Special Symposium: Multimorbidity

Multimorbidity and Evidence Generation

Carlos O. Weiss, MD^{1,2}, Ravi Varadhan, PhD^{2,3,4}, Milo A. Puhon, MD, PhD⁵, Andrew Vickers, PhD⁶, Karen Bandeen-Roche, PhD, MS^{3,4}, Cynthia M. Boyd, MD, MPH², and David M. Kent, MD, CM, MSc⁷

¹Department of Family Medicine, College of Human Medicine, Michigan State University, East Lansing, USA; ²Division of Geriatric Medicine and Gerontology, Johns Hopkins University, Baltimore, USA; ³The Johns Hopkins Center on Aging and Health, Baltimore, USA; ⁴Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA; ⁵Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA; ⁶Department of Epidemiology & Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, USA; ⁷Predictive Analytics and Comparative Effectiveness (PACE) Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, USA.

Most people with a chronic disease actually have more than one, a condition known as multimorbidity. Despite this, the evidence base to prevent adverse disease outcomes has taken a disease-specific approach. Drawing on a conference, *Improving Guidelines for Multimorbid Patients*, the goal of this paper is to identify challenges to the generation of evidence to support the care of people with multimorbidity and to make recommendations for improvement. We identified three broad categories of challenges: 1) challenges to defining and measuring multimorbidity; 2) challenges related to the effects of multimorbidity on study design, implementation and analysis; and 3) challenges inherent in studying heterogeneity of treatment effects in patients with differing comorbid conditions. We propose a set of recommendations for consideration by investigators and others (reviewers, editors, funding agencies, policymaking organizations) involved in the creation of evidence for this common type of person that address each of these challenges. The recommendations reflect a general approach that emphasizes broader inclusion (recruitment and retention) of patients with multimorbidity, coupled with more rigorous efforts to measure comorbidity and comorbidity burden and the influence of multimorbidity on outcomes and the effects of therapy. More rigorous examination of heterogeneity of treatment effects requires careful attention to prioritizing the most important comorbid-related questions, and also requires studies that provide greater statistical power than conventional trials have provided. Relatively modest changes in the orientation of current research along these lines can be helpful in pointing to and partially addressing selected knowledge gaps. However, producing a robust evidence base to support patient-centered decision making in complex individuals with multimorbidity, exposed to many different combinations of potentially interacting factors that can modify the risks and benefits of therapies, is likely to require a clinical research enterprise fundamentally restructured to be more fully integrated with routine clinical practice.

KEY WORDS: evidence-based medicine; chronic disease; guidelines; comorbidity; clinical trials.

J Gen Intern Med 29(4):653–60

DOI: 10.1007/s11606-013-2660-5

© Society of General Internal Medicine 2013

INTRODUCTION

Most older people afflicted by one chronic disease experience more than one. Considering just five major chronic diseases among older adults in the U.S. who have one, in fact 53–85 % have at least two of those five conditions, with the prevalence depending on the index chronic disease chosen.¹ Doctors and other clinicians caring for these patients need to address the patients' problems in their full complexity. This includes taking into account the seriousness and severity of each patient's conditions in the context of patient goals, attending to the burdens and risks of polypharmacy and medical testing, managing functional or cognitive disabilities that might impact a patient's ability to adhere to therapy, accounting for gastrointestinal, hepatic or kidney disease that impact the pharmacokinetics or risks of therapy or any other way chronic conditions interact, either directly or through their treatments. Yet clinical research methods are not generally designed to study the ways conditions interact. For example, a review of trials reported in the *New England Journal of Medicine* found only 27 % made it clear whether a subgroup analysis was prespecified or post hoc.² It should not be surprising, then, that the resultant "evidence-based" clinical practice guidelines (CPGs) may be absurdly unachievable, or potentially harmful, in patients with multiple chronic conditions.³

The large and growing footprint of multimorbidity demands an empirical evidence base that can directly inform practice through CPGs that take conditions other than the index condition into account.^{2,3} The goal of this paper is to identify challenges to the generation of such evidence, to support the care of people with multimorbidity and to advance a set of recommendations for consideration by investigators and others (reviewers, editors, funding agencies, policymaking organiza-

Received December 14, 2012

Revised April 10, 2013

Accepted September 4, 2013

Published online January 18, 2014

tions) involved in the creation of evidence for this common type of person. Here, we focus on improving the generation, analysis and reporting of primary data, while companion papers discuss evidence synthesis and integration,⁴ and the development of CPG recommendations themselves.⁵

METHODS

To provide a comprehensive list of challenges, a subset of investigators (CW, RV, DK) created an inventory of challenges raised by the presence of multimorbidity during the generation of evidence used by clinical practice guidelines, based on a review of investigator article libraries. The scope and terminology of the list was amended at an in-person research retreat by the nine-member Multimorbidity and CPG Study Group. Next, we performed a structured, limited literature review that examined papers from two symposia on multimorbidity^{6–10} and all articles citing papers from the symposia identified using ISI Web of Science.¹¹ Additional articles, including more recent literature, were included on an ad hoc basis. Our goal required that we survey the several topics described below, making a single systematic review unfeasible.

This inventory of challenges was reviewed and critiqued at a second in-person research retreat by all Improving Guidelines for Multimorbid Patients Study Group investigators (CB, SD, DK, BL, JS, TT, KU, RV, CW). After further editing, the list was sent to outside experts (AV, KB-R, MP) for review and critique. The revised inventory was discussed at a conference on Improving Guidelines for Multimorbid Patients (Baltimore, MD, October 2010) attended by researchers from various disciplines (medicine, public health, biostatistics), and stakeholders from government, other payors and industry. A goal of this conference was to encourage discussions that spanned the procedures inherent to evidence generation, evidence synthesis and guideline development. The final set of challenges and recommendations was created after incorporation of ideas discussed at the conference.

RESULTS

Challenges

Three broad categories of challenges were identified: 1) defining and measuring multimorbidity during evidence generation for clinical practice; 2) the effects of multimorbidity on study design and implementation in both randomized controlled trials and observational studies; and 3) reliably describing heterogeneity of treatment effect related to multimorbidity. These are summarized in Table 1 and discussed immediately below.

Defining and Measuring Multimorbidity During Evidence Generation for Clinical Practice. Perhaps the most basic

challenge in multimorbidity research is achieving clarity regarding the concept or concepts of multimorbidity that should be measured to address a particular question. Determining who has zero or just one chronic disease, versus more than one, depends on how the set of qualifying chronic diseases are selected. There is currently no standard list of diseases to be considered in a disease count measure of multimorbidity.^{1,12–14}

In addition, non-disease-specific conditions or measures of function (hearing, vision, gait, continence, etc.) are likely to be relevant for a given research question and may be desired as part of a multimorbidity measure^{15,16} or a complexity measure.¹⁷ Thus, substantial scientific judgment, which is to say subjectivity, is necessarily used to measure multimorbidity.

There are numerous multimorbidity measures that have been proposed to capture underlying sickness.^{18,19} As a group, they present a bazaar of indicators of disease severity, non-disease-specific conditions, functional status, receipt of treatment and utilization, and other items. It follows that the criteria used to validate these pre-tested

Table 1. Summary of Challenges to Evidence Generation for People with Multimorbidity

1. Definition and Measurement of Multimorbidity Status
There are several distinct concepts of multimorbidity: specific disease patterns, latent health status, or comorbidity.
Multimorbidity can be measured several ways: on presence of diseases alone or also using disease severity; conditions; measures of function; treatment or utilization.
Multimorbidity status or severity may be ascertained in various ways: self report, physician report, clinical examination, administrative, pharmacy, lab data, other.
Measurement of multimorbidity status may not be done comparably across studies and across research and clinical settings.
2. Multimorbidity-related Effects on Study Design and Implementation
Internal validity
Multimorbidity may lead to misclassification of treatment or inaccurate measure of treatment intensity (e.g. initiation—identification of ‘time zero’, duration, dosage).
Multimorbidity may affect selection of treatments or treatment intensity.
Multimorbidity may lead to confounding and interactions due to concomitant treatments.
Multimorbidity may lead to misdiagnosis of outcomes.
Multimorbidity may increase the likelihood of losses to follow-up.
External validity
Multimorbidity poses challenges for selection of participants from the at-risk population, especially recruitment.
Multimorbidity may affect adherence to treatment.
Multimorbidity may alter real-world effectiveness of treatments due to harms and competing risks.
3. Heterogeneity of Treatment Effect
People with multimorbidity are frequently excluded from trials—sometimes appropriately, sometimes inappropriately.
Even when people with multimorbidity are included in trials, summary treatment effects may not apply.
Multidimensionality is a fundamental problem when considering multimorbidity, as the number of potential multimorbidity-related questions are too large to be answered in any given clinical trial/trials.
A multiplicity of outcomes may be important when patients with multimorbidity are enrolled.
Subgroup analyses frequently yield false positive or false negative results due to multiple comparisons or low statistical power.
Multimorbidity-related multi dimensionality makes it challenging to pre-specify a small number of clinically important hypotheses related to subgroups and outcomes.

measures are also varied. They include prediction of mortality, health care utilization, institutionalization, functional status, quality of life and more. Thus, the selection of an appropriate measure should be driven by an explicit rationale for whether multimorbidity is being measured to study specific hypotheses, identify treatment effect modification, adjust for confounding, improve prediction or for another purpose.

We are unaware of any general framework that would help reduce the analytic complexity of the potential interactions between an index condition and comorbidities. In the case of diabetes, a framework has been proposed that categorizes comorbidities into typologies,²⁰ including clinically dominant conditions; concordant versus discordant conditions; and symptomatic versus asymptomatic conditions. While such typologies may ultimately prove useful for prioritizing care, it is at this point difficult to see how it might be applied to help reduce the analytic complexity of interaction effects potentially at play in comparative effectiveness research.

Additional technical challenges are also described in Table 1.

Multimorbidity-Related Effects on Randomized, Controlled Trial or Observational Study Design and Implementation.

While the trial is the most reliable method for obtaining a causal estimate, practical and ethical considerations may preclude the conduct of trials in certain types of people with multimorbidity, particularly after a narrowly focused trial has established the efficacy of a treatment.²¹ Therefore, it is essential to consider observational studies as a source of evidence for people with multimorbidity.

Observational studies are essential given the practical limitations of trials, but the absence of a focused design leaves them highly susceptible to multimorbidity-related sources of systematic bias, including measurement error and missing information on treatments, outcomes and prognostic variables.^{22,23} Further, the treatment selection process is extremely difficult to understand in observational studies, i.e. how treatment choices are made by or made for the person are only partially decipherable and the presence of multimorbidity can exacerbate confounding by indication. While prospective cohorts have the potential to address some of the concerns regarding measurement error and missing information that often mar other types of observational studies, it does not fully address treatment selection bias.

Multimorbidity can affect the internal and external validity of evidence generated by trials and observational studies. It can affect internal validity by increasing misclassification of treatment or inaccurate measure of treatment intensity; by influencing selection of treatment leading to confounding by indication; by contributing to misdiagnosis of outcomes; by interactions due to concomitant treatments; by modifying treatment responsiveness

and/or treatment effects; or by increasing the likelihood of losses to follow-up. In addition, multimorbidity provides challenges to external validity by affecting selection of participants from the at-risk population; by association with adherence to treatment; and by altering the real-world effectiveness of treatments due to harms and competing risks.

Heterogeneity of Treatment Effect. The fact that people with multimorbidity may have multiple chronic conditions, each with a spectrum of subtypes and severities as well as different treatments, means that people with multimorbidity have more ways to be dissimilar both to one another and to people without multimorbidity. When these differences meaningfully affect the risk of harm from a treatment and treatment benefit, it can be difficult to sort out this balance for an individual. For example, given just two chronic diseases and a single treatment for each, there are 16 possible exposure patterns (half of which may be relevant for CPGs) and four relevant interactions (disease A x disease B; disease A x treatment B; treatment A x disease B; treatment A x treatment B). Examples of interactions in each of these categories with the potential for clinically meaningful consequences are shown in Table 2 for chronic obstructive pulmonary disease. To account for all possible disease and treatment interactions in a real clinical population is analytically impractical if the goal is hypothesis testing rather than exploration. A comprehensive list for just a single condition would require a major effort, though the task might be somewhat simplified by studying the most prevalent chronic condition patterns in index conditions.²¹

Given the myriad ways that co-occurring chronic conditions might influence treatment effect, it should not be surprising that complex patients with multimorbidity are often excluded from clinical trials.^{24–26} Such people may be excluded by the protocol both directly (via condition-

Table 2. Sources of Multimorbidity-Related Heterogeneity of Treatment Effect (and Examples Using Chronic Obstructive Pulmonary Disease [COPD])

1. Disease A x Disease B
Example: Presence of congestive heart failure may exacerbate shortness of breath, altering measured effect of therapy on COPD.
2. Disease A x Treatment B
Example: COPD may be worsened by β -blocker medications for congestive heart failure or hypertension, altering the measured effect of therapy.
3. Treatment A x Disease B
Example: Steroids for COPD may worsen diabetic hyperglycemia or osteoporosis, altering the side effect profile.
4. Treatment A x Treatment B
Example: (index therapy affected by another therapy) Theophylline levels for COPD may be increased if co-administered with cimetidine, altering efficacy and safety.
Example: (index therapy affecting another therapy) steroids for COPD may increase international normalized ratio in people on warfarin.

specific exclusion criteria) or indirectly (e.g. through age-related or treatment-related exclusion criteria). They may also be excluded from trials through recruitment failure or by cautious investigators. For efficacy trials concerned with limiting toxicity for relatively untested therapies and with measuring treatment effect under conditions suited to optimal assay sensitivity, exclusion of multimorbid patients is often scientifically justified.

However, it cannot be assumed that treatment effects from highly selective efficacy trials apply to people with multimorbidity. Indeed, if such an assumption were valid, there would be little motivation or justification for excluding such people in the first place. There have recently been increasing calls for more pragmatic trials that enroll people more representative of those treated in clinical practice.^{27–29} It is widely believed that the results will be more generalizable to “real world” routine clinical practice. However, the results of such trials have to be interpreted carefully. Recent work has underscored that in the presence of clinically important heterogeneity of treatment effects, the summary result of a clinical trial can be misleading regarding many included patients.^{26,30,31} Including people with multimorbidity is likely to exacerbate the problem of within-trial heterogeneity of treatment effect, for the reasons discussed above. Indeed, the argument for including such people in clinical trials is precisely that the effect size for benefit or harm measured in healthier people may not apply to those with chronic diseases. Including subpopulations within a clinical trial where substantial heterogeneity of treatment effect is anticipated may thus yield a single summary result not as directly applicable to any individual as in a homogenous trial sample.³²

Heterogeneity of treatment effect within clinical trials is typically addressed through subgroup analysis. However, subgroup analysis is well known to be prone to false positive and false negative results due both to multiple hypothesis testing and low statistical power.^{33–35} Powering a study to test whether treatment effect is different according to the presence or absence of just a single chronic condition will require larger sample sizes in most cases, compared to powering for a main effect. Table 3 shows that the sample size required to test a

treatment effect interaction hypothesis increases by a factor of four to five if half the sample has a chronic disease of interest, and by a factor greater than ten if only one-fifth does. Testing more than one multimorbidity-related subgroup effect requires still larger samples.

Since people can have so many different combinations of multimorbidity and because any additional chronic condition or multimorbidity pattern can potentially influence the risks and benefits of treatment in different ways, the conundrum of multidimensionality is especially relevant when considering the evidence base for people with multimorbidity. While recommendations and frameworks for analyzing heterogeneity of treatment effects have been proposed^{36–39} to partially address these issues, as yet there is no clear consensus on how best to cope with heterogeneity of treatment effect in the context of clinical trials and in many—perhaps most—circumstances, it will simply not be possible to address every potentially clinically important multimorbidity-related question in the context of a randomized trial. This again implies that prioritization of multimorbidity-related questions, and other dimension reducing strategies, will be crucial. It further implies that we should anticipate gaps in the evidence that randomized controlled trials can produce and that observational studies and physician judgment will be needed to fill these.

Recommendations

Based in part on the barriers identified above, the conference on Improving Guidelines for Multimorbid Patients provided a dedicated opportunity to discuss the recommendations for how evidence generation might be modified in order to best inform the downstream processes of evidence synthesis and guideline development. Our initial set of recommendations focused on actions researchers should currently consider when performing a trial or observational study, to make their results more relevant to the multimorbid patient. These directly address the challenges identified above. However, we recognized that adhering to these recommendations was likely to have only

Table 3. Sample Sizes Required to Achieve Adequate Power for Main Effects, Versus Interaction Effects

Proportion of Sample with Multimorbidity	Subgroup Analyses (No.)	Sample Size Needed for Test of Main Treatment Effect (A)	Sample Size Needed for Test of Treatment Effect Modification by Multimorbidity (B)	Sample Size Multiplier to Study Multimorbidity Interaction (B / A)
0.0	n/a	800	n/a	n/a
0.2	1	800	9,700	12.1
	2	1,000	11,400	11.4
	3	1,100	12,800	11.6
0.5	1	1,300	6,400	4.9
	2	1,600	7,600	4.8
	3	1,800	9,100	5.1

Assumptions are: study power=0.8; the main treatment effect is a RR of 0.7; multimorbidity increases the outcome by a RR of 1.25; multimorbidity has a synergistic effect, increasing the treatment effect by a factor of 0.8 from 0.7 to 0.56; the baseline hazard rate is five events per 100 person-years; proportion of sample censored is 0.6. For two or three subgroup analyses, the family-wise type I error probability is kept at 0.05 according to a Bonferroni correction. Estimates are based on 10,000 simulations using the Cox proportional hazards model

a relatively modest impact on the evidence base. We therefore proposed research priorities to address critical methodological issues and a map that points to infrastructural changes necessary for more robust evidence generation. Even though these recommendations were derived from a conference of experts from a variety of fields, they are limited by the need to cover a number of relevant topics. Had it been possible to conduct formal systematic literature reviews for all of these topics, additional challenges and recommendations may have been identified.

Recommendations for Researchers. Recommendations for researchers are summarized in Table 4. These recommendations reflect a general approach that emphasizes the need for trials to be more broadly inclusive of patients with multimorbidity. At the same time, this more pragmatic approach will only be more useful for the care of multimorbid patients if accompanied by more rigorous measurement of comorbidities and multimorbidity and more rigorous estimation of comorbidity-related heterogeneity of treatment effects. In turn, more rigorous examination of heterogeneity of treatment effects requires careful and explicit attention to prioritizing the most important comorbid-related questions, and also requires studies that provide greater statistical power than conventional trials have provided. As discussed below, the need for greater statistical power implies that conventional randomized trials alone will need to be supplemented by other and newer approaches, including greater reliance on observational studies, cross-design synthesis and clinically integrated trials,⁴⁰ which might permit larger more representative samples to be included at lower costs.

Methodological Research Priorities. Future work should attempt to validate the recommendations in Table 4 through their use in worked examples across a variety of diseases, and their application in trials and observational studies. Other identified priorities for new areas of research included 1) to develop consensus and framework for the routine assessment of comorbidities and general measures of multimorbidity in clinical practice and clinical trials and in clinical practice; 2) to understand common multimorbidity patterns that may account for a large proportion of the multimorbidity population; 3) to develop a framework for the assessment and reporting of heterogeneity of treatment effects that yields credible subgroup results; and 4) to develop cross-design synthesis methods (i.e. methods that use information from different study designs) for the valid extrapolation of clinical trial results to excluded populations.

In order to ensure comparability across studies, researchers require guidance on which set of comorbidities to measure and how to measure comorbidity burden. Relevant regulatory or policy bodies (e.g. the Food and Drug Administration, the National Institutes of Health, the

Table 4. Provisional Recommendations for Developing Evidence for People with Multimorbidity

1. Definition and Measurement of Comorbidities and Multimorbidity Status
 - Identify, define, measure and routinely report the within-study prevalence of important and common comorbid conditions. As a general rule, this should also involve measuring and reporting summary metrics of multimorbidity.
 - Explicitly pre-specify the specific motivation for studying multimorbidity so that the most appropriate multimorbidity concept is measured.
 - Select a multimorbidity measurement instrument that best measures the most relevant aspect of multimorbidity for the research question of interest.
 - Based on specific motivation for studying multimorbidity, non-disease conditions should be considered.
 - Multimorbidity measurements should ideally be easily and reliably obtained in real world clinical practice.
 - Multimorbidity should be measured in ways that allow comparability of results across studies.
2. Multimorbidity-Related Effects on Study Design, Implementation and Analysis
 - Internal validity
 - Fully ascertain exposure to treatments (including onset, intensity and duration) and important multimorbidities with minimal bias and error.
 - Employ sound statistical methods, as well as external validation studies, to address treatment selection bias affecting internal validity.
 - Ascertain exposure to concomitant treatments and harms related to any treatments.
 - Design study procedures to avoid mis-diagnosis of outcomes in people with multimorbidity.
 - Pre-specify all likely treatment effect modifiers related to multimorbidity and implement study designs and procedures to deal with these.
 - Consider how multimorbidity will affect overall study power (especially outcome event rates) and whether study power should be calculated in light of important multimorbidities.
 - Design study to minimize losses to follow-up, especially in people with multimorbidity.
 - External validity
 - Study entry criteria should not explicitly or implicitly exclude people with important multimorbidity from trials unless necessary.
 - Design study procedures to recruit and retain people with important multimorbidity into trials.
 - Consider person-level and provider or environmental modifiers of treatment effect during study design and implementation, focusing on methods that will maximize applicability.
 - Monitor adherence to treatment rigorously to avoid misclassification according to multimorbidity status.
 - Consider distribution of competing risks in target populations and how it should be represented in study samples.
3. Heterogeneity of Treatment Effect
 - Identify and prioritize multimorbidity-related questions, obtaining expert clinical input where necessary.
 - Analyses should examine treatment effect modification across groups defined by important clusters of characteristics (e.g. by risk or by comorbidity burden), rather than only one-variable-at-a-time subgroup analysis.
 - Consider powering to support multimorbidity-related subgroup effects.
 - Consider stratification of randomization by important multimorbidity-related subgroups.
 - Consider multimorbidity-specific trials when indicated.
 - Collect relevant multimorbidity-related outcome data, particularly related to adverse events.
 - Define subgroups of interest for confirmatory, descriptive or exploratory analyses.
 - Subgroups for confirmatory analyses should be very few; fully pre-specified; explicitly justified; based on strong a priori pathophysiologic or empirical evidence of heterogeneity of treatment effects; adjusted for multiplicity if appropriate; and, always reported and labeled.

Institute of Medicine) should establish a minimum set of comorbidity measures, while being attentive to the costs

and burdens of collecting this information. These can be extended by disease-specific specialty societies and researchers would not be restricted from applying additional measures of comorbidity related to their specific hypotheses.

In addition, high priority was given to the need for better ways to tease out how treatment effects are modified by comorbid diseases and their treatments. An important concept is that patient-centered analysis of heterogeneity of treatment effect should consider how treatment effect is modified across patients defined by different clusters of characteristics, rather than in the conventional one-variable-at-a-time-fashion.³⁶ These clusters might be based on fundamental risk dimensions most likely to influence treatment effect (such as prognostic scores predictive of outcome risk or comorbidity scores predicting of competing risks).³⁶ Frameworks grouping comorbidities into categories according to how they may interact with the index condition may also prove useful,²⁰ if further defined. The main proposal emerging from this discussion was a classification scheme for subgroup analyses that clarifies how subgroup analyses may be used for different purposes. Confirmatory (or primary) subgroup analyses and exploratory subgroup analyses have previously been described.^{36,38,39}

Descriptive subgroup analyses describe a category that may be particularly relevant to address multimorbidity-related heterogeneity of treatment effect.^{38,39} Descriptive subgroups are those analyses that should be performed routinely.⁴¹ They are best defined through a consensus process such that the subgroups that constitute the descriptive set are chosen in order to promote evidence synthesis across studies, and are not based on study-specific hypotheses. These analyses are defined a priori, and are fully pre-specified (i.e. ideally at the start of the trial, but definitely before the analysis of any data). This category is distinguished from other exploratory analyses by virtue of always being intended for information synthesis, i.e. later meta-analysis, and derived through consensus regarding what should be done across related trials. It is expected that through this type of analysis, clinical experts and guideline developers will influence trial processes, at least by defining the *minimum* set of descriptive subgroups for a given index disease. Such descriptive subgroup analyses may also be deposited in an open access database as an alternative to conventional reporting in scientific journals. Thereby, investigators planning new trials or observational studies and meta-analysts could benefit from a complete evidence base with respect to potential effect modifiers.

Concluding Remarks and a Map to Generate Robust Evidence for People with Multimorbidity. Because the trial is adopted as a tool in medicine largely to develop and

evaluate new medical products, most trials are small phase 1 or phase 2 studies, and most phase 3 trials are industry-sponsored experiments specifically designed to satisfy regulatory requirements to support an efficacy claim and bring a product to market.⁴² Pragmatic designs will typically attenuate average treatment effect as protocols are liberalized to include patients, settings and care paradigms that deviate from the conditions that are suited to optimize assay sensitivity. It is unrealistic to expect pivotal efficacy trials to be optimized to measure treatment effects across all patients and settings, since a null trial, performed with an extreme pragmatic design, is not informative as to whether or not an intervention can work. For example, it may be that a treatment does not work in an average setting, but is effective when patients are better selected in terms of meticulous attention to diagnosis, adherence, disease severity or the absence of particular comorbidities or co-treatments, or in settings with better monitoring or a higher commitment to and/or more experience with the intervention.⁴³

However, prior to widespread dissemination and adoption of interventions beyond the efficacy population and setting, trials in populations where equipoise remains (defined by the absence of robust data) should be expected, demanded and encouraged or required by regulatory agencies (e.g. the Food and Drug Administration, the European Medicines Agency, etc.), the National Institutes of Health, journal editors and reviewers and professional societies. Because the purpose of these trials is not to establish an overall effect, but to evaluate modification of treatment effects across major subgroups or settings, these trials may need to be much larger than conventional trials. Such trials are only feasible when the clinical research and health delivery enterprise are more fully integrated, such that the incremental costs of research are minimized.^{40,44,45} Data collection, including outcome ascertainment, should be integrated with clinical practice through electronic health records.

Given the massive amount of data required for these analyses, this disruptive change to the clinical research enterprise⁴⁴ will need to be accompanied by new concepts for data ownership that facilitate and encourage data sharing.^{46–48} Thus, it is encouraging that there are recent examples of providing open access to clinical trial data,^{49–52} including directly by industry,⁵³ and also proposals that would make such data available on a routine basis.⁵⁴

That the current evidence-base for people with multimorbidity is deficient is broadly recognized.^{3,55} It seems clear that addressing this challenge requires a major commitment by the different organizations and individuals involved at each step in the generation and synthesis of evidence used in clinical practice guidelines.

Acknowledgements:

Developing Guidelines for Multimorbidity Core Group: Cynthia M. Boyd, Johns Hopkins Medical Institutes (JHMI), Sydney M. Dy, Johns Hopkins School of Public Health (JHSPH), David M. Kent, Tufts Medical Center (TMC), Bruce Leff, JHMI, Thomas A. Trikalinos, Brown University, Katrin Uhlig, TMC, Ravi Varadhan, JHMI, Carlos O. Weiss, JHMI.

Evidence Generation Work Group: Lawrence Appel, JHU, Karen Bandeen-Roche, JHU, Erica Breslau, NCI, Joseph Cappelleri, Pfizer, Dean Follmann, NIAID, Sherrie Kaplan, University of California-Irvine, David Kent (Co-PI), TMC, Todd Lee, University of Illinois at Chicago, Matthew McNabney, JHU, Erin Murphy, JHU, Milo Puhon, Johns Hopkins University (JHU), Sergei Romashkan, NIA, A. Sedrakyan, FDA/Cornell, Vincenza Snow, Pfizer, Ravi Varadhan (Co-I), JHU, Carlos Weiss (Co-I), JHU.

Funding: This manuscript was partially funded by grants R21 HS18597 and R21 HS017653 from the Agency for Healthcare Research and Quality and UL1RR025752 from the National Institutes of Health.

Conflict of Interest: The authors declare they do not have any conflicts of interest.

Corresponding Author: David M. Kent, MD, CM, MSc; Predictive Analytics and Comparative Effectiveness (PACE) Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, 800 Washington Street, Box 63 Boston, MA, 0211, USA (e-mail: Dkent1@tuftsmedicalcenter.org).

REFERENCES

- Weiss CO, Boyd CM, Yu Q, Wolff JL, Leff B. Patterns of prevalent major chronic disease among older adults in the United States. *JAMA*. 2007;298:1160–2.
- Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine—reporting of subgroup analyses in clinical trials. *N Engl J Med*. 2007;357:2189–94.
- Boyd CM, Darer J, Boulton C, Fried LP, Boulton L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA*. 2005;294:716–24.
- Trikalinos TA, Segal JB, Boyd CM. Addressing Multimorbidity in evidence integration and synthesis. *J Gen Intern Med*. 2013, doi: 10.1007/s11606-013-2661-4.
- Uhlig K, Leff B, Kent DM, Dy S, Brunnhuber K, Burgers JS, Greenfield S, Guyatt G, High K, Leipzig R, Mulrow C, Schumacher K, Schunemann H, Walter LC, Woodcock J, Boyd CM. A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med*. 2013, doi:10.1007/s11606-013-2659-y.
- Werner RM, Greenfield S, Fung C, Turner BJ. Measuring quality of care in patients with multiple clinical conditions: summary of a conference conducted by the Society of General Internal Medicine. *J Gen Intern Med*. 2007;22:1206–11.
- Lash TL, Mor V, Wieland D, Ferrucci L, Satariano W, Silliman RA. Methodology, design, and analytic techniques to address measurement of comorbid disease. *J Gerontol A Biol Sci Med Sci*. 2007;62:281–5.
- Boyd CM, Weiss CO, Halter J, Han KC, Ershler WB, Fried LP. Framework for evaluating disease severity measures in older adults with comorbidity. *J Gerontol A Biol Sci Med Sci*. 2007;62:286–95.
- Yancik R, Ershler W, Satariano W, Hazzard W, Cohen HJ, Ferrucci L. Report of the national institute on aging task force on comorbidity. *J Gerontol A Biol Sci Med Sci*. 2007;62:275–80.
- Karlamangla A, Tinetti M, Guralnik J, Studenski S, Wetle T, Reuben D. Comorbidity in older adults: nosology of impairment, diseases, and conditions. *J Gerontol A Biol Sci Med Sci*. 2007;62:296–300.
- Reuters T. Web of Science 2010. Available at: http://thomsonreuters.com/content/science/pdf/Web_of_Knowledge_factsheet.pdf. Accessed September, 2013.
- Guralnik JM. Assessing the impact of comorbidity in the older population. *Ann Epidemiol*. 1996;6:376–80.
- Guralnik JM, LaCroix AZ, Everett DF, Kovar MG. Aging in the Eighties: The Prevalence of Comorbidity and its Association With Disability. Hyattsville: National Center for Health Statistics. 1989;1989:170.
- Verbrugge LM, Lepkowski JM, Imanaka Y. Comorbidity and its impact on disability. *Milbank Q*. 1989;67:450–84.
- Crews JE, Campbell VA. Vision impairment and hearing loss among community-dwelling older Americans: implications for health and functioning. *Am J Public Health*. 2004;94:823–9.
- Guralnik JM. The impact of vision and hearing impairments on health in old age. *J Am Geriatr Soc*. 1999;47:1029–31.
- Safford MM, Allison JJ, Kiefe CI. Patient complexity: more than comorbidity. the vector model of complexity. *J Gen Intern Med*. 2007;22(Suppl 3):382–90.
- de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity. a critical review of available methods. *J Clin Epidemiol*. 2003;56:221–9.
- Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases—a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci*. 2011;66:301–11.
- Piette JD, Kerr EA. The impact of comorbid chronic conditions on diabetes care. *Diabetes Care*. 2006;29:725–31.
- Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. *BMC Med Res Methodol*. 2006;6:18.
- Green SB, Byar DP. Using observational data from registries to compare treatments: the fallacy of omnimetrics. *Stat Med*. 1984;3:361–73.
- Byar DP. Why data bases should not replace randomized clinical trials. *Biometrics*. 1980;36:337–42.
- Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA*. 2007;297:1233–40.
- Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: do they have external validity for patients with multiple comorbidities? *Ann Fam Med*. 2006;4:104–8.
- Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q*. 2004;82:661–87.
- Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials*. 2009;10:37.
- Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *CMAJ*. 2009;180:E47–57.
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003;290:1624–32.
- Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA*. 2007;298:1209–12.
- Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet*. 1995;345:1616–9.
- Bach PB, Gould MK. When the Average Applies to No One: Personalized Decision Making About Potential Benefits of Lung Cancer Screening. *Ann Intern Med*. 2012;157:571–3.
- Furberg CD, Byington RP. What do subgroup analyses reveal about differential response to beta-blocker therapy? The Beta-Blocker Heart Attack Trial experience. *Circulation*. 1983;67:198–101.
- Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey SG. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technol Assess*. 2001;5:1–56.
- Tannock IF. False-positive results in clinical trials: multiple significance tests and the problem of unreported comparisons. *J Natl Cancer Inst*. 1996;88:206–7.
- Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials*. 2010;11:85.
- Weiss CO, Segal JB, Boyd CM, Wu A, Varadhan R. A Framework to Identify and Address Heterogeneity of Treatment Effect in Comparative Effectiveness Research. Rockville: MD; 2010.

38. **Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO.** A framework for the analysis of heterogeneity of treatment effect in patient-centered outcomes research. *Journal of Clinical Epidemiology*. 2013. (In Press).
39. **Varadhan R, Stuart EA, Louis TA, Segal JB, Weiss CO.** Review of Guidance Documents for Selected Methods in Patient Centered Outcomes Research: Standards in Addressing Heterogeneity of Treatment Effectiveness in Observational and Experimental Patient Centered Outcomes Research. A Report to the PCORI Methodology Committee Research Methods Working Group. 2012. March 29, 2012.
40. **Vickers AJ, Scardino PT.** The clinically-integrated randomized trial: proposed novel method for conducting large trials at low cost. *Trials*. 2009;10:14.
41. **Wizeman TM IoM.** Sex-Specific Reporting of Scientific Research: A Workshop Summary. Washington DC: The National Academies Press; 2012.
42. **Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A.** Characteristics of clinical trials registered in ClinicalTrials.gov, 2007–2010. *JAMA*. 2012;307:1838–47.
43. **Kent DM, Kitsios G.** Against pragmatism: on efficacy, effectiveness and the real world. *Trials*. 2009;10:48.
44. **Califf RM, Filerman GL, Murray RK, Rosenblatt M, Merck, Co I.** The Clinical Trials Enterprise in the United States: A Call for Disruptive Innovation. Washington: Institute of Medicine; 2012. Discussion Paper.
45. **Sabin JE, Mazor K, Meterko V, Goff SL, Platt R.** Comparing drug effectiveness at health plans: the ethics of cluster randomized trials. *Hastings Cent Rep*. 2008;38:39–48.
46. **Vickers AJ.** Whose data set is it anyway? Sharing raw data from randomized trials. *Trials*. 2006;7:15.
47. **Vickers AJ.** Making raw data more widely available. *BMJ*. 2011;342:d2323.
48. **Ross JS, Krumholz HM.** Ushering in a new era of open science through data sharing: the wall must come down. *JAMA*. 2013;309:1355–6.
49. **Sandercock PA, Niewada M, Czlonkowska A.** The International Stroke Trial database. *Trials*. 2011;12:101.
50. **Specks U, Merkel PA, Seo P, et al.** Efficacy of remission-induction regimens for ANCA-associated vasculitis. *N Engl J Med*. 2013;369:417–27.
51. National Heart Lung and Blood Institute. NHLBI. Available at: <https://biolincc.nhlbi.nih.gov/home>. Accessed August 15, 2013.
52. National Institute of Diabetes and Digestive and Kidney Diseases. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Central Data Repository (CDR). Available at: <https://www.niddkrepository.org>. Accessed August 13, 2013.
53. GlaxoSmithKline. Clinical study requests—initial members of the Independent Review Panel. Available at: <https://clinicalstudydata.gsk.com/Members-of-the-Independent-Review-Panel.aspx>. Accessed August 15, 2013.
54. European Medicines Agency. Release of data from clinical trials. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topcis/general/general_content_000555.jsp&mid=WC0b01ac0580607bfa. Accessed August 15, 2013.
55. United States Department of Health and Human Services. Strategic Framework on Multiple Chronic Conditions, 2010.