Editorial

The potential value of discordant studies

Mark Doyle

Allegheny Health Network, Pittsburgh, PA, USA

Corresponding to: Mark Doyle, PhD. Division of Cardiology, CV MRI Center, Allegheny General Hospital, 320 East North Avenue, Pittsburgh, PA 15212-4772, USA. Email: mdoyle@wpahs.org.

Abstract: Many aspects of the clinical application of gated-single photon emission computed tomography (SPECT) have been well established by multiple trials and studies. However, its utility in the elderly (i.e., the Medicare population in the USA) remains unclear. This is an important population due to its rapid growth, coupled with the increasing prevalence of coronary artery disease with age. A paper in this issue, *Predictive value of exercise myocardial perfusion imaging in the Medicare population: the impact of the ability to exercise*, indicates that while gated-SPECT clearly directs the performance of interventions at the level of the coronary arteries in the elderly, outcomes are worse for those receiving an intervention *vs*. those receiving medical therapy. While some literature supports this observation, there are also well documented studies that indicate that the opposite is the case. As consumers of discordant studies, we find ourselves in the unenviable position of having to pull at the threads of evidence and follow them through in an attempt to reconcile the conflicting literature. This is reminiscent of the mythical Gregorian knot, a knot that was impossible to unravel by conventional means. However, it was "solved" by cutting it with a sword. In our case, the sword that we have is the removal of bias. It has been said that there are no unbiased studies, since we only measure what we believe and we tend to believe what we measure. This is further compounded in clinical practice since the Hippocratic Oath requires that the physician above all do no harm. Therefore it follows that whatever action is done is at least not detrimental to the patient. These are powerful belief systems that on the one hand allow us to rapidly discard "irrelevant" information and quickly get to the important point, but on the other hand they may inhibit us from seeing what is truly of value. Discordant and negative studies are important disruptors along the path to easy data assimilation, and force us to seek out sources of bias which otherwise may go unnoted. In the case of the above paper we might look past the perfusion data to the more important cardiac functional data which may contribute to changing the focus of diagnosis and treatment strategies, thus slicing through a little more of the knot.

Keywords: Cardiovascular; single photon emission computed tomography (SPECT); elderly; coronary artery intervention; bias

Submitted Feb 18, 2014. Accepted for publication Feb 19, 2014.
doi: 10.3978/j.issn.2223-3652.2014.02.10
Scan to your mobile device or view this article at: http://www.thecdt.org/article/view/3364/4202

In this issues' article by Kwon *et al*., *Predictive value of exercise myocardial perfusion imaging in the Medicare population: the impact of the ability to exercise* it is established that elderly patients undergoing exercise single photon emission computed tomography (SPECT), who have evidence of ischemia and undergo an intervention such as percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG), suffer worse outcomes compared to patients who received medical therapy (1). This is in contrast to younger populations where the value of identifying ischemic regions and differentiating them from fixed defects has important implications for considering a coronary artery intervention. It is known that maintaining homeostasis for an elderly individual generally requires increased use of reserve mechanisms even under resting conditions (2). Thus, any perturbation more easily pushes the elderly in to physiological trouble compared to the young. This might contribute to the adverse effects of

coronary artery intervention noted for the elderly. However, contradictory studies in the literature disrupt this line of reasoning. Necessarily, any comparative literature cited is almost guaranteed to relate to earlier studies, and thus changes in medication and procedures over time are bound to be suspected, along with difference in patient selection criteria between studies. In the above paper, differences are noted *vs*. an earlier major study involving 5,200 elderly patients by Hachamovitch *et al*. (3). It is speculated that medication may have not been comparable to contemporary standards and that the patient populations may have had different size ischemic regions as a group. Indeed, even contemporary use of medication in the elderly may not be optimally prescribed (4). However, in the absence of more detail information it is difficult to adjudicate the role of medication. Non-reporting of such details is understandable when reports were paper based, but as online reporting increases in dominance, these restrictions can surely be lifted. Another reason that such details were and continue not to be reported is that they may not have been thought to be relevant, being part of the accepted infrastructure. This represents a form of bias that is inherent in the performance and reportage of studies, i.e., that if the focus of the study is, say myocardial perfusion, then features concerned with cardiac function may only be superficially reported. This is unfortunate if it later transpires that cardiac function is more important (as appears to be the case here), making comparison between studies harder for future researchers.

The issue of bias in reporting of clinical data has a long history and causes great harm. Consider the historical response to hand-washing in surgery, widespread physician endorsement of smoking, historical beliefs concerning the cause of epilepsy, clinical trials of thalidomide, and more recently to the treatment of hypercholesterolemia (5-7). Some investigators questioned the validity of the 2002 guidelines concerning the treatment of hypercholesterolemia with statins, claiming that the studies cited in support of statin treatment were statin trials, as opposed to cholesterol trials (8,9). The recent overturning of those guidelines, citing lack of evidence for the previous guidelines, may only partially address some of the original concerns. Further, it is well established that positive, confirmatory, findings are more likely to be published, and that negative results tend to be not included in the abstract or conclusions, the sections most likely to be skimmed by a busy reader (10,11). There has been a succession of cardiovascular medication clinical trials where issues became less resolved as more (targeted and designed) trials were performed (12,13). In the Kwon paper, the divergence with previous literature is clearly stated, and, if true, and if sufficiently widely believed, could have a profound effect on the diagnostic and treatment path for the elderly, affecting longevity, costs, and resource usage. However, the unresolved conflict with the prior literature may inhibit physicians and policy makers from acting on this information. Even in cases of apparently supportive studies, there may be significant differences that invalidate direct comparison such as the Trial of Invasive versus Medical therapy in Elderly (TIME) study (14). Initially the TIME data seems to be partially supportive of the Kwon study, but it is confounded by many patients assigned to the medically treated group later receiving a vascular intervention, and by substantial cross over between invasive *vs*. medically treated groups. However, divergence from prior literature potentially represents an extremely important path for progress, since it shows where gains have been made, where understanding needs to be improved, and allows those gains and new understanding to form the focus for producing greater future benefit. However, none of this can legitimately take place until sources of confounding bias are removed, permitting the fundamental truths to be exposed.

In the two apparently contradictory studies above by Kwon and Hachamovitch, as well as others, it is clear that features connected with poor cardiac function carry an increased burden of adverse outcomes that is even more predictive than perfusion status. However, this major point of agreement is reported in the Kwon paper via the end systolic volume (ESV) and in the Hachamovitch paper via the ejection fraction (EF). Doyle *et al*. recently reported on the prognostic value of the cardiac functional measure of the internal energy utilization, calculated from knowledge of the ESV and systolic blood pressure, and which was demonstrated to be superior to EF and to myocardial perfusion variables (15). Importantly, the index of internal energy utilization of the heart was calculated separately for SPECT and for magnetic resonance imaging (MRI) and demonstrated that adverse events were almost absent for patients below a definable threshold. Thus, a more fundamental measure of cardiac function (i.e., modality independent) may be available to help resolve discrepancies. One of the aims of adding to the literature is to progressively expose fundamental truths. Admittedly, it is unrealistic that even extensive reporting will allow the exact combination of variables to be compared between publications. However, when seeking reasons for disparity, a fuller listing may be useful in pointing out a probable,

as opposed to merely a possible, candidate for further investigation. The practice of cardiovascular medicine has benefited from several decades of major clinical trials producing powerful medications and devices. However, the rate of advancement is rapidly declining for several reasons, including increasing trial costs, the increasing rate of negative trials, withdrawal of long established medications due to side effects, and competition from other major diseases. In this environment, we may have to rely even more on investigator initiated studies to advance knowledge, providing an even greater mandate to allow bias-free comparison between studies.

Bias is not only inherent in the performance and reporting of studies, but also in the manner by which one modality measures a physical property of a patient compared to another imaging modality. This source of bias is quite distinct from noise and artifact, and represents a more fundamental interaction between the patient and the imaging modality. Doyle *et al*. recently introduced an approach to measure and compensate for these sources of bias, in this case between MRI and SPECT, but at this early stage it is unknown how many aspects of an imaging modality contribute to bias (16). For instance in the Kwon paper, the study focused on exercise SPECT, with limited characteristics reported of the other patients who underwent pharmacologic stress testing. While it is recognized that pharmacologic stress patients have different characteristics compared to exercise patients (including worse outcome) what is unknown is if the modality-patient sources of bias are additionally different between exercise and pharmacologic testing for a given modality. Differences in the mix of exercise *vs*. pharmacologic stress patients was cited as being a possible confounder when comparing the Kwon *vs*. Hachamovitch papers, and made further comparison between studies difficult. Hopefully, as time progresses and modality-dependent sources of bias are systematically removed, comparison between studies and modalities will become more feasible. Again, it is impossible to anticipate every aspect, but if such corrections or normalizations become widely used, it will assist in making comparisons between studies, providing ever better opportunities to advance knowledge, especially in the case of divergent studies.

Given the background of multiple levels of bias, what can be learned from the Kwon study in the context of prior studies? Firstly, measures of cardiac function carry higher prognostic value *vs*. myocardial perfusion status. This might help redirect future resources to focus on cardiac function above perfusion status. Importantly, cardiac function may be assessed with less expense. Secondly, given the prominent role of cardiac function, more attention should be directed to medications that effectively treat this aspect. Potentially, this can be retrospectively evaluated, especially in an environment where the tarnish has been repeatedly rubbed off the unassailability of the prospective, randomized clinical trial. Thirdly, we should report the data as broadly as is feasible, given that the truth may reside in the periphery of the main study.

Acknowledgements

Disclosure: The author declares no conflict of interest.

References

1. Kwon DH, Menon V, Houghtaling P, et al. Predictive value of exercise myocardial perfusion imaging in the Medicare population: the impact of the ability to exercise. Cardiovasc Diagn Ther 2014;4:5-12.
2. Lang PO, Michel JP, Zekry D. Frailty syndrome: a transitional state in a dynamic process. Gerontology 2009;55:539-49.
3. Hachamovitch R, Kang X, Amanullah AM, et al. Prognostic implications of myocardial perfusion single-photon emission computed tomography in the elderly. Circulation 2009;120:2197-206.
4. Anzai A, Maekawa Y, Kodaira M, et al. Prognostic implications of optimal medical therapy in patients undergoing percutaneous coronary intervention for acute coronary syndrome in octogenarians. Heart Vessels 2014. [Epub ahead of print].
5. Fintel B, Samaras AT, Carias E. The Thalidomide Tragedy: Lessons For Drug Safety And Regulation. HELIX MAGAZINE. Available online: http://helix.northwestern.edu/article/thalidomide-tragedy-lessons-drug-safety-and-regulation (accessed February 14 2014) Jul 28, 2009.
6. Friedlander WJ. The rise and fall of bromide therapy in epilepsy. Arch Neurol 2000;57:1782-5.
7. Stone NJ, Robinson J, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the american college of cardiology/american heart association task force on practice guidelines. J Am Coll Cardiol 2013. [Epub ahead of print].
8. Abramson J, Wright JM. Are lipid-lowering guidelines

evidence-based? Lancet 2007;369:168-9.
9. National Cholesterol Education Program, National Heart, Lung, and Blood Institute, National Institutes of Health. Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report. NIH Publication No. 02-5215. Bethesda, MD: National Cholesterol Education Program, National Heart, Lung, and Blood Institute, National Institutes of Health, 2002.
10. Ioannidis JP. How many contemporary medical practices are worse than doing nothing or doing less? Mayo Clin Proc 2013;88:779-81.
11. Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. BMJ 2011;342:d2686.
12. Sander GE, Giles TD. ALLHAT and ANBP2: what have we learned from recent mega-trials? Am J Geriatr Cardiol 2003;12:267-71.
13. Prasad V, Vandross A, Toomey C, et al. A decade of reversal: an analysis of 146 contradicted medical practices. Mayo Clin Proc 2013;88:790-8.
14. Pfisterer M; Trial of Invasive versus Medical therapy in Elderly patients Investigators. Long-term outcome in elderly patients with chronic angina managed invasively versus by optimized medical therapy: four-year follow-up of the randomized Trial of Invasive versus Medical therapy in Elderly patients (TIME). Circulation 2004;110:1213-8.
15. Doyle M, Weinberg N, Pohost GM, et al. Left ventricular energy model predicts adverse events in women with suspected myocardial ischemia: results from the NHLBI-sponsored women's ischemia syndrome evaluation (WISE) study. Cardiovasc Diagn Ther 2013;3:64-72.
16. Doyle M, Pohost GM, Merz CN, et al. Improved diagnosis and prognosis using Decisions Informed by Combining Entities (DICE): results from the NHLBI-sponsored Women's Ischemia Syndrome Evaluation (WISE). Cardiovasc Diagn Ther 2013;3:216-27.

Cite this article as: Doyle M. The potential value of discordant studies. Cardiovasc Diagn Ther 2014;4(1):1-4. doi: 10.3978/j.issn.2223-3652.2014.02.10

