

Knowledge integration at the center of genomic medicine

Muin J. Khoury, MD, PhD^{1,2}, Marta Gwinn, MD, MPH^{1,3}, W. David Dotson, PhD¹ and Sheri D. Schully, PhD²

Three articles in this issue of *Genetics in Medicine* describe examples of “knowledge integration,” involving methods for generating and synthesizing rapidly emerging information on health-related genomic technologies and engaging stakeholders around the evidence. Knowledge integration, the central process in translating genomic research, involves three closely related, iterative components: knowledge management, knowledge synthesis, and knowledge translation. Knowledge management is the ongoing process of obtaining, organizing, and displaying evolving evidence. For example, horizon scanning and “infoveillance” use emerging technologies to scan databases, registries, publications, and cyberspace for information on genomic applications. Knowledge synthesis is the process of conducting systematic reviews using *a priori* rules of

evidence. For example, methods including meta-analysis, decision analysis, and modeling can be used to combine information from basic, clinical, and population research. Knowledge translation refers to stakeholder engagement and brokering to influence policy, guidelines and recommendations, as well as the research agenda to close knowledge gaps. The ultrarapid production of information requires adequate public and private resources for knowledge integration to support the evidence-based development of genomic medicine.

Genet Med 2012;14(7):643–647

Key Words: evidence-based medicine; genomic medicine; knowledge integration; management; synthesis; translation

Rapid discoveries in genomics and other “omic” fields are creating expectations that new tests and interventions will be developed for use in clinical practice and disease prevention.¹ Cancer has been at the forefront of clinical applications of these technologies, which offer the potential to use germ-line and tumor genomic data to develop personalized interventions.² Despite this promise, the information needed to move applications into clinical practice is often scarce and stakeholders sometimes disagree on how much evidence is needed.³

In this issue of *Genetics in Medicine*, three articles tackle the complexities of gathering, evaluating, and disseminating the evidence on genomic tests. Two articles stem from a National Cancer Institute initiative on comparative effective research in cancer genomics and personalized medicine.⁴ Comparative effective research arose as a result of increased interest in measuring patient-centered health outcomes and comparing alternative approaches to disease prevention and treatment.⁵ Genomic medicine provides an ideal opportunity to apply comparative effective research methods to compare genomic tools and applications with usual care in real-world settings.⁶

In the first article, Goddard et al.⁷ identify approaches to assessing genomic applications through literature reviews and demonstrate lessons learned from the National Cancer Institute initiative. Using case studies, they identify significant challenges in the conduct and evaluation of comparative effective research, including the rapid pace of innovation and data acquisition,

lack of oversight, and variable evidentiary thresholds for clinical and personal utility. They conclude that a variety of methodological approaches are needed to develop and synthesize the knowledge needed to ensure an effective translation of genomic discoveries in cancer. These approaches include a combination of comparative observational studies and randomized trials, as well as decision modeling and economic analysis of patient-centered outcomes.

In the second article, Deverka et al.⁸ explore how stakeholders view knowledge about genomic applications in cancer. Stakeholders include clinicians, insurers, test developers, advocates, policy makers, and others whose views can influence translation from research to practice. Using case studies, the authors present results of evidence synthesis to 22 diverse stakeholders who participated in a workshop to explore the evidence of cancer genomic tests for clinical practice and coverage decision-making. Describing how the stakeholders’ opinions on evidentiary thresholds diverged and changed during the workshop, the authors highlight the need for ongoing stakeholder engagement in unbiased settings. A common understanding of the existing evidence base should guide the development of evidentiary thresholds in genomic medicine.

In the third article, Wallace et al.⁹ explore a new way to rapidly update the evidence base on genomic applications in practice. Systematic reviews and meta-analyses, the principal tools for evidence synthesis, are labor intensive and time consuming;

¹Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia, USA; ²Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland, USA; ³McKing Consulting Corp, Atlanta, Georgia, USA. Correspondence: Muin J. Khoury (muk1@cdc.gov)

Submitted 23 February 2012; accepted 13 March 2012; advance online publication 3 May 2012. doi:[10.1038/gim.2012.43](https://doi.org/10.1038/gim.2012.43)

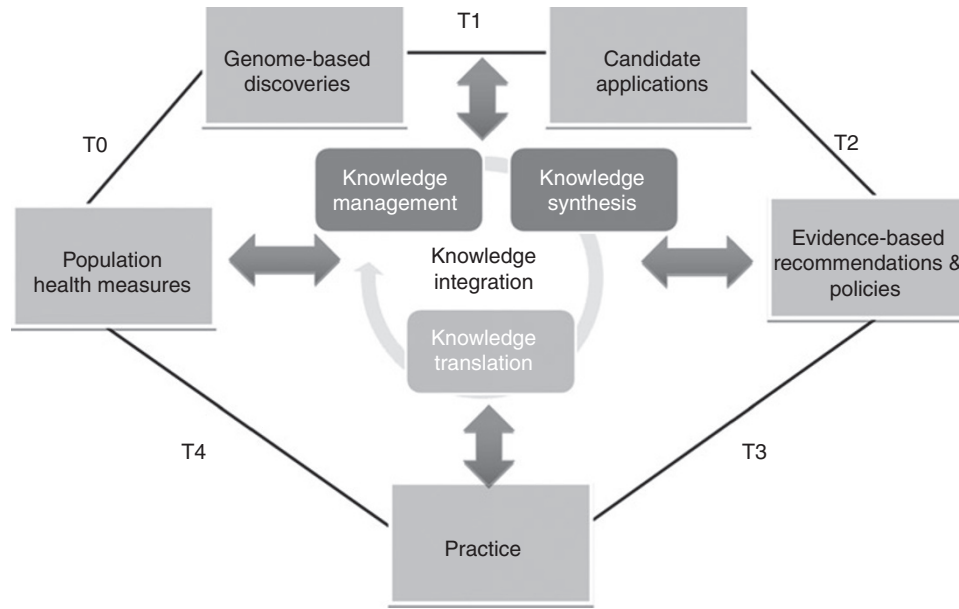


Figure 1 Components of knowledge integration in genomic medicine. Modified from Khoury *et al.*¹³

thus, the volume and rapid evolution of information in genomic medicine presents a substantial challenge. The authors demonstrate an approach to mining curated, online knowledge bases that can reduce the burden of updating systematic reviews. Their findings provide important impetus for the development and deployment of modern approaches based on text or data mining and other technologies to reduce the labor necessary to produce and maintain systematic reviews in the rapidly developing field of genomic medicine.

KNOWLEDGE INTEGRATION AT THE CENTER OF GENOMIC MEDICINE

Each of these three articles addresses a different component of the “knowledge integration” (KI) process, a term that has been used to mean different things in different contexts.^{10,11} Burke *et al.*¹² viewed KI within and across disciplines as the engine for the effective use of genomic information to improve health. They defined KI as “the process of selecting, storing, collating, analyzing, integrating and disseminating information both within and across disciplines for the benefit of population health. It includes methodological development, and is the means by which information is transformed into useful knowledge.”¹²

Because genomic discoveries result from basic, clinical, and population research, KI is at the center of translational activities in genomic medicine as illustrated in the T1–T4 translational pathway discussed elsewhere¹³ and elaborated here in **Figure 1**. Translational research proceeds in phases, using basic genome-based discoveries to develop promising applications such as tests and drugs (T1), evaluating efficacy of such applications and developing evidence-based recommendations (T2), implementing and disseminating evidence-based recommendations into clinical and public health programs (T3), and measuring effectiveness and cost-effectiveness of genomic applications at

the population level (T4). Most funded and published genomic research, even in cancer, remains either in the discovery or early translation phases¹⁴ and the evidence base for genomics in practice remains limited. The sheer volume and variety of information accumulating from primary research creates tremendous potential noise.¹⁵

A robust KI process is needed to digest this information and transform it into knowledge that drives policy, practice, and further research. Three essential components constitute the essence of KI: knowledge management (KM), knowledge synthesis (KS), and knowledge translation (KT; **Table 1** and **Figure 1**). These are common to all areas of biomedical research and genomics is no exception; however, the volume of new genomic information and the speed with which it is developing has the potential to affect all areas of medicine and public health.

KM: HORIZON SCANNING AND INFOVEILLANCE

The first component of KI is KM, which is a continuous process of horizon scanning to select, store, curate, and track relevant information from multiple disciplines and phases of translation. Bioinformatics tools can help automate the process of mining for information on genomic discoveries and their potential functional significance, as well as raise hypotheses about potential therapeutic or preventive interventions. The National Center for Biotechnology Information of the National Library of Medicine has organized a wealth of diverse information in open-access, online databases focused primarily on basic research; these include genomic sequencing databases, functional databases, and locus-specific databases.¹⁶

Developing methods for horizon scanning and surveillance of relevant information is part of the new field of “infoveillance.”¹⁷ An example of this approach is the Human Genome Epidemiology Navigator (HuGE Navigator¹⁸), a

Table 1 Components of knowledge integration in genomic medicine: definitions and examples

Component	Definition	Examples
Knowledge management	Process of scanning, selecting, storing, curating, and tracking genomic research information from multiple disciplines and phases of translation	Curated literature and Web-based databases, e.g., HuGE Navigator for genetic-epidemiology studies; ¹⁸ GAPPKB to find new genomic tests and link with available studies; ^{22,23} Genetic Testing Registry ²⁴
Knowledge synthesis	Systematic review of information from multiple disciplines to assess validity and utility of information; process usually employs methods of meta-analysis and can use modeling of value of information, using direct and indirect evidence	AHRQ evidence-based reviews; ²⁸ Cochrane reviews; ²⁷ CancerGEMKB; ²³ EGAPP reviews and recommendations; ²⁹ PharmGKB; ³⁰ PLoS Currents: evidence on genomic tests; ³² NIH consensus conferences; ³³ US Preventive Services Task Force; ³⁴ National Institute for Clinical Excellence ³⁵
Knowledge translation	Process of actively disseminating synthesized information to influence policy, guideline development, and research across the translation continuum, as well as clinical and public health practice. The process uses stakeholder engagement and knowledge brokering	IOM roundtable on translating genome-based research for health; ³⁹ the Genomic Applications in Practice and Prevention Network; ⁴⁰ Clinical Pharmacogenetic Implementation Consortium ⁴¹

AHRQ, Agency for HealthCare Research and Quality; CancerGEM KB, Cancer Genomic Evidence-based Medicine Knowledge Base; EGAPP, Evaluation of Genomic Applications in Practice and Prevention; GAPPKB, Genomic Applications in Practice and Prevention Knowledge Base; HuGE Navigator, Human Genome Epidemiology Navigator; IOM, Institute of Medicine; PharmGKB, Pharmacogenomics Knowledge Base.

continuously updated, curated knowledge base of published genetic association studies maintained by the Centers for Disease Control and Prevention. The number of such publications has quadrupled over the past 10 years, with >9,500 articles published in 2011 alone.¹⁹ The HuGE Navigator uses a combination of text-mining algorithms and human curation,²⁰ which is also the approach taken by Wallace et al.⁹ By placing a wealth of genetic association information at the fingertips of researchers and other users, such online applications can improve the efficiency and reduce the time required for KS.

The infoveillance approach can be applied to subsequent stages of translation. For example, the evaluation of genomic tests requires information on analytic performance, clinical validity, clinical utility, and ethical and legal issues.²¹ The Genomic Applications in Practice and Prevention Knowledge Base of the Centers for Disease Control and Prevention is a continuously updated database of emerging genomic tests proposed for use in clinical or public health practice.²² Between October 2009 and January 2012, >400 newly introduced tests were added to the database; approximately two-thirds of these tests were related to cancer.²² We have combined cancer-related information from the HuGE Navigator and the Genomic Applications in Practice and Prevention Knowledge Base into a trial version of CancerGEM KB,²³ which will be developed further in the next few years. This year, the National Institutes of Health plans to release the Genetic Testing Registry, which will provide public access to information submitted voluntarily by test developers.²⁴

KS: SYSTEMATIC REVIEWS, META-ANALYSES, AND MODELING

While KM is an essential first step in the KI process, a crucial second step is KS, which “makes sense” of the incoming information and transforms it into information that answers specific questions both within and across different scientific disciplines (Table 1 and Figure 1). For example, within the realm of genetic associations, it is important to know whether or not

a reported genetic association has been replicated across multiple populations and the magnitude of disease risks conferred by specific genetic variants in different populations. In the pre-genome-wide association studies era, candidate gene analyses often led to nonreplicable results.²⁵ The problem has been partially alleviated by the use of large-scale consortia and networks with sufficient sample size to conduct rigorous replication and meta-analysis. The National Human Genome Research Institute curates an online knowledge base of genome-wide association studies results, integrated with other information from the National Center for Biotechnology Information.²⁶

KS applies technical methods for systematic review of published and unpublished data using *a priori* rules of evidence. KS may include meta-analysis, decision analysis, and modeling to combine information from different study designs and different domains in basic, clinical and population research (e.g., Cochrane collaboration reviews;²⁷ Agency for Healthcare Research and Quality funded evidence-practice centers²⁸). The independent, multidisciplinary Evaluation of Genomic Applications in Practice and Prevention working group²⁹ has developed evidentiary rules for conducting comprehensive systematic reviews of the analytic validity, clinical validity, and clinical utility of genomic tests in specific clinical scenarios.

In the arena of pharmacogenomics, the Pharmacogenomics Knowledge base (PharmGKB, <http://www.pharmgkb.org/>) has for >10 years created a repository of primary data as well as tools to track associations between genes and drugs, and to catalog the location and frequency of genetic variations related to drug response.³⁰ With the explosion of new data over the past 10 years, PharmGKB now focuses on curating and synthesizing knowledge, and captures more complex relationships between genes, variants, drugs, diseases, and pathways.

In general, KS can be slow and tedious and is often criticized for this reason. As suggested by Goddard et al.⁷ and Veenstra et al.,³¹ additional tools could be used in the rapidly developing landscape of genomic medicine, including value of information analysis, decision analysis, and modeling. All these

tools can supplement but not replace synthesis of empirical data from observational studies and clinical trials. Methods for rapid review and display of existing information are in development. To accelerate KS for genomic tests, the Centers for Disease Control and Prevention recently launched an online publication through the open-access Public Library of Science. The collection, entitled Public Library of Science Currents: Evidence on Genomic Tests, Public Library of Science³² is intended as a rapid publication channel for synopses of knowledge on genomic tests. KS is often used by independent panels to develop evidence-based recommendations and practice guidelines.^{29,33–35} Much more methodological work is needed to accelerate the pace of KS.

KT: BROKERING INFORMATION TO INFLUENCE RESEARCH, POLICY, AND PRACTICE

The third component of KI is KT. Deverka *et al.*⁸ clearly illustrate that synthesized knowledge tends to be viewed differently by various stakeholders. If genomic medicine is to succeed, a strong KT process is crucial. KT is the active process of disseminating synthesized information to influence policy, guideline development, practice, and research across the translation continuum. This is the most “messy” component of KI because using information to influence research and practice requires “buy-in” from stakeholders with different perspective and can be an iterative process.

Evidence-based recommendations are necessary but not sufficient to move genomic medicine applications into practice. Many forces can affect their diffusion, adoption, and implementation. These forces often operate independently of knowledge synthesis and span a wide spectrum, including private investments in research and development, policy and legal frameworks, oversight and regulation, product marketing, coverage and reimbursements, consumer advocacy, provider awareness, access, and health services development and implementation. Deverka *et al.*⁸ demonstrate once again that payers generally require a higher level of evidence of clinical utility than genomic researchers or test developers. Issues around differential access and implementation lead to the phenomenon of “lost in translation” in clinical practice.³⁶ Furthermore, translation requires implementation sciences, health services, and outcomes research agendas (T3 and T4 research) that are currently underrepresented in funding and publications of genomics research.³⁷

An important component of KT is the convening of stakeholders around KS to address differences in evidentiary thresholds that can drive decision making.³⁸ Such collaborations link researchers and policy makers, facilitate interactions, understand goals and professional culture, and forge new partnerships to use evidence from existing knowledge and define areas for future research. This process, sometimes called “knowledge brokering,” is ultimately about developing and using evidence-based decision-making to deliver genomic medicine in clinical and public health settings. To be successful, KT needs to lead to evidentiary standards and empower

independent, transparent appraisal of the evidence. Examples of such collaborations include the Institute of Medicine roundtable on translating genomics into health³⁹ and the Genomic Applications in Practice and Prevention,⁴⁰ a collaboration of multiple groups including government agencies, the private sector, academia, consumers, and clinical and public health practice.⁵

In the arena of pharmacogenomics, the Clinical Pharmacogenetics Implementation Consortium was formed in late 2009 to develop and publish peer-reviewed guidelines with simultaneous posting to PharmGKB with supplemental information.⁴¹ The goal of Clinical Pharmacogenetics Implementation Consortium is to address some of the barriers to implementation of pharmacogenetic tests into clinical practice. Clinical Pharmacogenetics Implementation Consortium guidelines are designed to help clinicians understand how available genetic test results should be used to optimize drug therapy, rather than if such tests should be ordered, thus informing the stakeholder dialogue on evidence-based implementation of pharmacogenomics testing in clinical practice

CONCLUSION

The three papers in this issue^{7–9} illustrate why a robust KI process is needed to drive the growth and development of genomic medicine for years to come. Here, we have elaborated on the three components of KI: management, synthesis, and translation. Because of the rapid emergence of complex “omic” data from basic, clinical, and population research, we believe an adequately resourced KI process with as much automation as possible is needed to keep up with the avalanche of information. A multistakeholder KI enterprise should involve both the public and private sectors to ensure a rapid, transparent and credible process than can drive policy and practice.

ACKNOWLEDGMENTS

We thank Andrew Freedman and Deborah Winn for their comments on an earlier version of the manuscript.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

- Green ED, Guyer MS; National Human Genome Research Institute. Charting a course for genomic medicine from base pairs to bedside. *Nature* 2011;470:204–213.
- McDermott U, Downing JR, Stratton MR. Genomics and the continuum of cancer care. *N Engl J Med* 2011;364:340–350.
- Institute of Medicine Roundtable on Translating Genome-based Research for Health. Generating Evidence for Genomic Diagnostic Test Development: A Workshop Summary. National Academies Press: Washington, DC, 2011.
- National Cancer Institute. Comparative effectiveness research in genomics and personalized medicine. <http://cancercontrol.cancer.gov/od/phg/research.asp?type=CER>. Accessed 1 February 2011.
- Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med* 2009;151:203–205.
- Khoury MJ, Rich EC, Randhawa G, Teutsch SM, Niederhuber J. Comparative effectiveness research and genomic medicine: an evolving partnership for 21st century medicine. *Genet Med* 2009;11:707–711.

7. Goddard KAB, Knaus WA, Whitlock E, et al. Building the evidence base for decision making in cancer genomic medicine using comparative effectiveness research. *Genet Med*, this issue.
8. Deverka PA, Schully SD, Ishibe N, et al. Stakeholder assessment of the evidence for cancer genomic tests: insights from three case studies. *Genet Med*, this issue.
9. Wallace BC, Small K, Brodley CE, et al. Towards modernizing the systematic review pipeline in genetics: efficient updating via data mining. *Genet Med*, this issue.
10. Wikipedia. Knowledge integration. http://en.wikipedia.org/wiki/Knowledge_integration#mw-head. Accessed 1 February 2012.
11. Kerner JF. Knowledge translation versus knowledge integration: a 'funder's' perspective. *J Contin Educ Health Prof* 2006;26:72–80.
12. Burke W, Khoury MJ, Stewart A, Zimmern RL; Bellagio Group. The path from genome-based research to population health: development of an international public health genomics network. *Genet Med* 2006;8:451–458.
13. Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med* 2007;9:665–674.
14. Schully SD, Benedicto CB, Khoury MJ. How can we stimulate translational research in cancer genomics beyond bench to bedside? *Genet Med* 2012;14:169–170.
15. Ioannidis JP, Khoury MJ. Improving validation practices in "omics" research. *Science* 2011;334:1230–1232.
16. National Center for Biotechnology Information, National Library of Medicine. <http://www.ncbi.nlm.nih.gov/>. Accessed 1 February 2012.
17. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011;40(5 suppl 2):S154–S158.
18. Human Genome Epidemiology Navigator (HuGE Navigator). <http://www.hugenavigator.net/HuGENavigator/home.do>. Accessed 1 February 2012.
19. Khoury MJ, Gwinn M, Clyne M, Yu W. Genetic epidemiology with a capital E, ten years after. *Genet Epidemiol* 2011;35:845–852.
20. Yu W, Gwinn M, Clyne M, Yesupriya A, Khoury MJ. A navigator for human genome epidemiology. *Nat Genet* 2008;40:124–125.
21. Haddow JE, Palomaki GE. ACCE: a model process for the evaluation of genetic tests. In: Khoury MJ, Little J, Burke W (eds). *Human Genome Epidemiology*. Oxford University Press: New York, 2004:217–233.
22. Gwinn M, Grossniklaus DA, Yu W, et al. Horizon scanning for new genomic tests. *Genet Med* 2011;13:161–165.
23. Cancer Genomics Evidence based Medicine Knowledge Base CancerGEM KB. <http://www.hugenavigator.net/CancerGEMKB/home.do>. Accessed 1 February 2012.
24. National Center for Biotechnology Information: Genetic Testing Registry. <http://www.ncbi.nlm.nih.gov/gtr/>. Accessed 1 February 2012.
25. Siontis KC, Patsopoulos NA, Ioannidis JP. Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. *Eur J Hum Genet* 2010;18:832–837.
26. National Human Genome Research Institute. GWAS catalog. <http://www.genome.gov/gwastudies/>. Accessed 1 February 2012.
27. The Cochrane collaboration. <http://www.cochrane.org/>. Accessed 1 February 2012.
28. Agency for Healthcare Research and Quality: evidence practice centers. <http://www.ahrq.gov/clinic/epc/>. Accessed 1 February 2012.
29. Evaluation of Genomic Applications in Practice and Prevention (EGAPP). <http://www.egappreviews.org/>. Accessed 1 February 2012.
30. Thorn CF, Klein TE, Altman RB. Pharmacogenomics and bioinformatics: PharmGKB. *Pharmacogenomics* 2010;11:501–505.
31. Veenstra DL, Roth JA, Garrison LP Jr, Ramsey SD, Burke W. A formal risk-benefit framework for genomic tests: facilitating the appropriate translation of genomics into clinical practice. *Genet Med* 2010;12:686–693.
32. PLoS Currents: evidence on genomic tests. <http://knol.google.com/k/plos/plos-currents-evidence-on-genomic-tests/28qm4w0q65e4w/50#>. Accessed 1 February 2012.
33. NIH consensus development program. <http://consensus.nih.gov/>. Accessed 1 February 2012.
34. Agency for Healthcare Research and Quality. US Preventive Services Task Force. <http://www.ahrq.gov/clinic/uspstfix.htm>. Accessed 1 February 2012.
35. National Institute for Health and Clinical Excellence. <http://www.nice.org.uk/>. Accessed 1 February 2012.
36. Burke W, Burton H, Hall AE, et al.; Ickworth Group. Extending the reach of public health genomics: what should be the agenda for public health in an era of genome-based and "personalized" medicine? *Genet Med* 2010;12:785–791.
37. Khoury MJ, Clauser SB, Freedman AN, et al. Population sciences, translational research, and the opportunities and challenges for genomics to reduce the burden of cancer in the 21st century. *Cancer Epidemiol Biomarkers Prev* 2011;20:2105–2114.
38. Khoury MJ. Dealing with the evidence dilemma in genomics and personalized medicine. *Clin Pharmacol Ther* 2010;87:635–638.
39. Institute of Medicine roundtable on translating genome-based research for health. <http://iom.edu/Activities/Research/GenomicBasedResearch.aspx>. Accessed 1 February 2012.
40. Khoury MJ, Feero WG, Reyes M, et al.; GAPPNet Planning Group. The genomic applications in practice and prevention network. *Genet Med* 2009;11:488–494.
41. Relling MV, Klein TE. CPIC: The Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther* 2011;89:461–464.