1. Science Policy Roundtable on the Heterogeneity of Treatment Effects
   *Page i*

2. Faculty List
   *Page iv*

3. Faculty Disclosures
   *Page v*

4. Introduction
   *Pages S1-S2*
   Sheldon Greenfield

5. Heterogeneity of Treatment Effects: Implications for Guidelines, Payment, and Quality Assessment
   *Pages S3-S9*
   Sheldon Greenfield, Richard Kravitz, Naihua Duan and Sherrie H. Kaplan

6. Variability in Response to Antihypertensive Drugs
   *Pages S10-S20*
   Barry J. Materson

7. Potential Genetic Causes of Heterogeneity of Treatment Effects
   *Pages S21-S25*
   David B. Goldstein, Anna C. Need, Rinki Singh and Sanjay M. Sisodiya

8. Heterogeneity of Treatment Effects in Schizophrenia
   *Pages S26-S31*
   T. Scott Stroup
9. Healthcare Policy Implications of Heterogeneity of Treatment Effects

*Pages S32-S35*

Michael J. McLaughlin
Science Policy Roundtable on the Heterogeneity of Treatment Effects

GUEST EDITOR

Sheldon Greenfield, MD
Donald Bren Professor of Medicine
Executive Director
Center for Health Policy Research
Department of Medicine
University of California–Irvine School of Medicine
Irvine, California

This supplement is based on a symposium held March 9, 2006, in Washington, DC. The symposium and publication of this supplement were supported by an unrestricted educational grant from Pfizer Inc. Editorial support was provided by Peloton Advantage, Parsippany, New Jersey.

Statement of Peer Review: All supplement manuscripts submitted to The American Journal of Medicine for publication are reviewed by the Guest Editor(s) of the supplement, by an outside peer reviewer who is independent of the supplement project, and by the Journal's Supplement Editor (who ensures that questions raised in peer review have been addressed appropriately and that the supplement has an educational focus that is of interest to our readership).

Author Disclosure Policy: All authors contributing to supplements in The American Journal of Medicine are required to fully disclose any primary financial relationship with a company that has a direct fiscal or financial interest in the subject matter or products discussed in the submitted manuscripts, or with a company that produces a competing product. These relationships (e.g., ownership of stock or significant honoraria or consulting fees) and any direct support of research by a commercial company must be indicated on the title page of each manuscript. This information will be published in the frontmatter of each supplement.

Editor-in-Chief: Joseph S. Alpert, MD
Editor, Supplements: William H. Frishman, MD
Publishing Director: Pamela Poppalardo

Executive Supplements Editor: Brian Jenkins
Senior Production Editor: Mickey Kramer
Proof/Production Editor: Mary Crowell
Science Policy Roundtable on the Heterogeneity of Treatment Effects

Guest Editor

Sheldon Greenfield, MD
Donald Bren Professor of Medicine
Executive Director
Center for Health Policy Research
Department of Medicine
University of California–Irvine School of Medicine
Irvine, California

Faculty

Jill A. Berger, MAS
Health and Welfare
Marriott International
Washington, DC

Naihua Duan, PhD
Heath Services Research Center
Semel Institute for Neuroscience and Human Behavior
University of California–Los Angeles
Los Angeles, California

Robert Epstein, MD, MS
Medco Health Solutions
Franklin Lakes, New Jersey

Ronald Finch, PhD
Center for Prevention and Health Services
National Business Group on Health
Washington, DC

David B. Goldstein, PhD
Duke Institute for Genome Sciences and Policy
Duke University
Durham, North Carolina

Charles Ingoglia, MSW
National Council of Community Behavioral Healthcare
Rockville, Maryland

Sherrie H. Kaplan, PhD, MPH
Center for Health Policy Research
Department of Medicine
University of California–Irvine School of Medicine
Irvine, California

Richard Kravitz, MD, MSPH
Center for Health Services Research in Primary Care
Lawrence J. Ellison Ambulatory Care Center
University of California–Davis
Sacramento, California

Barry J. Materson, MD, MBA
Division of Nephrology and Hypertension
Department of Medicine
University of Miami Miller School of Medicine
Miami, Florida

Michael J. McLaughlin, MD
Peloton Advantage
Parsippany, New Jersey

Marie Michnich, DrPH
The Institute of Medicine
Washington, DC

Anna C. Need, PhD
Duke Institute for Genome Sciences and Policy
Duke University
Durham, North Carolina

Rinki Singh, MD
Department of Clinical and Experimental Epilepsy
University College London
London, United Kingdom

Sanjay M. Sisodiya, MD, PhD
Department of Clinical and Experimental Epilepsy
University College London
London, United Kingdom

T. Scott Stroup, MD, MPH
Department of Psychiatry
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina
Faculty Disclosures

The authors who contributed to this publication have disclosed the following industry relationships:

**Jill A. Berger, MAS**, is an employee of Marriott International.

**Naihua Duan, PhD**, has received research/grant support from Pfizer Inc.

**Robert Epstein, MD, MS**, is an employee of Medco Health Solutions.

**Ronald Finch, PhD**, is an employee of the National Business Group on Health.

**David B. Goldstein, PhD**, has received research support from GlaxoSmithKline; and serves as a consultant to Johnson & Johnson and Pfizer Inc.

**Sheldon Greenfield, MD**, has received research/grant support from Novo Nordisk and Pfizer Inc.

**Charles Ingoglia, MSW**, is an employee of the National Council of Community Behavioral Healthcare.

**Sherrie H. Kaplan, PhD, MPH**, has received research/grant support from Novo Nordisk and Pfizer Inc.

**Richard Kravitz, MD, MSPH**, has received research/grant support from Pfizer Inc; and served as a consultant to Forrester Research and Pfizer Inc.

**Barry J. Materson, MD, MBA**, has no financial arrangement or affiliation with a corporate organization or with a manufacturer of a product discussed in this supplement.

**Michael J. McLaughlin, MD**, is an employee of Peloton Advantage.

**Marie Michnich, DrPH**, is an employee of The Institute of Medicine.

**Anna C. Need, PhD**, has no financial arrangement or affiliation with a corporate organization or with a manufacturer of a product discussed in this supplement.

**Rinki Singh, MD**, has no financial arrangement or affiliation with a corporate organization or with a manufacturer of a product discussed in this supplement.

**Sanjay M. Sisodiya, MD, PhD**, is a member of the Speakers’ Bureau for Pfizer Inc; and has received research/grant support from UCB Pharma.

**T. Scott Stroup, MD, MPH**, has been a speaker at events sponsored by Eli Lilly and Pfizer Inc.
Introduction

The commonplace use of evidence-based medicine, clinical treatment guidelines, formulary entries, and quality-of-care measures applies average effects based on results in heterogeneous clinical trial populations to the treatment of individual patients. Yet individual patients and key subgroups receiving the same treatment often experience responses that can vary greatly, ranging from optimal resolution of the condition to detrimental or even lethal adverse events. These different responses to treatment are known as heterogeneity of treatment effects (HTE). Concerns about HTE are becoming a more prominent focus of consideration in the current healthcare environment.

Although HTE has always existed, this phenomenon has not yet been well characterized or investigated. The spectrum of effects reflects the numerous variables present within and acting upon every patient population. Key variables leading to HTE include factors such as illness severity and risk of poor outcome, age, sex, hepatic and renal function, use of concomitant medications, care setting, comorbidities, genetic variations, and diet; the list grows as our understanding of this phenomenon increases. HTE remains, even in well-designed clinical trials of investigational therapies in which attempts are made to control these confounding factors and variables.

A conference held in Washington, DC, on March 9, 2006, examined in depth the phenomenon of HTE and its implications for guidelines, payment, and quality-of-care assessment. The program began with several scientific and clinical presentations, followed by a policymaker roundtable discussion. The articles in this supplement to The American Journal of Medicine summarize the important information delivered in these presentations.

In the first article, my coauthors and I describe the sources of HTE within trials that can compromise the interpretation of results, the sources of HTE in the target population that limit the generalizability of trials, and strategies for understanding and managing HTE. We focus on 2 evolving phenomena that impair the ability to develop guidelines, payment rules, and quality-of-care measures based on randomized controlled trials. First, there is now a broader spectrum of illness severity inclusion, permitting patients with less severe disease, who are less likely to benefit from a drug or treatment, to be included in randomized controlled trials. These people are less likely to respond to an agent than are sicker patients, thereby reducing the power for the trial and yielding negative or null results for the trial. Second, although the general population is living longer with more chronic diseases, randomized controlled trials often exclude such longer-lived patients, only to have findings subsequently generalized from younger trial-eligible patients to these older, complex patients whose mortality from comorbid diseases reduces treatment effectiveness. Together, these phenomena impose challenges on the usefulness of the results of randomized controlled trials for clinical and policy applications.

In the second article, Dr. Barry J. Materson examines the presence of HTE in the treatment of hypertension. There are several layers of variables recognized in the measurement and treatment of hypertension. Use of blood pressure measurement guidelines and consistent techniques help to reduce the potential variability associated with clinician measurements. Intrinsic patient characteristics, such as age and race/ethnicity, can affect blood pressure and the efficacy and adverse events observed with antihypertensive medications. Dr. Materson also discusses clinical examples of mutations that affect antihypertensive response, including multiple polymorphisms within components of the renin-angiotensin-aldosterone system.

The third article, by Dr. David B. Goldstein, reviews the pharmacogenetic influences on HTE. Drug response may be dictated by variation in genes involved in both pharmacokinetic (PK) and pharmacodynamic (PD) pathways. Functional polymorphisms of PK genes can result in patients being poor, intermediate, efficient, or ultrarapid metabolizers of specific agents, thereby influencing efficacy and/or susceptibility to adverse drug reactions and necessitating individualized dosing. Variants of genes regulating PD pathways may alter drug target pathways, potentially affecting patient outcomes in a more pronounced manner. These PK and PD polymorphisms may act independently or in combination to affect drug response. Better understanding of these pharmacogenetic factors may help to clarify...
sources of HTE that were once considered intangible, thereby affecting patient treatment decisions.

The treatment of mental illness presents another opportunity to examine HTE. In general, outcome measures for psychiatric conditions are subjective, with symptomatology and treatment results varying greatly among patients. In the fourth article, Dr. T. Scott Stroup discusses HTE in the context of the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) research program sponsored by the National Institute of Mental Health (NIMH). The CATIE trial studied schizophrenia, a disease state inherently prone to HTE, and was designed with broad inclusion and minimal exclusion criteria to create a realistic and varied sample population. Dr. Stroup identifies some of the sources of treatment response variability within this diverse trial population. Collectively, the CATIE results highlight the extent of variable drug efficacy and tolerability response in the treatment of psychoses, demonstrating the need for individualized therapy for schizophrenia.

During the afternoon roundtable, healthcare policymakers discussed the clinical presentations on HTE and examined how such information could be incorporated into their decision-making process. In the final article, written on behalf of the HTE Policy Roundtable Panel, Dr. Michael J. McLaughlin presents their findings. The panel members agreed that HTE should be considered when determining healthcare policy. Their discussion highlights the implications of this phenomenon beyond patient–physician interactions, extending throughout seemingly disparate sectors of the healthcare system, e.g., government agencies, third-party payers, and employers. Some of the panel members have even taken steps within their own organizations to deliver individualized, quality healthcare in light of the existence of HTE. The consensus of the roundtable was that more data from clinical trials, patient databases, and similar sources should be made available to physicians and policymakers so that well-informed decisions can be implemented.

The implications of HTE for today’s medicine are extensive. By recognizing the factors associated with HTE, researchers can design clinical trials that better characterize those individuals and groups of individuals who will benefit from various therapeutic options. Clinicians and healthcare administrators can then make pragmatic use of the results by implementing policy changes to renovate healthcare in light of this significant phenomenon. More HTE-related clinical data, along with access to multiple pharmacotherapeutic options, appear to be the most promising ways to address the response variability relative to the delivery of quality healthcare.

Sheldon Greenfield, MD
Center for Health Policy Research
Department of Medicine
University of California–Irvine School of Medicine
Irvine, California, USA
Heterogeneity of Treatment Effects: Implications for Guidelines, Payment, and Quality Assessment

Sheldon Greenfield, MD,a Richard Kravitz, MD, MSPH,b Naihua Duan, PhD,c and Sherrie H. Kaplan, PhD, MPHa

aCenter for Health Policy Research, Department of Medicine, University of California–Irvine School of Medicine, Irvine, California, USA; bCenter for Health Services Research in Primary Care, Lawrence J. Ellison Ambulatory Care Center, University of California–Davis, Sacramento, California, USA; cHealth Services Research Center, Semel Institute for Neuroscience and Human Behavior, University of California–Los Angeles, Los Angeles, California, USA

ABSTRACT

Randomized controlled trial results are needed for developing guidelines, payment rules, and quality-of-care measures; however, 2 phenomena reduce the usefulness of randomized controlled trial findings. First, these studies now enroll patients with less severe disease, who are less likely to benefit from a drug or treatment. Second, patients are living longer but, as a result, have more chronic diseases. Although randomized controlled trials often exclude these older patients, trial findings continue to be generalized to them. Together, these phenomena impose challenges to the usefulness of the results of randomized controlled trials for clinical and policy applications. The convergence of these phenomena makes the current research paradigm underlying evidence-based medicine, guideline development and quality assessment fundamentally flawed in 2 ways. First, the “evidence” includes patients who may have a minimal benefit from the treatment being tested. This could reduce the power for the trial and yield negative or null results, leading to undertreatment of a group of patients with potential for a greater-than-observed benefit. Second, attempts to generalize the results from positive trials to patients who have been excluded from those trials may result in the overtreatment of those who could not benefit (e.g., because they will die from other causes before the benefit of treatment would occur) and therefore represents a parallel hazard. In this article, we describe sources of heterogeneity of treatment effects (HTE) within trials, which can compromise the interpretation and generalizability of results. We also examine strategies for understanding and managing HTE in practice, to increase the usefulness of trial results. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Health policy; Healthcare quality assurance; Heterogeneity; Population genetics; Practice guidelines

When applied in routine clinical practice, medical treatments frequently fall far short of the potential they exhibit when tested in randomized controlled trials.1,2 To demonstrate efficacy and maximize internal validity, randomized controlled trials are commonly limited to a relatively homogeneous group of patients, usually the most severely ill, who have strong potential to benefit from treatment. However, this practice might compromise the generalizability of study findings to patients who are less severely ill, for whom the treatment effect is likely smaller.3,4 To address the problems engendered by attempting to generalize results from randomized controlled trials to the “typical” patient seen in clinical practice (broadening external validity), more recent clinical trials have widened the spectrum of patients included.5–10 However, a number of recent and notable examples of such inclusive trials have demonstrated a potential weakness in this approach if applied without careful consideration of the study design.5,7,9 That is, those who could benefit only minimally from the treatment are included, diluting the average effect size and thereby reducing the power for the trial if the sample size does not accom-
moderate the diversity of the patients. These inclusive trials are potentially more generalizable to usual-care patients than were the more traditional restrictive trials that were limited to the most severely ill patients. However, the enhancement of generalizability comes at a price; as the pool of participants is allowed to be more heterogeneous, it is likely that the treatment effect will also be more heterogeneous. Therefore, the usual one-size-fits-all interpretation of randomized controlled trial data may no longer be valid.

An evolving divergence from the current research paradigm involves the recognition that, although a broader, more inclusive group of patients may be more representative of the target population, they may also have differential responses to the same treatment, resulting in heterogeneity of treatment effects (HTE), defined by Kravitz and colleagues as the “magnitude of the variation of individual treatment effects across a population.”[^4] This broad definition includes a more focused alternative definition of HTE as the interaction of treatment with patient individual characteristics. Understanding HTE will help researchers, clinicians, and policymakers identify those groups who could achieve the greatest benefit with the least risk for treatments under study.

Homogeneity of the patients included in trials and the inherent tension between internal and external validity of trials have long been debated. Why is understanding heterogeneity so important now? First, several recent null trials have produced confusing or counterintuitive results.[^5]–[^10] Whether attributable to the increase in willingness to publish such trials, to the increasing vigilance with which trials are monitored, or to the broadening of inclusion criteria to foster generalizability of results, such a trend underscores the importance of understanding HTE. Second, the increase in efforts to improve quality of care through evidence-based medicine, clinical guidelines, pay-for-performance, and quality-of-care assessment[^11]–[^21] highlights the importance of scrutinizing the synthesis of information from clinical trials for the purpose of guiding the standards of clinical practice. Synthesizing trial results with limited generalizability may produce suboptimal practice guidelines for important and prevalent subgroups in a clinical practice. This is especially problematic in aging adults, who generally experience more comorbid conditions. Such patients, commonly excluded from trials, now make up a substantial proportion of routine clinical practice, particularly of primary care providers. Often called complex patients, these individuals frequently take multiple medications (polypharmacy), may have problems with adherence, and because of their age and comorbid conditions, may die before they can benefit from the treatment being studied.[^22]–[^24]

In this article, we define HTE, examine the underlying principle of “average effects” as it relates to HTE, and address the sources of HTE, both within trials and as they affect the generalizability of trial results. Finally, we will discuss strategies for understanding and managing HTE.

**DEFINITION OF HETEROGENEITY OF TREATMENT EFFECTS**

As noted above, HTE refers to the differential response to the same treatment by different patients. More specifically, this definition includes different responses by patients with different characteristics.[^4] Those characteristics can include severity of the disease under study (severe versus milder forms of the same disease); sociodemographic characteristics, such as age, sex, and race/ethnicity; genetic characteristics; and health-related behaviors, such as adherence to treatment, alcohol consumption, and use of complementary or alternative medicine. Statistically, HTE occurs when there is an interaction between patient characteristics, such as those described above, and the treatment under study. These interactions moderate the treatment effect. If there were a substantial interaction between a treatment and specific patient characteristics, the average effect observed across patients in a trial would not apply to the subgroup of patients in the trial with different levels of those characteristics. It can be argued that the average effect is not a very useful construct, analogous to the metaphor of the average body temperature for a person with one foot in ice water and the other in boiling water.

HTE has substantial implications for the patient sample entered into trials[^24] (Figure 1), based on ideas developed by Longford[^25] and modified by Kravitz and colleagues.[^4] If the trial sample resembles sample 1 in the figure, the results lead to an average effect that is a good estimate of the average effect for the target population (patients who would be candidates for treatment in clinical practice). However, the small standard deviation for the individual treatment effects erroneously implies little HTE. Therefore, the conclusion appears to be that the treatment is good for the majority of patients, when in fact it is not. Sample 2 illustrates an average effect that is a poor estimate of the average effect for the target population. We infer this (but cannot know it with certainty) because the patients in the trial do not appear to be “average” (i.e., they have more severe disease than that commonly seen in routine practice). Sample 3 illustrates an average effect that is a good estimate for the target population (its mean and the population mean are identical). However, the large standard deviation around the mean for the outcome implies considerable HTE, and therefore a reduced power for detecting the average treatment effect.

Understanding and responding to the problems generated by HTE will require substantive changes in the design, conduct, timing, and analysis of clinical trials. However, the current reliance on average effects of trials, although compelling for simplicity, has substantial drawbacks.

**THE PROPERTIES AND PERILS OF AVERAGE EFFECTS**

To provide, on average, the most current, best-quality medical care, physicians are encouraged and have been conditioned to align their clinical practice with the results of...
well-conducted clinical trials. Although most physicians understand intuitively that patients vary in their individual response to treatment and in their experience of treatment side effects, the science of medicine fosters the pathocentric belief that there is a universality to pathophysiologic and pharmacologic mechanisms. That is, within a disease, treatments should follow the principle of standardization, as in one-size-fits-all. Treatment tailored to fit the unique needs of each patient, however, is largely dependent on the physician’s clinical experience, practice circumstances, and other variables inconsistent with the current principles of evidence-based medicine and quality-improvement efforts, such as pay-for-performance.

Efforts to distill clinical studies to help practitioners make the best “evidence-based” decisions in routine practice depend on, and are similarly disadvantaged by, the “averaging” of treatment effects across patients within trials, and “averaging the average effects” across trials. Within trials, not all patients benefit, or benefit in comparable magnitude. Evidence-based medicine, as it is usually practiced, is based on averaging or summarizing effects across trials. Without differentiating treatment effects for subgroups within trials, this higher-order averaging of effects is then the basis for the development of guidelines for clinical practice and for quality-of-care measures for the assessment of clinical practice.

The essential weaknesses underlying efforts to generalize results from clinical trials are 2-fold: the notion of the average patient and the changing nature of the target population. A number of patient characteristics, such as sex, race, or ethnic origin, are categorical and cannot be averaged. Extrapolating treatment effects to any of these categories of patients within trials may underestimate or overestimate results for any such patient subgroups that represent a deviation from the average patient in the trial. In addition, efforts to generalize clinical trial results to the majority of patients in the average clinical practice (an extrapolation of the average effect) inevitably involve those patients often excluded from trials, in some cases the very patients who might benefit most from treatment or who might be most at risk for an adverse response. Concurrently, the broader population is aging and has more and more complex comorbidities. These comorbidities can independently shorten patients’ longevity such that they will die before they could benefit from the treatment under study. Extrapolating trial results to these patients would result in overtreatment.

It should not be surprising, therefore, that the size of the average treatment effect in most clinical trials is small. Dramatic improvements in diagnostic technology, along with an emphasis on early detection and screening, have led to an increase in the number of patients diagnosed at much earlier, much less severe stages of disease. The result is that patients who previously would not yet have been diagnosed with a disease are now eligible for entry into clinical trials. Patients at the milder end of the severity spectrum are among those who have the least to gain from treatment, at least within the time frame of the typical randomized controlled trial. The result is that conventional trials enrolling a broad spectrum of patients are increasingly likely to be underpowered and to generate null results.

In parallel, the nature of the majority of the population (or the average patient) is changing. The aging of the population, along with the advent of life-prolonging therapies, has caused an increase in the proportion of patients within a diagnostic category who have multiple comorbidities, are treated with multiple medications, and therefore have multiple competing sources of mortality. Compared with 20 years ago, patients with prostate cancer are much more likely to have coexisting congestive heart failure, chronic obstructive pulmonary disease, and diabetes mellitus, and are much more likely to die from these conditions than from prostate cancer over the course of a few years. These patients therefore are much less likely to benefit from aggressive treatment. Two decision analyses have shown that intensive treatment of blood sugar in patients with diabetes...
who are aged >65 years has little impact on the reduction of complications in such individuals.\textsuperscript{27,28} In these patients, the effectiveness of aggressive treatment is substantially lower than that observed in clinical trials because these trials were conducted in younger patients without such comorbidities.

Two recently published studies further illustrate these issues. The Calcium and Vitamin D Supplementation trial did not show an overall reduction in fractures among patients on active treatment.\textsuperscript{9} However, in those who were aged >60 years, and thus deemed at higher risk for fracture, fractures were prevented. Subgroups were not prespecified, and the study was not designed with adequate power to test the findings in the subgroups. In the Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management and Avoidance Trial (CHARISMA), comparing clopidogrel plus aspirin with aspirin plus placebo for the prevention of cardiovascular disease,\textsuperscript{5} investigators observed that the overall rate of cardiovascular events did not differ significantly between the clopidogrel plus aspirin and aspirin plus placebo groups. However, asymptomatic patients with multiple risk factors showed a trend favoring the aspirin and placebo treatment, whereas those with clinically manifested atherothrombosis had a small, statistically significant reduction in the rate of the primary end point when given clopidogrel and aspirin. That is, the 2 subgroups had opposite findings. In this study, although 20 subgroup variables were prespecified, the study was not powered to detect differences between the 2 major subgroups (i.e., those with and without clinical atherothrombosis) that were created as a composite of the individual variables. The outcomes of the study could have been quite different if the study had been designed with adequate power for the low-risk group. When clinical studies are designed with the average effect as the primary goal, HTE is often overlooked and studies have inadequate power to detect between-group differences in treatment response. To understand HTE and address it in clinical practice, clinical studies should be designed to address HTE as a key design feature.

**Sources of Heterogeneity of Treatment Effects**

**Within-Trial HTE**

Given the inclusion of patients across an expanded spectrum of disease, and those with diverse characteristics to maximize the representativeness of the study sample, more patients are being included with differential risk for (1) poor outcomes from the disease under study, (2) potential response to treatment, and (3) treatment side effects. As noted above, the effects of a treatment for an individual patient depend on the disease, the patient, and the environment or context of care. With the expanded inclusiveness of patients in trials, each of these categories of variables can affect the magnitude and direction of response to treatment, including the experience of treatment side effects.

**Differential Risk for Poor Outcomes (Risk Without Treatment)**

The most important disease-related issue is the differential risk for incurring a disease-related poor outcome (such as a heart attack from coronary artery disease or blindness from diabetic retinopathy). Patients with atrial fibrillation and rheumatic heart disease are 18 times more likely to experience a stroke compared with those free of valvular heart disease.\textsuperscript{29} The potential benefit of taking blood thinners is therefore much greater for those with an abnormal heart valve. Hayward and colleagues\textsuperscript{30} reported that the average results of a clinical trial can be misleading by underestimating the benefit in high-risk study patients. When they assigned their hypothetical treatment with an overall relative risk reduction of 17\% for an adverse disease-related outcome, 1 such outcome was prevented for every 98 people treated. However, in this scenario, the higher-risk patients were actually 6 times more likely to benefit from treatment (number needed to treat [NNT] = 39) than were lower-risk subjects (NNT = 238).\textsuperscript{3,30}

**Differential Potential for Response to Treatment**

Some patients have greater potential for response to treatment than others because of genetic differences, environmental factors, or the interaction of both. As a result of variations in metabolism, density of drug receptors, and so forth, treatments are more potent in some individuals than in others. For example, recent findings indicate that 7\% of women with breast cancer lack the biochemical apparatus that allows the other 93\% to respond to tamoxifen.\textsuperscript{31} A similar percentage of individuals received no analgesic benefit from codeine because of a metabolic aberration.\textsuperscript{32} Owing to regional variations in the prevalence of penicillin-resistant pneumococci, patients with pneumonia who live in some areas in the United States are more likely to respond to penicillin than are those in other areas.\textsuperscript{33} In yet another example, in 3 randomized controlled trials for obsessive-compulsive disorder (OCD), social phobia, and panic disorder that compared medication, cognitive behavior therapies, and the combination of the 2 therapies with a placebo pill, patients with OCD were unexpectedly less likely to respond to placebo than were patients with either of the other 2 disorders.\textsuperscript{34} The differential potential for response to treatment is frequently not known before treatment is administered and therefore is not taken into account in the study design.

Another possible source of heterogeneity of response is individual variation, as suggested by Senn.\textsuperscript{35} One of the most researched areas of individual variations is that of blood pressure determination.\textsuperscript{36} Senn argues that an individual may respond at 1 time point and not at the next point in time. It is not clear how much of a role this inconsistency of response plays across the large number of topics studied, but investigators should be aware of it, and, as is done in hypertension research, trials should take account of it.
Differential Risk for Side Effects

Similarly, by virtue of their genetic makeup, the use of other concurrent therapies, and comorbid conditions, some individuals are uniquely vulnerable to the adverse effects of treatments while others are protected. When these adverse events are common enough (e.g., in the case of the interaction between sildenafil and nitrates in patients with coronary heart disease), they are detected in early clinical trials and used to inform practice. If the adverse events are rare or idiosyncratic to some patient subgroups, they may only become evident in detailed phase 4 or postmarketing clinical studies. Under current regulations and research practices, such studies are not performed routinely.

HTE and Generalizing Trial Results to the Target Population

The characteristics of the target population, primary candidates for the majority of treatments under study, are changing. Patients are older, have more and more complex comorbid diseases, are treated with multiple (often ≥6) medications (some of which interact unfavorably with the treatment under study), and have behaviors that can be at odds with effective treatment, such as nonadherence to prescribed treatment or health behaviors that could compromise effective treatment (such as use of alcohol or complementary or alternative medicines). These complex patients constitute an increasingly large proportion of the patients seen in clinical practice.

In addition, most, if not all, treatments are not just taken, they are delivered. The knowledge, attitudes, and skills of clinicians and the organizational resources of healthcare systems can profoundly alter treatment benefits and harms. Carotid endarterectomy, for example, is a life-saving and function-enhancing intervention in selected patients, but only if the complication rate (a function of patient factors and surgical skill) is <5%. Using a nonsurgical example, patients on warfarin, who are seen in specialized warfarin clinics, are more likely to have international normalized ratios within the therapeutic range and, therefore, to experience fewer strokes and bleeding events than patients in usual-care settings. Patients with specific forms of disease or no health insurance, or who are poor, may not be financially eligible for the study treatment, even though they could otherwise benefit.

The target population is therefore a moving target. With many prevalent patient subgroups excluded from trials, often for sound reasons, clinicians are left with a treatment dilemma when attempting to apply the clinical research literature to patient care. In reviewing this literature, clinicians are encouraged to assess the internal and external validity of studies to determine whether the patient being treated is similar enough to the average patient included in the trials to extrapolate trial results to that patient’s care. To be maximally useful to the practicing physician, and to avoid generating guidelines and quality measures with limited generalizability, evidence-based medicine advocates, clinical researchers, and policymakers must reassess their criteria for evaluating clinical trials.

That synthesis is daunting under the current research paradigm. An example that affects a large number of patients and their physicians concerns the treatment of prostate cancer. As noted above, the effectiveness of aggressive treatment, either by surgery or radiation, of early-stage prostate cancer is considerably different in patients who have moderate or substantial comorbidity. The body of evidence suggests that patients with these comorbidities (≥20% of patients) might not survive long enough to benefit from aggressive treatment of their prostate cancer. As noted above, aggressive treatment for patients with diabetes is also moderated by age and the presence of competing comorbidities. Studies that document HTE for these patients would substantially improve physicians’ ability to provide them with appropriate treatment options.

Even more disturbing than the implications of HTE for clinical practice is the averaging of average effects across trials, as is done in the current evidence-based medicine approach to quality improvement, including the development of clinical guidelines, quality-of-care measures, and pay-for-performance initiatives. These initiatives promote a sweeping uniformity of treatment that will almost certainly cause eventual harm.

HTE represents a serious problem for practicing physicians, clinical researchers, and policymakers. However, solutions can be found. The following methodologic and analytic approaches to clinical research attempt to address HTE and optimize the balance of risks and the balance of benefits for patients and physicians.

UNDERSTANDING AND ADDRESSING HETEROGENEITY OF TREATMENT EFFECTS

The problems associated with HTE are not new. What raises them to a level of national and international concern warranting attention by researchers and policymakers are 2 recent developments: (1) a growing awareness of the risks involved in applying clinical trial results to patients with complex comorbidities, who represent an increasing proportion of a physician’s practice; and (2) the advent of initiatives that rely on syntheses of current clinical research literature applied to clinical practice in an effort to improve quality (evidence-based medicine, treatment guidelines, quality-of-care assessment, and pay-for-performance). Addressing HTE will require modifying current approaches to clinical research. The following proposals are intended to foster thoughtful dialogue among clinicians, researchers, and policymakers about the directions such modifications could take.

Pretrial Identification of Risk Groups

Longitudinal observational studies identifying patients at risk for poor outcomes should be conducted and integrated into the planning and design of clinical trials. When conventional trial analysis is not sufficient, prior research...
can permit risk stratification. In 1 analysis, when the average baseline risk for poor outcome in the study population was 5.6%, the relative risk associated with improvement after treatment varied from a 41% improvement for those with no risk factors to a 36% decline in outcome for those with ≥4 risk factors.3,30 In the CHARISMA trial, noted earlier, 20 risk variables were identified and studied.5 These variables were likely to be correlated. There have been several successful attempts to combine such variables, using psychometric and other techniques, into composite measures that can capture and aggregate risk for poor outcomes.23,42–48

Redesign of Trials
Most importantly, trials must be powered adequately to address the intended aims, which should include assessment of key subgroup analyses and HTE analyses. Inclusive trials should have at least the sample number of severely ill patients who would have been enrolled in a restrictive trial, plus an adequate number of less severely ill patients to extend the knowledge base. In addition, various nonconventional designs might be attempted, including crossover designs, when applicable, and matched-pair designs, particularly when crossover designs cannot be conducted. The use of composite end points might improve precision and permit smaller sample sizes for key subgroups. Phase 4 trials might be performed in parallel or included as mandatory follow-up.

Posttrial Analysis of Subgroups
Subgroup analyses have been systematically discouraged by some statisticians associated with clinical trials.49 They argue appropriately that without a priori specification of subgroups, there will be insufficient power to draw valid and reliable conclusions from analyses of outcomes of treatment for subgroups. Furthermore, the likelihood of false-positive results increases with testing of treatment effects for increasing numbers of subgroups. However, learning from subgroup analyses of existing trials to generate hypotheses and inform the conduct of additional research can serve to inform the conduct of research that will enhance the assessment of treatment effects for target subgroups of patients that together represent a substantial proportion of routine clinical practice patient populations.50,51

Learning from Longitudinal Observational Studies to Inform Generalizability
As noted, there are increasingly large subgroups, particularly the elderly with multiple comorbidities, whose shortened life expectancy decreases the chance that they will benefit from treatment at the level achieved in randomized trials conducted in younger, less complex patients. In these patients, the reduced life expectancy gained from aggressive treatment must be balanced against the mortality, complications, and morbidity associated with aggressive treatment. In prostate cancer, subgroups of patients whose comorbidity has been shown to reduce life expectancy23 may choose “watchful-waiting” as the optimum choice of treatment. In type 2 diabetes, the reduced effectiveness of lowering the blood sugar to near-normal levels in patients with competing comorbidities must be weighed, in decision analyses, against the burden of intense treatment over long periods of time, with the possible crowding out of other more effective treatments (such as treating hypertension)24,52 in these subgroups. Observational studies that document prognosis in these various subgroups can form the bases for these analyses.23,28,52

SUMMARY
It is becoming evident that HTE exists, may be increasing, and must be addressed. Current randomized controlled clinical trials include patients who may have a minimum benefit from the treatment being tested, leading to negative or null results.

The continuing acceptance of average effects may have adverse effects on patients whose treatments are not paid for or recommended by evidence-based medicine guidelines, and on physicians who do not, for example, achieve a hemoglobin A1c level of 7% in subgroups of patients who are unlikely to benefit from achieving that level. Attempts to generalize positive trial results to patients who have been excluded from those trials may result in overtreatment of those who could not benefit or would likely die from other causes before benefit would occur. Strategies to overcome the problems caused by this heterogeneity should be pursued and will increase the usefulness of trial results.

References


24. Greenfield S, Kaplan SH, Nicolucci A. The diabetes patient’s “other” diseases lead to 5-year mortality. Presented at the Midwest Society of General Internal Medicine, Quality Summit II. September 29, 2006; Chicago, Illinois.


Variability in Response to Antihypertensive Drugs

Barry J. Materson, MD, MBA
Division of Nephrology and Hypertension, Department of Medicine, University of Miami Miller School of Medicine, Miami, Florida, USA

ABSTRACT

Heterogeneity of treatment effects (HTE) is a measure of the variations in individual treatment response to the same agent across a population. Hypertension affords an appropriate model for investigators of HTE. Use of blood pressure measurement guidelines and consistent techniques help to reduce the potential variability associated with clinician measurements. Patient characteristics such as age and race/ethnicity can affect blood pressure, including patient response and adverse events observed with antihypertensive medication. Through pharmacogenetic advances, potential underlying causes for such variation are emerging. The growing number of clinical examples of mutations that affect antihypertensive response includes multiple polymorphisms within the components of the renin-angiotensin-aldosterone system. The most prominent examples of these polymorphisms exist in the genes coding for angiotensinogen, angiotensin-converting enzyme, and the angiotensin II type 1 receptor. An understanding of the components of blood pressure variability and sources of HTE in antihypertensive therapy is important for analyzing published reports on this topic. It is also helpful when designing treatment protocols for individual patients with hypertension and in assessing their response to therapy. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Antihypertensive drugs; Biologic characteristics; Drug resistance; Heterogeneity; Population characteristics; Populations at risk

The existence of heterogeneity of treatment effects (HTE) is apparent when evaluating patient response to a drug in clinical trials and in clinical practice. Biologic and lifestyle-related factors, including disease severity, age, race/ethnicity, sex, hepatic and renal function, adherence, diet, exercise practices, smoking, and alcohol consumption, can at least partially account for HTE. Beyond these factors, genetic variation can be the underlying cause of the HTE phenomenon. Clinical trials are meant to reflect typical patient response to a given medication. Trial investigators studying a medication’s safety and efficacy calculate the minimum number of patients needed to obtain scientifically valid results across a population and then determine the sample size necessary to power the study adequately. This calculation does not aim at perfection; rather, it is based on the subgroup response, the average response observed in previous trials, and estimates of discontinuation of therapy, among other factors.

Despite this built-in level of uncertainty, such studies and the guidelines based on them can provide powerful information for clinicians. They will know what kind of results to expect in the average patient population, along with what adverse reactions might occur and how frequently. Nevertheless, they will not be able to predict with certainty a key subgroup’s or a particular patient’s response. In the case of nonsteroidal anti-inflammatory drugs (NSAIDs), it may take a trial of ≥2 chemically similar NSAIDs before one relieves a particular patient’s pain. Similarly, although some depressed patients do not respond to a given antidepressant, they may respond to another that is chemically similar. The reasons for these phenomena are poorly understood.

The treatment of hypertension provides an interesting example of HTE. Extrinsic causes of blood pressure measurement variability have been largely addressed through the evolution of measurement devices and the standardization of techniques. However, underlying patient characteristics and variable responses to individual situations con-
tinue to affect blood pressure, including patient response to antihypertensive drugs. Although many of the causes of such variation are still unknown, pharmacogenetic research continues to identify genetic differences that can affect blood pressure and medication response.

The purpose of this report is to describe HTE in the context of hypertension. Better understanding of the sources of HTE will affect the design of clinical trials, the development of treatment guidelines, and the care of patients with hypertension.

BRIEF HISTORY OF BLOOD PRESSURE MEASUREMENT

Blood pressure measurement has undergone an interesting and progressive evolution over the last several hundred years. Measurement devices and techniques have gradually progressed toward the noninvasive, convenient, and reliable form to which we are accustomed today.

Clergyman Stephen Hales is credited with the first documented measurement of intra-arterial pressure in 1733, when he inserted hollow tubing into the carotid artery of a horse and recorded blood rising to a height of almost 9 ft (270 cm) in a glass column. More than 150 years later (1896 to 1897), Scipione Riva-Rocci described the mercury sphygmomanometer, an ingeniously simple device that allowed indirect estimation of systolic blood pressure. The modern sphygmomanometer is considered a direct descendant of Riva-Rocci’s device.

Measurement technique has undergone improvements to maximize the recording applications and accuracy of the devices. Initial techniques used palpation of the radial pulse or visual mercury column fluctuations to determine the systolic blood pressure at which arterial flow could be occluded with external pressure. In 1905, however, Nikolai Korotkov described the 5 sounds that became the foundation for the auscultatory technique for determining blood pressure, allowing more precise measurement of both systolic and diastolic pressures. These distinct sound phases became known as the Korotkov sounds, with the first phase (appearance of sound) indicating systolic pressure and the fifth phase (disappearance) indicating diastolic pressure. Modern blood pressure measurement incorporating the use of an air-filled cuff and a stethoscope was born from the techniques developed by Riva-Rocci and Korotkov.

PRACTICAL APPLICATIONS: UPPER LIMITS OF “NORMAL”

Intra-arterial pressure is a continuous variable without natural cut points for normal or abnormal. For practical purposes, however, it is sometimes necessary to establish upper limits of normal blood pressure. The Federal Aviation Administration (FAA) has determined the acceptable upper limit for blood pressure in licensed pilots to be readings consistently not >155/95 mm Hg. This upper limit was determined in part from actuarial and epidemiologic data showing that consistently high blood pressure increases risk for cardiovascular complications. From a more pragmatic standpoint, the United States Air Force (USAF) has used an upper blood pressure limit, so that valuable flight training is not spent on pilot candidates with consistently higher blood pressure readings. The insurance industry has traditionally taken a much simpler stance, in that their actuaries have long recognized that higher blood pressure readings predict a greater risk of premature death. Similarly, to develop and execute a research or treatment protocol, one must establish arbitrary (traditionally diastolic) blood pressure points for dose titration and medical intervention.

Although these pragmatic cutoff points are commonly used, there is an epidemiologically approximated biologic point for blood pressure at which risk increases. Lewington and colleagues determined that 115 mm Hg systolic blood pressure is the approximate point above which the risk for cardiovascular disease increases markedly. These researchers determined that cardiovascular mortality risk doubles with each increase of 20/10 mm Hg blood pressure increment. For example, at 135/85 mm Hg, cardiovascular mortality risk doubles, whereas at 155/95 mm Hg the risk is 4 times greater, and at 175/105 mm Hg the risk is 8 times greater. Likewise, there are experimentally determined biologic cutoff points for blood pressure, above which the risk for target organ-specific damage increases. These points differ depending on the organ system and morbidity in question (e.g., myocardial infarction, stroke, end-stage renal disease).

Industries and clinicians attempt to determine or create such upper limits for blood pressure for both pragmatic and experimental use. These recognized upper limits, whether scientifically sound or arbitrarily assigned, are often applied to groups of people or even to individual patients, but they do not account for the variation that occurs between individuals or among subpopulations.

Normal Blood Pressure Variability

Better measurement technique has increased our understanding of individual blood pressure variations over time. David Ayman, an early investigator of clinical hypertension, performed placebo-controlled studies and meticulously measured blood pressure under standardized conditions. Using a mercury sphygmomanometer over periods ranging from 2 months to 2 years, Ayman observed and recorded the diastolic blood pressure of 76 patients with essential hypertension. To minimize variability, each patient was seated in a quiet room with 1 arm resting on an adjacent table. At each visit, readings were obtained when the patient first sat down, followed by recordings taken at 5- to 10-minute intervals over 5 to 45 minutes. Total diastolic blood pressure variations ranged from a minimum of 5 mm Hg to a maximum of 66 mm Hg, with an average variation of 30 mm Hg.

Perhaps more impressive is the natural blood pressure variation within an individual over the course of a single day. In 1968, Sir George Pickering used an intra-arterial
considerable variability in blood pressure measurement also depends on the technique used. Adherence to proper technique and guidelines is necessary for consistent assessment, and the evaluation of factors creating variability is necessary before determining the presence of HTE.

**American Heart Association Guidelines**

The American Heart Association (AHA) Guidelines for the Measurement of Blood Pressure in Humans and Experimental Animals state that the proper training of observers, positioning of the patient, and selection of cuff size are all essential to obtaining consistent and accurate blood pressure readings. The auscultatory technique by a trained observer using a mercury or calibrated aneroid sphygmomanometer continues to be the method of choice for blood pressure measurement in the office. The patient should be asked to remove all clothing covering the location of cuff placement. The patient should be comfortably seated, with legs uncrossed and the back and arm supported such that the middle of the cuff on the upper arm is at the level of the right atrium (the midpoint of the sternum). At the initial visit, blood pressure should be measured in both arms. The patient should be instructed to relax as much as possible and not talk during the measurement procedure; ideally, ≥5 minutes should elapse before the first reading is taken. The use of an appropriate cuff size is of paramount importance, especially in obese patients and in children. Errors in blood pressure measurement are greater when the cuff is too small relative to the patient’s arm circumference than when it is too large.

The AHA guidelines recognize potential limitations associated with in-office blood pressure measurements and acknowledge that they can be supplemented by self-measured readings taken with validated devices at home. It should be noted, however, that several home-use sphygmomanometers have been determined to be inaccurate when evaluated in clinical settings. Still, according to the AHA, 24-hour ambulatory monitoring gives a better prediction of risk than do office measurements. These monitors are useful for diagnosing “white-coat” hypertension and for identifying persistently high blood pressure during the night (so-called “nondippers”), a problem that may be associated with increased risk. Despite such attempts at perfecting and standardizing recording technique, the previously mentioned studies by Ayman and Pickering have demonstrated that an underlying natural variability in blood pressure exists.

**Physiologic Factors**

Several physiologic factors can also affect blood pressure measurement and contribute to confounding the assessment of HTE with regard to response to antihypertensive therapy. Pickering recognized that the time of day, season, ambient temperature, influence of consumed substances, and emotional state could affect blood pressure measurement. Although he did not coin the term “white-coat” hypertension, Ayman first identified the emotional component of blood pressure variability when he recognized the phenomenon of higher blood pressure readings when the recording was made in a physician’s office compared with blood pressure readings made at home. Ayman was also the first to formally recognize the placebo effect observed with most of the purported antihypertensive drugs of the early 20th century.
device. In this study, relative mean ± SD differences in systolic blood pressure were −0.6 ± 6.6 mm Hg and +1.1 ± 4.7 mm Hg in diastolic blood pressure. Absolute differences were +4.9 ± 4.4 mm Hg for systolic mm Hg and +3.7 ± 3.0 mm Hg for diastolic mm Hg, and were significantly associated with age, body mass index, ankle-brachial index (ABI), and hypertension. Hypertension, hypercholesterolemia, obesity, elevated hemoglobin A1c (HbA1c), and low ABI independently increased the risk for an absolute systolic blood pressure difference >10 mm Hg. In recognizing the interarm millimeters of mercury differences in patients, it is recommended that, to ensure accurate and consistent blood pressure readings, the arm with the higher systolic blood pressure measurement be used for subsequent measurements.17

Even when the physical measurement technique is standardized, the blood pressure of a patient during a single visit can be determined in several ways. Houweling and colleagues22 compared 4 methods for obtaining blood pressure measurements in 223 patients with type 2 diabetes mellitus, who were evaluated at 5 family practices in the Netherlands. Each patient’s blood pressure level was recorded as the first reading, as the mean of the first 2 readings, as the mean of the last 3 of ≥4 readings, or as the mean of the first 2 readings, with a maximum 5-mm Hg difference. Although the latter 2 methods resulted in the lowest blood pressure measurements, the smallest difference in blood pressure was noted between the first 2 methods (mean difference in systolic blood pressure, 1.6 mm Hg; P <0.001), and the largest difference was observed between the first and last methods (mean difference, 7.9/3.0 mm Hg; P <0.001). These differences are important when interpreting trial results and when making treatment decisions in practice.

Controlling Variability of Measurement
Good clinical trials are designed to minimize variability in blood pressure measurement results. Patient visits should be scheduled, if possible, for the same time of day, with the same observer each time. Only validated devices should be used, and they should be calibrated frequently. Careful attention should be paid to patient and arm position, and blood pressure should be measured in the same arm each time. Variation can be minimized by taking 12 readings, but this is impractical. Taking the average of 3 readings made by a trained observer and that do not vary by >5 mm Hg is a practical compromise. Some study protocols establish baseline blood pressure as the average of multiple readings or even multiple visits. Treatment effect determination may similarly account for multiple readings, sometimes taking the average of the last 2 visits.

There is an increasing trend to use ambulatory, automated blood pressure measuring devices, for daytime or for 24-hour measurement. This yields a large number of readings that can be averaged, with the degree of variation noted. The downside to this process is its expense and inconvenience.

HETEROGENEITY OF TREATMENT EFFECTS IN HYPERTENSION THERAPY
Even after minimizing the inherent sources of blood pressure variability, a review of the literature reveals that the treatment of high blood pressure with antihypertensive therapies is met with variable response as a result of HTE.

The placebo effect seen in clinical trials and in clinical practice contributes to HTE and may confound the assessment of individual treatment response. In the 1967 Veterans
Administration (VA) Cooperative Study, Freis and colleagues evaluated the effects of the administration of either active treatment (hydrochlorothiazide [HCTZ] plus reserpine plus hydralazine) or placebo in 143 men with severe hypertension (inhospital diastolic blood pressure range, 115 to 129 mm Hg). The researchers observed that the response to both active treatment and placebo varied greatly among the patients, with changes in systolic blood pressure ranging from $+11.0$ mm Hg to $-11.0$ mm Hg in the active treatment group and from $+22.0$ mm Hg to $-22.0$ mm Hg in the placebo group. Likewise, the change in diastolic blood pressure varied greatly in both, ranging from $+12.0$ mm Hg to $-12.0$ mm Hg in the active treatment group, and from $+28.0$ mm Hg to $-44.0$ mm Hg in the placebo group (Figure 2).

Response Based on Race/Ethnicity and Age

Several studies have pointed toward race/ethnicity as a determinant of antihypertensive therapy response. Self-identified race is only a rough surrogate for genetic determinants but can be useful clinically. The 1982 VA Cooperative Study demonstrated variable response by race/ethnicity to antihypertensive therapy. In this analysis, HCTZ monotherapy was compared with propranolol monotherapy in 683 men diagnosed with hypertension. In self-identified African Americans, HCTZ lowered mean systolic blood pressure by $20.3 \pm 14.3$ mm Hg compared with $8.2 \pm 12.2$ mm Hg for propranolol, and lowered mean diastolic blood pressure by $13.0 \pm 7.0$ mm Hg compared with $9.5 \pm 7.0$ mm Hg for propranolol. In whites, HCTZ lowered mean systolic blood pressure by $15.3 \pm 12.0$ mm Hg compared with $13.2 \pm 13.1$ mm Hg for propranolol, and lowered mean diastolic blood pressure by $10.9 \pm 5.7$ mm Hg compared with $12.6 \pm 6.6$ mm Hg for propranolol. As evidenced by these results, propranolol is as efficacious as HCTZ in whites, but HCTZ is more effective than propranolol in African Americans (Table 1).

In another analysis of antihypertensive effect based on race/ethnicity, the difference in effect between races was altered by the addition of supplementary antihypertensive therapy. In the initial analysis, the VA Cooperative Study Group observed that white patients responded better to low-dose captopril than did black patients: in white patients ($n = 170$), mean blood pressure was reduced from $14.7 \pm 1.1$ mm Hg to $10.7 \pm 0.6$ mm Hg, whereas in black patients ($n = 151$) it was reduced from $9.1 \pm 1.2$ mm Hg to $8.0 \pm 0.6$ mm Hg.
0.7 mm Hg. This difference was eliminated when HCTZ was added to low-dose captopril monotherapy (Figure 4).

Conversely, in the Quinapril Titration Interval Management Evaluation trial (ATIME), Mokwe and colleagues concluded that the variability observed in patient response to blood pressure treatment originates within, not between, racial/ethnic groups. In this study, crude systolic blood pressure responses to quinapril therapy averaged 4.7 mm Hg less and diastolic blood pressure responses 2.4 mm Hg less in black patients compared with white patients, but the response distributions largely overlapped. The multivariate linear regression models used were adjusted for variables in the study design and in measured participant characteristics, and reduced the racial/ethnic response in systolic blood pressure by 51% and in diastolic blood pressure by 21% (Figure 5). In this model, it was demonstrated that participant characteristics such as age, sex, body size, and pretreatment blood pressure severity contribute to the apparent racial/ethnic differences in treatment response. Unlike the VA single-drug therapy study, Mokwe and colleagues examined only race/ethnicity, not a race/ethnicity by age interaction.

The VA Cooperative Study Group included 1,292 men with a diastolic blood pressure measurement of 95 to 109 mm Hg and demonstrated that patient age, together with race/ethnicity, predicts patient response to antihypertensive monotherapy. This analysis of age by race/ethnicity evaluated the use of 6 antihypertensive drugs (HCTZ 12.5 to 50 mg/day, atenolol 25 to 100 mg/day, captopril 25 to 100 mg/day, clonidine 0.2 to 0.6 mg/day, a sustained-release preparation of diltiazem [diltiazem-SR] 120 to 360 mg/day, and prazosin 4 to 20 mg/day). The age by race/ethnicity differences observed in the study demonstrated that, at the end of 1 year of treatment, the extended-release formulation of the calcium antagonist diltiazem was most successful in younger (aged <60 years) and older (aged ≥60 years)
blacks, with HCTZ as a close second, whereas atenolol, captopril, and diltiazem-SR had the highest success rate in younger whites. All of the active medications had approximately equal efficacy in older whites (Figure 6).29 A report by Preston and colleagues30 demonstrated that plasma renin profiling was no more accurate than the age by race/ethnicity interaction. There is no need to perform renin testing before treatment of the patient with uncomplicated hypertension.

An association between race/ethnicity and adverse events from antihypertensive therapy has also been reported. McDowell and colleagues31 performed a meta-analysis of the cardiovascular medicine literature published before March 2005. In 5 studies meeting their assessment criteria, the researchers found a higher relative risk (RR) of angioedema from angiotensin-converting enzyme (ACE) inhibitors in black patients compared with other patients (RR, 3.0; 95% confidence interval [CI], 2.5 to 3.7). In pooled results from 2 studies, the relative risk for cough from use of ACE inhibitors was higher in East Asian patients compared with white patients (RR, 2.7; 95% CI, 1.6 to 4.5).
Response Based on Smoking Status

Smoking status has also been associated with variable response to antihypertensive pharmacotherapy. In an analysis of 108 smokers versus 232 nonsmokers, smokers had less reduction in diastolic blood pressure compared with non-smokers when treated with propranolol (−8.7 vs −10.9; P = 0.02). This observed effect is likely attributable to hepatic cytochrome P450 metabolism of propranolol, which is induced by polycyclic aromatic hydrocarbons. Further supporting this theory, smoking status demonstrated no impact on treatment effects with nadolol, likely because of its renal excretion rather than hepatic metabolism and conse-

Figure 6  (A) Success rates at 1 year in black and white patients aged <60 years. (B) Differences in success rates at 1 year between black and white patients aged ≥60 years. HCTZ = hydrochlorothiazide.
quent lack of influence from polycyclic aromatic hydrocarbons.\textsuperscript{33}

Pharmacogenetic Causes of Heterogeneity of Treatment Effects

Continuing evolution in the field of genetics presents distinct genetic causes of HTE in antihypertensive therapy, primarily stemming from polymorphisms in the renin-angiotensin-aldosterone system (RAAS). The RAAS plays an important role in regulating blood volume, arterial pressure, and cardiac and vascular function. The 3 major polymorphisms that have been identified are in genes that code for angiotensinogen, ACE, and the angiotensin II type 1 (AT\(_1\)) receptor.

The angiotensinogen gene polymorphism consists of a mutation from methionine to threonine at point 235. Normal individuals have an MM genotype, whereas TT is considered to be abnormal and a hybrid genotype of MT also exists. Although this polymorphism has been researched, data are conflicting regarding the association of the TT genotype with hypertension.

The ACE gene polymorphism consists of a mutant deletion at 190 base pairs (bp) (the D allele), a mutant insertion at 490 bp (the I allele), or both. In general, DD patients have higher levels of ACE and more pressor response to infused angiotensin II, and they are more responsive to ACE inhibitors. The blood pressure response to HCTZ of 87 patients with mild essential hypertension according to ACE gene I/D and \(\alpha\)-adducin Gly/Trp (at 460 bp) polymorphism, all with no prior treatment, was studied by Sciarrone and colleagues.\textsuperscript{34} \(\alpha\)-Adducin is a constitutive part of the cell membrane, and the 460Trp allele may confer increased sodium reabsorption in the renal tubules. In this trial, basal mean blood pressure was similar between various ACE and \(\alpha\)-adducin genotypes, but patients with \(\geq 1\) I allele of ACE and \(\geq 1\) 460Trp allele of \(\alpha\)-adducin showed the largest mean blood pressure decrease after treatment (\(-12.7 \pm 1.9\) mm Hg) (Figure 7).\textsuperscript{34} The odds ratio for these patients being responders to HCTZ was 15.75 compared with patients with the Gly460Gly/DD polymorphism, who showed the lowest decrease in mean blood pressure (\(-3.4 \pm 1.7\) mm Hg). This demonstrates that the \(\alpha\)-adducin and

![Figure 7](image_url) Angiotensin-converting enzyme (ACE) and \(\alpha\)-adducin polymorphism as markers of individual response to diuretic therapy. The ACE I allele and the \(\alpha\)-adducin tryptophan at 460 base pair (bp) of \(\alpha\)-adducin (Trp) polymorphisms are associated with a greater response to hydrochlorothiazide (HCTZ). BP = blood pressure; D = D allele of ACE; Gly/Gly = glycine/glycine at 460 bp of \(\alpha\)-adducin. (Reprinted with permission from Sciarrone et al. ACE and \(\alpha\)-adducin polymorphism as markers of individual response to diuretic therapy. \textit{Hypertension}. 2003;41:398–403.)

![Figure 8](image_url) Increasing pulse pressure (PP) associated with carriers of the C allele (C) of the angiotensin II type 1 (AT\(_1\)) gene, T = T allele of G298T nitric oxide synthase gene. (Reprinted with permission from Safar and Benetos. Factors influencing arterial stiffness in systolic hypertension in the elderly: role of sodium and the renin-angiotensin system. \textit{Am J Hypertens.} 2003;16:249–258, with permission from the American Journal of Hypertension, Ltd.)
ACE I/D polymorphism may be useful in predicting the degree of response to HCTZ between patients. The polymorphism for the AT1 receptor is characterized by the A1166C mutation. Normal individuals carry an AA genotype, whereas the mutant genotype is designated as CC. Carriers of the C allele may have greater aortic stiffness, which is correlated with increased pulse pressure with increasing age (Figure 8). Individuals with the mutant genotype also have a generally greater response to ACE inhibitors.

SUMMARY

Substantial HTE exists in patients receiving antihypertensive drugs. Understanding the variability of blood pressure, both within individuals and among individuals in the population, is a prerequisite for assessing HTE relative to antihypertensive therapy. Blood pressure has natural variation and may be influenced by physiologically associated factors, such as time of day, season, consumed substances, and emotional state. Physiologic reaction to stress and physical activity can also vary during the course of the day under natural circumstances. Measurement of blood pressure must be done in a standardized way to minimize methodologic differences, as well observer error and bias, and to ensure good patient care. This includes the patient position, arm used, and manual/auscultatory technique of the examiner, as well as the timing and method of identifying which recordings to use in an assessment.

Patient characteristics have also been shown to affect adverse event rates from antihypertensive therapy. Specifically, factors such as age, race/ethnicity, smoking, and genomic polymorphisms may play a role. For example, adverse events from use of ACE inhibitors such as angioedema (higher in blacks) and cough (higher in East Asians) have shown some correlation with race/ethnicity. Recent literature has also presented genetic sources for HTE in antihypertensive therapy, including mutations in the RAAS.

To accurately assess the degree of HTE in response to antihypertensive therapy, a clinician must carefully consider and factor in all of these variables. Race/ethnicity should be considered—but not as the only source of HTE in antihypertensive therapy—along with age as key drivers of variability. Pharmacogenomics presents an additional component for consideration, but no single source of HTE can be evaluated without looking at the other contributing factors. A more thorough understanding of all of the components of blood pressure variability and sources of HTE in antihypertensive therapy is imperative in determining which agents will effectively treat hypertension in individual patients and in assessing response to treatment and may help produce better research design.

References


Potential Genetic Causes of Heterogeneity of Treatment Effects

David B. Goldstein, PhD, Anna C. Need, PhD, Rinki Singh, MD, and Sanjay M. Sisodiya, MD, PhD

ABSTRACT

Nongenetic biologic and lifestyle-related factors, including age, sex, hepatic/renal function, diet/exercise practices, illness severity, smoking, and alcohol consumption habits can account for the heterogeneity of treatment effects (HTE). However, even when these factors are taken into account, considerable variation remains unexplained and could potentially be attributable to genetic differences between patients. Drug response may be dictated by variation in genes involved in both pharmacokinetic (PK) (absorption, distribution, metabolism, excretion [ADME]) and pharmacodynamic (PD) (receptors, ion channels, enzymes, immune system) pathways. Functional variants of the ADME genes can result in patients being poor, intermediate, efficient, or ultrarapid metabolizers of specific agents, thereby affecting efficacy and/or susceptibility to adverse drug reactions and necessitating individualized dosing. A well-documented example of ADME gene variation is the debrisoquine polymorphism, which is characterized by markedly different metabolism of numerous commonly prescribed drugs based on variants of the cytochrome P450 2D6 gene. Variants of genes regulating PD pathways cause altering of drug target pathways, which may affect efficacy in a more pronounced manner. Examples of gene variants affecting PD pathways include those coding for dopamine metabolism, synthesis, and transport. These gene variants may act independently, in combination with each other, and/or in combination with PK genes to affect drug response, for example to antipsychotic medications. Increased understanding of a patient’s genotype and its corresponding effect on drug response would be useful to the practicing clinician in choosing an effective drug and in optimizing the dose in a timely manner. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Biologic characteristics; Genetics; Heterogeneity; Population characteristics; Population genetics

The existence of the heterogeneity of treatment effects (HTE) becomes immediately obvious whenever response rates to most drugs are analyzed. For instance, Spear and associates1 presented the mean response to a major drug from each of 14 therapeutic areas and showed that the mean efficacy rate was approximately 51.5%. This included figures as low as 25% in oncology therapy and as high as 80% for cyclooxygenase-2 (COX-2) inhibitors. This heterogeneity in response to pharmacotherapy results partly from physiologic and environmental differences that affect individuals in the general population. Some examples of these factors include sex, age, race/ethnicity, hepatic and renal function, diet, concomitant medications, illness severity, and alcohol and tobacco consumption.2

In some cases, however, the heterogeneity of response to drug therapy cannot readily be explained by such factors. One such example is the differential response to the antiepileptic drug topiramate. This drug is often effective in treating refractory epilepsy, but it frequently causes cognitive side effects in clinical trials and in practice.3-7 There is currently no way to predict who will or will not experience these adverse events, and, consequently, this drug is often selected only after multiple other treatments have failed. In
In a collaborative effort with the Institute of Neurology, London, the authors have attempted to predict cognitive side effects by comparing patients receiving topiramate (Topamax; Ortho-McNeil Neurologics Inc., Titusville, NJ) therapy who do and do not experience cognitive side effects. Thus far, examination of sex, age at initiation of treatment, titration rate, maximum and maintenance dosages, dosage at the time of the adverse event, and concomitant antiepileptic drugs have revealed no significant differences between the 2 groups (Table 1). With a lack of explanatory demographic or clinical factors, it can reasonably be hypothesized that genetic differences between patients contribute to the heterogeneity of response to topiramate.

In cases such as these, where the causes of HTE cannot be easily explained by outwardly apparent influences, genetic factors should be considered as a possible source of treatment response variability. This concept is the basis for the field of pharmacogenetics, wherein clinicians seek to predict a patient’s genetic response to a specific drug. There are areas of therapeutics, particularly oncology, where pharmacogenetics is already making an impact on treatment decisions. However, many of the causes of HTE are not well understood, and pharmacogenetic research has not substantially enhanced our knowledge in this regard. Further advancement of the role of pharmacogenetics in clinical practice requires a better understanding of the 2 classes of genes that may harbor variants influencing drug response: pharmacokinetic (PK) and pharmacodynamic (PD) genes.

### PHARMACOKINETIC GENES

Pharmacokinetic genes are those directly involved in the metabolism of drugs in the body, i.e., the ADME (absorption, distribution, metabolism, and excretion) of drugs. These genes form the largest group of known pharmacogenetic factors. Variability in the expression of PK genes can either slow or accelerate the uptake, conversion, or excretion of drugs, which can result in parent compounds or their active metabolites being eliminated before achieving a therapeutic effect, or remaining too long or in too high a concentration so that the risk of adverse events increases.

---

### Table 1  Interim results of patients treated with topiramate who have and who have not experienced cognitive adverse events (AEs)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Topiramate Patients with Cognitive AEs (n = 102)</th>
<th>Topiramate Patients with No Cognitive AEs (n = 115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex, n (%)</td>
<td>54 (53%)</td>
<td>60 (52%)</td>
</tr>
<tr>
<td>Mean age at start of treatment, yr (range)</td>
<td>35.1 (15–58)</td>
<td>34.4 (23–61)</td>
</tr>
<tr>
<td>ILAE diagnosis (%)</td>
<td>IGE (10%)</td>
<td>IGE (6%)</td>
</tr>
<tr>
<td></td>
<td>JME (4%)</td>
<td>JME (3%)</td>
</tr>
<tr>
<td></td>
<td>LRC (39%)</td>
<td>LRC (43%)</td>
</tr>
<tr>
<td></td>
<td>LRS (39%)</td>
<td>LRS (42%)</td>
</tr>
<tr>
<td></td>
<td>Other (7%)</td>
<td>Other (6%)</td>
</tr>
<tr>
<td>Titration, n (%)</td>
<td>Standard: 54 (53%)</td>
<td>Standard: 55 (48%)</td>
</tr>
<tr>
<td></td>
<td>Faster*: 29 (28%)</td>
<td>Faster*: 45 (39%)</td>
</tr>
<tr>
<td>Other (noncognitive) AEs, n (%)</td>
<td>67 (66%)</td>
<td>74 (64%)</td>
</tr>
<tr>
<td>Maximum dose, mg (mean ± SD)</td>
<td>354 ± 209</td>
<td>333 ± 227</td>
</tr>
<tr>
<td>Maintenance dose, mg (mean ± SD)</td>
<td>269 ± 173</td>
<td>262 ± 169</td>
</tr>
<tr>
<td>Dose at time of AE (mean ± SD)</td>
<td>282 ± 181</td>
<td>274 ± 213†</td>
</tr>
<tr>
<td>Concomitant AEDs at time of AE (mean ± SD)</td>
<td>1.5 ± 0.75</td>
<td>1.7 ± 0.74</td>
</tr>
<tr>
<td>Total concomitant AEDs (mean ± SD)</td>
<td>2.3 ± 1.3</td>
<td>2.1 ± 1.1</td>
</tr>
<tr>
<td>Treatment completed, n (%)</td>
<td>78 (76%)</td>
<td>83 (72%)</td>
</tr>
</tbody>
</table>

AED = antiepileptic drug; IGE = idiopathic generalized epilepsy; ILAE = International League Against Epilepsy; JME = juvenile myoclonic epilepsy; LRC = localization-related cryptogenic; LRS = localization-related symptomatic.

*Titration rate ≥25-mg increments at 2-week intervals.
†Status at time of earliest occurring “other AE” was used for comparison.

### Table 2  Examples of specific enzymes coded by pharmacokinetic genes

- Uridine diphosphogluconosyltransferase
- Glutathione S-transferase
- N-acetyltransferase
- Thiopurine S-methyltransferase
- Sulphotransferase
- Dihydropyrimidine dehydrogenase
- Aldehyde dehydrogenase
- Thymidylate synthase
Examples of genes encoding enzymes involved in PK processes are shown in Table 2. Perhaps the best understood PK gene codes for cytochrome P450 2D6 (CYP2D6). The abundance of published literature on polymorphisms of CYP2D6–15 is partly owing to the fact that so many drugs in clinical use are metabolized, at least in part, by this enzyme (up to 25%).16 The “debrisoquine polymorphism” associated with this gene affects the rate of metabolism (poor, intermediate, efficient, or ultrarapid) of many commonly prescribed antipsychotics and antidepressants.9,11,12,17 Poor metabolizers are exposed to functionally higher doses of drugs dependent on CYP2D6 for clearance (e.g., perphenazine and paroxetine) and are therefore at greater risk for adverse events.17 In addition, poor metabolizers experience reduced benefit from drugs that require activation by CYP2D6 (e.g., codeine). Conversely, ultrarapid metabolizers require higher doses of drugs cleared by CYP2D6 to achieve a therapeutic effect.

Clinicians must recognize the HTE involved with this variation in metabolism and adjust dosages of specific drugs accordingly.17 Considering that many antidepressants and antipsychotics are metabolized via the CYP2D6 pathway, this information has sometimes been considered as of particular relevance for psychiatrists,18,19 but currently a strong evidence base is lacking.

There are currently >46 documented polymorphic alleles for CYP2D6, and these polymorphisms vary worldwide based on race/ethnicity.16 This variation in allele frequency observed between different ethnic groups results in HTE that appears to be stratified by race alone but is actually genetic in etiology. For example, white Europeans have a much higher frequency of the CYP2D6*4 variant allele20 compared with Asians, who have a higher frequency of the CYP2D6*10 variant allele,21 and with some African populations, who have a higher frequency of the CYP2D6*17 variant allele22 (Table 3). These genotype frequency differences may contribute to interracial or interethnic health disparities.23

### PHARMACODYNAMIC GENES

The PD genes are directly responsible for the structure of the molecule or receptor targeted by various medications. These genes may also influence the associated signaling pathway or interfere with another metabolic pathway involved in the manifestation of a disease. PD gene variants are likely causes of HTE when drug response seems to be independent of dose (e.g., cognitive adverse events related to topiramate treatment).

---

Table 3 Variations in allele frequency help explain the heterogeneity of treatment effects that occurs between different racial/ethnic groups*

<table>
<thead>
<tr>
<th>Ethnic/Regional Population</th>
<th>Most Common CYP2D6 Allele Variations</th>
<th>Frequency (%)</th>
<th>Metabolic Effect of Allele Variations†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>CYP2D6*10</td>
<td>51</td>
<td>Poor metabolizers</td>
</tr>
<tr>
<td>Ethiopian/Saudi Arabian</td>
<td>CYO2D6*2xn</td>
<td>10–16</td>
<td>Ultrarapid metabolizers</td>
</tr>
<tr>
<td>White European</td>
<td>CYP2D6*4</td>
<td>12–21</td>
<td>Poor metabolizers</td>
</tr>
<tr>
<td>Black African</td>
<td>CYP2D6*17</td>
<td>20–35</td>
<td>Poor metabolizers</td>
</tr>
</tbody>
</table>

*Ethnicity based on population in country of origin.
†With drugs susceptible to specified allele.

Table 4 Heterogeneity of treatment effects caused by functional variants in pharmacodynamic genes

<table>
<thead>
<tr>
<th>Pharmacodynamic Gene Variant</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catechol-O-methyltransferase (COMT) gene</td>
<td>The met allele produces a thermolabile protein with significantly lower activity</td>
</tr>
<tr>
<td>Dopamine transporter (DAT1) gene</td>
<td>A 40-base pair 3’UTR repeat significantly affects dopamine availability in the striatum, cerebellum, temporal lobe, and lymphocytes</td>
</tr>
<tr>
<td>Dopamine receptor D2 (DRD2)</td>
<td>A promoter insertion/deletion variant increases transcriptional efficiency</td>
</tr>
<tr>
<td>Monoamine oxidase A (MAOA) gene</td>
<td>MAOA transcription with 3.5- or 4-repeat alleles, showing 2–10 times more efficient transcription than those with 3-repeat alleles</td>
</tr>
<tr>
<td>Monoamine oxidase B (MAOB) gene</td>
<td>Intron 13 A/G single nucleotide polymorphism has been associated with different levels of enzyme activity in the platelets and brain</td>
</tr>
<tr>
<td>Tyrosine hydroxylase (TH) gene</td>
<td>A (TCAT)(n) repeat polymorphism in first intron has possible relation with catecholamine turnover rates</td>
</tr>
</tbody>
</table>
Possible PD effects may result from polymorphisms in genes associated with dopamine synthesis, transport, metabolism, and receptor subunits. These gene variants may act independently, in combination with each other, and/or in combination with PK genes to effect response, e.g., to DRD2 antagonists such as antipsychotic drugs.24–28 The interactions between PD gene polymorphisms and drugs tend to be highly complex and less understood than interactions observed between PK genes and drugs. Some examples of functional variants in PD genes resulting in documented cases of HTE29–34 are shown in Table 4; however, whether they act synergistically, independently, or antagonistically with each other is currently unclear. The authors recently demonstrated an interactive effect of genetic variants on obesity in the dopamine-metabolizing MAOA and MAOB genes.35

To identify clinical scenarios in which PD gene variants are the probable cause of HTE, the contribution of nongenetic factors and PK gene variants to HTE should be ascertained first.

**PHARMACOKINETICS VERSUS PHARMACODYNAMICS**

As mentioned previously, PK effects on treatment response have been much more systematically studied, primarily because of the well-defined, finite set of variants associated with PK interactions. However, although they are more complicated to investigate, initial studies have indicated that PD effects may be stronger than PK effects. The greater influence of PD effects on treatment response can be demonstrated in clinical examples of HTE with warfarin therapy.

Warfarin therapy provides an excellent example of HTE because there is a 10-fold interpatient variability in the dose required to attain a therapeutic response. The (S)-enantiomer (pharmacologically more active form) of warfarin is usually metabolized by the genetically variable enzyme CYP2C9, thereby creating an ideal situation to investigate the PK effects of CYP2C9 polymorphisms on treatment response. In the case of warfarin therapy, patients with CYP2C9*2 and CYP2C9*3 alleles have reduced warfarin metabolism and thus require a lower mean daily warfarin dosage; they are also at greater risk for bleeding with standard dosing.36

Vitamin K epoxide reductase (VKOR), which converts vitamin K epoxide back to vitamin K, is the target of warfarin anticoagulants and demonstrates the effects of PD gene variation.36 Polymorphisms in the VKOR complex subunit 1 gene (VKORC1) affect the dose of warfarin required to achieve therapeutic effect. Much greater than the effects produced by PK variation, the effects of PD-driven variation in VKORC1 account for a high proportion of interindividual variability in dosing with warfarin.

**ANALYZING PHARMACOGENETICS AS A COMPONENT OF HETEROGENEITY OF TREATMENT EFFECTS**

Several key considerations must be applied in a stepwise fashion to ensure an accurate analysis of pharmacogenetics as a component of HTE.19 Possible genetic sources of variability should be narrowed down by first identifying candidate genes and the functional variants of those genes.1,19,37,38 Tagging should be used to capture common variation within these genes.19,37,38 Alternatively, it is becoming increasingly affordable to use whole genome approaches even in larger cohorts. Either way, cohorts must be thoroughly phenotyped, in a manner that is consistent among different clinics, and examined for all possible non-genetic influences on variation before beginning the genetic study. These cohorts should be as homogenous and as large as possible, because there is no way to know in advance the size of the effect, and thus no way to make power calculations to guide sample size. In analyzing the data collected, there should be correction for multiple testing, taking into account nonindependence between single nucleotide polymorphisms.39 Consideration should also be given to drug–drug interactions, pathophysiologic factors, and environmental interactions, which can affect interindividual drug response.16

**SUMMARY**

The emerging field of pharmacogenetics provides a promising area for researchers investigating HTE. Heterogeneity in pharmacotherapy response results partly from physiologic differences in individual patients. When the more apparent sources of variability in the general population, such as age, race/ethnicity, and sex, fail to fully explain differences in therapeutic response, underlying pharmacogenetic differences may provide an explanation. Increasing numbers of PK and PD gene variants are being identified as sources of such differences. As more relevant pharmacogenetic polymorphisms are identified, the potential for the clinician to assess and adapt to HTE in the clinical setting becomes markedly greater. By applying a rational and stepwise approach to HTE analysis through pharmacogenetics, specific genetic sources of drug response variability can be identified.

The implications of these emerging tools are vast. Functional variants of PK genes can determine which patients will be poor, intermediate, efficient, or ultrarapid metabolizers of specific agents. With such additional information to help predict the degree of efficacy and the risk for adverse events, clinicians may be able to make more individualized treatment decisions. As a result, identification of genetic markers that affect drug response may provide the ability to determine more effective dosing of medications across a wide number of therapeutic areas. As such advances reach the clinic, they can lead to improved drug efficacy, reduced occurrence of adverse events, and enhanced regimen adherence, all combining to optimize patient outcomes.
The treatment of mental illness presents an opportunity to examine the heterogeneity of treatment effects. The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial was sponsored by the National Institute of Mental Health (NIMH) to evaluate the effectiveness of antipsychotic medications for schizophrenia in broad patient populations and in scenarios representative of standard clinical practice. Trial inclusion criteria were broad and exclusion criteria were minimal, allowing for a heterogeneous study population. The majority of patients in each phase 1 treatment group discontinued their randomly assigned treatment owing to inadequate efficacy, intolerable side effects, or other reasons. Phase 2 of CATIE featured 2 treatment pathways (efficacy and tolerability) with randomized follow-up medication based on the reason for discontinuation of the previous antipsychotic drug. Outcome differences between treatment groups and variable responses to medications across the study suggest why multiple medication trials are common and may be necessary in the treatment of schizophrenia. Collectively, the CATIE results highlight variable response in the treatment of schizophrenia and demonstrate the need for individualized therapy based on variations in drug efficacy and tolerability among patients. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Clinical Antipsychotic Trials of Intervention Effectiveness; First-generation antipsychotics; Heterogeneity of treatment effects; Schizophrenia; Second-generation antipsychotics; Treatment pathways

The outcome variability observed when a specific course of therapy is administered to a heterogeneous population, known as the heterogeneity of treatment effects (HTE), is apparent across numerous disease states, including schizophrenia. Although patient-related factors such as age, sex, and hepatic and renal function account in part for HTE, lifestyle-related factors (e.g., diet, exercise practices, smoking, alcohol consumption) may contribute as well. Consideration of these factors is important when designing and interpreting clinical trials and when applying results to everyday clinical practice.

Practical clinical trials are designed to closely mimic real-world settings to give an accurate representation of how treatments work when used in clinical practice. By conducting a study under conditions closely simulating typical medical practice, the ability to apply the results to the average clinical setting may be enhanced. Practical clinical trials seek to address questions faced by clinicians and policymakers and to compare clinically relevant alternative interventions. They should include a representative population of patients, be conducted at representative practice settings, and simulate clinical treatment conditions. The applicability of practical clinical trials benefits from researchers collecting data on a broad range of health outcomes that are clinically meaningful, rather than simply having a single end point.

Practical clinical trials highlight HTE by sampling a broad, heterogeneous population. This allows for heterogeneity within the sample (i.e., within trials) to accurately reflect heterogeneity in the population (i.e., patients outside of trials). Most trials examining antipsychotics for the treatment of schizophrenia feature extensive inclusion and exclusion criteria, limiting the variability observed in the outcomes as well as the generalizability of the results. This may be in part because clinical trials are often designed to test a hypothesis regarding a single drug and to isolate its effects.
Individual treatment response to antipsychotics is highly variable. As Fleischhacker and Widschwendter⁴ point out, not only do group results vary, but interindividual effects of antipsychotics vary as well. An antipsychotic may be effective in an individual patient, or it may be partially effective or poorly effective. Average results are variable and generally suboptimal. Because patient response to commonly prescribed antipsychotics varies greatly, discontinuation and switching are common.⁴,⁵

ALLOWING FOR HETEROGENEITY OF TREATMENT EFFECTS IN STUDY DESIGN
The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study, a multiphase practical clinical trial evaluating antipsychotics in the treatment of schizophrenia, provides an example within psychiatry to examine how drugs might work differently in different situations. The 57-site CATIE program, sponsored by the National Institute of Mental Health (NIMH), evaluated numerous antipsychotics. The trial design featured broad inclusion criteria, an 18-month length of duration, and a hybrid structure combining standard and pragmatic clinical trial characteristics. These components were incorporated to closely simulate clinical practice, thereby providing real-world results that are applicable to actual patient populations.⁶

The CATIE schizophrenia study used inclusion criteria that were broad and exclusion criteria that were minimal.⁶,⁷ The trial design allowed for the inclusion of patients aged 18 to 65 years with schizophrenia, as specified by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV).⁷ Concomitant medications, comorbid illnesses, and substance use disorders were allowed among the study population. The main exclusion criteria specified that the patients could not be experiencing their first psychotic episode and that they could not have treatment-refractory schizophrenia.⁶,⁷ These broad inclusion and minimal exclusion criteria allowed for a heterogeneous sample population that more accurately represented the population of patients with chronic schizophrenia (Table 1).⁶–⁸

To evaluate the discontinuation and switching of medications, which is common among patients receiving antipsychotic therapy, the CATIE study included multiple phases.⁶ Phase 1 (n = 1,460) was designed to determine efficacy and tolerability differences, if any, between 5 study antipsychotic drugs. The primary effectiveness measure was all-cause treatment discontinuation. Secondary outcome measures included reasons for discontinuation (e.g., efficacy, tolerability, patient decision), psychopathology, safety outcomes, service utilization and costs, neurocognition treatment adherence, comorbidity, substance abuse, and quality of life.⁶,⁷,⁹ Phase 1 comprised a double-blind, randomized design comparing treatment with the second-generation, or atypical, antipsychotics olanzapine, quetiapine, risperidone, and ziprasidone with perphenazine, a mid-potency first-generation antipsychotic agent. Despite haloperidol being more widely prescribed historically, perphenazine was chosen as the first-generation antipsychotic comparator because it was a mid-potency drug expected to cause only moderate rates of adverse effects, including extrapyramidal side effects and sedation.⁷ However, patients with tardive dyskinesia at baseline were excluded from random assignment to perphenazine to prevent the exacerbation of this condition.⁷ This design feature was meant to reflect typical clinical practice treatment decisions. Patients were treated with their initial antipsychotic drug for up to 18 months or until discontinuation or medication switch.

In phase 2, a follow-up of phase 1, patients who discontinued their initial antipsychotic drug were allowed to continue treatment in the study with another randomly assigned antipsychotic agent (Figure 1).¹⁰ Two separate study pathways—an efficacy pathway and a tolerability pathway—were designed to evaluate the use of a second antipsychotic in patients deciding to switch therapies.⁹,¹⁰ In both pathways, the primary outcome measure was time to discontinuation for any reason. Secondary outcomes measures included time to discontinuation because of lack of efficacy, lack of tolerability, or patient decision.

The efficacy pathway was designed for study participants who discontinued their phase 1 antipsychotic drug because of inadequate symptom control.¹⁰ Patients in this pathway (n = 99) were randomized to open-label treatment with clozapine or to blinded treatment with another common antipsychotic.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Patient demographics of the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>Patients (n = 1,460)</td>
</tr>
<tr>
<td>Age (yr), mean (SD)</td>
<td>40.6 (11.1)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td>Male 1,080 (74)</td>
</tr>
<tr>
<td>Female 380 (26)</td>
<td></td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>White 874 (60)</td>
</tr>
<tr>
<td>Black 513 (35)</td>
<td></td>
</tr>
<tr>
<td>Other 71 (5)</td>
<td></td>
</tr>
<tr>
<td>Spanish/Hispanic/Latino ethnicity, n (%)</td>
<td>170 (12)</td>
</tr>
<tr>
<td>Educational level, n (%)</td>
<td>Secondary 12.1 (2.3)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td>Married 167 (11)</td>
</tr>
<tr>
<td>Previously married 425 (29)</td>
<td></td>
</tr>
<tr>
<td>Never married 886 (59)</td>
<td></td>
</tr>
<tr>
<td>Employment status, n (%)</td>
<td>Unemployed 1,217 (85)</td>
</tr>
<tr>
<td>Exacerbation in prior 3 mo, n (%)</td>
<td>402 (28)</td>
</tr>
<tr>
<td>PANSS total score (30–210), mean (SD)</td>
<td>75.7 (17.6)</td>
</tr>
<tr>
<td>CGI-S score (1–7), mean (SD)</td>
<td>4.0 (0.9)</td>
</tr>
</tbody>
</table>

CGI-S = Clinical Global Impressions of Severity scale; PANSS = Positive and Negative Syndrome Scale.³

*Racial and ethnic categories as defined by the census. Data not available for 2 patients. “Other” includes American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and patients of ≥2 race ethnicities.³

†Score of 4 = moderately ill.
second-generation antipsychotic agent (i.e., olanzapine, quetiapine, or risperidone) that they had not received in phase 1. Clozapine was chosen as the comparison medication in the efficacy pathway because it is the only antipsychotic medication to demonstrate improved efficacy over other agents in controlling psychotic symptoms among patients with poor response in previous antipsychotic drug trials. However, clozapine has been associated with serious adverse events, such as agranulocytosis and myocarditis.

The tolerability pathway was designed for study participants who discontinued their phase 1 antipsychotic treatment because of drug-related adverse events. Patients in this arm (n = 444) were randomized to double-blind treatment with a second-generation antipsychotic drug they had not received in phase 1, including ziprasidone or olanzapine, quetiapine, or risperidone. Investigators chose ziprasidone as the comparison medication because the drug was the newest antipsychotic to be approved when CATIE began. Of primary significance in this consideration, the researchers expected many discontinuations in phase 1 to be attributable to weight gain, and ziprasidone was expected to cause little or no weight gain.

Although these 2 pathways were designed for patients discontinuing their first randomly assigned antipsychotic due to efficacy or tolerability, patients were permitted to choose either pathway, regardless of the reason for discontinuing their initial treatment. Patients were encouraged to enroll in the efficacy or tolerability pathways if discontinuing because of lack of efficacy or lack of tolerability, respectively, but choice was allowed to more accurately represent real-world conditions and clinical practice. The result was that most of the 99 patients who enrolled in the efficacy pathway (and who thus accepted possible randomization to clozapine) had discontinued their previous antipsychotic medication in phase 1 because of inadequate symptom control. However, 184 of the 318 patients who discontinued their phase 1 medication because of inadequate symptom control enrolled in the phase 2 tolerability pathway rather than the efficacy pathway.

**EVIDENCE OF HETEROGENEITY OF TREATMENT EFFECTS IN THE CLINICAL ANTIPSYCHOTIC TRIALS OF INTERVENTION EFFECTIVENESS**

For the purposes of our discussion on HTE, a brief overview of outcome results from phase 1 is presented in Figure 2.
Figure 2  Rates of discontinuation in phase 1 of the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia study attributable to (A) any cause, (B) lack of efficacy, and (C) intolerability.

Figure 3  Rates of discontinuation in the phase 2 efficacy pathway of the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia study attributable to (A) any cause, (B) lack of efficacy, and (C) intolerability.
ilar overviews from the phase 2 efficacy and tolerability pathways are shown in Figure 3 and Figure 4, respectively.9,10

**Phase 1**

Discontinuation rates in phase 1 of CATIE were relatively high, with 74% of patients discontinuing their study medication before 18 months (1,061 of the 1,432 patients receiving ≥1 dose). On the primary outcome measure, median times to discontinuation for any cause were as follows: olanzapine, 9.2 months; quetiapine, 4.6 months; risperidone, 4.8 months; perphenazine, 5.6 months; and ziprasidone, 3.5 months. The lowest rates of discontinuation for any cause (64%) or because of lack of efficacy (15%) were observed in patients receiving olanzapine.7 However, discontinuation of treatment because of lack of tolerability yielded different results. Discontinuation rates because of adverse events were lowest in the risperidone group (10%) and highest in the group receiving olanzapine (18%) (Figure 2). Although demonstrating decreased overall discontinuation rates compared with other antipsychotics, olanzapine was associated with substantially more weight gain and adverse changes in lipid metabolism compared with the other drugs. Surprisingly, perphenazine, the first-generation antipsychotic studied, was not associated with higher extrapyramidal side effects than the second-generation antipsychotics studied and it was generally as effective as the newer study drugs.7

In addition to the assigned drug, earlier discontinuation was predicted by higher baseline scores on the Positive and Negative Syndrome Scale (PANSS), younger age, longer duration since first using an antipsychotic agent, and which antipsychotic agent was taken before entry; patients taking either olanzapine or risperidone before entering the study were more likely to continue their study medications for a longer time.7 Although olanzapine was more effective than quetiapine and risperidone in the primary analysis of treatment discontinuation for any cause, no agent stood out as the definitive antipsychotic of choice. As is observed in clinical practice among individuals with chronic schizophrenia seeking a new treatment, discontinuation rates were high for all antipsychotic drugs studied.

**Phase 2**

In the efficacy pathway of phase 2, clozapine demonstrated the longest time to discontinuation for any reason (10.5 months), the lowest rate of discontinuation because of lack of efficacy (11%), the highest rate of continuation throughout the duration of the trial (44%), and the most pronounced decrease in PANSS scores compared with the other drugs studied.10

Time to treatment discontinuation in the tolerability pathway of phase 2 was longer for patients treated with risperidone (7.0 months) and olanzapine (6.3 months) compared with quetiapine (4.0 months) and ziprasidone (2.8 months).9 As mentioned previously, 184 patients who discontinued their phase 1 antipsychotic drug because of lack of efficacy were enrolled in the tolerability pathway. Among these patients, median time to discontinuation was 9.6
months with olanzapine and 6.4 months with risperidone compared with 4.2 months with quetiapine and 2.6 months with ziprasidone. In the subgroup of patients who discontinued the previous phase because of intolerability, risperidone demonstrated the longest median time until phase 2 treatment discontinuation. This variability in response to a phase 2 antipsychotic based on reason for discontinuation in phase 1 suggests that there is considerable interindividual variation in response to antipsychotic medications.

**PRACTICAL IMPLICATIONS**

As reflected in CATIE, the variable nature of responses to antipsychotic agents is a major issue for psychiatrists and individuals with schizophrenia. Overall drug discontinuation or switching rates were high, and these were related to inadequate efficacy, adverse events, and other causes. The CATIE results demonstrate that trying to identify the optimal treatment for a patient with schizophrenia is a major challenge.

The NIMH recognizes the considerable variability in patients’ response to antipsychotic medications and has stated that a “one-size-fits-all” policy for treating schizophrenia could be harmful. Instead, variable treatment response in schizophrenia suggests that healthcare providers and patients must be able to choose from an array of antipsychotic drugs. Treatments should be selected according to the specific history, needs, and preferences of individuals.

The complex interaction of factors involved in the treatment of schizophrenia demonstrates the need for further trials similar to CATIE. Pragmatic study environments, in which real-world variables are allowed to exert their influences, can illustrate the heterogeneous effects of drug treatments. Such trials are designed to improve the ability to generalize the study results and ultimately to contribute valuable information that will help clinicians provide safe and effective treatments.

**SUMMARY**

Practical clinical trials are intended to imitate typical treatment settings to produce results useful to clinicians, patients, and other decision makers. The overall effectiveness of antipsychotic drugs is limited by drug efficacy and tolerability, along with patient adherence to treatment regimens. The issue of HTE is an important consideration in the treatment of mental illness. Individuals with schizophrenia who do not respond well to one antipsychotic agent may do well on another. Antipsychotic drug choice should be based on an individual’s clinical situation, history, and preferences.

**References**


Healthcare Policy Implications of Heterogeneity of Treatment Effects

Michael J. McLaughlin, MD, for the members of the HTE Policy Roundtable Panel*

Peloton Advantage, Parsippany, New Jersey, USA

ABSTRACT

Heterogeneity of treatment effects (HTE) is a phenomenon wherein the same treatment produces different responses in different patients. Following scientific presentations at a conference on HTE in drug therapy, a roundtable panel discussed the policymaking implications of this phenomenon in the current healthcare environment. The presentation of evidence on HTE served as a backdrop for this more pragmatic, solutions-based discussion of how HTE should be addressed in light of the trend in healthcare toward use of evidence-based medicine along with professional society clinical practice guidelines for specific disease states. Overall, the panel concluded that a specific agent should be used when the clinician is equipped with sound data. However, in the absence of such data, care has to be individualized, using the clinician’s best judgment regarding available treatment options. The sharing of data across all levels of the healthcare infrastructure is crucial for policymakers seeking to ensure quality care while considering the phenomenon of HTE and, at the same time, keeping cost-effectiveness a major concern. © 2007 Elsevier Inc. All rights reserved.

Heterogeneity of treatment effects (HTE) is a phenomenon wherein patients exhibit different responses to the same treatment.1 This variability may exist within clinical trials, thus affecting interpretation of results. In addition, heterogeneity in the target population, including those not participating in trials, can limit the ability to generalize the results of clinical studies. A careful exploration is necessary to address the clinical and policy implications of HTE and to help identify patient groups that could achieve the greatest benefit with the least risk.

Until recently, HTE has been largely overlooked. However, this phenomenon is now garnering the attention of clinicians and policymakers as an area worthy of research and discussion.2–5 Following clinical presentations on HTE at a conference held in Washington, DC, on March 9, 2006, an open roundtable panel discussed the policymaking implications of HTE in the current healthcare environment. The following 4 key principles regarding HTE, which were identified in the presentations, served as the basis for the roundtable discussions: (1) within a given medical condition, a tremendous amount of heterogeneity exists in clinical practice; (2) specific sources of heterogeneity, such as environmental effects and different risk of poor outcomes and adverse events, can be addressed with better data and individually by clinicians once fully understood; (3) potential solutions to problems associated with HTE can be found if healthcare leaders are willing to address them; and (4) the field of genetics may potentially address many of the issues associated with HTE. The present article summarizes the substance of the discussion and conclusions of the roundtable participants.

APPROACHES TO HETEROGENEITY OF TREATMENT EFFECTS

Taking a pragmatic approach to the issue of HTE, the panel did not overlook the economic and political influences on healthcare policymaking. Marie Michnich, DrPH, of the Institute of Medicine, began by addressing these external factors, stating that she would “like to disabuse [the audience] of the notion that randomized clinical trials drive the decision-making process, at least at the federal level.” Dr.
Michnich explained that many factors besides scientific evidence influence decisions in healthcare today. “I think some of those realities and practicalities,” she continued, “might be the hallmark of what my colleagues here on the panel are going to discuss . . . about the significance of what we have heard and the implications for what we do for a living.”

The panel of policymakers recognized the existence of HTE, and some had already taken steps within their own organizations to address HTE-related issues. Jill A. Berger, MAS, of Marriott International, for example, discussed her company’s practice of providing individualized healthcare by reducing copayments for medically necessary drugs for employees with certain diagnosed conditions. Steps are then taken to encourage compliance. “We are starting off with the chronic diseases related mainly to diabetes mellitus, heart disease, and asthma,” she explained.

Ronald Finch, PhD, of the National Business Group on Health, expressed similar thoughts on providing tailored care for patients with mental health disorders, stating, “We addressed antidepressants and how they need to be prescribed, and we’re very concerned about a failed ‘first’ before going to a particular medication.” With regard to treatment options, Dr. Finch added, “If someone has been on an antidepressant in the past and it was effective for them . . . they probably should have the option of [taking] that antidepressant.”

Another psychiatric policy issue, addressed by Charles Ingoglia, MSW, of the National Council of Community Behavioral Healthcare, is the dilemma of evidence-based medicine versus funding policies. Pointing out that mental health services are typically inadequately supported, Mr. Ingoglia stated that lack of funding has led to many patients being cared for by nonphysician providers with only high school diplomas or bachelor’s degrees. “It is in that context that most mental health services are delivered,” he said. For instance, the average psychiatrist in one of his facilities has an extensive case load of patients. Additionally, tailoring treatment for individual patients is difficult, Mr. Ingoglia explains, because “if you look at most state Medicaid systems and mental health systems, there are not even codes to bill for the majority of the services . . . in evidence-based practices.”

Ms. Berger continued with the discussion of the customization of healthcare for individual employees and commented on the value of having access to patient data in implementing such programs. “The drugs are a lot easier to customize because we have figured out a way, through ActiveHealth Management, the vendor who scans our data to identify gaps in care and does disease management for Marriott, to identify all employees who should be on certain drug regimens, even those not on the drug. This allows us to do a more effective job in improving compliance. Because we can do this up front, when our folks go to the pharmacy, they get a copayment reduction right then and there.” This helps patients who would not be able to purchase the drug at their full copayment. However, access to medications is only one of many challenges, and having the data to help drive these decisions will help enormously.

The importance of having access to patient data was an underlying theme of the roundtable and was supported by Robert Epstein, MD, MS, of Medco Health Solutions. He stated that each time a patient walks into a pharmacy to fill a prescription, a safety-related message may come back to the pharmacist or physician about that prescription or the patient’s history. “If there was more good-quality scientific evidence around subgroups or HTE groups, or individuals who might benefit from a specific message,” he expounded, “we would love to put it in our system and get that out there, and that really does make a difference.”

Ms. Berger commented on the value of all types of data throughout the healthcare system. For instance, she supported more research to use in individual decision-making and in directing care into the hands of the best providers. “We need the outcomes data on the right providers,” she continued. “Which hospitals are doing a great job related to certain areas of care?”

Related to the availability of patient data is the issue of patient privacy, particularly in the area of individual genetic information. Dr. Epstein referred to a recent study in which patients with asthma, and then their pulmonary physicians, were interviewed to determine the proportion of people who may be concerned about insurance companies gaining access to their genetic information. “It was interesting,” commented Dr. Epstein. “The physicians were more concerned than the patients.” Dr. Epstein added that “the early use of this information would be better used mainly on safety than on reimbursement or cost-related issues. I think that would be a better place to start, and it would be less concerning to members.”

Ms. Berger summarized her thoughts on how information on HTE could be used in clinical practice by reiterating the importance of data collection and the proper use of data. She stated, “I think it is going to be use of data, and it is going to be better, more robust data collection that leads to better-quality healthcare.” She noted that many groups are currently working toward this, with some efforts being coordinated. “For instance,” she added, “health plans [that] develop high-performing networks of providers are not using just their data . . . they’re using the community data to do this.”

In support of Ms. Berger’s statements, Dr. Epstein gave examples wherein data on HTE are already being used in the area of plan design. “I think that as more evidence comes out about HTE issues, it will get incorporated into, at least, my world of plan design.” He then gave examples, first citing the area of hepatitis C, in which criteria currently available are based on the recognition that different treatment durations and dosing should be used, depending on gene type. Last year, the US Food and Drug Administration (FDA) approved BiDil (isosorbide dinitrate–hydralazine; NitroMed, Inc., Lexington, MA), indicated for use in
African Americans.7 “Healthcare policymakers,” explained Dr. Epstein, “do or do not use that information to say something about coverage.” He pointed to drugs that are proved to work in one sex and not the other, such as for the treatment of irritable bowel syndrome.8 According to Dr. Epstein, “Plan designers are taking that HTE difference into account. So, I think, as we get more and more information about what works and what does not, and in what groups, we are . . . getting more precise.”

**FUTURE STEPS**

Overall, the panelists recognized the existence of HTE and expressed a willingness to adapt their decision making based on such evidence, with emphasis on data collection and availability. With these data in hand, physicians can work with researchers and healthcare providers to further address the main issues identified in the clinical presentations.

As a first step, clinicians and policymakers must move beyond 2 widely held notions. The first is that if a drug has been shown to be safe and effective on average, it is safe and effective in all patients. A possible solution for correcting this mistaken belief would be to have the FDA implement different standards or approval levels for different drugs, with some therapies being granted approval for use in a broad spectrum of patients, with average effects, and other therapies being granted limited initial approval for use only in specific, narrow populations. The second notion is seen in the pay-for-performance arena. Trial results showing average effects are being used to formulate clinical practice guidelines, which then become pay-for-performance standards that do not allow for the HTE that clinicians know will occur.

Significantly greater attention should be paid to developing multivariate prediction indices for the 3 critical factors of HTE: risk/risk-prediction indices, responsiveness, and vulnerability to side effects. Post hoc subgroup analyses are no longer acceptable for policy purposes, and policymakers and clinicians should have access to numerous studies and clinical trials. This is particularly important for studies related to the development of clinical prediction indices, wherein patients populations are stratified, and this stratification data would help clinicians and policymakers formulate better, more individualized treatment strategies.

The field of genetics shows great promise relative to understanding and using data on HTE,2 and genetic correlates of responsiveness and vulnerability should be investigated. This is especially important in terms of adverse drug reactions. Yet the issue of cost-effectiveness cannot be overlooked. Even if reliable genetic tests are developed that can accurately predict which patients will or will not experience adverse effects, it may be too costly to use such tests. Further research, particularly to identify genetic polymorphisms associated with patient response and vulnerability to side effects, is needed to evaluate variations in heterogeneous populations.

In considering evidence-based reports, clinicians and policymakers should weigh different types of evidence relative to different treatments in their decision making. These reports often find heterogeneity in response to drugs within the same class. However, these conclusions should not immediately translate into payment decisions, particularly when that decision means that the least expensive drug will be paid for but the rest will not. Policymakers must consider different kinds of evidence than that used for gaining drug approval as a guide when making payment decisions.

Clinicians and policymakers are aware that patient care typically varies by location. A treatment received from a clinician at a given site may yield different results than the same treatment given by a different clinician at a different site. Some clinicians may be hesitant to use evidence-based treatment strategies because they are not sure the strategy will work in their own practice. Using a different approach, clinical evidence gathered from the literature is examined in combination with local data from sources such as electronic medical records. Using this approach will allow clinicians to more accurately assess the average effects of a treatment in their own community and formulate their own clinical care decisions in a more informed way.

**SUMMARY**

Taking the implications of HTE into consideration, clinicians and policymakers can collaborate in an information-sharing environment to better administer quality healthcare. The translation of clinical evidence to clinical practice and, ultimately, to public policy is necessary if HTE is to be dealt with in a pragmatic manner. Ultimately, access to a wide variety of treatment options is crucial to provide the highest possible standard of patient care, as the variability in response to different treatments is vast across any patient population.

**APPENDIX**

**Members of the HTE Policy Roundtable Panel: (Chair)** Marie Michnich, DrPH, Director, Robert Wood Johnson Fellowship, Institute of Medicine, Washington, DC. (Panellists) Jill A. Berger, MAS, Vice President for Health and Welfare, Marriott International, Washington, DC; Robert Epstein, MD, MS, Senior Vice President and Chief Medical Officer, Medco Health Solutions, Franklin Lakes, NJ; Ronald Finch, PhD, Director, Center for Prevention and Health Services National Business Group on Health, Washington, DC; and Charles Ingoglia, MSW, Director of Technical Assistance, National Council of Community Behavioral Healthcare, Rockville, MD.
References


