THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE
Volume 25, Supplement 1, 2019, pp. S21–S51
Mary Ann Liebert, Inc.
DOI: 10.1089/acm.2018.0499

JACM

REVIEW ARTICLE

Whole Systems Research Methods in Health Care: A Scoping Review

Nadine Ijaz, PhD,[1] Jennifer Rioux, PhD,[2] Charles Elder, MD, MPH,[3] and John Weeks[4]

Abstract

Objectives: This scoping review evaluates two decades of methodological advances made by ''whole systems research'' (WSR) pioneers in the fields of traditional, complementary, and integrative medicine (TCIM). Rooted in critiques of the classical randomized controlled trial (RCT)'s suitability for evaluating holistic, complex TCIM interventions, WSR centralizes the principle of ''model validity,'' representing a ''fit'' between research design and therapeutic paradigm.

Design: In consultation with field experts, 41 clinical research exemplars were selected for review from across 13 TCIM disciplines, with the aim of mapping the range and methodological characteristics of WSR studies. Using an analytic charting approach, these studies' primary and secondary features are characterized with reference to three focal areas: research method, intervention design, and outcome assessment.

Results: The reviewed WSR exemplars investigate a wide range of multimodal and multicomponent TCIM interventions, typified by wellness-geared, multitarget, and multimorbid therapeutic aims. Most studies include a behavioral focus, at times in multidisciplinary or team-based contexts. Treatments are variously individualized, often with reference to ''dual'' (biomedical and paradigm-specific) diagnoses. Prospective and retrospective study designs substantially reflect established biomedical research methods. Pragmatic, randomized, open label comparative effectiveness designs with ''usual care'' comparators are most widely used, at times with factorial treatment arms. Only two studies adopt a double-blind, placebo-controlled RCT format. Some cohort-based controlled trials engage nonrandomized allocation strategies (e.g., matched controls, preference-based assignment, and minimization); other key designs include single-cohort pre–post studies, modified n-of-1 series, case series, case report, and ethnography. Mixed methods designs (i.e., qualitative research and economic evaluations) are evident in about one-third of exemplars. Primary and secondary outcomes are predominantly assessed, at multiple intervals, through patient-reported measures for symptom severity, quality of life/wellness, and/or treatment satisfaction; some studies concurrently evaluate objective outcomes.

Conclusions: Aligned with trends emphasizing ''fit-for-purpose'' research designs to study the ''real-world'' effectiveness of complex, personalized clinical interventions, WSR has emerged as a maturing scholarly discipline. The field is distinguished by its patient-centered salutogenic focus and engagement with nonbiomedical diagnostic and treatment frameworks. The rigorous pursuit of model validity may be further advanced by emphasizing complex analytic models, paradigm-specific outcome assessment, inter-rater reliability, and ethnographically informed designs. Policy makers and funders seeking to support best practices in TCIM research may refer to this review as a key resource.

[1]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Canada.
[2]Integral Ayurveda and Yoga Therapy, Chapel Hill, NC.
[3]Kaiser Permanente Center for Health Research, Portland, OR.
[4]johnweeks-integrator.com, Editor-in-Chief, *JACM*, Seattle, WA.

Keywords: whole systems research, complementary therapies, integrative medicine, clinical trials as topic, pragmatic trials, complex interventions

Introduction

The adoption of ''fit-for-purpose'' clinical research designs has emerged in recent decades as a significant trend in health care. Policy makers increasingly formulate system-wide decisions informed by the combined results of ''pragmatic'' controlled trials, which rigorously investigate the real-world effectiveness of health care interventions (compared to their idealized ''explanatory'' efficacy).[1] More funders now commit to reducing health care costs by underwriting studies of complex interventions focused on preventive multidisciplinary care.[2] Researchers, in turn, widely augment measurements of objective biomarkers by evaluating patient-reported outcomes directly meaningful to those suffering ill health.[3] Finally, patients continue to demand evidence-informed care that reflects their values and priorities.[4]

Few would argue that the double-blind, placebo-controlled randomized controlled trial (RCT) continues to occupy pride of position at the top of evidence based medicine (EBM)'s methodological hierarchy of clinical trial designs. That said, researchers from multiple fields—including traditional, complementary, and integrative medicine (TCIM)—have critiqued the RCT's limitations and its disproportionate evidentiary dominance. The present work, a scoping review, represents a first retrospective analysis of almost two decades of research design advances made by scholars committed to rigorous, holistic clinical research designs that accurately represent the unique paradigmatic features of TCIM ''whole systems'' interventions.

Background

In 2003, Ritenbaugh et al.—researchers in the TCIM field—published a seminal article proposing a new branch of scientific inquiry, which they termed ''whole systems research'' (WSR).[5] WSR pioneers proposed to innovate clinical research designs to address the theoretical-methodological dissonance that may arise in using classical RCT designs—revered as the ''gold standard'' in biomedical research—to appropriately study TCIM care. TCIM ''whole systems'' paradigms (e.g., Chinese medicine and naturopathic medicine), they argued, exemplify several central features (detailed below) that distinguish them from conventional biomedicine. At the heart of WSR is the model validity principle, defined here as the ''fit'' between a study's design and the conceptual and clinical features of the studied intervention's underlying or originating paradigm.[6] WSR advocates envisioned the pursuit of model validity as a way to rigorously supplement (and reprioritize) existing approaches to achieving external and internal validity in clinical research.

The dominant RCT design, as critics had observed over the two decades prior,[7–9] seeks to study singular, isolated therapeutic components to ''determine the single best treatment for all patients.''[5] TCIM treatments, however, are typically complex (involving multiple synergistic treatment modalities or components) and individually tailored to the specific patient.[6,9] Classical RCTs were purpose developed to assess the causal effects of pharmaceutic treatments on particular physiologic pathways, under double-blinded, placebo-controlled conditions.[10,11] However, many TCIM interventions are behaviorally focused (with a ''salutogenic'' emphasis on lifestyle and disease prevention), rendering clinician and participant blinding difficult. Constructing credible, inert placebo controls for many TCIM treatments (e.g., acupuncture, chiropractic, and massage) had moreover proved notoriously challenging.[9] Finally, scholars working in the relatively-marginal TCIM field have characterized the high cost of conducting classical RCTs as a prohibitive barrier to research feasibility.[12]

WSR proponents in the TCIM field were certainly not alone in advocating for revisions to methodological conventions in clinical research; investigators in some biomedical fields (e.g., psychotherapy, surgery, and dietetics) had at the time articulated parallel concerns around the RCT's universal applicability.[6,13] However, WSR proponents additionally pointed to a unique set of research challenges arising from paradigmatic features of TCIM ''whole systems,'' in relation to which these differ substantively from conventional biomedical approaches.[5,6]

As detailed in Table 1, many whole TCIM systems rely on conceptual models and diagnostic approaches distinct from or in addition to biomedical science. Alongside an integrated (''whole person'') assessment of a patient's physical, mental, emotional, and psychosocial well-being, many TCIM occupations foundationally attend to patient preferences, priorities, and values in their treatment designs.[5,14] Classical RCTs engage objective measures at discrete endpoints to evaluate predetermined primary treatment outcomes related to a narrowly defined disease or dysfunction.[15,16] Conversely, TCIM providers—whose interventions are often multitarget or multimorbid in their aims—typically rely on subjective assessment modes to track progressive (and often long term) improvements in patient well-being alongside a range of inter-relating symptoms.[15,16] Finally, while RCTs classically evaluate an intervention's effects before it is being deployed in mainstream care, TCIM therapies are often in widespread usage before being formally trialed.[17]

For those advocating a WSR approach, the evaluation of singular, standardized TCIM modalities within classical RCT frameworks did not suffice as a means by which to evaluate these therapies' effects. Rather, they insisted that model validity must be sought.[6] Mirroring a growing chorus of biomedical researchers, WSR advocates heralded the ascent of ''pragmatic'' RCT designs which—they noted—might rigorously compare the real-world effectiveness of complex individualized interventions with ''usual'' biomedical care, with reference to diverse rather than homogenous populations.[6,18–20] They called for engagement with modified RCT designs (e.g., patient preference, factorial and n-of-1 trials; matched or waiting list controls)[6,19] and recommended adoption of more efficient and equally-rigorous design-adaptive allocation alternatives to randomization (e.g., minimization).[21] Advocating for mixed methods study designs, they argued that qualitative methods could not only ''assist in the development of appropriate outcome

Table 1. Characteristics of Clinical Whole Systems Paradigms

Paradigm	*Conceptual model*	*Diagnostics*	*Treatment modes*
Ayurvedic medicine[29,30]	Typological assessment of constitution (*prakruti*) and disequilibrium (*vikruti*) in relation to whole person[a] parameters (three *doshas*: *kapha*, *vata*, *pitta*), metabolic function (*agni*), toxin load (*ama*), bodily essences (*tejas*, *ojas*, *prana*), qualities, disease stages, and locations.	Narrative case-taking, physical examination, palpation, pulse and tongue assessment, detoxification, and rejuvenation therapies.	Diet and lifestyle counseling, herbal medicine, manual therapies, enemas and purgation, nasal treatments, yoga and meditation, breathing exercises, music and mantra, and self-awareness activities.
Anthroposophic medicine[31]	Biomedical assessment + typological assessment of whole person[a] constitution and disequilibrium in relation to four levels of formative forces (physical, etheric, astral, ego) and threefold structural/functional systems (nerve-sense, motor-metabolic, rhythmic).	Biomedical diagnostics + additional narrative case-taking.	Anthroposophic medication (homeopathic, herbal), diet and lifestyle counseling, art therapy, rhythmical massage therapy, Eurythmy movement therapy, biographical counseling, psychotherapy ± usual biomedical care.
Chinese medicine[32]	Typological assessment of whole person[a] constitution (vital substances) and disequilibrium (pathogenic factors, stagnations) in relation to six divisions of *yin* and *yang*, system function/interaction (five elements), stages and levels of disease.	Narrative case-taking, physical examination, palpation, pulse, and tongue assessment.	Acupuncture, moxibustion, herbal medicine, tuina massage, cupping, *guasha* scraping, *t'ai chi*, *qi gong*, dietary and lifestyle counseling.
Chiropractic medicine[33,34]	Assessment of biomechanical disorders based on biomedical concepts of musculoskeletal/ nervous systems; controversially/historically conceptualized in nonbiomedical terms as ''vertebral subluxation.''	Biomedical diagnosis + physical examination, palpation, functional assessment, diagnostic imaging, laboratory testing.	Spinal and soft tissue manipulation, physical modalities, home care, and counseling on diet, exercise, and stress reduction.
Complementary/ integrative medicine[35]	Inclusion of treatments originating from a range of whole system/whole practice paradigms within the auspices of conventional/ preventive biomedical health care delivery.	Biomedical diagnostics + optional system/ modality-specific diagnostics.	Usual biomedical care plus single or multiple treatment approaches from one or more whole systems/whole practice paradigms (including approaches not listed here).
Energy medicine[36]	Assessment of whole person[b] energetic field.	Intuitive/energetic observations.	A range of on-body (e.g., healing touch, Reiki) and off-body treatments.
Homeopathic medicine[37,38]	Typological assessment of remedy signature (*simillimum*) of whole person[a] constitution and disequilibrium, possibly in relation to disease miasm (e.g., *psoric*, *sycotic*, *syphilitic*) and/or kingdom (plant, animal, mineral).	Narrative case-taking.	Homeopathic dilutions of a range of plant, animal, and mineral substances.
Midwifery[39]	Woman-centered perinatal care in which birth is normalized as a healthy event, and the midwife's role is to holistically support and facilitate the individual woman's birthing choices.	Biomedical, narrative case-taking, physical examination.	Case-load based (continuous), as well as midwife-led/team-based (noncontinuous) pre-, intra-, and postpartum birth-related care (including counseling related to diet, lifestyle, and infant feeding/care) with option of home- or hospital birth.

(*continued*)

Table 1. (Continued)

Paradigm	*Conceptual model*	*Diagnostics*	*Treatment modes*
Naturopathic medicine[40]	Biomedical assessment reinterpreted through a whole person[a] lens + optional Chinese medicine/homeopathic assessments.	Biomedical diagnostics + additional narrative case-taking, physical examination, laboratory and bioelectrical testing, palpation + optional Chinese medicine/homeopathic diagnostics.	Diet, physical activity, stress management counseling; herbal medicine, nutritional supplementation, acupuncture, homeopathy, hydrotherapy, manual therapies, physical modalities, and instruction in mind/body techniques.
Preventive/restorative biomedicine[c]	Biomedical assessment with preventive/ restorative and physiologic/psychosocial lens.	Biomedical.	Individual/group-based behavioral interventions geared to preventing/rehabilitating chronic disease, including diet, nutrients, exercise, sleep, stress management, psychosocial supports, mind/body techniques, ±usual biomedical care.
Swedish massage therapy[41]	Biomedical assessment with an emphasis on soft tissue and musculoskeletal disorders.	Biomedical diagnostics + physical examination, palpation, functional assessment.	Manual therapy, including friction, *effleurage*, *petrissage*, vibration, *tapotement*, and skin-rolling.
t'ai chi[42]	See Chinese medicine, above	See Chinese medicine, above	Ritualized movement sequence incorporating breathwork, mindfulness, imagery, physical touch, and social interaction.
Yoga therapy[43,44]	Management, reduction, or elimination of symptoms that produce suffering; enhancement of function; illness prevention; and salutogenesis.	Assessment of person as ''multidimensional system'' that includes interconnections of body, breath, intellect, mind, and emotions.	Therapeutic movement (*asana*) and breathwork (*pranayama*) with dietary, meditation, mantra, mudra, chanting, ritual, and self-awareness/lifestyle practices.

[a]This table provides an overview of selected whole systems paradigms, studies from which are evaluated in this review. It is not meant to be an exhaustive representation of all clinical whole systems—there are many others.

[b]Whole person parameters concurrently address physiologic, psychologic/mental, emotional, spiritual, social, intergenerational, and environmental factors as part of a holistic conceptual paradigm. In other words, these factors are understood as fundamentally interconnected and mutually generative in relation to health and well-being.

[c]The term ''preventive/restorative biomedicine'' is provisionally used here to characterize studies with a set of unique paradigmatic features, led by conventional medical doctors.

measures'' before a clinical trial but also gather ''unique physical and psychosocial context'' within it and subsequently help to ''explain the trial results.''[22]

Going further, WSR proponents argued that diverse research modes—prospective and retrospective; experimental, quasi-experimental, and observational; qualitative and quantitative; and holistic and reductive—be equally valued for their distinct contributions and rigorously applied as contextually appropriate.[6,23] Asserting that EBM's ''prescriptive evidence hierarchies of research methods'' should be supplanted,[20] TCIM scholars variously conceptualized evidentiary frameworks (e.g., ''evidence matrix,''[24] ''evidence house,''[25] and ''circular model''[19]) in which a range of research designs might synthetically contribute to assessing a particular intervention's efficacy, effectiveness, and other contextual dimensions.

Taking on model validity with respect to intervention selection and design, WSR advocates favored the evaluation of ''whole systems, or 'bundles' of therapies'' rather than ''single…modalities'' alone.[6] They envisioned studies in which patients would undergo ''double classification'' using biomedical diagnostics, as well as diagnosis from within the relevant TCIM paradigm, and receive care that was individualized on this basis.[6] Research teams, they advised, should include insiders from within the paradigms in which the interventions originated[26,27]; and study recruitment strategies should address persons with complex, multifactorial health conditions,[28] as well as patient treatment preferences.[14]

WSR leaders equally envisioned study outcome assessment in relation to the model validity principle.[6] At a time when validated, patient-reported outcome measures (PROMs) were just beginning to be widely used in conventional research, WSR proponents characterized subjective and paradigm-adherent quantitative outcome measures[15,22] as key evaluative tools, alongside qualitative methods.[22] Measurables, they proposed, should be multiple (addressing the therapeutic techniques applied, patient–practitioner relationship, and range of health/wellness impacts[15,45]) and at more frequent intervals and over a longer period than in conventional trials.[16,23] ''Innovative statistical methodology''[6]—including ''participant-centered'' approaches[46]—would be needed to synthesize the voluminous data generated.[6] They called for ''complex conceptual models''[16] to evaluate a whole system's combined effects ''over and above its components''[6] and variously proposed methodological engagement with network science, complexity science and nonlinear dynamical systems,[23,47–49] action research,[7,16] and program theory[16] to this end.

Since 2003, the dominant landscape of clinical research has transformed significantly. Although the classical RCT continues to be prioritized in EBM's evidentiary hierarchy, pragmatically designed comparative effectiveness studies and ''fit-for-purpose'' research designs[1] have become more widely accepted as important clinical and policy-making resources.[50] Usage of PROMs, clinical trial guidelines, and quality assessment tools has become more widespread; and, ''following considerable development in the field,'' the Medical Research Council's framework for trialing complex interventions will once again be renewed in 2019.[2] Within the TCIM world, WSR principles have been increasingly taken up,[43,51–55] although more conventional research designs still predominate.[56] To date, however, no comprehensive retrospective analysis of WSR advances has been undertaken; that is thus the present work's aim.

Methods

This article is a scoping review of the methodological features of WSR studies, with reference to the model validity principle. Scoping reviews ''map the literature on a particular topic or research area and provide an opportunity to identify key concepts; gaps in the research; and types and sources of evidence to inform practice, policymaking, and research.''[57] Scoping reviews ''differ from systematic reviews as authors do not typically assess the quality of''[58] nor ''seek to 'synthesize' evidence or to aggregate findings from different studies.''[59] They also diverge from ''narrative or literature reviews in that the scoping process requires analytical reinterpretation of the literature.''[58] ''Not linear but iterative'' in character, scoping reviews primarily take a qualitative analytic approach, supported by numerical representation of the ''extent, nature, and distribution'' of key findings.[59]

The present review adopts Arksey and O'Malley's six-step scoping study framework, involving: (1) research question identification; (2) study identification; (3) study selection; (4) data charting; (5) result collation, summary, and reporting; and (6) (optional) consultation with area experts to validate findings.[59]

Research question identification

The primary question driving this review is twofold, interrogating: (1) the range and characteristics of WSR clinical studies and (2) the ways in which these studies engage the *model validity* principle.

Study identification

WSR-type studies have been undertaken in multiple health care paradigms, and the methodological terminology used across them varies. Thus, a broad initial keyword-based literature search (e.g., ''whole systems research,'' ''complex,'' ''individualized,'' ''complementary medicine,'' and ''model validity'') helped to locate many relevant *methodological* publications, but proved insufficient to identify a representative set of *clinical WSR exemplars*. A group of field experts (listed in the study acknowledgments) was therefore assembled by one coauthor (J.W.) to share WSR exemplar citations. In addition to reviewing the relevant historical literatures, the primary review author (N.I.) reviewed each of these recommended studies and scrutinized their reference lists for additional candidate exemplars. The other coauthors (J.R. and C.E.), as WSR field experts, further supplemented this initial list. As the review process progressed, study identification through additional literature searches continued iteratively with study selection and data charting (below).

Study selection

To be eligible for inclusion, studies were required to directly report clinical outcomes with respect to an intervention based in a defined therapeutic whole system marked by a conceptual and/or diagnostic model distinct from conventional biomedical care. Studies adopting complex, individualized, salutogenic, and/or multimorbid/multitarget modes of care were prioritized. Only peer-reviewed studies (with one or more associated publications) were included; and all demonstrated a strong emphasis on model validity in at least one of the following: adopted research method(s),

intervention selection or design, and outcome assessment. Studies were not required to refer directly to the model validity principle nor to use WSR terminology. Pilot/feasibility designs were included; unfulfilled study protocols were not. No attempt was made to exhaustively assemble all published studies meeting study inclusion criteria; rather, the emphasis was on assembling a diverse subset of such studies.

Addressing a long-standing debate in the WSR field,[20] the multicomponent, stand-alone disciplines of yoga therapy and *t'ai chi* were defined as distinct whole systems, despite their respective historical and conceptual connections to the Ayurvedic and Chinese medicine systems. Studies from midwifery (a discipline not always included under the TCIM ''umbrella'') were determined eligible for inclusion based on: (1) the profession's uniquely holistic, woman-centered paradigm, distinct from conventional obstetrics and (2) its historical roots in traditional/indigenous health care. Studies from a field provisionally termed ''preventive/restorative biomedicine'' were also included, recognizing that: (1) such studies diverge paradigmatically from conventional therapeutic norms and (2) that the multimodal, behaviorally focused studies led in particular by Ornish et al. in the 1980s[60,61] provided early methodological inspiration for whole systems researchers.

Study selection (i.e., identification, charting, and culling) continued iteratively until: (1) ''theoretical saturation''[62] was reached, in that review of additional candidate publications failed to reveal new WSR methodological features; and (2) a wide range of clinical whole systems paradigms were represented within the dataset. Study results were *not* taken into account during the selection process.

About 90% of studies recommended by at least one subject area expert were included; and approximately two-thirds of the included studies had been directly recommended by at least one WSR expert (including coauthors J.R. and C.E.); the remainder was identified in literature searches undertaken by the primary author (N.I.). Four expert-recommended studies were excluded because they: (1) did not meet the study inclusion criteria ($n=1$) or (2) were methodologically very similar to other selected exemplars, providing little added value to the review ($n=3$). The final selection of studies deliberately over-represents *traditional* (i.e., Ayurvedic and Chinese) medicine systems, to thoroughly address the paradigm-specific diagnostic, intervention, and outcome design considerations that arise in these contexts.

Data charting

Focused around three primary analytic categories—Study Design, Intervention Selection, and Outcome Evaluation—the primary author (N.I.) summarized and evaluated each candidate study using an emergent set of tables and charts. Through a constant comparative approach that reviewed each study in relation to all others,[63] a set of analytic subparameters and conceptual frames progressively emerged. This process permitted a finalized study selection and a detailing of each study's distinct and nondistinct methodological features.

Expert validation of findings

While analysis and reporting were undertaken primarily by the primary author (N.I.), a subset of categorizations related to ''dual diagnostics'' and paradigm-specific outcomes was independently corroborated by another coauthor (J.R.). All coauthors (J.R., C.E., J.W.) contributed insights as to the emerging conceptual categories as the project progressed and provided input on the final analyses before this work's peer review by other WSR field experts.

Result collation, summary, and reporting

Results are synthetically presented and discussed in what follows using both narrative and graphical reporting. To facilitate reading ease, in-text WSR exemplar references name first authors only; full citations may be found in the reference list. To provide context and language to facilitate nuanced reporting of the WSR field's features, two novel theoretical frameworks are presented below.

Theory

Model validity framework

The model validity principle has been conceptualized as central to WSR; and, as noted earlier on, various scholars have suggested ways in which this principle may be enacted within clinical research contexts. What remained implicit in much (although not all) of the early WSR methodological literature is that WSR itself may be understood as part of an ''integrative medicine'' movement[64] geared to transforming dominant health care systems such that TCIM therapies may be more broadly integrated alongside or as an adjunct to conventional biomedical care. Clinical research is in this scenario envisioned as a necessary but insufficient tool to help to dismantle barriers to integration.[17]

Several scholars, critiquing the integrative medicine project, have however suggested a potential for the distinct paradigmatic features of and practices with origins in nonbiomedical therapeutic systems to be co-opted, appropriated, or assimilated in such a process.[65] Model validity, as a theoretical construct, represents a commitment to actively preserving these paradigms and practices in their own right, an approach aligned with the concept of a clearly ''articulated,''[66] equitable medical pluralism,[67] rather than an assimilative mode of integration.[65]

What WSR pioneers were not able to fully apprise in advance was how and to what degree future WSR methods might ultimately align or diverge with conventional research strategies in pursuit of model validity. To facilitate analysis of these points in the present scoping review, it is proposed that the model validity principle be theoretically differentiated into three co-embedded categories as seen in Figure 1: *paradigm compatibility*, *paradigm consistency*, and *paradigm specificity*. These categories are not mutually exclusive, that is, a single study may concurrently include different aspects (e.g., method, intervention design, and outcome measures) marked by one or more of the identified characteristics.

Paradigm compatibility, model validity's driving concept, is conceptualized as a category that includes two others—*paradigm consistency* and *paradigm specificity*—the second of which is embedded within the first. *Paradigm compatible* research methods are those typically associated with dominant biomedical clinical research, but which also readily lend themselves to the study of whole systems clinical interventions. *Paradigm-consistent* methods differ in key

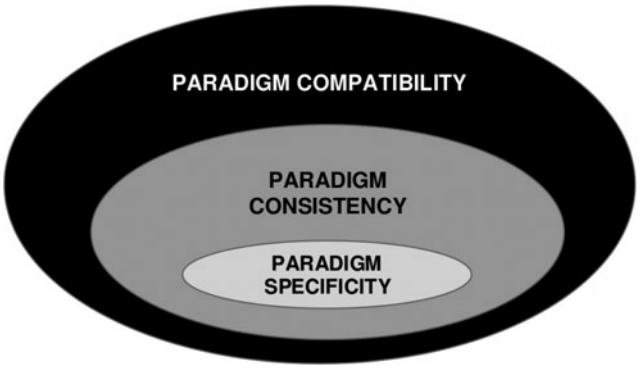

FIG. 1. Model validity framework.

ways from conventional research approaches, but are distinctly suited to evaluating a wide range of whole systems interventions. *Paradigm-specific* research methods differ or diverge from conventional research approaches and are furthermore uniquely tailored to one specific clinical whole system or paradigm.

Individualization spectrum

Individualized care represents a core therapeutic principle across TCIM whole systems and is thus a key consideration with regards to WSR model validity. Strategies for individualizing care in clinical research contexts have been explored over the last two decades, in particular in the field of biomedical psychotherapy. Researchers in that field have unfolded what has come to be known as ''manualization,'' in which formal treatment manuals specify a predetermined set of intervention parameters, within which study clinicians are granted scope to individually tailor treatments.[68] It should be similarly noted that Chinese, Ayurvedic, and other traditional medicine systems have for many centuries used semistandardized treatment protocolization as a structure within which to personalize patient treatments.[30,32] In such traditional systems, generalized treatment parameters (e.g., dietary recommendations, herbal formulations, and acupuncture point combinations) are detailed in relation to particular primary diagnostic or constitutional patterns, providing clinicians with a framework within which to further tailor care. However, in other TCIM paradigms (e.g., naturopathic medicine and chiropractic), individual clinicians commonly individualize treatments with fewer defined constraints.

To facilitate a nuanced representation of the range of approaches to intervention individualization evident in the WSR exemplars reviewed, a theoretically-novel *individualization spectrum* is presented in Figure 2. This spectrum differentiates the broad range of approaches to treatment personalization under three broad categories: *general standardization*, *manualization with tailoring*, and *unconstrained individualization*. Toward the left of the spectrum—''general standardization''—are interventions involving predefined inflexible interventions, uniformly delivered to all participants. At the spectrum's right are treatments characterized by their ''unconstrained individualization,'' in which providers have discretion to uniquely treat each patient within the breadth of their clinical scope. At the spectrum's center are ''manualization with tailoring'' approaches, in which clinicians have autonomy to personalize treatments in adherence to prespecified intervention parameters. As seen in Figure 2, the spectrum's three base categories are neither rigid nor mutually exclusive; rather, features of one approach may be evident in an intervention or study dominated by another.

Results Overview

This scoping review evaluates a total of 41 WSR studies from across the paradigms of anthroposophic,[69–71] Ayurvedic,[29,72–76] Chinese,[77–86] chiropractic,[33] complementary/integrative,[87–91] energy,[36] homeopathic,[92] naturopathic,[85,93–96] and preventive/restorative[60,97–99] medicines, as well as midwifery,[100] Swedish massage,[101,102] *t'ai chi*,[103] and yoga therapy.[29,104] The whole systems interventions reported across these studies range in size from one[74] to almost three thousand[98] patients and in duration from 1 day[87] to several years.[69,71] Conducted across several continents, these studies address many areas of clinical focus, including: acute[87] and chronic[94] anxiety; adjunct oncology care[83,88,89,91]; acute,[79] as well as chronic,[71] illness (including headache,[36] rheumatoid arthritis,[69] heart disease,[60,72,95,98,99] and diabetes[93,105]); insomnia,[92] obesity,[29] and tinnitus[82]; musculoskeletal pain[33,70,75,85,86,90,96,101,103]; reproductive[60,74,77,81,98] and respiratory[73,78,80] conditions; and medically unexplained symptoms.[84] Rather than treating ''disease'' conditions *per se*, a number of studies focus primarily on well-being,[76] quality of life (QoL),[88,91] social and emotional skills,[104] prevention and rehabilitation,[60,97–99] clinical care dynamics,[102] and patient satisfaction with clinical care.[100]

Several of the reviewed studies have secondary associated publications detailing qualitative research[33,84,93,101] or economic outcomes.[95,96,106] Other secondary publication types include: stand-alone study protocols,[33,75] earlier pilot/feasibility studies[80,101]; methodological works[33,86,101]; and articles detailing additional/follow-up outcomes.[60,70,83,92,98,100]

FIG. 2. Spectrum of clinical individualization strategies.

Table 2. Methodological Overview of Whole Systems Research Studies

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Attias 2016, Israel[87]	**Complementary/ integrative medicine** Preoperative anxiety	**Randomized controlled trial** Open label, multiarm design comparative effectiveness design; usual care and intraparadigmatic comparators. $N=360$, 1 day.	**Arms I–V:** Usual care (pharmaceutical) + **I:** standardized guided imagery; **II:** individualized guided imagery; **III:** individualized reflexology; **IV:** individualized guided imagery + individualized reflexology; **V:** individualized acupuncture. **Arm VI:** Usual care (pharmaceutical).	**Primary:** Symptom severity PROM (anxiety VAS).
Azizi 2011, China[77]	**Chinese medicine** Menopause-related symptoms	**Randomized controlled trial** Open label, multiarm design; usual care and intraparadigmatic comparators. $N=57$, 2 months.	**Arm I:** Standardized traditional herbal mixture. **Arm II:** Standardized herbal mixture + standardized acupuncture. **Arm III:** Usual care (pharmaceutical).	**Primary:** Symptom severity PROM (Kupperman index). **Secondary:** Hormonal bloodwork (estradiol); number of symptoms.
Bell 2011, United States of America[92,108,109]	**Homeopathic medicine** Coffee-induced insomnia	**n-of-1 Series** Dynamically allocated, patient-blinded, placebo-controlled, two-period, comparative effectiveness (AB^1/AB^2) design. $N=54$; 4×1 week phases.	**Intervention A:** Homeopathic placebo. **Intervention B^1:** Homeopathic remedy I **Intervention B^2:** Homeopathic remedy II	**Primary:** Functional sleep biomeasures (Polysomnography); Symptom severity PROM; Mental health PROMs.
Ben-Arye 2018, Israel[88]	**Complementary/ integrative medicine** Chemo-induced taste disorder	**Pre–post cohort study** Single arm, chart review design. $N=34$, ≤12 weeks.	**Single arm:** Individualized complementary/integrative medicine + usual care (biomedical oncology).	**Primary:** QoL PROM (MYCaW); Symptom severity PROM.
Bradley 2012, United States of America[93,110]	**Naturopathic medicine** Type II diabetes	**Controlled pre–post cohort study** Open label, comparative effectiveness design[93]; usual care comparator. Qualitative substudy[110]. $N=40$ + $N=329$ (control). 6–12 months. In-depth interviews: $N=5$.	**Arm I:** Individualized whole system naturopathic + usual care. **Arm II:** Usual biomedical care data from electronic health records **Qualitative aims:** Exploration of patient-reported experiences receiving first-time naturopathic care.	**Primary:** Adherence PROM; mental health PROMs (PHQ-8), self-efficacy PROM, emotional wellness PROM; blood lipids; blood pressure. **Secondary:** Treatment satisfaction PROM, health service utilization; medication usage.
Bredesen 2016, United States of America[97,111]	**Preventive/restorative biomedicine** Alzheimer's disease	**Retrospective case series** $N=10$, 5–24 months, ≤3.5 year follow-up.	Manualized/tailored preventive diet/ lifestyle protocol.	**Primary:** Functional/disease progression testing (quantitative MRI, neuropsychologic testing); narrative case reporting.
Brinkhaus 2004, Germany[78]	**Chinese medicine** Seasonal rhinitis	**Randomized controlled trial** Placebo-controlled, patient-blinded design; intraparadigmatic comparator. $N=57$, 2 months.	**Arm I:** Standardized traditional herbal mixture + manualized/tailored herbal mixture + manualized/ tailored acupuncture. **Arm II:** Herbal placebo + Sham acupuncture.	**Primary:** Symptom severity PROM (VAS). **Secondary:** Symptom severity PROM, QoL PROMs (SF-36); placebo credibility scale; medication usage; bloodwork for adverse effect testing.

(*continued*)

TABLE 2. (CONTINUED)

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Cooley 2009, Canada[94]	**Naturopathic medicine** Moderate/severe anxiety	**Randomized controlled trial** Open label, placebo-controlled, comparative effectiveness design; usual care comparator. $N = 81$, ≥8 weeks.	**Arm I:** Manualized/tailored diet/exercise counseling + standardized multivitamin and herb + standardized deep breathing education **Arm II:** Placebo (multivitamin) + standardized deep breathing education + usual care (psychotherapy).	**Primary:** Symptom severity PROM. **Secondary:** QoL PROM (SF-36); Fatigue PROM, patient-generated PROM (MYMOP); adherence PROM; and patient satisfaction PROM.
Dubroff 2015, Canada[72]	**Ayurvedic medicine** Coronary heart disease	**Pre–post cohort study** $N = 19$, 3 months.	Manualized/tailored Ayurvedic dietary counseling, herbal formulation + standardized yoga, meditation, and breathwork.	**Primary:** Arterial pulse wave velocity. **Secondary:** Anthropometrics (BMI), blood pressure, blood lipids; medication usage.
Elder 2006, United States of America[105]	**Ayurvedic medicine** Type II diabetes	**Randomized controlled trial** Open label, comparative effectiveness design; usual care comparator. $N = 60$, 18 weeks, 6-month follow-up.	**Arm I:** Standardized Ayurvedic intervention (dietary counseling, meditation, herbal supplement) + individualized exercise. **Arm II:** Standard diabetes education classes + usual care.	**Primary:** Blood glucose. **Secondary:** Blood lipids; blood pressure; pulse; anthropometrics (weight); adherence; qualitative adherence barriers/facilitators, perceived benefits.
Elder 2018, United States of America[33,112–115]	**Chiropractic medicine** Chronic neck/back pain	**Controlled pre–post cohort study** Open label, comparative effectiveness design with propensity score matched controls. Associated cross-sectional survey and electronic medical record review[114]; qualitative substudy[115]. $N = 70$, $N = 139$ (control), 6 months. Interviews/focus groups: $N = 90$ (patients); $n = 25 + n = 14$ (health care providers).	**Arm I:** Individualized whole system chiropractic care. **Arm II:** Usual care (biomedical). **Qualitative Aims:** Initial—To explore patient/practitioner experiences with and decision-making around A/C use for chronic MSK pain; Iterative—To characterize practical issues faced by patients seeking alternatives to opioid chronic MSK pain management.	**Primary:** Symptom severity PROM. **Secondary:** Sleep quality PROM; Mental health PROMs; QoL PROM; direct health costs.
Flower 2012, United States of America[79]	**Chinese medicine** Urinary tract infection	**Prospective case series** Feasibility/pilot design. $N = 14$, 6 months.	Individualized herbal mixture + standardized ''acute'' herbal mixture.	**Primary:** Symptom severity PROM (unvalidated). **Secondary:** Symptom severity and wellness PROMs; medication usage.
Forster 2016, Australia[100,116]	**Midwifery** Patient satisfaction/C-section rates	**Randomized controlled trial** Comparative effectiveness design; usual care comparator. $N = 2314$, perinatal care +2 months postpartum.	**Arm I:** Manualized *caseload* ''continuous'' midwifery perinatal care. **Arm II:** Usual care (noncaseload midwifery, junior obstetric, or general practitioner).	**Primary:** Patient satisfaction PROM (unvalidated); cesarean birth. **Secondary:** Medication usage; instrumental/induced births; maternal perineal trauma; infant anthropometrics.

(*continued*)

Table 2. (Continued)

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Hamre 2007, Germany[70,117]	**Anthroposophic medicine** Low back pain	**Controlled Pre–post cohort study** Open label, preference allocated, comparative effectiveness design; usual care comparator. $N=62$, 12 months; 2-year follow-up.	**Arm I:** Individualized whole system anthroposophic care. **Arm II:** Usual care (biomedical).	**Primary:** Symptom severity PROM (symptom score); QoL PROM (SF-36).
Hamre 2013, Germany[71,107]	**Anthroposophic medicine** Chronic diseases	**Pre–Post Cohort Study** $N=1510$, 4 years.	Individualized whole system anthroposophic care.	**Primary:** Patient-generated symptom severity PROM (Symptom Score). **Secondary:** QoL PROM (SF-36); Mental health PROM; Patient satisfaction.
Hamre 2018, Germany[69]	**Anthroposophic medicine** Rheumatoid arthritis	**Controlled pre–post cohort study** Open label, preference allocated, comparative effectiveness design; enhanced usual care comparator. Associated synthetic overview of 21 related publications.[107] $N=251$, 4 years.	**Arm I:** Individualized whole system anthroposophic medicine + Usual care (corticosteroid and NSAIDs) **Arm II:** Enhanced usual care (corticosteroids, NSAIDs + DMARDs).	**Primary:** Symptom severity PROM; C-reactive blood protein; disease progression test (Ratingen Score).
Huang 2018, China[80,118]	**Chinese medicine** Bronchiectasis	**n-of-1 Series** Randomized, double-blind, six-period crossover comparative effectiveness (ABABAB) design. $N=17$, 3 phases (2×4 weeks); 3 week washouts.	**Intervention A:** Manualized/tailored herbal mixture. **Intervention B:** Standardized herbal mixture.	**Primary:** Patient-generated symptom severity PROM **Secondary:** Sputum volume; bloodwork for adverse event assessment.
Hullender Rubin 2015, United States of America[81]	**Chinese medicine** IVF outcomes	**Retrospective post-only study** Multiarm, comparative effectiveness design; usual care and intraparadigmatic comparators. $N=1231$.	**Arm I:** Individualized whole system Chinese medicine + IVF **Arm II:** Standardized acupuncture + IVF **Arm III:** Usual care (IVF alone)	**Primary:** Health event (Live birth). **Secondary:** Health events (biochemical pregnancy: singleton/twin/triplet/ectopic/aborted; gestational age)
Jackson 2006, United States of America[82]	**Chinese medicine** Tinnitus	**n-of-1 Series** Open label, two-period (AB) design. $N=6$, 14 days treatment, two (pre–post) 14 day evaluation phases,	**Period A:** Individualized traditional acupuncture, 10 treatments. **Period B:** Post-treatment control period.	**Primary:** Symptom severity PROM (unvalidated). **Secondary:** patient-generated PROM (MYMOP).
Joshi 2017, India[73]	**Ayurvedic medicine** Asthma	**Controlled pre–post cohort study** Multiarm proof-of-concept design; intraparadigmatic and healthy comparators. $N=115 + n=69$ (control), 6 months.	**Arms I + II:** Manualized/tailored whole system Ayurvedic medicine (**I:** vata-dominant asthma; **II:** kapha-dominant asthma.) **Arm III:** Healthy control group comparator.	**Primary:** Blood IgE levels, eosinophil counts, spirometry, cytokines; Lung function. **Secondary:** Symptom severity PROM.
Kessler 2015, Germany[74]	**Ayurvedic medicine** Infertility	**Single Case Report** $N=1$, 12 months.	Individualized whole system Ayurvedic medicine.	**Primary:** Health event (live birth); narrative case reporting.

(*continued*)

TABLE 2. (CONTINUED)

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Kessler 2018, Germany[75,119]	**Ayurvedic medicine** Knee osteoarthritis	**Randomized Controlled Trial** Open-label, comparative effectiveness design; usual care comparator. $N = 151$, 12 weeks; 6- and 12-week follow-up.	**Arm I:** Manualized/tailored whole system Ayurvedic medicine **Arm II:** Usual care, including occupational/manual therapy, home exercise, dietary counseling, and pain medication.	**Primary:** Symptom severity PROM. **Secondary:** QoL, mood, pain, sleep PROMS (SF-36); medication usage.
Litchke 2018, United States of America[104]	**Yoga therapy** Social/emotional skills in autism	**Pre–Post Cohort Study** Exploratory design. $N = 5$, 4 weeks.	Standardized biweekly yoga therapy instruction.	**Primary:** Psychosocial scales observer reported; Narrative notes.
McCulloch 2011, United States of America[83,120]	**Chinese medicine and Vitamin therapy** Lung/colon cancer survival	**Retrospective Controlled Post-Only Study** Matched control, comparative outcomes design with propensity score and marginal structural analytic methods; usual care comparator. Lung cancer: $N = 239 + n = 12{,}754$ (control); Colon cancer: $N = 193 + n = 13{,}665$ (control).	**Arm I:** Manualized/tailored whole system Chinese medicine + Vitamin therapy + Usual care (biomedical oncology) **Arm II:** Usual care (biomedical oncology)	**Primary:** Health event (Survival).
Mills 2016, United States of America[76]	**Ayurvedic medicine** Well-being	**Quasi-Randomized Controlled Trial** Open label, comparative effectiveness design; usual care comparator. $N = 69$, 6 days.	**Arm I:** Standardized Ayurvedic mind–body residential group program **Arm II:** Residential vacation at same site	**Primary:** Wellness PROMs for spirituality, gratitude, self-compassion, psychologic well-being, mental health. **Secondary:** Blood pressure; anthropometrics (height, weight).
Ornish 1998, United States of America[60,121]	**Preventive/restorative biomedicine** Cardiovascular rehabilitation	**Randomized Controlled Trial** Open label, comparative effectiveness design; usual care comparator. $N = 48$, 1 year; $n = 35$, 5-year follow-up.	**Arm I:** Standardized lifestyle program (diet, exercise, smoking cessation, stress management) with biweekly support group. **Arm II:** Usual care (biomedical).	**Primary:** Angiography (coronary artery lesion characteristics); blood lipids.
Paterson 2011, United Kingdom[84,122]	**Chinese medicine** Medically unexplained symptoms	**Randomized controlled trial** Open label design with waiting list crossover control. Qualitative substudy[122]. $N = 80$, 6 months, 1-year follow-up. $N = 20$ pre- and postinterviews.	**Arm I:** Individualized traditional acupuncture + usual care. **Arm II:** Wait list + usual care.	**Primary:** Patient-generated PROM (MYMOP). **Secondary:** Wellbeing and QoL PROMs; medication usage; health service utilization.
Perlman 2016, United States of America[41,101,123,124]	**Swedish massage therapy** Knee osteoarthritis	**Randomized controlled trial** Multiarm open label comparative effectiveness design; protocol development[41], dose-finding,[124] and qualitative[123] substudies. $N = 222$, 8 weeks, 52-week follow-up. $N = 18$ interviews.	**Arm I:** Manualized/tailored Swedish massage. **Arm II:** Standardized light touch bodywork. **Arm III:** Usual care (biomedical).	**Primary:** Symptom severity PROM. **Secondary:** Functional tests (walk test, range of motion); health expenditure; medication usage.

(*continued*)

TABLE 2. (CONTINUED)

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Rioux 2014, United States of America[29]	**Ayurvedic medicine** Obesity	**Pre–post cohort study** Feasibility/pilot design. *N* = 17, 3 months, 6- and 9-month follow-ups.	Manualized/tailored Ayurvedic diet/lifestyle counseling + standardized yoga therapy instruction and home practice.	**Primary:** Anthropometric (weight, BMI, body fat %, waist/hip circumference/ratio); Paradigm-specific PROMs (unvalidated); Psychosocial PROMs (Bandura); Adherence.
Ritenbaugh 2008, United States of America[85]	**Chinese and Naturopathic medicine** Temporomandibular disorders	**Randomized controlled trial** Multiarm, open label comparative effectiveness design; interparadigmatic and usual care comparators. *N* = 160, 6 months (Chinese medicine), 8 months (naturopathic medicine).	**Arm I:** Manualized/tailored whole systems Chinese medicine + standardized relaxation tapes. **Arm II:** Manualized/tailored naturopathic medicine + standardized nutritional supplement. **Arm III:** Specialty dental care; physical therapy/psychologic referrals.	**Primary:** Symptom severity PROM.
Ritenbaugh 2012, United States of America[86,125,126]	**Chinese medicine** Temporomandibular disorders	**Dynamically-allocated controlled trial** Open label, stepped care comparative effectiveness design; usual care comparator. *N* = 168, 1 year.	**Arm I:** Manualized/tailored whole systems Chinese medicine. **Arm II:** Psychosocial self-care education.	**Primary:** Symptom severity PROM. **Secondary:** QoL, mental health, wellbeing PROMS; Medication usage.
Seely 2013, Canada[95,127]	**Naturopathic medicine** Cardiovascular disease prevention	**Randomized controlled trial** Open label, comparative effectiveness design; usual care comparator. Economic substudy.[127] *N* = 246, 1 year.	**Arm I:** Manualized/tailored diet/exercise counseling, nutritional supplementation + Usual care (biomedical). **Arm II:** Usual care (biomedical). **Economic Evaluation:** Direct + indirect costs to employer, society.	**Primary:** Cardiovascular risk (Framingham algorithm) based on blood lipids and glucose, blood pressure, and anthropometrics (waist circumference). **Secondary:** QoL PROM (SF-36); patient generated PROM (MYMOP).
Shalom-Sharabi 2017, Israel[89]	**Complementary/integrative medicine** Cancer gastrointestinal/QoL	**Controlled pre–post study** Comparative effectiveness design, pragmatically- assigned control; usual care comparator. *N* = 175, 6 weeks.	**Arm I:** Individualized multidisciplinary complementary/integrative medicine + usual care (biomedical oncology) **Arm II:** Usual care (biomedical oncology)	**Primary:** QoL PROMs, including patient-generated PROM (MYCaW).
Silberman 2010, United States of America[98,128–130]	**Preventive/restorative biomedicine** Cardiovascular rehabilitation	**Pre–post cohort study** Time series design. *N* = 2974, 1 year.	Standardized diet/lifestyle and stress management program with psychosocial group support.	**Primary:** Functional testing (treadmill); blood lipids and sugar; mental health PROMs; adherence.
Sutherland 2009, United States of America[36]	**Energy medicine** Chronic headaches	**Pre–post cohort study** Qualitative design *N* = 13 interviews, 3–7 weeks, 3-month follow-up.	Individualized healing touch sessions.	**Qualitative outcomes:** Interview for Symptom severity; wellness/QoL.

(*continued*)

Table 2. (Continued)

Study/location	*Paradigm/focus*	*Design*	*Interventions*	*Outcomes*
Szczurko 2007, Canada[96,131]	**Naturopathic medicine** Chronic low back pain	**Randomized controlled trial** Open label, comparative effectiveness design with optional cross-over; usual care comparator. Economic substudy.[131] $N = 75$, ≥ 8 weeks.	**Arm I:** Standardized acupuncture, deep breathing, diet/exercise counseling. **Arm II:** Educational physiotherapy booklet + live exercise/relaxation instruction. **Economic evaluation:** Direct + indirect costs to individual, employer, society.	**Primary:** Symptom severity PROM (Oswestry); QoL PROM (SF-36). **Secondary:** Symptom severity PROM; anthropometrics (BMI, weight); functional testing (range of motion); medication usage; health service utilization; adherence; work absenteeism.
Wang 2016, United States of America[103]	***t'ai chi*** Knee osteoarthritis	**Randomized controlled trial** Open label, comparative effectiveness design; usual care comparator. $N = 204$, 12 weeks.	**Arm I:** Standardized *t'ai chi* group instruction + home practice. **Arm II:** Usual care (physical therapy).	**Primary:** Symptom severity PROM. **Secondary:** Mental health, QoL, and self-efficacy PROMs (SF-36); functional testing (walk); outcome expectation PROM.
Wayne 2018, United States of America[90]	**Complementary/ integrative medicine** Low back pain	**Controlled pre–post study** Open label, preference allocated, comparative effectiveness design; usual care comparator. $N = 309$, 12 months.	**Arm I:** Individualized multidisciplinary complementary/ integrative medicine plus usual care (biomedical) **Arm II:** Usual care (biomedical)	**Primary:** Symptom severity PROM. **Secondary:** QoL PROM; health service utilization; medication usage; treatment satisfaction.
Witt 2015, Germany[91]	**Complementary/ integrative medicine** Breast cancer QoL	**Randomized controlled trial** Open label, comparative effectiveness design; usual care comparator. $N = 275$, 6 months.	**Arm I:** Individualized, complex complementary/integrative medicine + Usual care (biomedical oncology). **Arm II:** Usual care (biomedical oncology).	**Primary:** QoL PROM; patient-generated PROM (unvalidated); health event (survival); medication usage; health service utilization; treatment satisfaction.
Welch 2013, United Kingdom[102]	**Swedish massage therapy** Integrative care dynamics	**Ethnography** Qualitative-dominant, action research design drawing on perspectives from clinicians, patients, and staff in an integrative care clinic. $N = 11$ (physicians); $n = 33$ (staff); $n = 22$ (patients); $n = 1$ (massage therapist); 12–13 weeks.	Individualized Swedish massage therapy. **Qualitative aims:** To evaluate contextual factors at play in an integrative care setting, with attention to the interfacing of clinicians from multiple disciplines.	Mixed-methods questionnaires; interviews; field notes; reflexive journals.
Zeng 2013, United States of America[99]	**Preventive/restorative biomedicine** Cardiovascular rehabilitation	**Controlled pre–post cohort study** Multiarm, open label comparative outcomes design with matched control comparator and economic evaluation. $N = 461 + 1796$ (control); 1 year +3 year follow-up.	**Arm I:** Standardized diet, lifestyle, and stress reduction program (Ornish) with group component **Arm II:** Standardized diet, lifestyle, and stress reduction program (Benson-Henry) with group component **Arms III and IV (matched pairs):** Traditional cardiovascular rehabilitation; No cardiac rehabilitation.	**Primary:** Health service utilization (hospitalization); health event (mortality); direct costs (institutional).

Headings bolded for emphasis.

BMI, body mass index; DMARD, disease-modifying anti-rheumatic drug; IgE, immunoglobulin E; IVF, in vitro fertilization; MRI, magnetic resonance imaging; MYCaW, Measure Yourself Concerns and Wellbeing; MYMOP, Measure Yourself Medical Outcome Profile; NSAID, nonsteroidal anti-inflammatory drug; PROM, patient-reported outcome measure; QoL, Quality of life; SF-36, Short-Form 36; VAS, visual analog scale.

Some studies[33,71,92,101] feature multiple associated publications; a synthesis article[107] related to one mixed methods study in particular[71] details 21 inter-related peer-reviewed publications.

What follows is a synthetic analytic report of the major methodological features of the reviewed WSR studies, presented in three parts. Part I (Study Design) addresses the primary methodological modes selected by whole systems researchers. Part II (Interventions) reviews the main characteristics of and strategies used in defining WSR interventions across the reviewed exemplars. Part III (Outcome Assessment) elaborates the range of approaches to outcome assessment adopted in each of the WSR exemplars and across the field as a whole. At the end of each of these three sections, findings are discussed with reference to the model validity principle and with a view to practical considerations relevant for researchers in the WSR field. A subsequent Discussion/Conclusion segment synthetically integrates findings from all three sections, positioning them in a broader health systems context.

Table 2 provides a detailed overview of the 41 reviewed studies' methodological features. Additional Tables and Figures are used throughout this review to detail and summarize findings. Where data are clearly represented with citations in Tables and/or Figures, a note to this effect is made in the review text; direct in-text citations are provided for more detailed findings not represented in graphical form.

Part I: study design

The reviewed WSR studies engage a cross-section of prospective and retrospective study types, including various controlled and uncontrolled, experimental, quasi-experimental, and observational designs (Figure 3). Figure 4 presents a detailed overview, by study, of major research design features and will be repeatedly cited in the text to assist readers in identifying exemplars with particular characteristics. As elaborated in what follows and is summarized in Figure 5, open label, prospective comparative effectiveness designs with usual care comparators and randomized allocation represent the most common WSR approach; placebo controls and double blinding are rarely applied in the reviewed studies. On the whole, quantitative methods dominate across almost all reviewed exemplars. That said, one-third are mixed methods studies, most of which incorporate qualitative methods (Fig. 6), and a few with economic evaluations (Fig. 7).

Comparative/controlled trials. Twenty-seven reviewed studies, including two with retrospective designs, involve interventions whose clinical outcomes are contrasted head-to-head with at least one control/comparator arm (most often ''usual care''). Seven of these trials have three or more arms (Fig. 4). Pre–post designs are evident in all prospective studies, whereas the two retrospective studies evaluate postoutcomes only. As elaborated in what follows, the reviewed comparative/controlled studies implement various statistical and pragmatic approaches to participant allocation, use controls/comparators that are largely active/positive, and—while typically open label—apply assessor blinding methods in several cases.

Statistical allocation. Of the 27 evaluated controlled studies, 16 engage statistical approaches in allocating

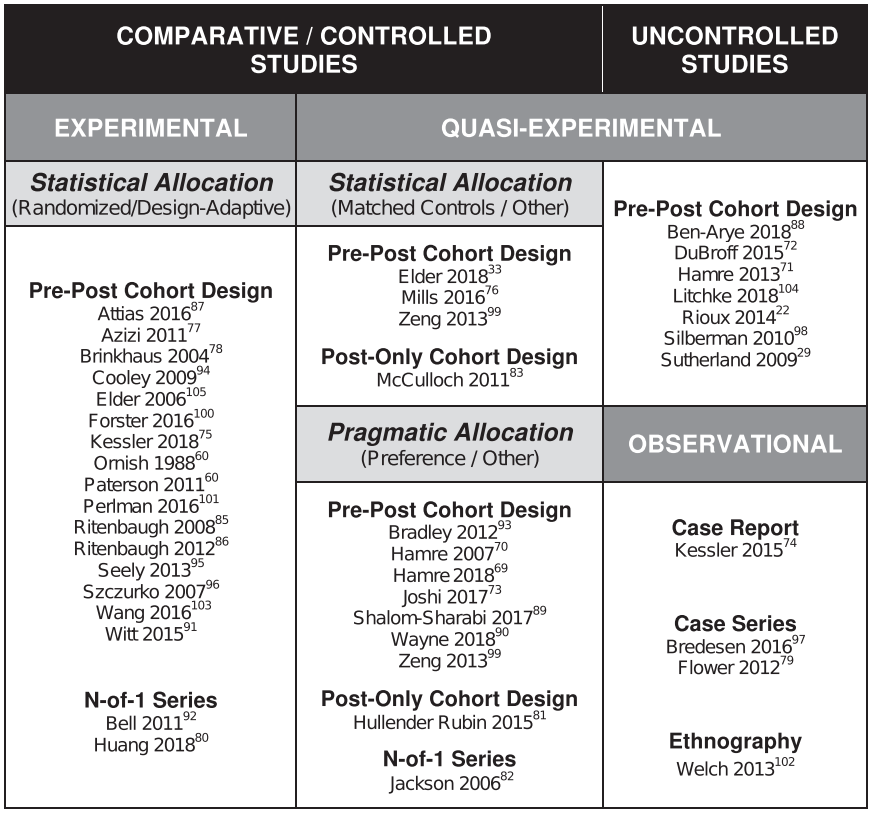

FIG. 3. Typology of whole systems research designs.

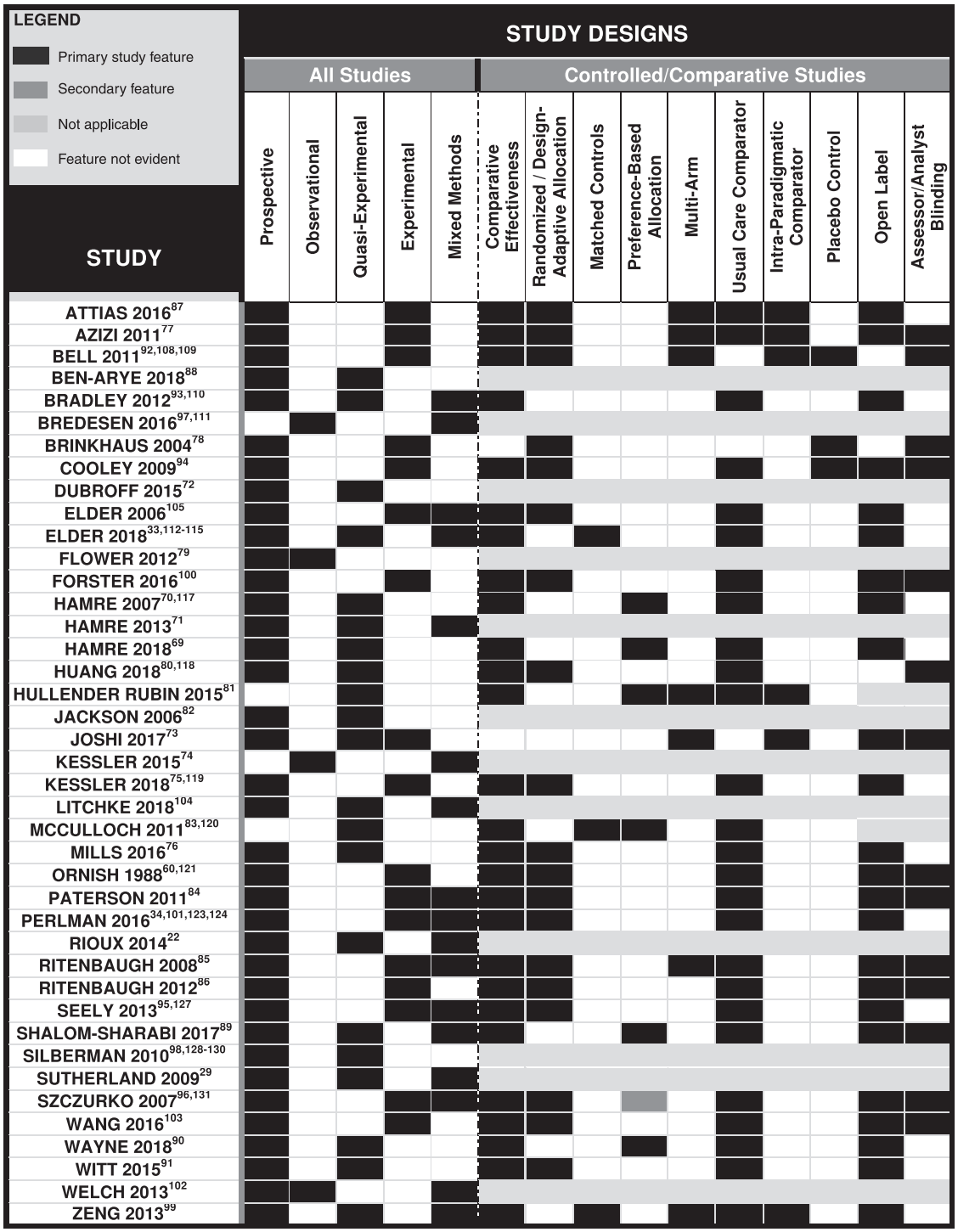

FIG. 4. Study designs in whole systems research.

patients to particular treatment arms. Randomization is the dominant approach, although some studies use design-adaptive allocations (e.g., *minimization*) or matched control designs (Fig. 3).

Simple randomization[76–78,84,91,94,96,100] ($n=8$; $n=4$ with demographic stratification[78,91,94,100]) and *block randomization*[75,87,95,103] ($n=4$; $n=2$ with stratification[75,95]) are at times applied alongside additional elements. Szczurko et al.'s simply randomized study, for instance, implements an optional, preference-based crossover.[96] Attias et al.'s six-armed study uses block randomization to first allocate for individualized versus standardized care, subsequently assigning intervention-arm patients to receive a particular complementary care approach based on the clinician type scheduled to work on the week day of their scheduled surgery.[87]

Two randomized trials provide no additional details on their allocation designs,[101,105] although a few others use distinctive randomization variants. Ornish et al. engage a *randomized*

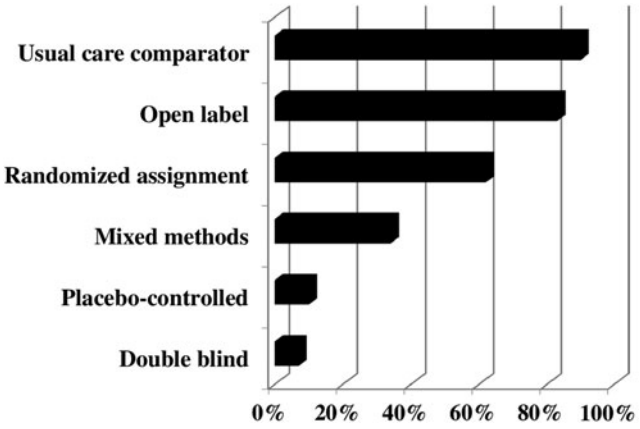

FIG. 5. Controlled/comparative whole systems research designs.

invitational design[60] aimed at reducing disappointment-related attrition, by asking participants to ''agree to be tested'' without being advance-apprised of the active intervention's specifics.[132] After using simple randomization to assign the first 20 (of 80) participants, Paterson et al. use *minimization*—a design adaptive allocation strategy—to allocate the remaining patients.[84] Ritenbaugh et al.'s study also applies a design adaptive randomization approach, with reference to several balancing factors.[85]

In a nonrandomized, design-adaptive approach, Ritenbaugh et al.'s study[86] implements a *stepped care* (triaged) method, using *minimization* to dynamically allocate those with the most severe symptoms. The researchers automatically assign those with lesser symptoms to standard care and, after a period of treatment, reassign standard care recipients with continued ''substantial pain'' either to intervention or control.

Three studies use nonrandomized, statistical allocation methods to create *matched control* groups—composed of two[33] or more[83,99] similar concurrent controls per intervention patient—from electronic medical records (EMRs). Two of these studies also apply *propensity score* methods,[33,83] in one case in a particularly innovative manner[114] and in the other alongside additional statistical methods (including *marginal structural models*) to further adjust for confounding.[83]

Pragmatic allocation. Of the 10 controlled trials that use (primarily) nonstatistical allocation strategies, four[69,89,90] are prospective *patient preference trials*, in which similar-sized groups of intervention and control patients concurrently select their favored treatments (Fig. 4). Two such studies use EMRs, in one case to recruit intervention and control arm patients[89] and in the other for a comparator group alone.[90] In the other two studies,[69,70] both led by Hamre et al.,[69] patients self-select to begin condition-specific care in anthroposophic and conventional care clinics, respectively.

In Hullender Rubin et al.'s retrospective preference-based study, patients in three arms receive *in vitro* fertilization (IVF) alone, IVF plus same-day acupuncture, or IVF plus whole systems Chinese medicine, respectively.[81] Study analysts quasi-experimentally ''adjust for covariates …[via]

LEGEND: Usage / Non-usage	DATA GENERATION						ANALYTIC APPROACH			
STUDY	Clinical Interactions	Questionnaire Items	Focus Groups	Participant Interviews	Ethnography	Participant Observation	Clinical Analysis	Thematic / Content Analysis	Thematic Frequencies	Quantitative Triangulation
ALI 2017[123,a]				■				■		■
BRESEDEN[97,111,b]	■						■			■
ELDER 2006[105,c]		■						■	■	
KESSLER 2015[74,b]	■						■			■
LITCHKE 2011[104,c]						■		■		■
OBERG 2012[110,d]			■	■				■	■	
PATERSON 2011[84,e],RUGG 2011[122,e]		■		■			■	■		■
PENNEY 2016[115,f]			■	■				■		
RIOUX 2014[29,c]	■			■			■	■	■	■
RITENBAUGH 2008[85,c]		■								
EAVES 2014[126,g]				■	■			■		
SHALOM-SHARABI 2017[89,c]		■								
SUTHERLAND 2009[36,b]				■				■	■	
WELCH 2013[102,b]	■	■		■	■	■		■		

Stand-alone qualitative publication linked to: [a]Perlman 2012[101]; [d]Bradley 2012[93]; [f]Elder 2018[33], [g]Ritenbaugh 2012[86].
[b]Qualitative methods as dominant research approach.
[c]Qualitative methods and analysis embedded in quantitative clinical outcomes publication.
[e]Qualitative results dually presented in stand-alone qualitative and quantitative papers.

FIG. 6. Qualitative methods in whole systems research.

	ECONOMIC PERSPECTIVES			COSTS MEASURED	
STUDY	Individual	Institutional	Societal	Direct	Indirect
ELDER 2018[33,a]		■		■	
HAMRE 2006[106,b]		■	■	■	■
HERMAN 2008[131,c]	■	■	■	■	■
HERMAN 2014[127,d]		■	■	■	■
ZENG 2013[99,a]		■		■	

[a]Economic evaluation embedded in clinical outcomes publication.
Stand-alone publication linked to: [b]Hamre 2013[71]; [c]Szczurko 2007[96]; [d]Seely 2013[95].

FIG. 7. Economic evaluations in whole systems research.

multivariable logistic regression analysis'' to ''minimize potential bias'' related to baseline intergroup differences. Bradley et al.'s prospective trial—conversely marked by intervention patients' *lack* of experience with or preference for naturopathic medicine ($n=40$)—uses EMRs to assemble a substantially-larger ($n=329$), demographically-similar (quasi-matched) control group.[93] Finally, Joshi et al. contrast intervention outcomes with those from a similarly-sized, demographically-similar *healthy control* group from the ''general population.''[73]

Positive controls. All but one[78] of the controlled studies reviewed have active (positive) comparator groups, in almost all cases with a *usual care* arm (Fig. 4). Several among these[73,74,76,86,94,99,103] engage complex, individualized *time-attention* controls. For instance, Kessler et al.'s osteoarthritis usual care control mirrors the study's multimodal primary Ayurvedic intervention with an equivalent number of individualized physiotherapy sessions paired with home exercises, dietary counseling, and medication.[75] One study, conversely, has notably low time-attention matching (validated educational booklet vs. multimodal naturopathic intervention).[96] Some usual care comparators are innovative (e.g., a residential ''vacation'' to control for a mind–body retreat[76]). Others more simply represent *real-world* usual care (e.g., conventional obstetric care compared with a caseload midwifery intervention[100]). Several studies' primary interventions, reflecting the normative context of biomedical care, are furthermore designed as *adjunctive* to control, that is, they include the same usual care as received by the comparator group[81,83,84,87,89–91,93–95,119] (e.g., complementary/integrative cancer care that includes conventional treatment[83,89,91]).

One usual care-controlled study engages a crossover, *waiting list* controlled design.[84] Those with multiple intervention arms[73,77,81,85,87,92,99] almost universally implement *intra-paradigmatic*, *factorial* comparator group designs (Fig. 3) to trial a subset of or variation upon the primary intervention (e.g., herbal mixture vs. herbal mixture plus acupuncture[77]). Finally, the single actively controlled study without a usual care comparator includes two distinct (nonfactorial) *intra-paradigmatic* intervention arms plus an untreated *healthy control* group.[73]

Placebo/sham controls. Just two of the reviewed cohort-based controlled studies apply placebo and/or sham controls (Fig. 4). Cooley engages a multimodal naturopathic medicine design in which a multivitamin placebo forms part of a complex, open-label, active usual care comparator.[94] Brinkhaus et al. trial *verum* versus *sham* acupuncture and an *active* versus *nonspecific* herbal mixture.[78]

Blinding. Almost all controlled cohort-based WSR studies have *open label* designs in which both patients and interventionists are alert to participants' treatment allocations; assessor/analyst blinding is however almost universally applied in these same studies (Fig. 4). Brinkhaus et al.'s randomized, placebo-controlled study is the only cohort-based study that implements full patient blinding.[78]

Uncontrolled studies. Seven of the reviewed studies apply prospective, uncontrolled cohort designs. Four such studies are relatively small, with fewer than 20 participants[29,36,72,104]; 2 are notably large, with well over 1000 patients each.[71,98] Aside from the absence of comparator arms, most of these studies do not differ substantially in intervention or outcome design from the comparative/controlled trials discussed above. That said, a few have distinct methodological features. Silberman et al.'s study uniquely adopts a time series design to evaluate outcomes at (and between) intervals[98]; Sutherland et al.'s study uses qualitative interviews (rather than quantitative measures) as its primary data generation approach[36]; and Ben-Arye et al.'s study derives most outcomes prospectively from patient charts[88] rather than using quantitative outcome measures alone.

n-of-1 series. Of the three n-of-1 series trials reviewed, all represent adaptations of the classical n-of-1 single patient crossover design. Huang et al.'s three-phase (ABABAB) comparative effectiveness design evaluates individualized (A) versus standardized (B) Chinese herbal mixtures for bronchiectasis,[80] using randomization to determine the order of treatment versus control in each phase, between washouts. Bell et al.'s placebo-controlled, dynamically allocated, two-phase (AB) design (A: placebo, B: treatment) comparatively trials two different homeopathic insomnia remedies with intermittent washouts.[92] Both of these studies report patient blinding, with clinician blinding additionally applied by Huang. Jackson et al.'s acupuncture/tinnitus trial by contrast uses a quasi-experimental, open label two-period (AB) design (A: treatment, B: no treatment), reporting individual and combined outcomes from two-week pre- and postintervention measurement periods.[82]

Case study and case series. This review includes one case study and two case series, each of which presents a detailed narrative account of the effects of a particular complex

treatment approach on specific individuals. Like Kessler et al.'s single case study,[75] Bredesen et al.'s case series[97] is retrospective, detailing exceptional clinical outcomes from a particular whole systems intervention. Kessler et al.'s[75] study provides considerable detail about the Ayurvedic treatment approach applied, well beyond the level of detail given in cohort-based studies within the same paradigm. Flower and Lewith's uniquely prospective case series,[79] designed as a preliminary clinical outcomes trial, tracks common Chinese medicine diagnostic patterns and other informative participant data to inform future study designs.

Ethnography. One reviewed study applies ethnographic methods (e.g., participant observation, interview, questionnaire, and so on) within an action research framework to equally give voice to the perspectives of patients, clinicians, and staff, while also reporting clinical outcomes for a Swedish massage therapy intervention.[102]

Mixed methods designs. Seventeen reviewed studies engage mixed methods research designs, either incorporating qualitative alongside quantitative methods ($n=13$), economic evaluations alongside clinical outcomes ($n=5$), or in one complex design,[33] both.

Qualitative methods. As shown in Figure 6, the 14 studies incorporating qualitative methods use open-ended questionnaire items, focus groups, and/or participant interviews to investigate qualitative questions relating to treatment outcomes,[29,84,89,105,123] treatment choices,[115] patient experiences,[86,110,122] and protocol compliance.[105] One study also engages participant observation to document outcomes.[104] Content analysis, with multianalyst corroboration of thematic results, represents the most common qualitative analytic approach, at times with numeric frequency calculations and/or quantitative corroboration. Most studies present ''thick descriptive'' results, using narrative and/or table-based formats, and report their qualitative findings either in stand-alone publications or alongside quantitative results in mixed-methods clinical outcome articles.

In just four studies, qualitative methods dominate. Kessler et al.'s case study[74] and Bredesen et al.'s case series[97,111] provide narrative accounts of specific patients' therapeutic trajectories, secondarily referring to quantitative data. Sutherland et al. uses in-depth interviews to explore clinical and methodological questions relating to a healing touch intervention.[36] Welch et al.'s ethnographic study uses multiple qualitative methods to study stakeholder perspectives and outcomes in an integrative medicine setting.[102] The remaining nine studies deploy qualitative methods secondarily; two do not report their qualitative results.[85,89]

The subordination of qualitative to quantitative methods across most studies might initially appear to convey a positivist or post-positivist orientation[133] consistent with the general ethos of biomedical clinical research. That said, many studies use inductive data analytic approaches within their qualitative subcomponents, suggesting a pragmatic approach to mixed methods analysis that accommodates constructivist perspectives.[133] In Ritenbaugh et al.'s study,[86] for instance, study participants were repeatedly interviewed, in an ethnographic mode, over a year-long period. The researchers' initial intention to ''relate…qualitative narratives to quantitative data on outcomes'' was ultimately abandoned in light of the ''complexity of participants' [narratives, which]…precluded a simplistic comparison between these two disparate types of data.''[126]

Economic evaluations. Of the five reviewed studies that include economic evaluations (Fig. 7), two report their economic results within quantitative clinical outcomes articles and three in stand-alone publications. All report on direct *institutional* expenditures associated with the interventions (vs. comparators) under study; the specified institutions include the public purse,[99,106] a corporate employer,[127,131] and a nonprofit health maintenance organization.[105] Those economic evaluations published as stand-alone publications[106,127,131] additionally address indirect health-related costs (e.g., work absenteeism and health related QoL) and report from multiple expenditure vantage points beyond the institutional (e.g., individual, societal/total).

Model validity and practical considerations in WSR design selection. Overall, the research designs selected by whole systems researchers are similar to those used by biomedical researchers, at times with minor adaptations to enhance their model validity. It is unclear whether this emphasis on conventional *paradigm compatible* (and to a lesser degree *paradigm consistent*) designs reflects these researchers' preferences or is perhaps conversely indicative of the available financial support. Regardless, novel (i.e., *paradigm specific*) designs are—on the whole—not evident among the reviewed exemplars. That said, just one of the reviewed cohort-based studies follows the classical RCT model in its concurrent use of randomized allocation, participant and clinical blinding, and placebo controls. Echoing early WSR critiques of the classical RCT's model validity in TCIM contexts, Brinkhaus et al. explicitly recognize that neither of their study's two adopted placebos is ''entirely inactive.''[78]

All other controlled cohort-based studies elect to either: (1) select among a set of established, modified RCT, or non-RCT research designs (that demonstrate greater paradigm compatibility than the classical RCT) or (2) implemented adaptations of such conventional study designs (to render them more paradigm consistent). Open label designs appear preferable, although assessor/analyst blinding does not appear to compromise paradigm compatibility. On the whole, comparative effectiveness designs with active ''usual care'' comparators show strong paradigm compatibility in WSR contexts. Multiarm designs with intra-paradigmatic, factorial comparators also appear useful for comparing whole/complex versus singular/isolated TCIM practices. Randomization remains the most common allocation approach across controlled WSR studies, with some form of stratification applied in most cases to increase balance. The allocation alternatives engaged in a few studies (e.g., matched controls; preference-based allocation; and design adaptive assignment) are neither novel nor *uniquely* designed to address paradigmatic considerations in the TCIM field; however, they each appear to have distinct advantages (and potential disadvantages) for the WSR researcher.

Matched control designs, used in three studies, have the potential to produce results with internal validity similar to randomized trials at a lower cost; and as McCulloch

et al.'s studies[83,120] demonstrate may be fruitfully used in retrospective designs that rely on existing patient data. Preference-based designs explicitly recognize patients' differential choices of TCIM versus biomedical treatments, strengthening studies' external validity. In one such study, baseline demographic characteristics differed significantly between preference-allocated cohorts, compromising internal validity.[90] This was however not the case in other preference-allocated WSR studies reviewed, two[69,81] of which designed specific strategies to prevent such confounding.

As exemplified in Ritenbaugh et al.'s trial, design adaptive allocation (such as minimization) may match or exceed randomization's rigor while permitting implementation of innovative experimental frameworks (e.g., ''stepped care''[86]). Design adaptive assignment is furthermore cost-effective and ''socially responsible,'' using ''the smallest possible number of study participants to reach definitive conclusions about therapeutic benefits and harms.''[21] However, as Aickin notes, ''the cultural bias in favor of randomization will probably outlast the failure to defend it on rational grounds.''[21] Researchers may thus be challenged to access funding for such designs, which may moreover be excluded from ''meta-analyses and structured evidence reviews.'' As such, there may remain ''a good argument…for employing design-adaptation with a ''randomization'' feature'' in WSR studies,[21] such as in two of the reviewed exemplars.[84,85]

Uncontrolled, quasi-experimental pre–post designs, both large and small, do not differ significantly from the controlled trials aside from the absence of comparator arms. Such paradigm-compatible designs may be more cost-effective than controlled studies, particularly when based in existing clinical settings, and generate pragmatic outcomes while exploring controlled trial feasibility. Scaled versions of such studies, exemplified by Hamre et al.'s anthroposophic chronic disease trial,[71] may themselves generate valuable effectiveness data. Large retrospective comparative designs have similar evidentiary potential, whether reliant on concurrent active control groups[81] or electronically matched cohorts.[83]

Adaptations to increase conventional study designs' paradigm consistency are evident in the three n-of-1 trials reviewed. Conventional n-of-1 designs, study authors observe,[80,82] readily accommodate interventions geared to rapidly palliating symptoms, but they fail to account for progressive onset and extended carryover of treatment effects associated with TCIM whole systems' emphasis on root causes. TCIM researchers may thus prudently consider n-of-1 design adaptations, a point scant raised in previous related literature.[134,135] Huang et al.'s *actively*-controlled n-of-1 design furthermore addresses challenges in recruiting patients to *placebo*-controlled trials in the Chinese national context[80]; ethno-culturally situated considerations such as these warrant greater attention by WSR scholars, given TCIM's globalized context.

Case series and case studies remain important WSR designs in their more explicit detailing of paradigm-specific treatment considerations than is generally evident in other study types. Like n-of-1 trials, they may draw attention to TCIM therapies' potential when ''usual care'' falls short[74] and to understudied interventions with significant outcomes.[97] As Flower and Lewith's study furthermore suggests, prospective case series may serve as feasibility models for larger trial designs.[79]

As proposed by early WSR advocates, mixed methods designs significantly increase studies' paradigm consistency. Qualitative methods across the reviewed exemplars amplify participant and clinician perspectives and suggest parameters for better outcome assessment tools. However, in light of many TCIM whole systems' qualitative underpinnings, the dominance of quantitative methods across most WSR studies reinforces the biomedically-dominant contexts in which TCIM researchers seek model validity in their research designs.

Although some early RCT critics (e.g., Heron[7]) had proposed participatory, ethnographic designs as optimal modes of TCIM research, the ethnographic research modes adopted in just one study[102] (and suggested in two others[29,126]) are indeed unusual in biomedical clinical research contexts. These studies move boldly from *paradigm consistency* toward *paradigm specificity*, and their methodological propositions warrant careful attention. How ethnographically-informed hybrid designs may fruitfully enrich established clinical research approaches remains to be seen, as whole systems researchers carefully balance the pursuit of model validity with funding limitations and their own resistance of TCIM's biomedical co-optation.

Part II: interventions

This review undertakes a granular approach to analyzing the features of WSR interventions, both in terms of their general traits and in terms of the diagnostic and individualization strategies engaged. As summarized in Figure 8 and detailed in Figure 9, interventions across the reviewed studies are typically complex, multimorbid or multitarget in focus, behaviorally-focused, and in some cases multidisciplinary. About half of the reviewed studies implement dual (multiparadigmatic) diagnoses; and most treatments involve some form of individualization, representing a range of approaches across the individualization spectrum.

Complex interventions. All but one[92] of the reviewed WSR studies implement complex (i.e., multimodal and/or multicomponent) interventions (Fig. 9); treatments delivered within particular paradigms in some cases exhibit distinct traits. In all studies of anthroposophic, Ayurvedic, and

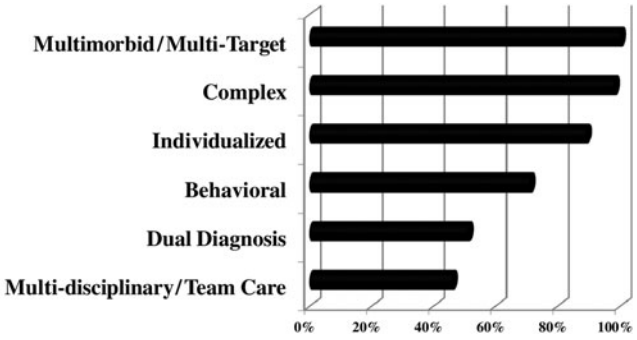

FIG. 8. Primary features of whole systems research interventions.

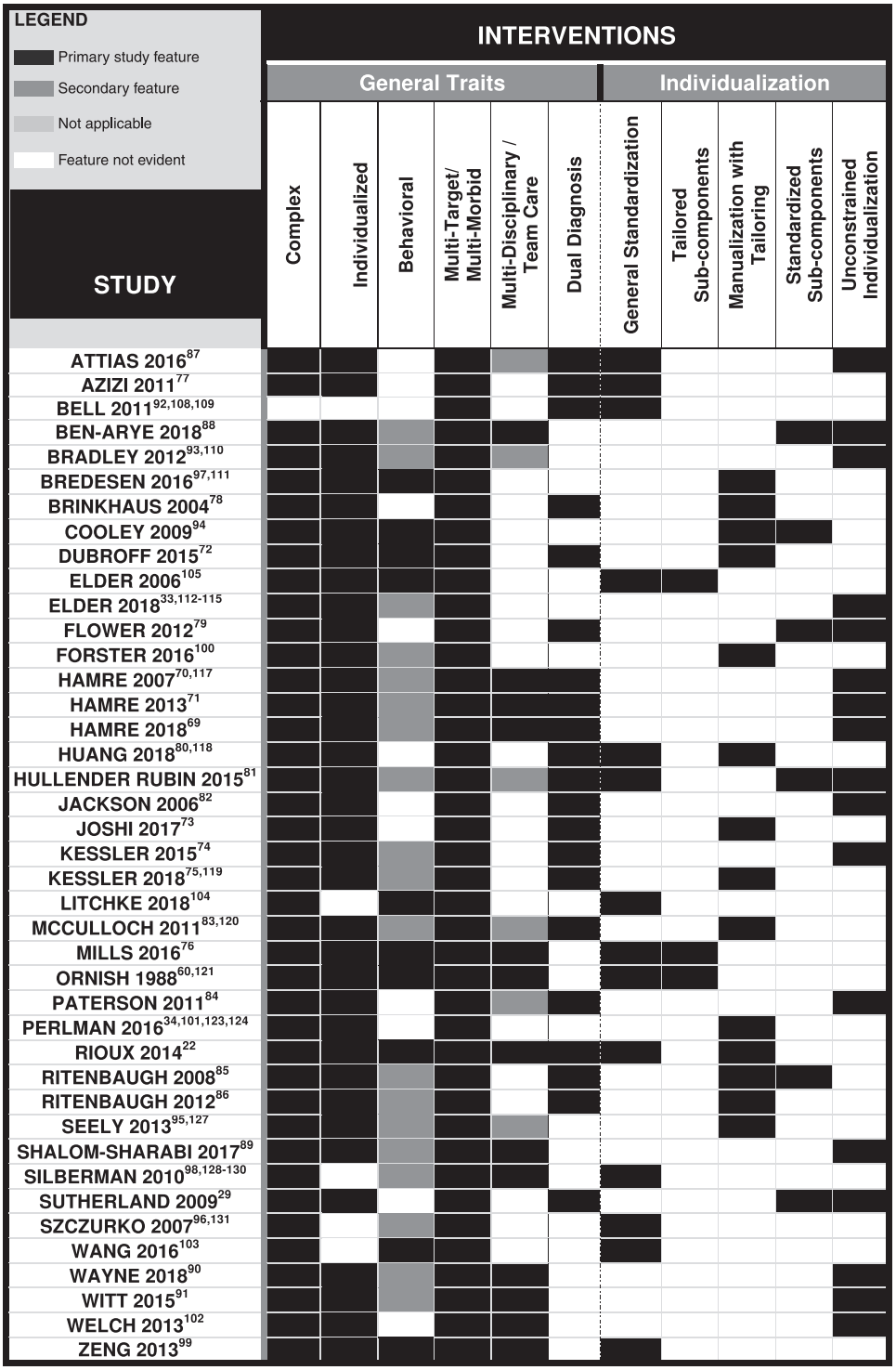

FIG. 9. Interventions in whole systems research.

naturopathic medicine care, and in almost all Chinese medicine studies, interventions reflect the full range of multimodal treatments that typify these paradigms (Tables 1 and 2). All studies reporting on complementary/integrative medicine interventions include ''usual'' biomedical care as an adjunct to treatment from at least one additional whole systems paradigm. Participants assigned to preventive/restorative biomedical study interventions all received combined instruction or counseling in nutrition, exercise, and stress-reduction practices.

Three Chinese medicine studies are not clearly multimodal in character, but their treatments include multiple components (e.g., acupuncture *with* moxibustion[82]; *multi*-herb mixtures[79,80]). Multicomponent interventions are also evident in studies centralizing manual therapies (e.g., multiple types of chiropractic adjustments[33] or various massage techniques[41,101]), as well as movement-based therapies (e.g., yogic poses + breathwork + visualizations[104]; multimovement, *t'ai chi* series that concurrently target ''physical function, balance, and muscle strength''[103]). Midwifery care

in Forster et al.'s study includes pre-, intra-, and postpartum care components.[100]

Behavioral interventions. Behavioral interventions (Fig. 9)—designed to facilitate patient implementation of salutogenic or preventive activities in their own lives—feature in a significant majority of all studies reviewed. About a quarter of these studies centralize behavioral approaches—such as diet, exercise, stress management, mind–body practices, and/or movement-based therapies—as *primary* intervention(s),[29,60,72,76,94,97–99,103–105] at times alongside a standardized nutritional supplement or herbal product.[94,105] Such behavioral interventions are either delivered in a group setting,[29,76,98,99,103,104,121] one-on-one with a clinician,[72,94,97,105] or both.[29] In another group of studies, similar types of behavioral interventions are delivered *secondarily* as part of an individualized whole systems treatment package constituted within the paradigms of anthroposophic,[69–71] Ayurvedic,[74,75] Chinese,[81,83–86] chiropractic,[33] complementary/integrative,[88,90] or naturopathic[93,95,96] medicine.

Individualization. The vast majority of interventions in the studies reviewed—whether prospective or retrospective—include some form of individualized treatment (Fig. 9).

Generally-standardized designs are evident in all six reviewed studies involving group-based interventions,[76,98,99,103,104,121] regardless of paradigm, as well as in several nongroup based studies.[77,80,87,92,96,105] Exemplars whose interventions are distinguished by their *unconstrained individualization* include all of the reviewed anthroposophic[69–71] and complementary/integrative[87–91] medicine trials, each of which also involves team care; the single chiropractic[33] and energy medicine[36] studies, one[93] (of five) naturopathic, one[74] (of eight) Ayurvedic, and four[79,81,82,84] (of ten) Chinese medicine trials analyzed. *Manualized/tailored* studies include three (of five) naturopathic,[85,94–96] five (of ten) Chinese medicine,[78,80,83,85,86] and five (of seven) Ayurvedic[29,72,73,75,105] trials, as well as the single massage therapy[101] and midwifery[100] studies reviewed. Some studies falling generally under one category's auspices concurrently include features of another[29,36,77,79,87,88,92,105] (i.e., tailored/standardized subcomponents). Flower and Lewith's Chinese medicine study, for instance, delivers a standard herbal formulation for participants' ''acute'' urinary tract infection usage, alongside individualized (patient-specific) ''preventative'' herbal formulations.[79] Rioux et al.'s study[29] similarly implements a standardized yoga therapy component alongside manualized/tailored Ayurvedic diet and lifestyle counseling.

Manualized protocol development—as exemplified in Ali et al.'s stand-alone publication[41]—generally occurred across studies through expert consensus, informed by paradigm-specific and peer-reviewed literatures. Various manualization approaches are moreover evident among the reviewed exemplars. Ritenbaugh et al.'s trial, for instance, defines acupuncture point lists and ''base herbal formulas'' for each of 12 Chinese medicine diagnostic categories, furthermore articulating optional subsets of pattern-specific acupoints and herbal additions for tailoring.[86] Cooley et al.'s naturopathic study, by contrast, more simply elaborates a set of predefined parameters for tailored diet and lifestyle counseling.[94]

Dual diagnosis. In 21 of the 41 reviewed studies, patients are diagnosed both from a biomedical perspective and from within another paradigm(s) (Fig. 9). These include each of the homeopathic and energy medicine studies, 5 of 7 Ayurvedic, all 3 anthroposophic, 1 of 5 complementary/integrative, 1 of 4 naturopathic, and 8 of 10 Chinese medicine studies reviewed.

In 7 of these 21 studies, little detail is provided beyond a general indication that multiparadigmatic diagnostics have taken place.[36,69–71,75,81,84] For instance, Hullender Rubin et al. note that each ''patient was assessed according to TCM [Traditional Chinese Medicine] theory,'' providing the basis for a ''detailed WS [whole systems]-TCM treatment plan.''[81] Similarly, Hamre et al. refer to a set of anthroposophy-specific principles (''formative force systems''), a paradigm-specific ''constitutional'' diagnostic process, and a set of distinct anthroposophic ''medications and nonmedication therapies,'' but do not detail the specific anthroposophic diagnoses made for study patients.

The remaining 14 of the 21 identified dual diagnosis studies explicitly identify the *primary* paradigm-specific diagnoses given to participants. All patients in Jackson's Chinese medicine study are, for example, ''diagnosed with a mixture of two predominant syndromes: Liver Qi Stagnation and Kidney Deficiency.'' In each of these 14 studies, patient treatments are individualized on the basis of paradigm-specific diagnoses. A few moreover detail (typically in table- or appendix format) specific treatment protocols related to such diagnoses.[29,72,78,80,83,85,86] Brinkhaus et al., for instance, delineate a core set of acupuncture points and base herbal formulation for all study patients, specifying additional points and herbal additions for each of five specific Chinese medicine diagnoses.[78] In four cases,[72,73,79,80] study authors furthermore provide a detailed breakdown of *all* patients' paradigm-specific diagnoses. Study inclusion criteria in another four[29,73,77,92] studies rely on paradigm-specific diagnoses. Participants in Rioux et al.'s study, for instance, all exemplify one of two (*kapha*-aggravated) Ayurvedic constitution/imbalance profiles; persons with other Ayurvedic diagnostic profiles are designated ''ineligible,'' as paradigm-specific etiology ''for these individuals would… entail a causally distinct trajectory.''[29]

In addition, four studies *explicitly* address intra-trial consistency in the subjective determination of paradigm-specific diagnoses. Kessler et al.'s study relies on a team of four Ayurvedic practitioners to reach consensus on diagnostic and treatment parameters for ''the first 30 patients.''[75] Similarly, two Chinese medicine physicians ''independently assessed'' each patient in Huang's 2018 trial, calling on a third ''distinguished veteran doctor of TCM'' to resolve any controversy between them. A secondary publication[125] associated with Ritenbaugh et al.'s Chinese medicine study[86] details usage of a standardized questionnaire, accompanied by a clinician training process, to enhance inter-rater reliability. Azizi et al.'s study notes its reliance on a single diagnostician ''to ensure uniform diagnosis''[77]; other studies[29,72,82] also have just one diagnostician, but do not link this point to the issue of paradigm-specific diagnostic consistency.

Multitarget/multimorbid interventions. All of the reviewed studies have clinical foci, outcome measures, and/or intervention designs that are clearly multimorbid, multitarget, or both (Fig. 9).

Some studies explicitly address more than one biomedical diagnostic category (e.g., cardiovascular disease *and* depression[98]; multiple chronic illnesses[71]) or nonbiomedical diagnoses for complex comorbid pathologies (e.g., a Chinese medicine diagnosis of ''Damp-Heat in the Bladder,'' compounded in some patients with ''Spleen Qi deficiency and Liver Qi stagnation'' and/or ''Kidney deficiency''[79]). Other studies set aside a singular disease-based emphasis in favor of multitarget conceptions of wellness, implied by constituting (for example) ''medically-unexplained symptoms''[84] or health-related QoL[89,91] as their primary clinical foci.

Moreover (as detailed further on and shown in Figs. 10 and 11), almost two-thirds of the reviewed studies use modes of outcome assessment designed to evaluate QoL and/or psychosocial wellness parameters. Such tools—which typically assess for such health concerns as ''pain, fatigue, nausea, depression anxiety, drowsiness, shortness of breath, appetite, sleep, and feeling of well-being'' as well as ''physical, role, emotional, cognitive, and social functioning''[89]—are clearly multitarget in their focus.

Even among the small number of studies that focus on a singular biomedical diagnosis and use no QoL-related, psychosocial, or qualitative outcome measures,[69,72,101,105,121] the *interventions* studied are not only multimodal but also behavioral in design, suggesting a broadly conceived (i.e., multitarget) salutogenic focus.

Multidisciplinary/team care. Twelve reviewed studies report on *team-based* interventions in which practitioners from across more than one discipline deliver *bilaterally coordinated* care to participants (Fig. 9). Team care interventions take place intraparadigmatically in three anthroposophic,[69–71] two Ayurvedic,[29,76] and three preventive/restorative biomedical studies.[98,99,121] In other words, in these studies, disciplinarily diverse providers from within a single paradigmatic system deliver different aspects of care (e.g., anthroposophic physician care with referrals to anthroposophic art, movement, and/or massage therapists). Conversely, in four[35,89–91] (of five) complementary/integrative medicine studies, and the one study involving concurrent Ayurvedic/yoga therapy care,[29] teams are composed of providers representing more than one health care paradigm.

In three additional studies,[81,87,120] two of which are retrospective,[81,120] nonbiomedical health care providers *unilaterally coordinate* their interventions with biomedical treatment (e.g., Hullender Rubin et al.'s study practitioners time their Chinese medicine infertility treatments to coincide with IVF).[81] Three other studies deliver *un-coordinated* multidisciplinary care, in which Chinese medicine[84] or naturopathic[93,95] care act as independent adjuncts to ''usual'' biomedical treatment.

Model validity and practical considerations in designing WSR interventions. Across exemplars, the evaluated interventions are generally *paradigm-specific*, representing complex, real-world practice rather than isolated components thereof. A group of intervention traits furthermore emerges as *paradigm-consistent* in WSR contexts as shown in Figure 8: WSR interventions are almost universally multimorbid/multitarget, complex, and individualized; often include salutogenic behavioral therapies and multiparadigmatic diagnoses; and at times feature multidisciplinary care. Excepting dual diagnoses, these individual characteristics are not necessarily uncommon in complex clinical trial designs across other health care disciplines. It is that these traits appear repeatedly *together* in a single study that distinguishes WSR interventions from those in other fields.

Through diverse approaches to therapeutic individualization, WSR studies furthermore implement *paradigm-specific* research interventions. Some individualization modes appear specifically relevant to particular TCIM paradigms, producing tension between model validity and research rigor more broadly conceived. Traditional (Chinese and Ayurvedic medicine) exemplars commonly engage *manualization with tailoring* approaches to align patient care with paradigm-specific diagnoses and associated treatment parameters. However, in the context of (for instance) naturopathic medicine, manualized/tailored protocols limit clinicians' treatment decisions ''to a greater degree than is typical'' in routine practice,[94] threatening

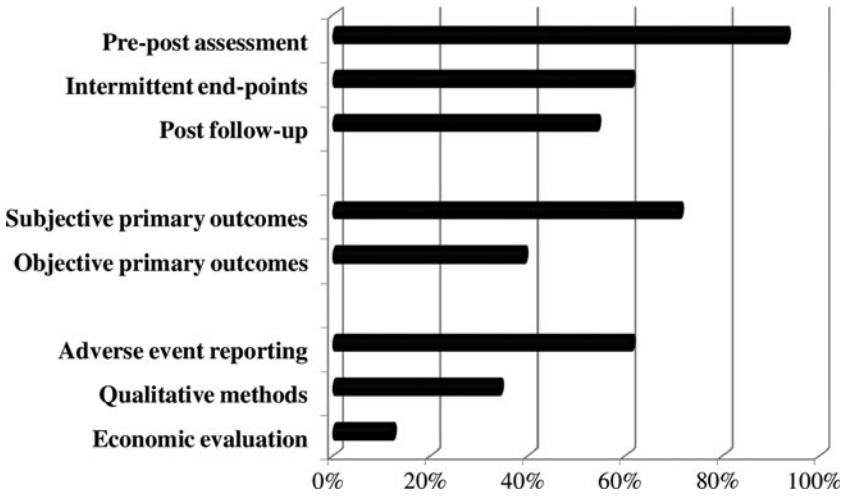

FIG. 10. Outcome assessment trends in whole systems research.

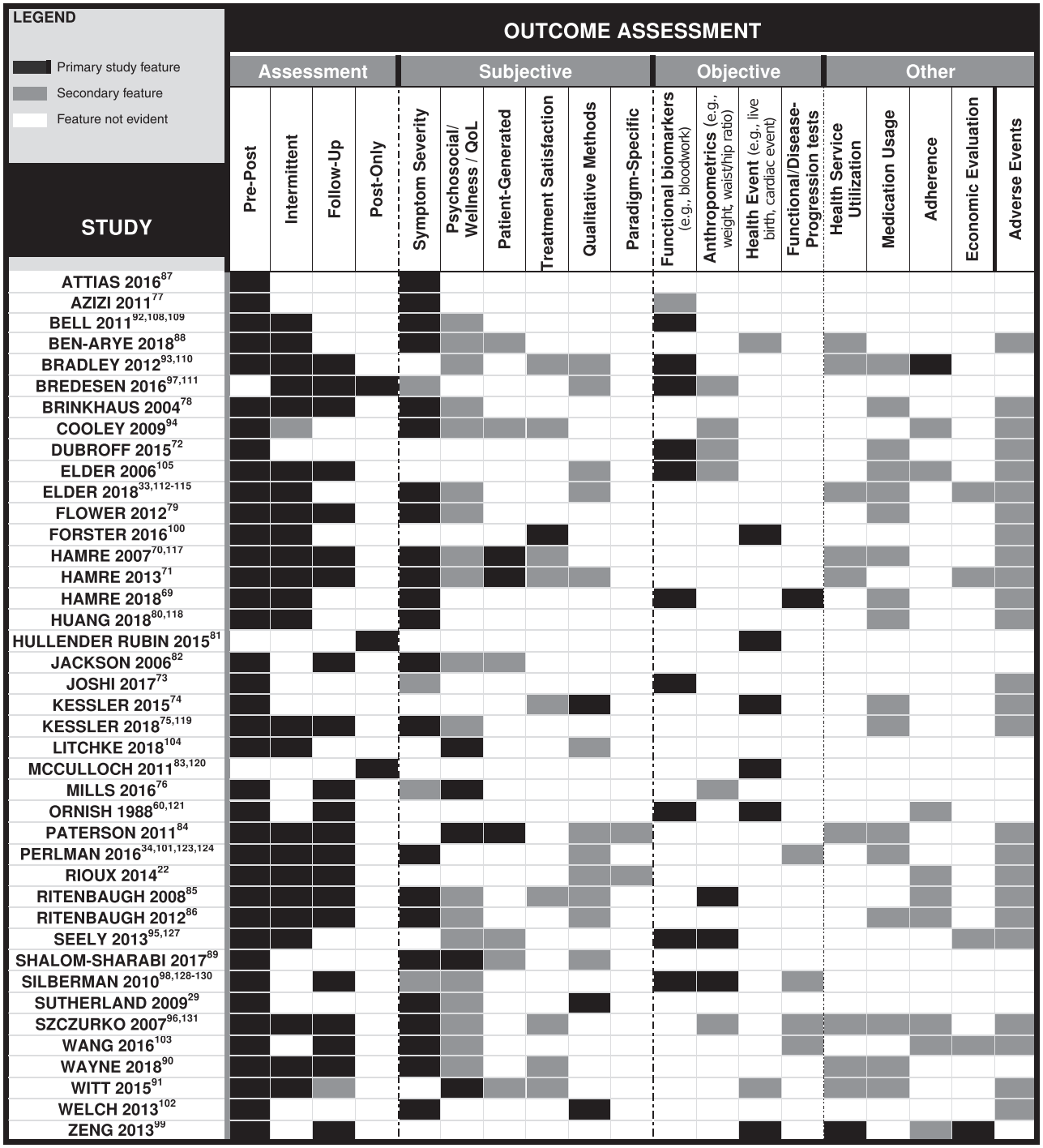

FIG. 11. Outcome assessment in whole systems research.

model validity. *Unconstrained individualization* is arguably a more suitable approach here and is also repeatedly engaged in anthroposophic and complementary/integrative medicine exemplars. While standardized and manualized designs lend themselves readily to replicability and generalizability (key markers of external validity), this proves more challenging when clinicians' treatments are unconstrained.

Regardless, it should be emphasized that dual diagnostics emerge as a unique design feature across a significant proportion of WSR exemplars, clearly distinguishing WSR from conventional biomedical research. Studies that apply manualized/tailored protocols tend to more explicitly detail the paradigm-specific diagnoses engaged. Such detailing may enhance external validity by facilitating study replication. Strategies to promote inter-rater reliability furthermore emerge as significant vis-a-vis paradigm-specific diagnoses. In addition to the approaches used in a few reviewed exemplars, whole systems researchers may refer to a growing methodological literature in this area.[26,125,136]

Multidisciplinary care is evident in several WSR exemplars, some of which implement ''usual care plus'' designs in which TCIM care serves as a biomedical adjunct. Such designs accurately represent the broader context of biomedical dominance and are typical features of real-world practice for many TCIM clinicians; therefore, ''usual care plus'' designs may enhance some studies' external validity. In terms of model validity, however, team care interventions which study multidisciplinary care from within a single[70,71] or two compatible TCIM paradigms[29] are significant in their ''articulation''[66] of TCIM whole systems as distinct autonomous disciplines.

Part III: outcome assessment

Across the WSR studies reviewed, a range of quantitative (and, to a lesser extent, qualitative) measurement instruments were used to evaluate outcomes, at various intervals. As summarized in Figure 10, the majority of studies used pre- and postmeasures of treatment impacts, often alongside intermittent and follow-up assessments. Primary outcome measures were more frequently subjective than objective, and adverse event reporting was common. Figure 11 provides a detailed graphical representation of primary and secondary outcome measure type and usage, discussed and contextualized in what follows; actual study results receive no attention in this analysis.

Reporting intervals. Most of the reviewed prospective studies implement concurrent evaluations of several primary and secondary outcomes, with measurements taking place both before and after the intervention. About two-thirds secondarily report outcomes as measured at intermittent intervals during the intervention period; two-thirds report ''follow-up'' outcomes from posttreatment measurements; and just under one-third do both (Fig. 11). The four reviewed retrospective studies report postoutcomes only,[81] although the single case report[75] and one case series[111] furthermore elaborate on treatment progress over the intervention period.

Primary and secondary outcome measures. Over 70% of the prospective studies reviewed adopt subjective measures—and, more specifically, PROMs—to evaluate their primary outcomes (Fig. 10). About one-third by contrast apply *objective endpoints*—such as blood-based biomarkers, anthropometrics such as weight, or health outcomes like survival or live birth rates—as primary outcomes, in some cases alongside PROMs (Fig. 11). Of the range of PROMs used to evaluate primary outcomes, condition-specific *symptom severity scales* dominate across studies; almost all of these are validated scales developed with reference to biomedical health/disease conceptualizations. Validated PROMs measuring QoL and *wellness-related* scores also appear in most studies as secondary outcome measures and in three studies as a primary measure. The aforementioned outcome types also serve as secondary (or co-primary) measures in some studies, as do the following:

- *Patient-generated* outcome measures in which participating patients individually define the health- and wellness-related parameters being measured, at times with clinician support.
- PROMs to measure *treatment expectation* and *treatment satisfaction.*
- Quasi-objective, clinician-assessed *tests of physical function* (e.g., walking or spinal flexion tests) or *disease progression* (e.g., radiologic tests for rheumatoid arthritis progression).
- Health and/or economic outcomes, including *medication usage*, *health service utilization*, and *work absenteeism* (Fig. 11).

All studies that include a standardized behavioral intervention specifically track patient adherence.

Notably, two specific sets of validated QoL and wellness measurement PROMs appear in multiple studies. These are:

(1) the Short-Form 36 (SF-36)[70,71,75,96,103] and an abbreviated version thereof, the SF-12[90,91]: generic, predetermined scales designed to gather QoL- and wellness-related data from patients[137,138]; and
(2) the patient-generated quantitative outcome measures known as ''MYMOP''[139] (Measure Yourself Medical Outcome Profile)[79,82,88,94,95] and ''MYCaW'' (Measure Yourself Concerns and Wellbeing),[89] the latter of which also gathers qualitative data from patients in the form of an open-ended questionnaire item.[140]

Finally, the reviewed *retrospective* studies generally use objective health events (e.g., live birth and death/survival), alongside other subjective and objective assessment approaches, to express their outcomes.

Adverse event reporting. Most reviewed studies include adverse event reporting, monitoring for which occurred through questionnaire/survey, live during interventions, by telephone and/or online (Fig. 11). Several herbal medicine studies also sampled blood and/or urine at baseline, during, and after the intervention, as a safety monitoring mechanism.[78–80,86,105,118]

Paradigm-specific outcome assessment. Paradigm-specific instruments to measure study outcomes appear in just two of the reviewed studies. Rioux et al. uses custom-designed tools ''to capture data in five lifestyle-related areas identified by Ayurveda as potential contributors or impediments to weight loss.''[29] Forster et al.'s midwifery study similarly uses a custom-modified PROM that emphasizes dimensions of care uniquely central to the midwifery paradigm,[100] noteworthy given an elsewhere-identified absence of such tools in that field.[141]

That said, four additional studies produce paradigm-specific[72,73,77,109] outcomes using biomedically developed instruments to evaluate symptom scores and other outcomes associated with singular paradigm-specific diagnoses (e.g., ''kidney and liver *yin* deficiency accompanied by liver yang hyperactivity''[77]). The ensuing results are uniquely relevant to those working within or evaluating the tenets of a study's driving paradigm. Bell et al.'s 2012 use of nonlinear dynamical analyses to reinterpret objective study outcomes also produces results that uniquely refer to homeopathic medicine's explanatory tenets.[108]

Complex outcome evaluation models. Aside from Bell et al.'s[108] use of complexity theory described above, just a few studies employ distinct outcome analytic models that address the multidimensional data generated. Consistent

with biomedical research approaches, six mixed methods studies actively triangulate qualitative with quantitative findings (Fig. 6); and a few studies with standardized behavioral interventions[29,60,121] correlate adherence with treatment effectiveness measures. Many studies concurrently report on a variety of outcome measures in a single publication, but do not directly draw connections between them. Some studies—such as Forster et al.'s midwifery RCT[100,116]—use separate publications to report upon different sets of measured outcomes (e.g., patient satisfaction vs. cesarean section rates).

Three additional studies engage with ethnographically-informed modes of outcome assessment, which deliberately draw attention to multiple clinical outcomes and/or contextualize participants' experiences *over the course of* (rather than at discrete endpoints in) a whole systems intervention.

Aiming to evaluate relationships between separately measured outcomes, Rioux et al.'s Ayurvedic/yoga therapy weight loss study begins to model the mixed-methods concept of a ''topographical data set,'' informed by the ''anthropological notion of thick description.''[29] To this end, Rioux et al.[29] graphically plot an overview of 15 distinct clinical ''data collection measures'' alongside each measure's specific ''time points for collection.'' The 2014 publication referenced in this study reports on anthropometric and adherence outcomes, as well as some qualitative results, while complete outcomes from the trial, including paradigm-specific measures, are published in this JACM whole-systems special issue for the first time. Welch et al.'s study, in turn, explicitly uses ethnographic methods to report on contextual factors from the clinical environment, reporting minimally on treatment outcomes.[102] While ''thick description'' is similarly evident across most studies using qualitative methods, the substudy associated with Ritenbaugh and colleagues trial uniquely engages trial participants in a *series* of qualitative interviews at intervals during the study, ethnographically theorizing *process-related* findings regarding patients' treatment ''expectations and hopes.''[126]

Model validity and practical considerations in WSR outcome assessment. Aligned with conventional biomedical research norms, most reviewed studies engage quantitative outcome measures to report their results; and subjective rather than objective measures dominate as primary assessment tools. Measuring outcomes of direct relevance to patients is of course no longer atypical in *pragmatic* biomedical trials outside of the WSR world. Further suggesting *paradigm compatibility*, most primary PROMs used in the reviewed studies had been developed in biomedical contexts. However, a set of complex outcome measurement trends emerged in common across multiple studies, indicating a *paradigm consistent* approach distinct from clinical research norms.

Exemplars commonly use symptom severity PROMs alongside QoL/psychosocial measures, with reference to multiple endpoints (i.e., pre-, post-, intermittent, and follow-up). Such an approach—complemented in a quarter of exemplars with treatment satisfaction measures—clearly reflects the patient-centered, salutogenic underpinnings of TCIM paradigms and an emphasis on progressive, enduring treatment impacts. Repeated usage of some QoL/wellness PROMs (e.g., SF-12, SF-36, MYMOP, and MYCaW), some of which have been developed by TCIM researchers, suggests that these particular tools may be considered particularly *paradigm consistent*.

Objective outcome measures are certainly not absent among WSR exemplars, but rarely appear to the exclusion of concurrent PROMs. Moreover, about half of all objective study outcomes refer to considerations of direct significance to patients (e.g., weight change, live birth, and survival), rather than being concerned primarily with biomedically conceptualized disease causation. Only one reviewed exemplar uses objective primary outcomes with the explicit aim of establishing biomedical mechanisms of action.[73]

Bell et al.'s use of objective measures to assess primary homeopathic outcomes is noteworthy in the context of a research paradigm routinely dismissed in biomedical contexts as physiologically implausible.[92] Other studies that engage objective primary outcomes appear to do so to render their results comparable with conventional biomedical trials addressing the same chronic health conditions (Type II diabetes, cardiovascular disease): a consideration reasonably geared toward external validity.

In contrast to the widespread engagement of *paradigm-specific* interventions across the WSR studies reviewed, relatively few reviewed studies engaged *paradigm-specific outcome* measures. Some scholars have advised that *paradigm-specific* outcome measures be avoided as *primary* variables in TCIM research as they may limit studies' external validity within biomedically dominant health systems.[142] Regardless, paradigm-specific outcome measures tools—not presently in widespread WSR usage—may usefully differentiate the impacts of TCIM interventions delivered on the basis of paradigm-specific diagnoses, as now discussed.

Conventional PROMs are certainly useful in gathering outcomes from the patient's perspective; patient-*generated* outcome measures have further potential to capture effects not preconceptualized by researchers. Such tools, however, are not designed to evaluate changing pathologies with reference to a particular TCIM system's indigenous concepts.

PROMs custom developed to align in paradigm-consistent and paradigm-specific ways with TCIM systems' distinct conceptions of health and disease may begin to fill this gap.[143] Such tools—which will ultimately require rigorous validation—may be based on qualitative research outcomes, as proposed in Sutherland et al.'s exemplar,[36] purpose innovated as in Rioux et al.'s[29] and Forster et al.'s[100] studies, and/or formulated from the rich bodies of paradigm-specific literature that inform TCIM care.[143,144] The Self-Assessment of Change tool,[145] a validated, paradigm-consistent, patient-centered outcome measure developed by a group of whole systems researchers in 2011,[146,147] was not used in any of the reviewed exemplars. Aligned with previous research on ''whole person healing,''[148] and informed by the lived experiences of TCIM patients, this PROM aims to evaluate the ''emergent'' effects of therapeutic interventions.

> ''beyond those [effects] associated with…specific treatment goals, including unanticipated outcomes and multidimensional shifts in overall well-being, energy, clarity of thought, emotional and social functioning, lifestyle patterns, inner life, and spirituality.''[146]

Elsewhere applied,[149,150] use of this tool may notably improve WSR studies' reporting of ''whole person'' patient outcomes[146,147] moving forward. *Clinician-reported*, paradigm-specific outcome measures[144]—combined with inter-rater reliability strategies—will also likely prove important. Inspiration to renew a centralized open repository of validated, paradigm-compatible and paradigm-specific outcome measures for WSR, informed by previous work by Canada's INCAM Research Network,[151] might be further drawn from the biomedical PROMIS[3] project.

Furthermore, the application of complex evaluation models will prove critical in bringing the WSR imperative to fruition in line with pioneers' vision of holistically contextualized outcomes. Although applications of program theory have begun to be explored in TCIM clinical research contexts,[152] uptake of complex system science in WSR has been not as readily undertaken as anticipated, despite publication of multiple theoretical works on the subject. Securing funding for such complex designs remains a considerable challenge in this regard. Designs that emphasize the study of ''process'' rather than ''outcomes'' remain to be fully implemented,[16,53] although the relationships between the two may fruitfully be studied through Rioux et al.'s ''topographical'' dataset proposition.[29] Methods that further interrogate ''individual differences rather than group averages''[53] will also likely prove important, as whole systems researchers seek to integrate the multiple synergistic aspects of holistic clinical interventions.

Discussion and Conclusions

This scoping review of WSR methods represents a first synthetic consolidation of over 15 years of advances in a distinctive field of scientific inquiry. At first glance, WSR has much in common with conventional clinical research. Its range of study designs—whether controlled or uncontrolled—generally represent adaptations upon (rather than reinventions of) established research methods; and its predominantly quantitative outcome measurements echo those applied in biomedical research.

On the whole, WSR designs align with established norms surrounding the evaluation of *complex* clinical interventions.[2] Related features include: the application of ''appropriate methodological choices''; the use of relevant randomization alternatives; identification of a ''coherent theoretical basis'' for intervention design; the engagement of multiple rather than singular primary and secondary outcomes; and, at times, the inclusion of economic evaluations.[2]

Reviewed post-facto in light of the PRECIS-2 pragmatic/explanatory study continuum's nine domains,[1] most comparative/controlled WSR studies also exhibit considerably more pragmatic design features, geared toward evaluating the real-world *effectiveness* of particular therapeutic interventions. This is evident across studies in: the enrolment of patients and clinicians in existing clinical settings; broad inclusion of multimorbid participants; high levels of intervention flexibility (i.e., individualization); and primary outcome measures directly relevant to patients (e.g., symptom severity and QoL).

As this review equally demonstrates, WSR is distinguished by a set of unique features. Studies centralize the epistemological and practical features of health care paradigms distinct from conventional biomedicine. Many WSR studies rely on dual diagnoses, supplementing, reframing, or replacing biomedical concepts of health and disease with paradigm-specific diagnostic and etiologic concepts. Complex salutogenic interventions are commonly tailored to the patient on this basis, using various individualization strategies. Whole systems researchers, as this work makes evident, have successfully innovated a range of strategies for achieving a paradigmatic-methodological fit, that is, ''model validity.''

Such strategies variously include alignment with specific, established research designs (''paradigm compatibility''), modification of conventional methods (''paradigm consistency''), and/or innovation of novel research strategies (''paradigm specificity''). As summarized in Figure 12 and elaborated throughout this work, model validity's dimensions appear differentially relevant to study design selection, interventions delivered, and outcomes evaluated in WSR contexts.

Although some of WSR's key features are not themselves unique, taken together as a synergistic set of design features, they become notable for their holistic patient-centered orientation. These features include: recruitment of multimorbid participants; delivery of multitarget therapies; centralization of subjective, patient-reported outcomes; diversified and multiple measurements of treatment effects; and concurrent engagement of mixed (quantitative and qualitative) methods. Reflecting on the vision articulated by WSR pioneers just after the turn of the century, it is clear that the field has significantly advanced; and TCIM researchers now have a body of WSR exemplars from which to learn.

Challenges of course remain. At a 2010 roundtable discussion, WSR leaders debated how to: contend with large bodies of quantitative and qualitative data; implement designs addressed to complexity; undertake trials of sufficiently powered size to reach meaningful conclusions; accommodate interpractitioner differences in practice style; provide training for new researchers; locate publication venues for multidimensional studies; and address scientific skepticism about the field.[53] Echoing some of these issues, the current review additionally calls for greater emphasis on ethnographically-informed designs, inter-rater reliability, and paradigm-specific outcomes.

It is hoped that this review will serve as a primary resource for researchers, practitioners, funders, and policymakers interested in the rigorous evaluation of TCIM as widely practised. The previous absence of a synthetic analysis of the field's advances has perhaps presented a barrier to WSR's centralization in strategic plans at core TCIM hubs, such as the U.S. National Center for Complementary and Integrative Health (NCCIH, formerly NCCAM). A principal element of the enabling statute from the U.S. Congress to that agency was to examine the integration of these ''systems and disciplines with conventional

	Paradigm Compatibility	Paradigm Consistency	Paradigm Specificity
STUDY DESIGN	■	■	
INTERVENTIONS		■	■
OUTCOMES	■	■	■

FIG. 12. Model validity in whole systems research.

medicine and as a complement to such medicine and into the health care delivery systems.''[153]

Regardless, as recently as 2016, former NCCIH leadership resisted calls to support WSR, on the premise that it was not yet clear what types of methods might be appropriate for this purpose. ''Protocols to domesticate the wildness of integrative personalization'' in the context of complex TCIM care would be needed, NCCIH leadership argued at the time.[154] As this work clearly documents, rigorous WSR methods do indeed exist; further, they have been successfully implemented.

Securing funding to conduct innovative WSR studies is certainly a prominent challenge that researchers in this field continue to face.[53] Researchers may elect to align with established methods, such as ''pragmatic,'' ''complex,'' ''comparative effectiveness,'' and ''mixed methods'' to solicit support for their work, and are wise to adhere to established guidelines in these areas. It is however important to recall that WSR as a maturing scholarly discipline extends *beyond* the aforementioned approaches. The inconsistent use of ''WSR'' and ''model validity'' terminology across the reviewed studies suggests that the field as it stands could benefit from greater cohesion. ISCMR, an active global organization of TCIM scholars whose founding mission was to advance WSR,[155] might advantageously renew its role in this regard.

Discussion of the WSR field, in which individualized care comprises a vital component, would not be complete without reference to the emerging trend toward ''personalized'' biomedical treatment. In contrast to TCIM providers' holistic reliance on paradigm-specific diagnoses, patient preferences, and contextual factors to personalize care, objective genomic testing is rapidly becoming the primary driver of individualization in biomedicine. As Mazer, a biomedical doctor, astutely observes: ''[t]he rise of 'personalized medicine' is, ironically, a continuation of [a] reductionist mode…that deconstructs an individual into her faceless genetic components.''[156]

WSR is ultimately a hybrid phenomenon that stretches the boundaries of biomedical research to better accommodate diverse, holistic health care approaches. At a historical moment when TCIM providers find their long-held values—personalized patient-centered care; salutogenesis and prevention; complex interventions; and patient-reported outcomes—to have become buzzwords within biomedicine's highest echelons, the potential for co-optation is significant. Despite evident challenges, WSR advocates and leaders who seek to advance the field must continue to insist that the multiple dimensions of health cannot be reduced to an objective set of biomarkers and that the whole is far more sophisticated than the sum of its most evidenced parts.

Acknowledgments

The authors wish to acknowledge the following whole systems research area experts, who each contributed to the list of clinical exemplars reviewed in this work: Heather Boon, Scott Mist, Barb Reece, Cheryl Ritenbaugh, and Dugald Seely. Many thanks also to Katya Korol-O-Dwyer for her research assistance with the project. The development of this article was supported, in part, through Grant #4221 from the Lotte and John Hecht Memorial Foundation.

Author Disclosure Statement

No competing financial interests exist.

References

1. Loudon K, Treweek S, Sullivan F, et al. The PRECIS-2 tool: Designing trials that are fit for purpose. BMJ 2015; 350:h2147.
2. Medical Research Council. Developing and evaluating complex interventions [homepage on the Internet]. 2018. Online document at: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance, accessed October 21, 2018.
3. Northwestern University. PROMIS [homepage on the Internet]. 2018. Online document at: http://healthmeasures.net/explore-measurement-systems/promis, accessed November 13, 2018.
4. Rathert C, Wyrwich M, Boren S. Patient-centered care and outcomes: A systematic review of the literature. Med Care Res Rev 2012;70:351–379.
5. Ritenbaugh C, Verhoef M, Fleishman S, et al. Whole systems research: A discipline for studying complementary and alternative medicine. Altern Ther Health Med 2003;93:186–189.
6. Verhoef M, Lewith G, Ritenbaugh C, et al. Complementary and alternative medicine whole systems research: Beyond identification of inadequacies of the RCT. Comp Ther Med 2005;13:206–212.
7. Heron J. Critique of conventional research methodology. Comp Med Res 1986;1:12–22.
8. Dossey L. How should alternative therapies be evaluated: An examination of fundamentals. Altern Ther Health Med 1995;1:6–10; 79–85.
9. Carter B. Methodological issues and complementary therapies: Researching intangibles? Complement Ther Nurs Midwifery 2003;9:133–139.
10. Kiene H. A critique of the double-blind clinical trial: Part 2. Altern Ther Health Med 1996;2:59–64.
11. Kiene H. A critique of the double-blind clinical trial: Part 1. Altern Ther Health Med 1996;2:74–80.
12. Kienle G, Albonico H-U, Fischer L, et al. Complementary therapy systems and their integrative evaluation. Explore (NY) 2011;7:175–187.
13. Bothwell L, Greene J, Podolsky S, Jones D. Assesing the gold standard—lessons from the history of RCTs. N Engl J Med 2016;374:22.
14. Elder C, Aickin M, Bell I, et al. Methodological challenges in whole systems research. J Altern Complement Med 2006;12:843–850.
15. Long A. Outcome measurement in complementary and alternative medicine: Unpicking the effects. J Altern Complement Med 2002;8:777–786.
16. Paterson C, Baarts C, Launso L, Verhoef M. Evaluating complex health interventions: A critical analysis of the 'outcomes' concept. BMC Complement Altern Med 2009; 9:18.
17. Fonnebo V, Grimsgaard S, Walach H, et al. Researching complementary and alternative treatments—the gatekeepers are not at home. BMC Med Res Methodol 2007;7:7.
18. MacPherson H. Pragmatic clinical trials. Complement Ther Med 2004;12:136–140.
19. Walach H, Falkenberg T, Fonnebo V, et al. Circular instead of hierarchical: Methodological principles for the evaluation of complex interventions. BMC Med Res Methodol 2006;6:1–9.

20. Boon H, MacPherson H, Fleishman S, et al. Evaluating complex healthcare systems: A critique of four approaches. Evid Based Complement Altern Med 2007;4:279–285.
21. Aickin M. Beyond randomization. J Altern Complement Med 2002;8:765–772.
22. Verhoef M, Casebeer A, Hilsden R. Assessing efficacy of complementary medicine: Adding qualitative research methods to the ''gold standard''. J Altern Complement Med 2002;8:275–281.
23. Bell I, Koithan M. Models for the study of whole systems. Intregr Cancer Ther 2006;5:293–307.
24. Lewith G. The use and abuse of evidence-based medicine: An example from general practice. Complement Ther Med 1996;4:144.
25. Jonas W. The evidence house: How to build an inclusive base for complementary medicine. West J Med 2001;175:79–80.
26. Mathie R, Roniger H, van Wassenhoven M, et al. Method for appraising model validity of randomised controlled trials of homeopathic treatment: Multi-rater concordance study. BMC Med Res Methodol 2012;12:1–9.
27. Mathie R, van Wassenhoven M, Jacobs J, et al. Model validity and risk of bias in randomised placebo-controlled trials of individualised homeopathic treatment. Complement Ther Med 2016;25:120–125.
28. Bornhoft G, Maxion-Bergemann S, Wolf U, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. BMC Med Res Methodol 2006;6:1–13.
29. Rioux J, Thomson C, Howerter A. A pilot feasibility study of whole-systems Ayurvedic medicine and yoga therapy for weight loss. Glob Adv Health Med 2014;3:28–35.
30. Lad V. Textbook of Ayurveda, Volume I: Fundamental Principles. Albuquerque, NM: Ayurvedic Press, 2002.
31. Kienle G, Albonico H-U, Baars E et al. Anthroposophic medicine: An integrative medical system originating in Europe. Glob Adv Health Med 2013;2:20–31.
32. Maciocia G. The Foundations of Chinese Medicine. London, United Kingdom: Churchill Livingstone, 1989.
33. Elder C, DeBar L, Ritenbaugh C, et al. Comparative effectiveness of usual care with or without chiropractic care in patients with recurrent musculoskeletal back and neck pain. J Gen Intern Med 2018;33:1469–1477.
34. Souza TA. Differential Diagnosis and Management for the Chiropractor: Protocols and Algorithms, 5th ed. Burlington, MA: Jones and Bartlett Learning, 2016.
35. Ben-Arye E, Livne Aharonson M, Schiff E, Samuels N. Alleviating gastro-intestinal symptoms and concerns by integrating patient-tailored complementary medicine in supportive cancer care. Clin Nutr 2015;34:1215–1223.
36. Sutherland E, Ritenbaugh C, Kiley L, et al. An HMO-based prospective pilot study of energy medicine for chronic headaches: Whole- person outcomes point to the need for new instrumentation. J Altern Complement Med 2009;15:819–826.
37. Hahneman S. Organon of Medicine, 6th ed. (Translated by Kunzli JNA, Pendleton, P). Boston, MA: J.P. Tarcher, Inc., 1982.
38. Sankaran R. The Substance of Homeopathy. Mumbai, India: Homoeopathic Medical Publishers, 1994.
39. Marshall J, Raynor M. Myles Textbook for Midwives. London, United Kingdom: Churchill Livingstone, 2014.
40. Wardle J, Sarris J. Clinical Naturopathy: An Evidence-Based Guide to Practice, 2nd ed. Sydney, Australia: Churchill Livingstone, 2014.
41. Ali A, Kahn J, Rosenberger L, Perlman A. Development of a manualized protocol of massage therapy for clinical trials in osteoarthritis. Trials 2012;13:1–6.
42. Wayne P, Kaptchuk T. Challenges inherent to t'ai chi research: T'ai chi as a complex multicomponent intervention. J Altern Complement Med 2008;14:95–102.
43. Rioux J. Yoga therapy research: A whole-systems perspective on comparative effectiveness and patient-centred outcomes. Int J Yoga Therap 2015;25:9–19.
44. International Association of Yoga Therapists. Educational standards for the training of yoga therapists [homepage on the Internet]. 2017. Online document at: https://cdn.ymaws.com/www.iayt.org/resource/resmgr/accreditationmaterials/2017_11_Updates-Ed_Stds/2017_IAYT_Educational_Standa.pdf, accessed November 16, 2018.
45. Miller M, Crabtree B, Duffy B, et al. Research guidelines for assessing the impact of healing relationships in clinical medicine. Altern Ther Health Med 2003;9:A80–A95.
46. Aickin M. Participant-centred analysis in complementary and alternative medicine comparative trials. J Altern Complement Med 2003;9:949–957.
47. Koithan M, Bell I, Niemeyer K, Pincus D. A complex systems science perspective for whole systems of complementary and alternative medicine research. Forsch Komplmentarmed 2012;19:7–14.
48. Rioux J. A complex, nonlinear dynamic systems perspective on Ayurveda and Ayurvedic research. J Alt Comp Med 2012;18:709–718.
49. Howerter A, Hollenstein T, Boon H, et al. State-space grid analysis: Applications for clinical whole systems complementary and alternative medicine research. Forsch Komplmentarmed 2012;19(suppl 1):30–35.
50. Andrews GJ, Boon H. CAM in Canada: Places, practices, research. Complement Ther Clin Pract 2005;11:21–27.
51. Baars E, Hamre H. Whole medical systems versus the system of conventional biomedicine: A critical, narrative review of similarities, differences, and factors that promote the integration process. Evid Based Complement Alternat Med 2017;2017:4904930.
52. Ostermann T, Beer A, Bankova V, Michalsen A. Whole-systems research in integrative inpatient treatment. Evid Based Complement Alternat Med 2013;2013:962729.
53. Ritenbaugh C, Aickin M, Bradley F, et al. Whole systems research becomes real: New results and next steps. J Altern Complement Med 2010;16:131–137.
54. Zick S, Schwabl H, Flower A, et al. Unique aspects of herbal whole system research. Explore (NY) 2011;5:97–103.
55. Hawk C, Khorsan R, Lisi A, et al. Chiropractic care for nonmusculoskeletal conditions: A systematic review with implications for whole systems research. J Altern Complement Med 2007;13:491–512.
56. Mills P, Patel S, Barsotti T, et al. Advancing research on traditional whole systems medicine approaches. J Evid Based Complementary Altern Med 2017;22:527–530.
57. Daudt H, van Mossel C, Scott S. Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013;13:1–9.
58. Levac D, Colquhoun H, O'Brien K. Scoping studies: Advancing the methodology. Implement Sci 2010;5:1–9.
59. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 2003; 8:19–32.

60. Ornish D, Scherwitz L, Billings J, et al. Intensive lifestyle changes for reversal of coronary heart disease. JAMA 1988;280:2001–2007.
61. Ornish D, Scherwitz L, Doody R, et al. Effects of stress management training and dietary changes in treating ischemic heart disease. JAMA 1983;249:54–59.
62. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: Exploring its conceptualization and operationalization. Qual Quant 2018;52:1893–1907.
63. Fran S. The constant comparative analysis method outside of grounded theory. Qual Rep 2013;18:1–25.
64. Baer H. Toward an Integrative Medicine: Merging Alternative Therapies with Biomedicine. Walnut Creek, CA: AltaMira Press, 2004.
65. Hollenberg D, Muzzin L. Epistemological challenges to integrative medicine: An anti-colonial perspective on the combination of complementary/alternative medicine with biomedicine. Health Sociol Rev 2010;19:34–56.
66. Weeks J. Colonialism, health justice, and (re)examination of the value of traditional practitioners and licensed integrative health professionals. J Altern Complement Med 2018;24:301–303.
67. Ijaz N, Boon H. Medical pluralism and the state: Regulatory language requirements for traditional acupuncturists in English-dominant diaspora jurisdictions. SAGE Open 2018;April–June:1–15.
68. Scaturo D. The evolution of psychotherapy and the concept of manualization: An integrative perspective. Prof Psychol 2001;32:522–530.
69. Hamre H, Pham V, Kern C, et al. A 4-year nonrandomized comparative phase-IV study of early rheumatoid arthritis: Integrative anthroposophic medicine for patients with preference against DMARDs versus conventional therapy including DMARDs for patients without preference. Patient Prefer. Adherence 2018;12:375–397.
70. Hamre H, Witt C, Glockmann A, et al. Anthroposophic vs. conventional therapy for chronic low back pain: A prospective comparative study. Eur J Med Res 2007;12:302–310.
71. Hamre H, Kiene H, Glockmann A, Ziegler R, Kienle G. Long-term outcomes of anthroposophic treatment for chronic disease: A four-year follow-up analysis of 1510 patients from a prospective observational study in routine outpatient settings. BMC Res Notes 2013;6:269.
72. DuBroff R, Lad V, Murray-Krezan C. A prospective trial of Ayurveda for coronary heart disease: A pilot study. Altern Ther Health Med 2015;21:52–62.
73. Joshi K, Nesari T, Dedge A, et al. Dosha phenotype specific Ayurveda intervention ameliorates asthma symptoms through cytokine modulations: Results of whole system clinical trial. J. Ethnopharmacol 2017;197:110–117.
74. Kessler C, Stapelfeldt E, Michalsen A, et al., The effect of a complex multi-modality Ayurvedic treatment in a case of unknown female infertility. Forsch Komplmentarmed 2015;22:251–258.
75. Kessler C, Dhiman K, Kumar A, et al. Effectiveness of an Ayurveda treatment approach in knee osteoarthritis - a randomized controlled trial. Osteoarthritis Cartilage 2018; 26:620–630.
76. Mills P, Wilson K, Pung M, et al. The self-directed biological transformation initiative and well-being. J Altern Complement Med 2016;22:627–634.
77. Azizi H, Liu YF, Du L, et al. Menopause-related symptoms: Traditional Chinese medicine vs. hormone therapy. Altern Ther Health Med 2011;17:48–53.
78. Brinkhaus B, Hummelsberger J, Kohnen R, et al. Acupucture and Chinese herbal medicine in the treatment of patients with seasonal allergic rhinitis: A randomized-controlled clinical trial. Allergy 2004;59:953–960.
79. Flower A, Lewith G. A prospective case series exploring the role of Chinese herbal medicine in the treatment of recurrent urinary tract infections. Eur J Integr Med 2012; 4:e421–e428.
80. Huang H, Yang P, Wang J, et al. Investigation into the individualized treatment of traditional Chinese medicine through a series of n-of-1 trials. Evid Based Complement Alternat Med 2018;2018:5813767.
81. Hullender Rubin L, Opsahl M, Wiemer K, et al. Impact of whole systems traditional Chinese medicine on in-vitro fertilization outcomes. Reprod Biomed Online 2015;30: 602–612.
82. Jackson A, MacPherson H, Hahn S. Acupuncture for tinnitus: A series of six n = 1 controlled trials. Complement Ther Med 2006;14:39–46.
83. McCulloch M, Broffman M, van der Laan M, et al. Lung cancer survival with herbal medicine and vitamins in a whole-systems approach: Ten-year follow-up data analysed with marginal structural models and propensity score methods. Intergr Cancer Ther 2011;10:260–279.
84. Paterson C, Taylor R, Griffiths P, et al. Acupuncture for 'frequent attenders' with medically-unexplained symptoms: A randomised controlled trial. Br J Gen Pract 2011; 61:e295–e305.
85. Ritenbaugh C, Hammerschlag R, Calabrese C, et al. A pilot whole systems clinical trial of traditional Chinese medicine and naturopathic medicine for the treatment of temporomandibular disorders. J Altern Complement Med 2008;14:475–487.
86. Ritenbaugh C, Hammerschlag R, Dworkin S, et al. Comparative effectiveness of traditional Chinese medicine and psychosocial care in the treatment of temporomandibular disorders—associated chronic facial pain. J Pain 2012;13:1075–1089.
87. Attias S, Keinan Boker L, Arnon Z, et al. Effectiveness of integrating individualized and generic complementary medicine treatments with standard care versus standard care alone for reducing preoperative anxiety. J Clin Anesth 2016;29:54–64.
88. Ben-Arye E, Doweck I, Schiff E, Samuels N. Exploring an integrative patient-tailored complementary medicine approach for chemotherapy-induced taste disorders. Explore 2018;14:289–294.
89. Shalom-Sharabi I, Samuels N, Lavie O, et al. Effect of a patient-tailored integrative medicine program on gastrointestinal concerns and quality of life in patients with breast and gynecologic cancer. J Cancer Res Clin Oncol 2017;143:1243–1254.
90. Wayne P, Eisenberg DM, Osypiuk K, et al. A multidisciplinary integrative medicine team in the treatment of chronic low back pain: An observational comparative effectiveness study J Altern Complement Med 2018;24: 781–791.
91. Witt C, Ausserer O, Baier S, et al. Effectiveness of an additional individualized multi-component complementary medicine treatment on health-related quality of life in breast cancer patients: A pragmatic randomized trial. Breast Cancer Res Treat 2015;149:449–460.
92. Bell I, Howerter A, Jackson N, et al. Effects of homeopathic medicines on polysomnographic sleep of young

adults with histories of coffee-related insomnia. Sleep Med 2011;12:505–511.
93. Bradley R, Sherman K, Catz S, et al. Adjunctive naturopathic care for type 2 diabetes: Patient reported and clinical outcomes after one year. BMC Complement Altern Med 2012;12:44.
94. Cooley K, Szczurko O, Perri D, et al. Naturopathic care for anxiety: A randomized controlled trial. PLoS One 2009;4:e6628.
95. Seely D, Szczurko O, Cooley K, et al. Naturopathic medicine for the prevention of cardiovascular disease: A randomized clinical trial. CMAJ 2013;185:E409–E416.
96. Szczurko O, Cooley K, Busse J, et al. Naturopathic care for chronic low back pain: A randomized trial. PLoS One 2007;2:e919.
97. Bredesen D, Amos E, Canick J, et al. Reversal of cognitive decline in Alzheimer's disease. Aging (Albany NY) 2016;8:1250–1258.
98. Silberman A, Banthia R, Estay I, et al. The effectiveness and efficacy of an intensive cardiac rehabilitation program in 24 sites. Am J Health Promot 2010;24:260–266.
99. Zeng W, Stason W, Fournier S, et al. Benefits and costs of intensive lifestyle modification programs for symtopmatic coronary disease in Medicare beneficiaries. Am Heart J 2013;165:785–792.
100. Forster D, McLachlan H, Davey M-A, et al. Continuity of care by a primary midwife (caseload midwifery) increases women's satisfaction with antenatal, intrapartum and postpartum care: Results from the COSMOS randomised controlled trial. BMC Pregnancy Childbirth 2016;16:28.
101. Perlman A, Gould-Fogerite S, Keever T, et al. Exploring massage benefits for arthritis of the knee (EMBARK). Osteoarthritis Cartilage 2016;24:S534.
102. Welch P, Thomas C, Bingley A. Working at the coalface: Using action research to study 'integrative medicine' in the NHS. Eur J Integr Med 2013;5:75–82.
103. Wang C, Schmid C, Iverson M, et al. Comparative effectiveness of tai chi versus physical therapy for knee osteoarthritis. Ann Intern Med 2016;165:77–86.
104. Litchke L, Liu T, Castro S. Effects of multimodal mandala yoga on social and emotional skills for youth with autism spectrum disorder: An exploratory study. Int J Yoga 2018; 11:59–65.
105. Elder C, Aickin M, Bauer V, et al. Randomized trial of a whole-systems ayurvedic protocol for type 2 diabetes. Altern Ther Health Med 2006;12:24–29.
106. Hamre H, Witt C, Glockmann A, et al. Health costs in anthroposophic therapy users: A two-year prospective cohort study. BMC Health Serv Res 2006;6:1–8.
107. Hamre H, Kiene H, Ziegler R, et al. Overview of the publications from the anthroposophic medicine outcomes study (AMOS): A whole system evaluation study. Glob Adv Health Med 2014;3:54–70.
108. Bell I, Howerter A, Jackson N, et al. Nonlinear dynamical systems effects of homeopathic remedies on multiscale entropy and correlation dimension of slow wave sleep EEG in young adults with histories of coffee-induced insomnia. Homeopathy 2012;101:182–192.
109. Brooks A, Bell I, Howerter A, et al. Effects of homeopathic medicines on mood of adults with histories of coffee-related insomnia. Forsch Komplmentarmed 2010;17:250–257.
110. Oberg E, Bradley R, Hsu C, et al. Patient-reported experiences with first-time naturopathic care for type 2 diabetes. PLoS One 2012;7:e48459.
111. Bredesen D. Reversal of cognitive decline: A novel therapeutic program. Aging 2014;6:707–717.
112. DeBar L, Elder C, Ritenbaugh C, et al. Acupuncture and chiropractic care for chronic pain in an integrated health plan: A mixed methods study. BMC Complement Altern Med 2011;11:118.
113. Elder C, DeBar L, Ritenbaugh C, et al. Acupuncture and chiropractic care: Utilization and electronic medical record capture. Am J Manag Care 2015;21:e414–e421.
114. Johnson E, Dickerson J, Vollmer W, et al. The feasibility of matching on a propensity score for acupuncture in a prospective cohort study of patients with chronic pain. BMC Med Res Methodol 2017;17:1–11.
115. Penney L, Ritenbaugh C, DeBar L, et al. Provider and patient perspectives on opioids and alternative treatments for managing chronic pain: A qualitative study. BMC Fam Pract 2016;17:1–15.
116. McLachlan H, Forster D, Davey M-A, et al. Effects of continuity of care by a primary midwife (caseload midwifery) on caesarean section rates in women of low obstetric risk: The COSMOS randomised controlled trial. BJOG 2012;119:1483–1492.
117. Hamre H, Witt C, Kienle G, et al. Long-term outcomes of anthroposophic therapy for chronic low back pain: A two-year follow-up analysis. J Pain Res 2009;2:75–85.
118. Huang H, Yang P, Xue J, et al. Evaluating the individualized treatment of traditional Chinese medicine: A pilot study of n-of-1 trials. Evid Based Complement Alternat Med 2014;2014:148730.
119. Witt C, Michalsen A, Roll S, et al. Comparative effectiveness of a complex Ayurvedic treatment and conventional standard care in osteoarthritis of the knee—study protocol for a randomized controlled trial. Trials 2013;14: 149.
120. McCulloch M, Broffman M, van der Laan M, et al. Colon cancer survival with herbal medicine and vitamins combined with standard therapy in a whole-systems approach: Ten-year follow-up data analysed with marginal structural models and propensity score models. Integr Cancer Ther 2011;10:240–259.
121. Ornish D, Brown S, Scherwitz L, et al. Can lifestyle changes reverse coronary heart disease? Lancet 1990;336: 129–133.
122. Rugg S, Paterson C, Britten N, et al. Traditional acupuncture for people with medically unexplained symptoms: A longitudinal qualitative study of patients' experiences. Br J Gen Pract 2011;61:e306–e315.
123. Ali A, Rosenberger L, Weiss T, et al. Massage therapy and quality of life in osteoarthritis of the knee: A qualitative study. Pain Med 2017;18:1168–1175.
124. Perlman A, Ali A, Njike V, et al. Massage therapy for osteoarthritis of the knee: A randomized dose-finding trial. PLoS One 2012;7:e30248.
125. Mist S, Ritenbaugh C, Aickin M. Effects of questionnaire-based diagnosis and training on inter-rater reliability among practitioners of traditional Chinese medicine. J Altern Complement Med 2009;15:7097.
126. Eaves E, Ritenbaugh C, Nichter M, et al. Modes of hoping: Understanding hope and expectation in the context of a clinical trial of complementary and alternative medicine for chronic pain. Explore 2014;10:225–232.
127. Herman P, Szczurko O, Cooley K, Seely D. A naturopathic approach to the prevention of cardiovascular disease:

Cost-effectiveness analysis of a pragmatic multi-worksite randomized clinical trial. J Occup Environ Med 2014;56: 171–176.
128. Daubenmier J, Weidner G, Sumner M, et al. The contribution of changes in diet, exercise and stress management to changes in coronary risk in women and men in the multisite cardiac lifestyle intervention program. Ann Behav Med 2007;33:57–68.
129. Govil S, Weidner G, Merritt-Worden T, Ornish D. Socioeconomic status and improvements in lifestyle, coronary risk factors, and quality of life: The Multisite Cardiac Lifestyle Intervention Program. Am J Public Health 2009; 99:1263–1270.
130. Frattaroli J, Weidener G, Merritt-Worden T, et al. Angina pectoris and atherosclerotic risk factors in the multisite cardiac lifestyle intervention program. Am J Cardiol 2008; 101:911–918.
131. Herman P, Szczurko O, Cooley K, Mills E. Cost-effectiveness of naturopathic care for chronic low back pain. Altern Ther Health Med 2008;14:32–39.
132. Schultz A, Chao S, McGinnis H. Integrative Medicine and the Health of the Public: A Summary of the February 2009 Summit. Washington, DC: National Academy of Sciences, 2009.
133. Onwuegbuzie A, Johnson R, Collins K. Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. Int J Mult Res Approach 2009;3:114–139.
134. Johnston B, Mills E. n-of-1 randomized controlled trials: An opportunity for complementary and alternative medicine evaluation. J Altern Complement Med 2004;10:979–984.
135. Ulbrich-Zurni S, Teut M, Roll S, Mathie R. The n-of-1 clnical trial: A timely research opportunity in homeopathy. Homeopathy 2018;107:10–18.
136. Weinschenk S, Gollner R, Hollman M, et al. Inter-rater reliability of neck reflex points in women with chronic neck pain. Forsch Komplmentarmed 2016;23:223–229.
137. Rand Corporation. 36-Item short form survey instrument [homepage on the Internet]. 2018. Online document at: https://rand.org/health/surveys_tools/mos/36-item-short-form/survey-instrument.html, accessed September 25, 2018.
138. Rand Corporation. 12-Item short form survey (SF-12) [homepage on the Internet]. 2018. Online document at: https://rand.org/health/surveys_tools/mos/12-item-short-form.html, accessed September 25, 2018.
139. University of Bristol. Questionnaires [homepage on the Internet]. 2018. Online document at: http://bris.ac.uk/primaryhealthcare/resources/mymop/questionnaires, accessed September 25, 2018.
140. University of Bristol. Measure yourself concerns and wellbeing [homepage on the Internet]. 2018. Online document at: http://bris.ac.uk/primaryhealthcare/resources/mymop/sisters, accessed September 25, 2018.
141. Smith V, Daly D, Lundgren I, Eri T, Benstoem C, Devane D. Salutogenically focused outcomes in systematic reviews of intrapartum interventions: A systematic review of systematic reviews. Midwifery 2014;30:e151–e156.
142. Levin J, Glass T, Kushi L, et al. Quantitative methods in research on complementary and alternative medicine: A methodological manifesto. Med Care 1997;35:1079–1094.
143. Chen R, Wong C, Lam T. Construction of a traditional Chinese medicine syndrome-specific outcome measure: The kidney deficiency syndrome questionnaire (KDSQ). BMC Complement Altern Med 2012;12:73.
144. Zhang Y, Liu B, Zhang R, et al. Clinician-reported outcomes in traditional Chinese medicine: A critical review. OA Alt Med 2013;1:15.
145. Ritenbaugh C. Welcome to the self-assessment of change website [homepage on the Internet]. 2018. Online document at: http://selfassessmentofchange.org, accessed December 18, 2018.
146. Thompson J, Kelly K, Ritenbaugh C, et al. Developing a patient-centered outcome measure for complementary and alternative medicine therapies II: Refining content validity through cognitive interviews. BMC Complement Altern Med 2011;11:136.
147. Ritenbaugh C, Nichter M, Nichter A, et al. Developing a patient-centered outcome measure for complementary and alternative medicine therapies I: Defining content and format. BMC Complement Altern Med 2011;11:135.
148. Koithan M, Verhoef M, Bell I, White M, Mulkins A, Ritenbaugh C. The process of whole person healing: ''unstuckness'' and beyond. J Altern Complement Med 2007;13:659–668.
149. Lillehei A, Halcon L, Gross C, Savik K, Reis R. Well-being and self-assessment of change: Secondary analysis of an RCT that demonstrated benefit of inhaled lavender and sleep hygiene in college students with sleep problems. Explore 2016;12:427–435.
150. Johnson M, Bertrand S, Fermon B, Coleman J. Pathways to healing: Patient-centered responses to complementary services. Glob Adv Health Med 2014;3:8–16.
151. Kania A, Verhoef M, Dryden T, Ware M. IN-CAM outcomes database: Its relevance and application in research and practice. Int J Ther Massage Bodywork 2009;2:8–16.
152. Skovgaard L, Bjerre L, Haahr N, et al. An investigation of multidisciplinary complex health care interventions—steps towards an integrative treatment model in the rehabilitation of people with multiple sclerosis. BMC Complement Altern Med 2012;12:50.
153. Weeks J. How NCCAM's ''real world'' congressional mandate is optimal for NCCAM's 2010–2015 strategic plan [homepage on the Internet]. 2010. Online document at: http://theintegratorblog.com/index.php?option=com_content&task=view&id=606&Itemid=189, accessed November 20, 2018.
154. Weeks J. Into the light: Reflections on whole-systems research after a case series finds reversal of Alzheimer's. J Altern Complement Med 2016;22:585–587.
155. ISCMR. About ISCMR [homepage on the Internet]. 2016. Online document at: https://iscmr.org/content/about-iscmr, accessed November 16, 2018.
156. Mazer B. Do physicians dream of electric sheep: Examining the conflict between compassion and technology. Yale J Biol Med 2018;91:343–344.

Address correspondence to:
Nadine Ijaz, PhD
Leslie Dan Faculty of Pharmacy
University of Toronto
144 College Street
Toronto, ON M5S 3M2
Canada

E-mail: nadineijaz@gmail.com