Are Randomized Controlled Trials Overrated? Time to Revisit the Evidence Pyramid

by Dr Bud Wiederman MD, MA, Evidence eMended Editor, Grand Rounds

Tom Frieden, the former director of the CDC, published a nice review article earlier this month. He highlights a long-recognized problem in assigning too much legitimacy to the gold standard randomized controlled trial, at the expense of other study designs that might actually provide better evidence for clinical management.


Welcome to a 5th Tuesday on Evidence eMended, an opportunity for me to go rogue and comment on an article not included in AAP Grand Rounds. I'm especially pleased to highlight what is essentially a review article/editorial by Dr. Frieden that gives us much food for thought.

Frieden’s basic premise is that too much deference is given to the randomized controlled trial (RCT) study design, perhaps overlooking important aspects of how study results can (or cannot) be applied in real life. Even a very well designed and conducted RCT may fail to impact everyday health care delivery, most often because either the study conditions can’t be reproduced in daily medical practice. This might happen, for example, if the RCT employed a small army of research nurses who helped study subjects comply with a complex treatment and follow up regimen, a situation that isn't feasible in clinical practice. Alternatively, the patient population under study might have been too narrow for the study findings to be generalized beyond a relatively small group of patients - e.g. the study lacks generalizability (aka external validity).

At this point, it is useful to recall the so-called "evidence pyramid." As I've mentioned previously, this pyramid I think has been touted too highly, overlooking the original premise of the hierarchy of study designs it is intended to represent. What I said back in 2014 was that "... the pyramid is frequently misunderstood..., but all it says is that some study designs (e.g. randomized controlled trials), if performed well, are more likely to stand up to scrutiny years later than are other design types... Stated differently, a well-performed case-control study is more likely to be proven incorrect in subsequent studies than is a well-performed randomized controlled trial.”

Frieden elaborates on the growing body of literature describing limitations of RCTs, as well as some great successes from other study designs. He mentions the case-control studies of Sudden Infant Death Syndrome (SIDS) that suggested a link between prone sleeping and SIDS. Subsequently, a movement to educate parents to have their infants sleep on their backs has been associated with a decreased rate of SIDS internationally. I would add the similar phenomenon of the case-control studies showing an association of aspirin with Reye Syndrome; subsequent campaigns to eliminate aspirin as a fever treatment in children (in the face of great opposition from aspirin manufacturers, by the way) led to the virtual disappearance of Reye Syndrome.

Frieden mentions 5 other examples from the public health arena of non-RCT studies that led to improvements in healthcare: 1) post-marketing analysis of influenza vaccine efficacy led to suspension of the use of live attenuated influenza vaccine spray when poor (even near zero!) efficacy was demonstrated; 2) implementing directly observed therapy for tuberculosis, without the aid of an RCT, turned out to be a practical approach to managing treatment of large populations of infected individuals; 3) population analyses of sodium intake and cardiovascular health led to reduced salt intake and lowered heart disease rates (though still subject to some controversy); 4) rare disease registries, for illnesses like multiply-resistant tuberculosis infections too rare to allow for an RCT to be performed, have led to breakthroughs in novel antibiotic treatments; and 5) using novel infrastructure study designs to demonstrate superiority of thiazolidinediones over sulfonylureas for type 2 diabetes.
He includes 1 very large table in his article that I’m sure will be used by medical educators and medical and other healthcare students across the world to help remember the differences, strengths, and weaknesses of various study designs.

As I said in 2014, if you haven't heard of GOBSAT, or you just can't recall what it means, Google it!

Frieden closes with a comment on “big data” that can be mined now that most medical care is documented digitally. He seems perhaps a bit premature in his enthusiasm about this innovation, in my opinion. Particularly with regard to the electronic health record, which is overly (again, my opinion) focused on the billing aspects of medical care, I find the quality of the actual medical record to be worse than the bad old days of handwritten notes. (At least the notes are legible, however!) There are simply too many instances of clicking on template statements and copying old notes forward into new ones, all creating an incomplete or even inaccurate picture of the patient, for me to be comfortable that the data can be mined at that level.

Frieden's main premise, elevating RCTs to the extent that truly valuable non-RCT studies are overlooked, is very much on target. We should keep in mind the hierarchy of the evidence pyramid, but also remember its limitations.