Finding the Evidence in Real-World Evidence: Moving from Data to Information to Knowledge

Richard L Schilsky, MD, FACP, FASCO

Information to guide the delivery of optimal cancer care has traditionally been developed through the conduct of prospective clinical trials sponsored by the National Cancer Institute or by commercial entities, ie the pharmaceutical industry. Results of these trials form the evidence base for clinical practice guidelines, treatment pathways, drug compendia listings, and other care standards that support reimbursement policies and, therefore, access to care. Where evidence from clinical trials is lacking, reimbursement policies and clinical decision-making are typically supported by data from observational studies, tumor registries, or analysis of outcomes derived from insurance claims data. As treatment options for cancer patients continue to expand, all of these mechanisms will likely be necessary to understand how treatments compare, and which treatment works best for which patients. Collectively, the data obtained from sources outside of traditional clinical trials are often referred to as real world data, and the evidence derived from aggregation and analysis of such data as real world evidence (RWE). Real world data typically display the characteristics of "big data," namely, volume, velocity, variety, and veracity, with the latter often presenting the greatest challenge for evidence generation.

CONTRASTING RANDOMIZED CLINICAL TRIALS AND REAL WORLD EVIDENCE

Randomized clinical trials (RCTs) provide the highest level of evidence to establish the efficacy of the intervention being studied. Oncology RCTs conducted by both the academic research community and commercial sponsors have provided data to support the regulatory approval of new drugs or new indications for existing drugs that can potentially cure or improve survival of cancer patients; refine the methods of delivery, scheduling, and dosing of oncology drugs; identify subpopulations of patients who are most likely to benefit (or be harmed) from a specific therapy; and establish the utility of combining different therapeutic modalities to treat patients.[1,2]

Although RCTs have clearly advanced the care of cancer patients, they have significant limitations. An RCT is costly to develop and conduct. The process of developing and activating an RCT is slow and is plagued by a burdensome infrastructure and substantial regulatory oversight.[3] These trials often require large numbers of patients to identify modest differences between treatments and can take years to accrue and reach the primary endpoint being studied. An RCT typically requires complex protocols and collection of large amounts of patient data and documentation, which increases the work load and costs for participating sites. Recent studies suggest that a substantial proportion of phase III oncology trials are never completed, wasting both financial and patient resources.[4] As the treatment of cancer advances and new findings are discovered, the delays in start-up and completion of RCTs may lead to results that are no longer relevant by the time they are reported due to changing standards of care. Furthermore, all RCTs have eligibility criteria in order to define the patient population necessary to address the trial's objectives. Eligibility criteria, by their nature, limit the applicability of the trial results. Therefore, critics of RCTs argue that the patient population studied often does not reflect the "real world" practice of medicine because the inclusion criteria may lead to selection of only the healthiest patients and may exclude patients with medical comorbidities or borderline organ function. So, although an RCT may adequately assess the efficacy of an intervention (ie what can work); the "real world" effectiveness that is seen once the intervention is deployed in community practice (ie what does work) may be substantially different.

In addition, RCTs often evaluate therapies under idealized clinical conditions, including protocol-specified dose modifications and toxicity management; therefore, the results generated from an RCT may not be replicated when the therapy is translated to general practice settings and to real world patients. Furthermore, the efficacy endpoints traditionally used in cancer clinical trials may not reflect outcomes that are most important to patients, such as relief of symptoms, improvement in quality of life, or

Disclosure Information: Nothing to disclose.

Presented at the American College of Surgeons 102nd Annual Clinical Congress, Washington, DC, October 2016.

Received October 3, 2016; Accepted October 3, 2016.

From the Executive Office, American Society of Clinical Oncology, Alexandria, VA.

Correspondence address: Richard L Schilsky, MD, FACP, FASCO, American Society of Clinical Oncology, 2318 Mill Rd, Suite 800, Alexandria, VA 22314. email: richard.schilsky@asco.org

http://dx.doi.org/10.1016/j.jamcollsurg.2016.10.025
ISSN 1072-7515/16

Abbreviations and Acronyms

CER = comparative effectiveness research
EGFR = epidermal growth factor receptor
RCT = randomized clinical trial
RWE = real world evidence

achievement of personal goals. Better measures of these patient-reported outcomes are urgently needed and must be incorporated in clinical trials to better assess the impact and value of a new treatment.

Tumor heterogeneity also challenges the ability to develop new cancer treatments through traditional prospective clinical trials. Because common tumors are often divided into rare molecular subtypes, it is increasingly challenging to identify eligible patients and complete recruitment to clinical trials in a timely fashion. Rates of enrollment of adult cancer patients in clinical trials remain stagnant, at no more than 3% to 5%. With more tumor types, more drugs, fewer eligible patients, and strained research budgets, it is no longer possible to learn everything that still needs to be learned in cancer treatment through the conduct of conventional, prospective clinical trials. These limitations have given rise to interest in the use of RWE to fill knowledge gaps that simply cannot be addressed by conventional clinical trials.

A challenge is to extract information from real world data that provides clinically meaningful and reliable insights that can be applied in patient care. Figure 1 provides an example of a knowledge hierarchy to illustrate progression of data to information to knowledge to wisdom. Detection of the L858R mutation in the epidermal growth factor receptor (EGFR) provides a piece of data that signifies a genomic alteration. The information attributed to this data element is that it represents a DNA mutation that sensitizes tumor cells to EGFR tyrosine kinase inhibitors. The knowledge of how to use these inhibitors comes from clinical studies that demonstrate clinical benefit for lung cancer patients whose tumors harbor these mutations, but the wisdom associated with their use derives from the recognition that not all patients will benefit, and for those who do, the benefit will likely be transient.

It is useful to contrast the strengths and weaknesses of RCTs and RWE, and the advantages and disadvantages of each are depicted in Tables 1 and 2. Advantages of RCTs include collection of data that are complete, accurate, unbiased, and standardized. Disadvantages are the long time and high expense typically required to complete an RCT and the lack of generalizability of the trial data to populations not eligible for study participation. By contrast, RWE has the advantage of capturing the outcomes of patients in the usual practice setting. But studies that rely on RWE are also subject to bias, incomplete or inaccurate data, and use of data elements and outcomes measures that are not standardized across study sites, all of which contribute to concerns about the reliability of the information obtained from analysis of such data sets. The many potential sources of bias in RWE are

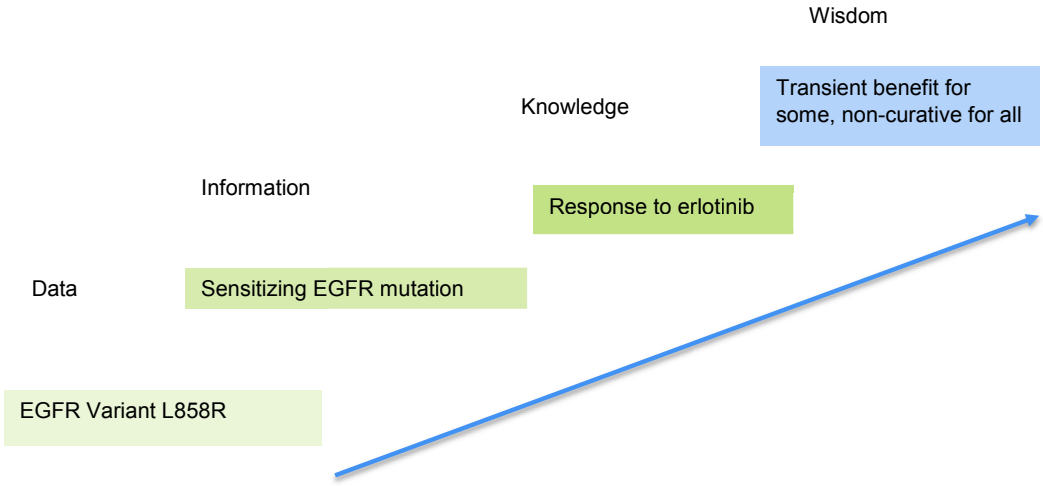

Figure 1. Knowledge hierarchy. Detection of the L858R mutation in the epidermal growth factor receptor (EGFR) provides a piece of data that signifies a genomic alteration. The information attributed to this data element is that it represents a DNA mutation that sensitizes tumor cells to EGFR tyrosine kinase inhibitors. The knowledge of how to use these inhibitors comes from clinical studies that demonstrate clinical benefit for lung cancer patients whose tumors harbor these mutations, but the wisdom associated with their use derives from the recognition that not all patients will benefit, and for those who do, the benefit will likely be transient.