading Rationale™

Level of Evidence Grade	Criteria
A (Strong Scientific Evidence)	Statistically significant evidence of benefit from >2 properly randomized trials (RCTs), OR evidence from one properly conducted RCT AND one properly conducted meta-analysis, OR evidence from multiple RCTs with a clear majority of the properly conducted trials showing statistically significant evidence of benefit AND with supporting evidence in basic science, animal studies, or theory.
B (Good Scientific Evidence)	Statistically significant evidence of benefit from 1-2 properly randomized trials, OR evidence of benefit from ≥1 properly conducted meta-analysis OR evidence of benefit from >1 cohort/case-control/non-randomized trials AND with supporting evidence in basic science, animal studies, or theory.
C (Unclear or conflicting scientific evidence)	Evidence of benefit from ≥1 small RCT(s) without adequate size, power, statistical significance, or quality of design by objective criteria, OR conflicting evidence from multiple RCTs without a clear majority of the properly conducted trials showing evidence of benefit or ineffectiveness, OR evidence of benefit from ≥1 cohort/case-control/non-randomized trials AND without supporting evidence in basic science, animal studies, or theory, OR evidence of efficacy only from basic science, animal studies, or theory.
D (Fair Negative Scientific Evidence)	Statistically significant negative evidence (i.e., lack of evidence of benefit) from cohort/case-control/non-randomized trials, AND evidence in basic science, animal studies, or theory suggesting a lack of benefit.
E (Strong Negative Scientific Evidence)	Statistically significant negative evidence (i.e. lack of evidence of benefit) from ≥1 properly randomized adequately powered trial(s) of high-quality design by objective criteria.
Lack of Evidence	Unable to evaluate efficacy due to lack of adequate available human data.

quality influenced the results. For some immune outcomes, the desired direction of change of each outcome will vary depending on the disease being studied. For grading purposes, a statistically significant positive outcome was defined as a *p*-value of less than .05 and in the direction of change hypothesized by the investigators.

RESULTS

Study Characterization

A total of 914 studies were selected for review. Three hundred and forty were duplicates, 216 were excluded due to outcome measures, 80 due to design, 145 due to intervention, and 19 due to mind-body intervention being movement-based. In addition, one could not be located, and three could not be interpreted. One hundred and eleven studies with a total 4,777 combined subjects were reviewed to provide evidence regarding the state of research on mind-body medicine and immune outcomes (included studies are listed after references). The studies were published between 1964 and 2007, with 47 studies published after 2000. Seventy-three percent (81) were RCTs, 13% (14) were non-randomized controlled trials (NRCTs), 10% (11) were pre-post studies, and 4% (5) were cross-over controls. Seventy-seven percent incorporated some method of blinding (excluding pre-post studies). The three largest intervention type categories were Relaxation Training (23%), CBSM (20%), and Hypnosis (19%) (Table 2). Sixty-seven percent incorporated one modality and the remaining used two or more (e.g., relaxation with visualization). The average number of subjects in each study was 52±46 (range 5-303). Of the studies that reported mean ages (73%), the total mean age for participants across all intervention types was 36 ±14 years. Five studies were with children (under 18 yrs of age). Forty-five percent involved healthy subjects, 52% involved patients with chronic

disease, and 3% involved patients with acute disease. Among the subjects with chronic and acute disease, a diverse range of conditions was examined (Table 3).

Table 3. Number of Studies by Condition

Condition	Studies
Healthy	50
HIV	18
Cancer (breast, prostate, ovarian, malignant melanoma)	13
Allergy (urticaria, asthma, eczema)	7
Mental health (anorexia, depression, insomnia, panic disorder)	5
Upper respiratory infections	4
Herpes Simplex Virus	2
Multiple sclerosis, Rheumatoid Arthritis	2 each
Chronic Fatigue Syndrome, Facial Pain, Coronary Artery Disease, Tinnitus, Ulcerative Colitis, Decreased WBC, Hemopoietic Stem Cell Transplantation	1 each

Forty-eight percent of the studies used group interventions rather than individual therapy (40%) and 12% did not report how the intervention was administered. Reporting of the actual intervention varied. Seventy-five percent of the studies reported intervention session length, weekly frequency, and duration. Mean sessions were 78 minutes (range 9-360 minutes) with sessions ranging from 1 to 7 days per week, and interventions lasting 1 to 52 weeks long (mean 10.75±11.48). Fifty-one percent incorporated home practice

which occurred between intervention sessions, half of which reported home practice details (mean session time 26 ± 15 minutes).

Ninety percent of the studies were controlled with a varying number of arms (74 with two arms, 22 with three arms, 2 with four arms, and two with five arms). The different types of control groups included no treatment (37%), treatment with a different therapy (active control) (31%), waitlist (19%), standard of care (5%), and 8% did not report control type. A majority of the studies with three arms had a no treatment and an active control group, while the remaining studies had two different active control groups. Fifteen stud-

ies with an active control group did not report control session details in terms of frequency and duration. Of those that did, the time for the active control matched the time for the intervention.

Quality Assessment

The quality assessment items are listed in Table 4, along with the number of studies that included each criterion. The mean total quality scores representing different study designs were pre-post: 67 ± 11 (range 43-82), cross-sectional: 67 ± 10 (52-76), NRCT: 69 ± 13 (4-78), RCT: 73 ± 08 (46-96), and 70 ± 10 (4-96) for all studies. When analyzed using a one-way

Table 4. Quality Assessment Items and Number of Studies Addressing Each Criterion

Quality Assessment Item	# of Studies
1. The statement of purpose of the study was given.	111
2. The outcome variables for therapeutic effects defined prior to study.	108
3. Magnitude of difference in outcome between groups specified prior to study.	6
4. Data collection planned prior to T/M of subjects; data collected prospectively under specified conditions.	111
5. Subjects selected prior to T/M and evaluated prospectively.	111
6. Sample size was predetermined.	11
7. The total number of subjects was specified.	111
8. Adequate number of subjects were enrolled to detect magnitude of T/M differences under investigation or sufficient hazards identified to preclude further study.	12
9. Demographic criteria of subject population were adequately described including age, race, sex, disease/health status.	87
10. Inclusion and exclusion criteria were described.	76
11. Subjects selected for the study were suitable for questions posed by researchers.	108
12. Randomization claimed and documented.	81
13. Use of either prognostic stratification prior to study entry or retrospective stratification during data analyses. Did they match groups on important confounders? Were differences between groups found on baseline characteristics and if so did they adjust for this with statistical analysis?	61
14. Treatment groups were randomized.	81
15. Informed consent was obtained.	74
16. T/M protocol was adequately described.	89
17. T/M reasonable and appropriate to answer questions posed by researchers.	104
18. Blinding claimed and appeared realistic.	20
19. Description of all subjects or their records which were lost or dropped.	65
20. Laboratory and other outcomes appear standardized and consistent.	108
21. Treatment compliance assessed.	38
22. Evaluation methods (outcomes) adequately described.	108
23. Prospective evaluation of important hazards or toxicity (adverse events) described.	4
24. All comparisons involve same number of subjects or any discrepancy is explained.	94
25. Descriptive measures identified for all important variables.	85
26. Statistical tests used for comparisons involving important variables and are clearly identified, appropriately used and identified.	104

ANOVA, these scores were significantly different ($F(3,110) = 3.69, p < .02$). Bonferroni post-hoc analysis attributes these to differences between RCT's and NRCT's. Other study design category differences were not significant. The mean quality score for studies published after 2001 was 74 ± 7 , whereas before 2001 the figure was 69 ± 10 ($p < .02, t = 2.46$).

All assessed studies stated the purpose of the study, recruited subjects and collected data prospectively, and gave the total number of subjects. Ninety-seven percent of the studies defined the outcome variables prior to the study, asked suitable research questions, had standardized and consistent laboratory and other outcomes and described the evaluation methods adequately. Ninety-three percent of studies used an appropriate treatment to answer the research questions, and clearly identified and appropriately used statistical tests. Eighty percent adequately described the treatment, 66% obtained informed consent, 34% reported collecting data on treatment compliance, 11% reported conducting sample size calculations or having adequate numbers of subjects to detect differences, and 4% reported on adverse events.

Immune Outcomes

Out of the 357 immune outcomes assessed in all the studies, 38 or 10% were assessed by the immunologist as not appropriate for the specific study. These markers included lymphocyte reactivity to mitogen ($n=25$), neutrophil adherence ($n=3$), immunoglobulin G ($n=3$), immunoglobulin M ($n=3$), suppressor T cells ($n=2$), T lymphocytes ($n=1$), IFN- γ ($n=1$) and IgA ($n=1$). These particular studies were eliminated for a variety of reason. Some studies measured outcomes at inappropriate times, some had inappropriate outcomes for the disease studied, and others had protein levels that were significantly outside normal ranges suggesting mislabeled unit measures, or operator error. Sixteen were assessed as unclear if appropriate and the rest were assessed as appropriate.

The most common immune outcomes assessed in the mind-body intervention publications were natural killer cell outcomes ($n=61$), CD4 T lymphocyte ($n=38$), immunoglobulin A (IgA) ($n=24$), CD8 T lymphocyte ($n=23$), and delayed-hypersensitivity skin tests ($n=21$). Using the Natural Standard evidence-based validated grading rationale™ (Table 2), immune outcomes were assessed for strength of scientific evidence to be affected by mind-body medicine. Evaluating immune outcomes regardless of intervention type, only IgA demonstrated strong scientific evidence for positive effects resulting from mind-body interventions (Table 5). IgA included both serum and saliva measures. When examined separately, only salivary IgA had positive evidence (7 positive, 1 negative, high quality studies) whereas, serum IgA had negative evidence (1 positive, 3 negative, high quality studies) (Table 5D). All other outcomes were scored unclear, conflicting, lacking evidence, or with negative evidence according to the grading criterion. Because the health of the subjects probably plays a large role in immune outcomes, grading was repeated separating healthy subject and patient population studies. Also, grading was repeated for all the studies, regardless of study quality. These additional assessments did not change the strength or direction of the evidence.

The effect of interventions on all immune outcomes combined revealed that relaxation training had strong scientific evidence. Biofeedback, Humor, and Meditation lacked adequate data to grade due to a small number of studies included and all others had unclear or conflicting data (Table 6).

DISCUSSION

Study Characterization

It was not surprising that relaxation therapy, CBSM, and hypnosis were the most studied as they have been practiced for a longer duration. Music, disclosure, and humor were included as mind-body medicine based on the concept that their effects are mostly likely mediated through the mind however, there were limited studies on these interventions. During the literature search, studies on humor and music were unexpectedly found, although these search terms were not specifically included. There was discussion within the team on whether to include the studies in the review. It was decided to include the studies although additional searches were not conducted to search for humor and music studies because it was not part of the original study design.

A majority of the studies were small RCT's. As expected, most of the studies did not incorporate double-blinding in the traditional sense, where both the subject and investigator were blinded. The nature of mind-body interventions precludes blinding the subject to their group assignment. Efforts were made by most investigators to include some sort of blinding through data entry, laboratory personnel, and assessment. Surprisingly, most studies did not include adverse events reporting. Whether this is because of a reporting failure or lack of adverse events is unknown. Even though mind-body medicine is a low-risk therapy, reporting adverse events data is essential. Approximately half of the studies utilized healthy subjects, which often results in negative trials. Interpreting efficacy from a negative trial using healthy subjects is problematic because the immune system response may be different in a healthy versus patient participant. Regardless, no difference in evidence grades was found when healthy subject studies were graded separately from patient population studies.

Dosing parameters in mind-body medicine are still as of yet undetermined [21]. Often details of the actual intervention and home practice were not reported, nor were compliance measures assessed. Total exposure time may influence results, is important data to capture, and should be reported in future studies.

It was encouraging that 90% of the studies used a control group. Many used an active and non-active control and when incorporated the active control group exposure time matched the intervention group time. For mind-body studies where placebo may play a pivotal role, both an active and non-active control group should be included [22, 23]. The active control and non-active control group protocols should duplicate the time, attention, and home practice of the experimental group. In this way, non-specific effects like placebo and expectancy can be assessed.

Quality Assessment

As expected, the RCT's had a higher mean quality score than the other study designs. Poor reporting was the major

Table 5. Strength of Evidence for Immune Outcomes

A	Blood Cell Count	p<.05	p>.05	Grade**	B	Cytokines Soluble Factors	p<.05	p>.05	Grade
	Basophil	0 (2)*	0 (1)	L		IFN-γ	2 (4)	1 (1)	C
	Eosinophil	1(1)	0 (2)	C		IL-10	1 (1)	0 (3)	C
	Granulocytes	1 (1)	2 (2)	C		IL-1β	2 (2)	2 (3)	C
	Lymphocytes	4 (8)	8 (10)	D		IL-2	2 (3)	1 (2)	C
	Macrophage	1 (1)	0 (0)	C		IL-4	1 (1)	0 (2)	C
	Monocyte	1 (3)	2 (8)	D		IL-6	0 (0)	2 (4)	D
	White blood cell	2 (4)	4 (10)	D		TNF-α	1 (3)	2 (4)	D
	Neutrophil	1 (7)	2 (4)	D		Neuropeptide Y	0 (0)	1 (0)	C
	Thrombocytes	0 (0)	0 (1)	L		C Reactive Protein	1 (1)	2 (2)	D
C	Cell Activation	p<.05	p>.05	Grade	D	Immunoglobulin	p<.05	p>.05	Grade
	CD3	3 (3)	5 (6)	D		IgA	7 (16)	5 (8)	A
	CD4 T Lymph	6 (12)	14 (25)	F		IgE	1 (1)	0 (0)	C
	CD4/CD8 Ratio	2 (3)	3 (7)	D		IgG	0 (2)	1 (2)	C
	CD8 T Lymph	3 (7)	9 (15)	F		IgM	2 (2)	0 (1)	C
	CD19 (B cells)	2 (2)	3 (5)	D					
	CD23	1 (1)	0 (0)	C					
	CD56 (NK cells)	7 (25)	16 (33)	F					
E	Reactivity to Antigen						p<.05	p>.05	Grade
	Lymphocyte reactivity-ConA, Candida, CMV, MLR, PHA, PWM, VZ***						4 (12)	9 (18)	F
	Flu frequency, duration						0 (2)	1 (1)	C
	Viral titer						3 (7)	6 (7)	D
	Skin Delayed Hypersensitivity						0 (10)	2 (11)	D
	Erythrocyte Sedimentation Rate						0 (0)	1 (1)	C
	Colitis Activity Index						0 (0)	1 (1)	C

*Numbers outside of parentheses represent number of outcomes in each category with a quality assessment score greater than 72. Numbers in parentheses represent the number of outcomes in each category for all studies regardless of quality assessment score.

**Grades were made according to the Natural Standard evidence-based validated grading rationale™ as depicted in Table 1. (A=strong scientific evidence, B=good scientific evidence, C=unclear or conflicting scientific evidence, D = fair negative scientific evidence, F = strong negative scientific evidence, L = lack of evidence)

***ConA= Concanavalin A; CMV= Cytomegalovirus; MLR= Mixed lymphocyte reaction; PHA = phytohemagglutinin; PWM = Pokeweed Mitogen; VZ = Varicella Zoster Virus.

contributor in most quality score deductions. Surprisingly, only 66% of the studies noted obtaining informed consent. It is assumed that consent was obtained but was not reported. However, this is an unnecessary omission. Additionally, many studies did not report power calculations or rationale behind subject number. When a study does not report whether a power calculation was done or that adequate subject numbers were present to detect differences between groups, we must infer that these trials were not adequately powered or that they failed to report power calculations. Either way, the reader is left with uncertainty. Pre-clinical studies not attempting to definitively assess efficacy may not need to report power calculations, yet the objectives of the study as a pilot should be clearly stated.

The quality assessment scores significantly improved for studies published post-2001 after the release of the Miller meta-analysis [19]. The Miller review examined psychological interventions' effect on immune outcomes, with a comprehensive review of the therapies, and discussion of recommendations for future trials including subject selection, choosing appropriate immune system measures, designing methodologically rigorous studies, and testing mechanistic pathways. The Miller and Cohen review reported similar findings in immune system outcomes and the need for improved methods in these trials [19]. Regardless of whether the improvement in study quality was a direct consequence of the publication or some other guidelines in quality study design and reporting, the results are hopeful. This study

Table 6. Strength of Scientific Evidence by Intervention

Intervention Code	p<.05	p>.05	Grade**
Biofeedback	0 (4)*	0 (8)	L
CBSM***	19 (25)	26 (33)	F
CBT	2 (6)	28 (37)	F
Disclosure	1 (5)	4 (13)	F
Humor	0 (3)	0 (0)	L
Hypnosis	15 (28)	16 (38)	D
MBSR	8 (9)	22 (22)	F
Meditation	0 (1)	0 (2)	L
Music	1 (1)	0 (1)	C
Relaxation training	16 (44)	14 (39)	A
Support	5 (13)	5 (12)	C
Visualization	5 (12)	4 (10)	C

*Numbers outside of parentheses represent number of outcomes in each category with a quality assessment score greater than 72. Numbers in parentheses represent the number of outcomes in each category for all studies regardless of quality assessment score.

**Grades were made according to the Natural Standard evidence-based validated grading rationale™ as depicted in Table 1. (A=strong scientific evidence, B=good scientific evidence, C=unclear or conflicting scientific evidence, D = fair negative scientific evidence, F = strong negative scientific evidence, L = lack of evidence).

*** CBSM-cognitive based stress management; CBT-cognitive based therapy; MBSR-mindfulness-based stress reduction.

builds upon the Miller review by examining studies conducted through October 2007, expanding the intervention categories, using alternative grading criteria, and contributing additional recommendations for future trials.

Immune Outcomes

The variability in intervention application is one of the main limitations to accurately synthesizing data regarding mind-body modalities' effects on the immune system. Meta-analytic methods were not used for this study because of heterogeneity not only between the intervention groups but also within the intervention groups. The intervention's implementation consisted of different session lengths, frequency and duration and thus could not be directly compared. For example, relaxation training was held three times a week for 45 minutes for three weeks in one study and once a week for 20 minutes for four weeks along with focused breathing in another. Even if these two studies had identical immune measures, the results could not be combined for meta-analysis because of the application differences. Comparisons may have been conducted if effect sizes were uniformly reported, but they were not.

The study variability also highlights the fact that mind-body medicine research has a paucity of pre-clinical trials where dose response, optimal dose, and preliminary efficacy are established. Investigators often attempt to conduct a Phase 3 definitive assessment of therapy efficacy in an under-powered RCT. Unfortunately, these studies undermine the field because they often yield negative results. Pre-clinical studies must be conducted to move the field forward. Most of the studies reviewed were small RCT's and could be considered pilots. Most studies did not include power calcu-

lations and thus it is uncertain whether the studies could be considered definitive.

Only IgA showed strong evidence for being affected by mind-body medicine. This measure may not be ideal for every intervention or patient group but has shown strong evidence of effects resulting from mind-body interventions. Salivary IgA had positive evidence whereas serum IgA did not possibly reflecting the faster rate of change of salivary IgA and the less stressful collection method. Salivary IgA may be more reliable for mind-body intervention studies. The Miller and Cohen review reported similar findings on IgA [19]. Overall, relaxation training demonstrated the strongest evidence for a mind-body intervention to influence immune outcomes overall. Incorporating some type of relaxation training into mind-body medicine therapies may help improve health outcomes through immune system mediation.

Interpreting immune outcome results includes multiple factors to consider. The direction of change of the immune outcomes can be different for the same outcome with different populations. For example, an increase in IFN γ may be positive in a population of people with a viral infection. The same increase in IFN γ would be considered negative in a population with a Th1-mediated autoimmune disease. Also, the immune outcome must be relevant to the research question and be able to be changed within the time-frame of the intervention. Immune outcome changes may differ in healthy versus patient populations and must be considered when making conclusions. Another issue in interpreting these findings is in the sensitivity, reliability, and validity of immune markers. Are the results of these studies truly negative due to lack of effect on immunity or because the markers employed

lack sensitivity or are improperly used? Some immune outcomes reliability and validity are not well-established and thus using these markers may not be viable. Immune markers are also influenced by nutrition, exercise, caffeine, sleep, and pharmaceuticals. The depth of controlling for or reporting these variables was limited in many of the evaluated studies. Further research is required to assess appropriate, sensitive, reliable, and valid immune outcome measures in mind-body medicine. Additionally, the immune outcome choice may not be relevant to the disease studied or sensitive to the intervention.

One issue we experienced in conducting the review was the definition of mind-body medicine and which modalities should be included. Movement-based practices such as yoga and tai chi were excluded although some may argue that they should have been included in the study. However, the results would have been inconclusive because the immune changes may have been a result of the increased movement rather than the change of mental state. Also, some modalities such as music and humor may not be considered mind-body medicine. Although they may change mental states for some, it arguable whether they are actually a mind-body medicine.

Various biases must be considered when reviewing these results. There is a language bias in the study because although we attempted to include all languages we were unable to translate three of the articles and thus did not include them in the study. There is also a possibility of publication bias as we only included published papers. We were unable to conduct a funnel plot analysis because the gradings were qualitative. Publication bias may be present although usually publication bias presents itself as greater positive trials being published rather than negative ones as was evidenced in this review (149 positive and 208 negative outcome measures).

Another limitation of the study is its qualitative rather than quantitative nature. Ideally, a traditional meta-analysis would have been conducted. However, the extent of mind-body research is not yet vast enough to include multiple studies of similar design to allow for grouped analysis. Because of this, the results from this study must be viewed with a cautionary note that these are observed qualitative trends rather than conclusions.

In order to help improve future mind-body-immune studies, the following recommendations are made:

- 1) Follow CONSORT guidelines for study design and manuscript preparation even in NRCT or uncontrolled trials. A new set of guidelines have recently been created for non-pharmacological treatments and is applicable to mind-body interventions [24].
- 2) Create a dialogue with investigators studying similar interventions and attempt to create consensus on intervention session length, frequency, and duration through pre-clinical studies examining dosing parameters.
- 3) Combine resources between investigators to conduct larger, possibly multiple site studies.
- 4) Use appropriate control groups to account for non-specific effects.
- 5) Conduct and report on power calculations for definitive studies and/or report study is exploratory.

- 6) Continue assessing appropriate, sensitive, reliable, and valid immune outcomes.

ACKNOWLEDGEMENTS

The author would like to thank Susan Norris, MD, Mary McDonagh, MD, Andrew Hamilton, MSLIS, Erin Connelly, Khivan Singh, Tracy Edinger, ND, Helfgott Research Institute at the National College of Naturopathic Medicine, and Oregon Health & Sciences University for their support of this project. This publication was made possible with support from the Oregon Clinical and Translational Research Institute (OCTRI), UL1 RR024140 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research, and NIH T32 AT002688.

SUPPORTIVE/SUPPLEMENTARY MATERIAL

Supplementary material can be viewed at: <http://www.bentham.org/open/toaltmedj>.

REFERENCES

- [1] Barnes PM, Powell-Griner E, McFann K, Nahin RL. Complementary and alternative medicine use among adults: United States, 2002. *Adv Data* 2004; 343: 1-19.
- [2] Mind-Body Medicine: an overview. NCCAM Publication No D239. [updated May 2007] Available from: <http://nccam.nih.gov/health/backgrounds/mindbody.htm>
- [3] Ernst E, Pittler MH, Wider B, Boddy K. Mind-body therapies: are the trial data getting stronger? *Altern Ther Health Med* 2007; 13(5): 62-4.
- [4] Wahbeh H, Elsas SM, Oken BS. Mind-body interventions: applications in neurology. *Neurology* 2008; 70 (24): 2321-8.
- [5] Rutledge JC, Hyson DA, Garduno D, Cort DA, Paumer L, Kappagoda CT. Lifestyle modification program in management of patients with coronary artery disease: the clinical experience in a tertiary care hospital. *J Cardiopulm Rehabil* 1999; 19(4): 226-34.
- [6] Astin JA, Shapiro SL, Eisenberg DM, Forsys KL. Mind-body medicine: state of the science, implications for practice. *J Am Board Fam Pract* 2003; 16(2): 131-47.
- [7] Mundy EA, DuHamel KN, Montgomery GH. The efficacy of behavioral interventions for cancer treatment-related side effects. *Semin Clin Neuropsychiatry* 2003; 8(4): 253-75.
- [8] Irwin MR. Human psychoneuroimmunology: 20 years of discovery. *Brain Behav Immun* 2008; 22(2): 129-39.
- [9] Ader R, Cohen N. Behaviorally conditioned immunosuppression. *Psychosom Med* 1975; 37(4): 333-40.
- [10] Robinson FP, Mathews HL, Witek-Janusek L. Issues in the design and implementation of psychoneuroimmunology research. *Biol Res Nurs* 2002; 3(4): 165-75.
- [11] Pedersen BK, Toft AD. Effects of exercise on lymphocytes and cytokines. *Br J Sports Med* 2000; 34(4): 246-51.
- [12] Radom-Aizik S, Leu SY, Cooper DM, Zaldivar F, Jr. Serum from exercising humans suppresses t-cell cytokine production. *Cytokine* 2007; 40(2): 75-81.
- [13] Gleeson M, Bishop NC. The T cell and NK cell immune response to exercise. *Ann Transplant* 2005; 10(4): 43-8.
- [14] Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ* 2004; 328(7430): 39-41.
- [15] Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001; 323(7303): 42-6.
- [16] Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; 282(11): 1054-60.
- [17] Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics* 1989; 84(5): 815-27.
- [18] Deeks JJ, Dinnes J, D'Amico R, *et al*. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003; 7(27): iii-x, 1-173.

- [19] Miller GE, Cohen S. Psychological interventions and the immune system: a meta-analytic review and critique. *Health Psychol* 2001; 20(1): 47-63.
- [20] Jadad AR, Moore RA, Carroll D, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996; 17(1): 1-12.
- [21] Caspi O, Burleson KO. Methodological challenges in meditation research. *Adv Mind Body Med* 2005; 21(1): 4-11.
- [22] Oken BS, Zajdel D, Kishiyama S, *et al.* Randomized, controlled, six-month trial of yoga in healthy seniors: effects on cognition and quality of life. *Altern Ther Health Med* 2006; 12(1): 40-7.
- [23] Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technol Assess* 1999; 3(3): 1-96.
- [24] Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008; 148(4): 295-309.

Received: October 16, 2008

Revised: January 28, 2009

Accepted: February 14, 2009

© Wahbeh *et al.*; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.