European Urology Oncology xxx (xxxx) xxx

available at www.sciencedirect.com
journal homepage: euoncology.europeanurology .com

European Association of Urology

Facts and Myths About Stage Migration: Should the Will Rogers Phenomenon Ride off into the Distance?

The need for accurate diagnosis is a fundamental tenet of medical practice. Over the years the diagnostic process has progressed from simple observation to increasingly complex paradigms involving investigative procedures extending from blood tests to advanced imaging technologies. Each independent piece of information refines the diagnosis and contributes to guiding therapeutic choices. Nevertheless, health technology assessment agencies, supported by some clinical opinion-leaders, maintain that new technologies should only be approved if it can be shown that they improve patient outcomes, by which they generally mean survival as though this is the only beneficial outcome of an improved diagnosis [1,2].

This type of argument has recently played out in discussions about the benefits of prostate-specific membrane antigen positron emission tomography (PET)/computed tomography in various clinical settings for which these techniques have been shown to be significantly more accurate than conventional investigations [3] and are recommended by expert clinical bodies [4]. It has also been argued that when an improvement in survival is suggested, it should be dismissed as being a statistical artifact engendered by stage migration and the so-called Will Rogers phenomenon. This line of argument has led to the suggestion that before test substitution can be recommended, the criteria for introducing new imaging techniques in oncological practice should be more stringent and require results from randomized controlled trials comparing the effectiveness of the new test to alter outcomes compared with the current diagnostic standard [5]. Such a recommendation qualifies as an instance of the ethical principle of nonmaleficence, which prescribes that patients should not be worse off as a result of the medical interventions that ensue from a given diagnosis. In fact, it is supposed that "seeing more" without being able to "do more" results in overmedicalization of patients and may limit access to appropriate care [6].

Although we share the preoccupation that patients' overall well-being should be the highest-ranking priority for physicians, and that nonmaleficence ought to be respected, we would like to argue here that there might be a logical confusion in the above line of argument. The confusion concerns the Will Rogers effect: while we acknowledge that this is a statistical reality in some cases, it is turned into a specious argument in many situations. We believe that the specious argument rests on failing to distinguish between changes in conceptual structure, biases in study design, statistical descriptions, and the causal structure of the phenomenon itself.

1. A change in the conceptual structure

When a new and more accurate diagnostic test is introduced to replace an old one, patients that belonged to a certain group according to the old test may happen to migrate to another diagnostic group as an effect of the new diagnosis impacts the prognosis of both groups advantageously without altering the prognosis of the population as a whole. This phenomenon is also known as stage migration. Migration on the basis of a change in definition is ubiquitous in science. Pluto, formerly a planet of the solar system, migrated to the group of dwarf planets in 2006 as an effect of a change in the working definition of planets and of more accurate measurements of its mass, and therefore altered the number of planets in the solar system and the furthest distance of a planet from the sun without the solar system changing in any material way. Migration can thus be a change in our description of reality and in the way in which we conceptualize it. Technological progress is one factor and is especially relevant to the case of oncological imaging, but it is just one of many, along with conventions, expert decisions, and all the reasons that bring about changes in diagnostic and classification systems in medicine. Migration is not good or bad per se: it does not even make sense to consider this as a question. If we accept that diagnostic tests and classification procedures can change with time (or place, or both), migration happens. The question that should rather be asked is whether the new conceptual tool—be it a diagnostic test or a classification system—represents reality better than the previous tool. Given the importance of accuracy in diagnostic tests for defining the presence and extent

https://doi.org/10.1016/j.euo.2021.12.005

Please cite this article as: S. Fanti, E. Lalumera and R. Hicks, Facts and Myths About Stage Migration: Should the Will Rogers Phenomenon Ride off into the Distance?, Eur Urol Oncol (2022), https://doi.org/10.1016/j.euo.2021.12.005

of disease, once this is settled, the issue of migration relates primarily to its ability to appropriately guide management. By better detecting or characterizing disease before treatment, the ability to assess the efficacy of those treatments will also be potentially enhanced.

2. A bias in study design

If a new test is proven to be more accurate, as PET has very consistently been shown to be, why have many of the medical community been persuaded by the idea that stage migration is a bad thing to have in medical research? One reason may be that it may raise issues that force re-evaluation of treatment paradigms developed on the basis historical studies. Judging outcomes against historical controls becomes fraught in this setting. For example, the use of fluorodeoxyglucose PET scanning in non–small-cell lung cancer significantly improved outcomes in stage I–III disease treated with curative intent compared to a historical cohort treated in a similar manner on both univariate and multivariate analysis [7]. How much of this improvement was achieved by excluding patients with occult metastatic disease and how much was due to better targeting of the radiation or selection of higher-risk patients for combined modality treatment is unclear, but it is often dismissed by sceptical clinicians as simply an effect of stage migration.

We acknowledge that use of a control group, or part of a cohort, that dates back to a point in time before the diagnostic change and any consequent stage migration took place may introduce a chronological bias. A bias is a systematic error in study design that leads to incorrect associations between interventions and outcomes. Biases should be avoided, and can be avoided, to a certain extent, if we recognize their risk in advance. Thus, the chronological bias that may be caused by stage migration should be avoided, and can be avoided, by recognizing its risk in advance. But this is a contingent methodological problem that can be solved [8].

3. A statistical effect

The Will Rogers effect is a statistical anomaly whereby migration of some members of a group to another raises the average value of both groups. The name of the effect comes from the comedian Will Rogers, to whom this witty remark is attributed: "When the Okies left Oklahoma and moved to California, they raised the average intelligence level in both states."

In oncology, the effect most often relates to survival as assessed by Kaplan-Meier analysis. For example, if the dominant effect is upstaging of disease, as it tends to be with a more sensitive test such as PET imaging, when patients with the worst prognosis assigned to an intermediate disease stage on conventional analysis transition to a group previously assigned as having the worst prognosis, the prognosis of the group they leave is enhanced. Simultaneously, since their survival is, in turn, likely to be superior to that of patients previously in that group by virtue of having a smaller burden of disease, the prognosis of the group they join is correspondingly also improved. The major impact is not on overall survival of the group, which is not necessarily impacted by the scan result, but rather on the proportion of the cohort in each prognostic group, with more patients represented in each progressively worse prognostic group [9].

The corollary of this that fewer patients may get an attempt at cure, but those who do have a higher chance of being cured, while the impact of systemic treatments in those with metastatic disease may appear to be superior to historical controls. Improved specificity works in the opposite direction, potentially migrating patients with false-positive results to a less advanced stage, and if their prognosis was, nevertheless, worse than that for the group they joined, they would worsen the prognosis of both groups. When a test is both more sensitive and more specific, and correspondingly more accurate, individual patients may migrate in both directions, diluting the prognostic effects of this stage migration and in such cases the net effect will be influenced by the dominant direction of migration.

Is the Will Rogers effect good or bad? It is neither: it is simply a consequence of applying statistics to phenomena that reflect averages within groups. In fact, the title of a famous article published in 1985 already made it clear: it is a source of misleading statistics for survival in cancer [10]. Just like ignoring that fact that stage migration can lead to a chronological bias in study designs, ignoring the Will Rogers effect may lead to misinterpretation of the significance of study results. But again, this is a problem of study design, not a problem of diagnostic tests.

4. The structure of reality

We have shown that the introduction of new and more accurate imaging techniques brings about a change in our conceptual framework in the way in which we represent reality. However, it does not change reality. Classifying Pluto as a dwarf planet did not cause it to shrink. Analogously, the migration of many asymptomatic COVID-19 patients from the nondiseased group to the diseased group after the introduction of new molecular tests did not make COVID-19 more or less deadly than before. Similarly, more frequent diagnosis of metastatic disease does not change the fact that the metastases were already present and that the eventual outcome for the patient will be determined by the efficacy of the treatments delivered when metastatic disease eventually becomes clinically apparent. This is the level of reality, not to be confused with the conceptual description or statistical representation. The problem to focus on is not stage migration or the Will Rogers effect, but rather the efficacy of treatments. Will Rogers often played a Western cowboy role in movies, and we suggest that his legacy in the medical literature should also ride off into the sunset.

Conflicts of interest: Stefano Fanti has received honoraria from and participated in advisory boards, speaker bureaus, and sponsored talks for AAA, Amgen, Astellas, Bayer, GE, IBA, Janssen, Novartis, Sofie, and Telix; and has received grant support for trials from Amgen, Janssen, and AMNI. Rodney Hicks holds shares in Telix Pharmaceuticals. Elisabetta Lalumera has nothing to disclose.

Please cite this article as: S. Fanti, E. Lalumera and R. Hicks, Facts and Myths About Stage Migration: Should the Will Rogers Phenomenon Ride off into the Distance?, Eur Urol Oncol (2022), https://doi.org/10.1016/j.euo.2021.12.005

References

[1] Sundahl N, Gillessen S, Sweeney C, Ost P. When what you see is not always what you get: raising the bar of evidence for new diagnostic imaging modalities. Eur Urol 2021;79:565–7. https://doi.org/10.1016/j.eururo.2020.07.029.

[2] Ware RE, Francis HW, Read KE. The Australian government's review of positron emission tomography: evidence-based policy-making in action. Med J Aust 2004;180:627–32. https://doi.org/10.5694/j.1326-5377.2004.tb06125.x.

[3] Sullivan R, Peppercorn J, Sikora K, et al.. Delivering affordable cancer care in high-income countries. Lancet Oncol 2011;12:933–80. https://doi.org/10.1016/S1470-2045(11)70141-3.

[4] Hofman MS, Lawrentschuk N, Francis RJ, et al.. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. Lancet 2020;395:1208–16.

[5] Mottet N, Cornford P, van den Bergh RCN, et al.. EAU guidelines: prostate cancer. Arnhem, The Netherlands: European Association of Urology; 2021. , https://uroweb.org/guideline/prostate-cancer/.

[6] Lalumera E, Fanti S, editors. Philosophy of advanced medical imaging. Cham, Switzerland: Springer; 2021.

[7] Mac Manus MP, Wong K, Hicks RJ, Matthews JP, Wirth A, Ball DL. Early mortality after radical radiotherapy for non-small-cell lung cancer: comparison of PET-staged and conventionally staged cohorts treated at a large tertiary referral center. Int J Radiat Oncol Biol Phys 2002;52:351–61. https://doi.org/10.1016/s0360-3016(01)02673-6.

[8] Catalogue of Bias Collaboration, Spencer EA, Heneghan C. Chronological bias. In: Catalogue of bias 2017. https://catalogofbias.org/biases/chronological-bias/.

[9] Kondylis FI, Moriarty RP, Bostwick D, Schellhammer PF. Prostate cancer grade assignment: the effect of chronological, interpretive and translation bias. J Urol 2003;170:1189–93. https://doi.org/10.1097/01.ju.0000085675.96097.76.

[10] Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. N Engl J Med 1985;312:1604–8. https://doi.org/10.1056/NEJM198506203122504.

Stefano Fanti [a,b]*
Elisabetta Lalumera [c,]*
Rodney Hicks [d]

[a] *Department of Nuclear Medicine, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy*
[b] *University of Bologna, Bologna, Italy*
[c] *Department of Life Quality Studies, University of Bologna, Rimini, Italy*
[d] *Department of Medicine, St. Vincent's Medical School, University of Melbourne, Melbourne, Australia*

*Corresponding author. Department of Life Quality Studies, University of Bologna, Corso d'Augusto 237, Rimini, RN 47921, Italy. Tel. +39 333 409 5100.
E-mail address: elisabetta.lalumera@unibo.it (E. Lalumera),

Accepted 13 December 2021