

Complementary and Integrative Medicine (CIM) Stakeholder Symposium

Meeting Summary

November 9, 2009
World Trade Center Baltimore
Maryland Room, 21st Floor
Baltimore, MD

BACKGROUND

The **Center for Medical Technology Policy (CMTP)** and the **Institute for Integrative Health (TIIH)** are collaborating to create a framework for designing comparative effectiveness research (CER) studies on complementary and integrative medicine (CIM). In November 2009, the two non-profit organizations hosted a working symposium for diverse stakeholders to initiate important conversations around topics including current evidence gaps in CIM, optimal study designs, how to utilize CER innovations from allopathic medicine studies for CIM, and preferred outcomes to better inform decision-making. This meeting engaged patients, public and private health plans representatives, researchers in both CIM and CER, clinicians, and government employees from funding and payment organizations in a process of sharing perspectives on how to improve the quality of research on complementary and integrative medicine. Desired outcomes from this collaboration include a white paper to summarize the discussion at the November 2009 meeting and elucidate next steps, and the development of a guidance document (or documents) to provide concrete recommendations for the design of comparative effectiveness research studies for CIM interventions.

INTRODUCTION

Dr. Brian Berman of TIIH opened the meeting with a brief history of research in complementary and integrative medicine. He stressed that while the quantity and quality of research has come a long way, there is still demand for better clinical evidence. Integrative medicine is not just the use of a particular alternative therapy or group of therapies; it encompasses a holistic approach to care. The CIM community has spent a great deal of time debating the merits of “gold-standard” RCTs as appropriate tools for the evaluation of integrative medicine interventions, but because RCTs are best suited to determine a single intervention’s net benefit under closely controlled circumstances, it is becoming increasingly apparent that they are not well suited to provide the best evidence for integrative options. Therefore, the CIM research community needs to explore other research approaches to provide high quality data for decision-makers about the effectiveness of CIM interventions in actual practice. The goal of the current symposium is to move toward agreement on key evidentiary principles of research that balance internal validity with important considerations of relevance, timeliness and feasibility.

Dr. Sean Tunis from CMTP followed with a short presentation on the current state of comparative effectiveness research in the national health policy agenda. While over 18,000 RCTs are published every year, there are still significant evidence gaps for decision-makers and end users of research in all clinical areas and specialties. The gaps (as seen by decision-makers) include highly selected patients that don’t reflect real populations, research settings that are not typical of settings where most care is delivered, the use of surrogate outcomes that have not been validated, and missing or incorrect comparators, among others. Essentially, there is a structural defect in the research enterprise, with no mechanism for decision-makers (including patients, clinicians, and payers) to communicate their needs and questions to those who generate evidence. CER has emerged as a potential approach for addressing this need, and some form of federal support for it certainly will be included in the new health reform package. However, basic aspects of CER including valid study designs, appropriate outcomes, and infrastructure are not yet established- and it is not well known how to best apply CER methods to research questions in various clinical areas to answer different questions.

Dr. Tunis also stressed that the varied perspectives of the group attending this symposium will prove crucial to reaching a practical consensus on how to approach evidence generation for decision-makers. For example, asking payers like CMS to communicate their evidentiary needs to researchers at the beginning of the research process is a new way to approach the generation of useful clinical knowledge. The CIM community can contribute much to the growing conversation on CER, given that integrative medicine embodies the type of complex, multi-modal, real-world decisions that CER is intended to address, and recommendations emerging from this group will prove very useful to researchers and clinicians in both complementary and allopathic care.

SESSION 1: EVIDENTIARY CHALLENGES CONFRONTING CIM

Presentation - The CIM Research Agenda: Dr. George Lewith led a discussion of current gaps in the evidence and the future research agenda for CIM. The use of complementary and integrative medicine in clinical practice is very common, especially in chronic, non-malignant conditions. Most of the decisions surrounding CIM in the UK are made by primary-care physicians, and they are influenced by a number of things other than published evidence, including opinion leaders, other physicians, product representatives, their own experience, and their patients. To make published evidence more useful and influential in primary care, we need to begin thinking about effectiveness, not efficacy, and move to research models that reflect this.

Dr. Lewith then reviewed the work of the King's Fund, a health care policy organization in the U.K., on a report released in 2008 regarding the challenges of investigating complementary medicine. This report stressed that new studies need to better consider factors such as the context of the intervention in question, understanding the nonspecific (placebo) effects vs. therapy-specific effects, and the importance of pragmatism in new research endeavors. Patients seek CIM because it is a patient-centered approach to treating a multitude of diseases. Therefore, it is even more important to incorporate the concerns of the public and patients when considering the research agenda for CIM. Dr. Lewith asserted that equipoise is often an important issue and may influence outcomes from RCTs. He also reviewed a trial in which patients were assigned to treatment along with an empathic consultation or no consultation, resulting in those receiving the consultation experiencing significant health improvements. The conclusion from this study is that we need to be very careful about what assumptions we make in regard to mechanisms of effect. If we create contextual changes for people, they may create their own physiological change. We are too reliant on thinking about how to use drugs to change molecules, and instead need to empower patients to change their own molecules. Essentially, we want to (ethically) maximize the non-specific effect, not minimize it. Finally, Dr. Lewith stressed that since CIM interventions are complex, whole systems, we need to think about how they may differ from conventional practice.

GROUP DISCUSSION: The group was interested in Dr. Lewith's discussion about maximizing the non-specific effect, with one participant suggesting that the real question might be to see if "normal care" is really a "nocebo" and is damaging a natural healing response. When patients begin to dread going to the doctor, their response to care worsens. When doctors use a manualized approach to interacting with patients, it is very rarely effective. This means that traditional NIH requirements for research funding make it difficult to incorporate understanding the real benefits of good doctor-patient interaction in CIM research. Another researcher agreed, asserting that the most important difference between PCTs and conventional trials is that PCTs leave room for that type of natural, spontaneous interaction. One clinician pointed out that changes in the health care system are making it more difficult to devote an adequate amount of time to any single patient, so it is going to be even more difficult to harness the potential power of these interactions. Additionally, patients are often reluctant to disclose or admit to pain, and don't get early and appropriate treatment for pain.

One participant noted that one of the most difficult parts of interpreting CIM research is that there is no well-established explanation for variability, either by a biological or psychological model. If the CIM community is required to establish explanatory models, some efficacy work may need to occur before moving on to clinical effectiveness research. One difficulty with enrollment in CIM studies is that people enroll in trials because they *want* a CIM therapy to work, which is likely to affect results. Conversely, disappointment in randomization to the conventional therapy arm may result in less therapeutic efficacy for those patients. Another participant suggested that the unit of analysis should be the health encounter, rather than the individual patient.

Symposium participants also discussed how to get new CIM research funded, especially by agencies like NCCAM, NIMH, and NCI, each of which holds a very different perspective on research criteria and design. When asked if it is going to be possible to create a set of guidelines that can be given to a grant review committee to use to evaluate research, participants from funding agencies tentatively agreed. However, some asserted that this still isn't going to change practice given that most CIM practitioners won't believe a negative result, and most allopathic practitioners won't believe a positive result. However, others contended that this is because the whole CIM research enterprise is currently set up to use RCTs for conventional medicine as a baseline for evaluation, and this is an opportunity for the CIM community to distinguish itself in the emerging CER movement. There is a risk that if the CIM community tries to assimilate itself with the allopathic CER movement too much, and becomes too similar to conventional medicine, it may lose its identity. Some researchers still asserted that the efficacy/RCT model does have real value, so we should not dismiss it entirely, despite recognizing the potential power of CER.

SESSION 2: THE STATUS OF CER IN ALLOPATHIC MEDICINE

The purpose of this session was to review current methods and study designs in CER, with a focus on how they have been used in allopathic research thus far, and how they might be applied to CIM research.

Presentation - Overview of CER: Dr. Hal Sox began the session by reviewing the work of the Institute of Medicine (IOM) Committee on Comparative Effectiveness Research to establish a list of initial national priorities for CER. While many agencies and researchers claim that they have been doing CER for a long time, there has never been a national dialogue before, and the efforts of the IOM committee will shape future work in the field. Dr. Sox reviewed the IOM's working definition of CER, emphasizing unique aspects of CER including 1) it incorporates both the generation of new research and the analysis of existing research, 2) it involves head-to-head comparisons, looking at benefits and harms of different treatment methods, 3) it is intended to help patients, clinicians, payers, and policy-makers make informed decisions, and 4) study populations must reflect real-world clinical practice. He noted that 3 of the 100 topics chosen by the IOM were focused on CIM interventions and that many others could include comparisons of CIM treatments to allopathic treatments. Finally, he reviewed a number of other recommendations that the IOM CER committee made regarding CER infrastructure, including more research and innovation in CER methods, more work on dissemination of CER findings, and the need to engage the public at all levels of CER.

Presentation - Pragmatic Clinical Trials: Dr. Sean Tunis followed with a presentation on methods for pragmatic (or practical) clinical trials (PCTs). The central notion of PCTs is that their purpose is to inform a particular decision, whereas the purpose of an explanatory trial is to determine the reason for an effect, usually under optimal conditions. Dr. Tunis then outlined the "pragmatic-explanatory continuum indicator summary" (PRECIS) framework for designing PCTs to inform health care decisions. PRECIS is a tool to guide research designers to consider such trial design factors as eligibility criteria, patient adherence, and practitioner expertise, among others, along a scale from "explanatory" to "pragmatic" during protocol development.

Presentation - Bayesian and Adaptive Trials: Dr. Bryan Luce then discussed the use of Bayesian and adaptive trial design methods in CER. He began by noting that while conventional trials are costly, take a lot of time, and often chase the wrong answer, pragmatic trials are going to be unsustainable without systematic changes in the operations and methods used for generating evidence for decision making. He then reviewed the “evidence hierarchy” for decision making, starting with 1) RCTs and meta-analyses of RCTs, 2) analytical observational studies, 3) non-analytical observational studies, 4) non-analytical studies, and then 5) expert opinion as the lowest in quality. However, he argued that the traditional reliance on the evidence hierarchy is flawed because it leads to an over-simplistic assessment of evidence quality. Bayesian trials provide an appealing approach to generating evidence. They are designed to combine prior evidence with new data, and they can produce predictive probabilities of trial success and individual patient results with different treatments. Furthermore, they are well-suited to answering questions that are more complex than a simple yes-or-no, and instead are able to answer questions about whom, when, and under what conditions therapies are effective. Ultimately, his point was that for coverage and other policy decisions, we may only need enough targeted information to “tip” a decision, a goal for which Bayesian trials may be very appropriate.

Dr. Luce also described adaptive trials, which are trials that change aspects of design based on prospective rules and the information that accrues as a trial takes place. These aspects may include adaptive sample sizes, randomization, accrual rates, and the ability to adapt to responding sub-populations, among others. The advantages of adaptive trials include the fact that they are often smaller, faster, and less costly, they have more targeted protocols, and they are able to closely mimic real-life decision making. In general, to develop useful CER evidence to inform decision-making, he would opt for “messy” CER effectiveness over the “pristine” RCT efficacy model.

Presentation - Registries and Observational Studies: The last presentation in this session was given by Dr. Rich Gliklich on the use of registries and observational studies in CER (OCER). There is significant skepticism about OCER given heterogeneity in the quality of studies and the fact that decision-makers are not yet familiar enough with observational research methods to discern those studies with greater or lesser risk of bias. A registry is defined as an organized system that uses observational study methods to collect uniform data, and most importantly, serves a predetermined decision-making purpose. Key characteristics of registries include collection of data in a naturalistic manner uniform for every patient, with specific and consistent data definitions. Some recent progress has occurred defining methods for good OCER and registries, particularly the release of a guide through AHRQ’s Effective Health Care program and the GRACE Initiative’s principles. Important considerations include whether the study plan was specified in advance of conducting the study, how valid the interpretation of CE for the population of interest is, and if the study was conducted, analyzed, and reported in a manner consistent with good practice. Dr. Gliklich also reviewed the value of cluster randomization, in which clusters of individuals rather than independent individuals are randomly assigned to a treatment- this is useful when natural variation will not produce the desired treatment groups.

GROUP DISCUSSION: The group discussion began with a number of questions about the use of registries. Noting that the IOM report specifically cited that a registry model should be used for studies of back pain, one participant queried why and how that was chosen as an appropriate model. Dr. Sox and Dr. Tunis, both of whom were on the IOM committee explained that it was the choice of particular committee members, and acknowledged that since the committee was given such a short amount of time to prepare the report, some of the choices would seem a little ad hoc. Dr. Tunis also acknowledged that there are numerous questions about back pain, and we need to do more work on determining optimal methods for studying different types of questions. Dr. Gliklich volunteered that a registry model would probably be useful in this case because it would allow for the study of variations in treatments that might be affected by study sites and referral patterns. Dr. Sox voiced concern that journal editors,

who are often responsible for deciding if data are valid, may not know how to interpret and incorporate registry trials given the large amount of missing data and poor follow up that frequently occurs. Dr. Gliklich countered that registries should establish the degree of follow-up a priori, and if they follow through, it would be considered an acceptable study design. It was also confirmed that practice-based research networks (PBRNs) would serve as a good source for developing a registry.

Symposium participants asked for a more detailed explanation of cluster randomization trials. Dr. Tunis explained it as randomizing at the level of the clinician or clinical practice, rather than the individual. One of the main benefits is that you don't necessarily need to control for the placebo effect, because the patient doesn't have to agree to be randomized. Additionally, it broadens the number of practices that can be included because it doesn't require each study site to perform all of the procedures. One of the patient representatives supported this technique, because patient recruitment is very difficult when there is randomization. In addition, methods that don't require patient consent for randomization would maximize the integrity of mind-body connection in CIM practice, which is an important aspect of treatment. Others pointed out disadvantages of cluster randomization, namely that it requires a larger sample size, it is hard to control for the quality of the doctor-patient interaction, and that payers may not accept it as a valid method for coverage and reimbursement decisions. Dr. Tunis replied to this last comment by acknowledging that payers are always going to default to RCTs if the community doesn't provide a consensus on an alternative with a united front. We need to understand what payers view as best methods for different types of questions, and the IOM's Clinical Effectiveness Research Innovation Collaborative (CER-IC) may be able to help with that. Symposium members also debated the value of adding biostatisticians to this type of conversation, but did not reach a conclusion.

Conversation then turned to the topic of how end-users actually make decisions based on evidence. One participant noted that patients tend to greatly value considerations of safety and risk, and PBRNs can be used to answer these types of questions. However, some participants argued that different funding agencies and organizations (FDA, IOM, AHRQ, etc.) use different standards for safety, and these need to be pulled together in a systematic way. Ultimately, the barrier to registries is funding, not that people doubt the value of them, because they will help us identify important questions to further study. One of the practicing clinicians also voiced concerns that the generation of new evidence is not sufficient to change practice patterns, and wondered if there is more research on dissemination and implementation other than just using coverage and reimbursement as a blunt policy tool. Dr. Tunis acknowledged that more dissemination research is needed. Some examples of other tools, such as computer programs to aid physician decision-making, were discussed, but it was largely agreed that more work is necessary. One of the payers also acknowledged lingering concerns that even evidence-based guidelines can lead to poor results, given some of the mistakes that have been made in the past. Dr. Sox steered the conversation back to the question of evidence needed for reimbursement and coverage. Health plan representatives stated that it really depends on what is being treated- for example, chronic pain is an important issue and health plans tend to be a lot more liberal about it. When asked, one health plan representative stated that they do not rely on actuarial data, and that they do their own research to minimize bias.

SESSION 3: DEFINING THE CONTEXT

This session was intended to present the contextual changes driving the need for new methods and evidence and CIM, and where the research community stands currently.

Presentation – Evidence Based Practice: Dr. Larry Green led a discussion on the value of “practice-based evidence”, or the idea that evidence generation needs to be centered on contexts and processes as much as on the intervention itself. He argued that current RCT-derived evidence is largely ungeneralizable because subgroup analysis fell out of favor in the 1980s and has been relatively discouraged since then. While current research concentrates on provider outcomes and variables, studies should emphasize what he terms “mediator” and “moderator” variables. Mediator variables are mechanisms through which the intervention influences the outcomes. Examples include behavior and environment. Moderator variables are those that alter the effect of the intervention on the mediator variables, or the mediator variables on the outcome variables. Examples include age, sex, SES, cultural values, and previous experience. Essentially, these are factors on which the intervention has little or no influence, but must be taken into account for the intervention to yield the intended effect. Dr. Green also discussed the need to reform the process for producing and vetting research, most notably in the priority-setting and peer review of federal research grants, because this process limits what gets published and included in guidelines, which often miss the mark of for decision-makers. We also need to develop a new kind of “practice (or policy) impact factor” for journals, so they can indicate how research will influence decision-making.

Presentation – SPICER: Dr. Anne McCaffrey then spoke about her work on a NIH Challenge Grant to assess the use of CIM in managing common medical conditions, called “Studies in Patient-based Informatics for CER (SPICER).” The study is using a new method of data collection from electronic medical records to look at useful clinical outcomes in IM. The four conditions of back pain, fibromyalgia, irritable bowel syndrome, and perimenopausal symptoms were chosen because they are commonly seen at integrative care centers, and while the project will produce important information on treating these four conditions, it is intended to be a first step toward a much larger enterprise, changing how clinical research is done. Essentially, SPICER is pioneering a novel method of data collection, which is typically done in a lockstep measurement pattern that makes it easy to analyze but limits the kind of information we can collect, and thus influences how studies are designed. The traditional approach to measurement does not reflect real life patient trajectories, so the SPICER investigators are pioneering a new data collection method called event-stream analysis (ESA). ESA allows data collection to parallel life experiences, which are not uniform and need to be interpreted in the light of what other events happened nearby in time. Dr. McCaffrey outlined the process for ESA, describing it as an iterative process in which the initial plan is tested to determine utility and then is continually refined. The first step is to identify what outcomes are to be measured and the variables needed to do this, and the second step is to use data from the EMRs to create a “fulcrum data set”. This is a way to organize data in which the unit of measure is the event, not the patient, and provides the raw materials for asking questions having to do with timing of various types of events, tracking specific values over time, and more. The third step is to determine the specific analysis needed to answer the research questions, use event stream statistical routines to carry out the analysis, and then repeat as necessary.

Advantages of ESA include the fact that it separates complex data extraction from data analysis, it can be used to combine information from many sources, and “events” do not need to be determined at the start of the study. ESA is new, but it is not complicated. In fact, it greatly simplifies data management and analysis, while producing possibilities for complex time-related analyses that are all but impossible in a lockstep system. ESA holds potential to create methodology that can develop neglected aspects of health and healing, such as patient choice and patient-practitioner interactions. The main disadvantage is that a higher level of programming is needed for ESA, but to the study will have a flexible plan for data extraction and analysis with an expert committee to sort out the issues inherent in a pilot project. The long-range goal of SPICER is to create a sustainable model for collaboration across multiple clinical centers for conducting focused observational comparative-effectiveness studies within and across multi-modality clinics, including more complex health care delivery systems and a very broad range of modalities.

GROUP DISCUSSION: One participant began the discussion by asking for a good example of how practice-based research has influenced conventional medicine. Dr. Green cited the example of tobacco use reduction via taxes, which started in California and Massachusetts, and then spread to other states once the practice-based evidence demonstrated that it was effective. Others had a number of questions for Dr. McCaffrey about the use of ESA, especially about how events might be analyzed in a manner similar to “episodes of care” as is done at RAND. Dr. McCaffrey confirmed that a fulcrum data set can be turned into episodes rather than events. Dr. Hal Sox pointed out that people will probably want to see study results analyzed in both the traditional way and by ESA until we better understand ESA. Others noted that this pre-supposes a large data set, but wanted to better understand at what point you can decide when to start the analysis. One participant noted that there are really two kinds of practice-based evidence, one from research evidence reflected in the practice setting, and the other in clinicians’ knowledge based on their practice. They tend to rely more on the latter than the former, and won’t necessarily adopt EBM, unless we can incorporate their experience in the research. However, ESA would be a great way to work with decision makers to find the most meaningful outcomes and ways to analyze them.

One of the patient representative pointed out that the patients themselves are often the most active decision-maker in CIM- they are usually the ones seeking out more information about alternatives, and physicians are not often doing enough to help them. One of the researchers mentioned that this is a difficult issue because we still don’t understand why patients seek what they seek, and given that patients are not required to use evidence there is still a fundamental disconnect between the patient world and the research world. Dr. McCaffrey mentioned that the ESA method allows researchers to put qualitative information into the data set, which would help with incorporating the patient perspective. Finally, one of the clinicians mentioned that it is going to be difficult to get clinicians to think about uncertainty when making decisions based on evidence, and a number of participants agreed that they consider biostatistics to often be misleading. Overall, the group did not come to consensus on the appropriate role of biostatisticians in analyzing research to help inform decision-making.

SESSION 4: BREAKOUT GROUPS

The beginning of the afternoon was devoted to breakout groups, in which participants were asked to consider the main features of CER trials for a specific clinical topic in CIM. Participants were asked to identify evidence gaps, consider the use of patient-reported outcomes, and talk about study designs. As Dr. Tunis explained, this group is well-positioned to come to a consensus about the best practices and a set of principles for generating knowledge for decision-makers, and the breakout session exercises would be helpful as a starting point for generating a set of concrete study design recommendations.

1. ACUTE PAIN – FACILITATOR RICHARD HAMMERSCHLAG

The group began with a discussion of the difference between acute and chronic pain, with some agreeing that anything after 4-6 weeks of pain can be considered chronic. Participants also addressed the possibilities of researching multi-modal treatments of acute pain. It needs to be established how the combined use of opiates and acupuncture might work. For example, opiates may negate the effects of acupuncture and disrupt the non-specific effects of treatment. Herbs were eliminated as a possibility because they are rarely used for acute pain. Group members also distinguished between pre-hospital and post-operative acute pain. Potential modalities to look at for pre-hospital treatments include touch, talking (especially because conversation gives patients the opportunity to make choices in their treatment), and breath induction. Since EMTs are the health providers that usually encounter patients with acute pain, we need to see what kind of skills they use and can be taught. For post-operative pain, studies should focus on techniques like hypnosis, physical therapy, and the reduction of pain medications.

Participants then addressed a number of issues related to researching CIM treatments for acute pain. Members agreed that placebos are not useful, and usual care would serve as the best control, with additional treatments as the variables. Group members also discussed potential study designs like crossover studies, adaptive and Bayesian modifications, and tracking every patient (via a journal) to compile a directory with precise treatment details (i.e. exact acupuncture points), and to clearly classify the responders and non-responders. The research community will need to address questions about how to adapt a study when a new treatment comes on the market, what ancillary providers see as acceptable endpoints, what kind of pre-even and post-even issues need to be reported, and what patient reported outcomes might look like. Other group members also suggested the utility of creating a new pain scale because the individual mean will vary, and so the research community needs a tool to evaluate the varied cultural responses to pain, and the family history of coping with pain. It would also be important to report adverse effects throughout all types of studies.

2. ARTHRITIS – FACILITATOR ANDY AVINS

The breakout group began with a discussion of outcome measures, because it is an aspect of trial design that ought to be common in trials looking at conventional or integrative treatments. Members agreed that they should be assessed and weighted primarily from the patient viewpoint, and that both single outcomes and surrogate outcomes have little use in these trials. Researchers should include a broad range of patient reported outcomes, assessing both quantitative and qualitative measures. They also felt that the PRO scale needs to be expanded further, and one researcher mentioned that there is some work being done for this. It is going to be important to use electronic medical records as much as possible, but will probably require work to improve the capture system for new qualitative tools. Another participant pointed out that it might be worthwhile to study the unanticipated, positive “side effects” of integrative treatments like increased peacefulness or centeredness, potentially using methods like qualitative interviews or developing new tools. CIM-related outcomes should include some measure of general wellness, beyond the simple absence of pain. Some also thought it would be worthwhile to look at mediators of effect, and self-efficacy was identified as one of the best predictors of outcomes. Finally, group members mentioned that assessing provider satisfaction also would be important.

Discussion then turned to the topics of potential treatment arms and research designs that could be used to study CIM-related therapies for arthritis. Among treatment arms, one researcher asserted that trials should allow for the possibility of staging treatments and the possibility of switching treatments, and will have to include analgesics as well as CIM therapies, because it can be very difficult to recruit individuals to a trial without pain medications. Control arms should reflect real-world practices, and should not involve placebos. Participants agreed that there should be an element of patient choice or preference, and that it is additionally important to track the frequency and purpose of visits. It will also be important to do careful work on subgroup analysis given that smaller groups tend to respond strongly, and are worth identifying. Other group members brought up the issue of patients using other treatments outside of trial assignments, and mentioned that the seeking of outside treatments could be assessed as an outcome in the trial. They also suggested that patients should keep a diary to record their daily activity levels, pain response, and any differences in pain during night and day. However, a proxy individual would probably have to provide additional information to corroborate, as patients can often over or under-estimate their changes.

3. BACK PAIN – FACILITATOR DAVID EISENBERG

Participants began by agreeing that there is no need for new trials of monotherapies for back pain, and looking towards integrative models (with either multiple CIM therapies or traditional therapies) would be more useful. This would best reflect real-world practices because more than half of people who use CIM therapies use more than one. Essentially, single-CIM therapy research is not going to be a pragmatic research model for the users who employ multiple techniques. The alternative, whole systems approaches, are multidisciplinary, complex treatment modalities that exist in practice and need to be reflected in research. However, participants questioned whether researchers will be able to work around the fact that variance will be difficult to explain due to multiple mechanisms of treatment. Many participants noted a risk of having iterative studies with multiple CIM outcomes, and that we need to consider how to use both passive and active therapies. Others pointed out that it was important not to “manualize” the treatment because it might reduce the effect size, and one participant suggested the use of new media to change how the CIM community disseminated new treatments without the risk of over-manualizing anything. There is a need to balance the collection of information that allows for replication, without gutting the core essence of the treatment.

Participants then discussed some of the issues surrounding dosing. As we add types of therapy, perhaps it reflects more time with valued providers, and so maybe the most valuable “intervention” is the chance to meet multiple providers and find one with whom a patient connects. However, group members were not sure how to measure this. It is also important to keep in mind the value of the role of the patient themselves, their ability to be participatory, and the therapy that comes from the provider-patient relationship in CIM. All of these may be best captured through “preference-centered” trials. Breakout group members then spent some time discussing potential study design methods. One researcher suggested creating a registry or database to observationally track outcomes, with a large group of researchers to pool data to avoid having to duplicate previous efforts. Some wondered if it would be possible to recruit and train members of the CIM community to maintain a registry, and it was agreed that it would probably have to be through the practice networks. It was also noted that there exists significant previous research of value, so it is going to be important to apply this in a meaningful way, especially when it comes to distinguishing subtypes of back pain with multiple unique variables.

Outcome measures are also a complicated aspect of designing a study. One group member raised the possibility of including global outcomes in addition to pain, quality of life, and other variables. It was not confirmed whether primary outcomes adequately reflect meaningful changes in a patient’s life. If patient satisfaction is distinct from pain, the research community needs to develop better outcome measures from the perspective of the patient. This would also allow the patient to become a member of the treatment team and harness the important psychological aspect of CIM therapies. Finally, the group addressed the use of studying cost effectiveness as well, because it would be valuable information for agencies like the VA and the DoD, but they anticipated a challenge in how to analyze it in combined therapies rather than monotherapy.

4. OTHER CHRONIC PAIN – FACILITATOR MICHAEL BAIME

Group members began by discussing endpoints to study in treating chronic pain. Potential tools for measurement include an analog pain scale, measurements of occupational performance or global performance, and economic measurement (both medical and societal cost), among others. It was also noted that there are important and relevant quality-of-life assessments beside tools to measure pain. Participants also agreed that it would be useful to find a way (probably by using medical anthropology methods) to assess the doctor-patient interaction, given that it may be as important as the treatment itself in affecting outcomes. Additionally, one member cited the use of quantitative measures derived from psychology research to assess mindfulness. Group members agreed that sham or placebo controls would not be acceptable, and potential alternatives would include different doses or different periods of extinction. These could be used to arrive at a mechanistic understanding of the treatment effect.

Participants then further discussed the possibilities of assessing the doctor-patient interaction, and how it might inform coverage decisions. Researchers agreed that it would involve a lot of provider and patient self-measurement, which could employ empowerment scales. Other participants mentioned that many of the patients who seek CIM therapies are more motivated than the average patient, and so it will be important to measure a patient's ability to "self-activate", and the provider's ability to facilitate that. The patient representative also mentioned that most current studies of chronic pain don't measure outcomes at important intervals, and a longer study would be useful. Based on earlier elements of the conversation, one group member proposed that a new trial should start with observational cohort studies to generate a set of hypotheses about what elements of a caring relationship are important, and then move into clinical studies. Another participant stated that since you can't know everything a priori, it would be important to have a staged research process, with iterative cooperative opportunities to redesign. This allows for flexibility in terms of where eventual study design will go and teases out the mechanistic questions to create a truly pragmatic study. As such, collaborative studies are also going to be important, and this method would allow researchers to maintain internal validity. Seed studies would most likely be funded through a purchaser, and not the government. However, group members could not agree on the utility of cluster randomization, and the payer representative cited particular concern about the amount of variability that might be unexplained. As another participant noted, under the current system, there is still a lot of unexplained variability. Finding a way to quantify it would be a significant contribution to the entire CER field.

CONCLUDING DISCUSSION

Dr. Margaret Chesney began the afternoon group discussion with a brief overview of common themes she observed throughout the day. She reaffirmed the belief that the CIM community has a great opportunity to utilize and advance the CER movement, given that both groups currently are engaged in similar dialogues. Some of the most specific recommendations that she noted addressed the need for studies that truly integrate usual care techniques and complementary techniques, so that trials should compare usual care with integrative medicine. This will allow researchers to look more at mechanism and process, and develop models for what we see in practice. She also stressed to the research community that CER is an opportunity for them to be truly innovative in their study designs, not only by integrating usual and complementary techniques in the comparator arms, but by finding ways to 1) do preference trials, 2) maximize non-specific effects, 3) study the beliefs, expectations, and interactions of patients and providers, and 4) include outcomes like cost, biological measures, and patient and provider reported outcomes, among others.

Dr. Hal Sox noted that a new CER institute will have an approximate budget of \$400 million a year, which is not very substantial, so the research community will probably have to do a lot of work with observational studies. This will require finding a way to embrace imperfect evidence. He also asserted that it will be important to keep in mind that outcomes are really a complex function of a number of variables- including characteristics of the patient, the doctor, their interactions, etc. and the distinction between the doctor-patient interaction in CIM vs. allopathic medicine may not be possible to isolate and study. He suggested that the group consider the role of longitudinal cohort studies, doing a phased investigation using observational methods to identify key variables, and then studying further using appropriate methods. Participants agreed that we have to acknowledge up front that there is not going to be a single “magic bullet” and that we cannot isolate each element of the interaction, and so we need methods that allow us to look at CIM practice as comprehensive and synergistic. Dr. Tunis agreed that this would be very different than the pharmaceutical/RCT model, but currently, payers tend to apply only that model to their policy and coverage decisions, even for complex clinical areas like CIM. One of the payer representatives acknowledged this, agreeing that it is hard for them to extricate themselves from this model, but they recognize that it needs to change, and discussions such as this symposium will aid in the process.

One of the researchers then pointed out that it will be expensive to test out these models until we find the optimal ones, and wondered who will pay for multi-modal experiments. Clinical researchers who use new CER methods are going to be marginalized until some federal agency or Congress steps up to advocate for additional funding. A participant from a federal funding agency cautioned the group that they will have to expand their network of collaborators because very few people know how to do the types of observational CER that have been discussed, and they need to be persuaded to collaborate on grant proposals for funding. Dr. Tunis countered that despite this, it is worth doing the work up front, and then proving its worth, calling this the “field of dreams” model of research. As presented today, there exists a broad range of CER methodologies, and the research community needs to develop an intelligent way of matching them to appropriate research questions. In particular, the CIM community needs to demonstrate the ability to do this well. Another researcher cautioned the group to not go from one narrow approach to another, but to be sure that the research community stays in touch with reality, focusing on where the real problems are in the evidence base. Others seconded this, and further advocated the need to view patients as valued collaborators rather than as study subjects, an area where CIM research could lead the field- and the use of practice-based and community-based participatory research networks would be an excellent place to start. Other participants suggested the use of social media and networking to systematically study the viewpoints of patients and consumers, and their interactions with both hospital and non-hospital providers.

Drs. Brian Berman and Sean Tunis then closed the conversation with a brief review of the broader goals of this collaboration and what they hoped to achieve with the participants at the symposium. This meeting was lauded as a rich conversation, and just the beginning of how this group can capitalize on the window that has opened up in the past year to really improve CER, find methods that get to the heart of the clinical encounter, and create a new paradigm for research and health that incorporate patient values and clinical expertise. They expressed the wish that the symposium participants have been similarly inspired, and will participate further as a working group to create a document or documents to provide evidence guidance for CER in CIM. Dr. Tunis proposed the next step of this project to be a draft white paper to capture the day’s conversation to share and get feedback on, but the evidence guidance documents will be the ultimate goal to guide research using methodologies that reflect the principles of patients, providers, payers, researchers, and research funders.