European Journal of Internal Medicine 63 (2019) 15–18

Contents lists available at ScienceDirect

European Journal of Internal Medicine

journal homepage: www.elsevier.com/locate/ejim

Special Article

Appropriateness, inappropriateness and waste of resources: Unfulfilled expectations?

ARTICLE INFO

Keywords:
Appropriateness
Inappropriateness
Choosing wisely

1. Introduction

Recent years have been characterised by an increased awareness in the fact that a high number of overused diagnostic imaging, medications and procedures provide no benefit to patients and may even cause harm. Moreover, many studies have addressed the problem of the correct approach and implementation of the concepts of appropriateness and inappropriateness in daily medical practice. All health workers, as well as patients and the public, should know the consequences of inappropriate use of tests, procedures and treatments, such as damage due to radiation exposure, adverse drug reactions and complications due to diagnostic or therapeutic procedures. Inappropriateness may also have relevant economic consequences, particularly important in a historical period characterised by increased growth in health costs and by the problem of the sustainability of our health systems.

This article focuses on three problems: 1) the complexity and the variability of the concept of appropriateness; 2) the weakness of data regarding the economic consequences (and potential savings) of inappropriate choices in clinical practice; 3) the importance but also the limitations, of the efforts for the reduction of inappropriate tests and treatments.

2. Appropriateness, a complex and variable entity

In health care the primary concept for the definition of appropriateness, is the balance between risk and benefit of a treatment: “usually not appropriate”, where the harms of doing the treatment outweigh the benefits; and “usually appropriate” where the benefits of doing a treatment outweigh the harms or risks [1]. However, such a narrow conceptualization of appropriateness is limited and cannot ensure that high quality care is provided with a responsible use of limited medical resources, in order to control the increasing health care costs. Indeed the evaluation of appropriateness should include not only risk benefit ratio but also the magnitude of the effects, the individuals or groups with specific characteristics to whom the finding apply (on the basis of the principles of evidence-based medicine), considering cost-effectiveness and ethical principles. Therefore appropriate care should be effective (based on valid evidence); efficient (cost-effectiveness) and consistent with the ethical principles of the community or society. Regrettably, there is currently insufficient evidence to make such evaluations across a broad spectrum of potential clinical indications. The problem of appropriateness is complicated especially if we also consider that some clinical characteristics of patients (such as comorbidities), are sometimes the complex drivers for specific test ordering. Moreover potential purposes of test include the detection or exclusion of disease, risk stratification and the evaluation of therapeutic efficacy. Sometime also appropriate tests may not result in any active change in the management plan (variation in diagnostic workup, therapeutic decisions, or follow-up planning) in a relevant proportion of cases [2] and therefore the concept of appropriateness and the concept of usefulness may diverge considerably. Acknowledging this inferential caveat, especially for diagnostic tests, appropriate indications will not mean their use leads to a secure patient benefit [3]. At the same time, however, in the absence of comparative data is also difficult to understand whether merely changes in action plans equate with clinical usefulness or whether a test is still of value when it reassures that the current management plan is fine.

Appropriateness is a complex problem and a variable entity whose criteria and characteristics may change over the course of time and be influenced by different clinical settings. Indeed the concept of appropriateness is strictly connected with the cultural evolution and the acquisition of new data and evidence. These are in turn implemented in guidelines which are endorsed by the scientific societies. Advances in knowledge may create apparently paradoxical changes, for example the use of beta blockers for heart failure treatment were considered as highly inappropriate in the past, but are now a mainstay of treatment.

In the literature, its frequency is highly variable due to many factors such as data collection methods and the characteristics of the population studied. Studies range from experiences of a few centres with small populations and short periods of observation [2,4,5] to huge national studies with more than a million patients included [6]. Moreover, “unmappable cases” (39–46%) and “uncertain cases” (20–23%) are often present in retrospective studies making it difficult to determine the real frequency with precision [7].

Appropriate use of tests and procedures also change according to

https://doi.org/10.1016/j.ejim.2019.03.016
Received 26 January 2019; Received in revised form 26 March 2019; Accepted 27 March 2019
Available online 19 April 2019

the clinical situations in which they are applied. For example, appropriateness for the percutaneous coronary interventions is very high in acute coronary syndromes (> 95%) [7,8], while in stable angina it is much lower (about 50%) [6,8]. The qualification of reference centres may also account for variations in the level of appropriateness and sometimes there are relevant differences between hospitals (with range of inappropriateness from 0% to 55%) [8].

Appropriateness is also influenced by the professional background of the physicians requesting a specific test or procedure, probably reflecting a greater familiarity with its role and limitations. In fact, tests and procedures when prescribed by specialists are frequently more appropriate than when prescribed by non-specialists [9], for example the appropriateness for cardiac computed tomography examinations was 58.5% for cardiologists, in contrast to 29.7% for non-cardiologists ($p < .01$) [9]. Finally, the insurance status of patients is also not negligible, as it may cause differences in the rate of appropriate prescriptions of procedures [10].

Thus, inappropriateness is the consequence of many components: ranging from ignorance of the correct indications, the defensive approach, but also, in some health systems, ease of access to tests and procedures.

Inappropriateness is not without risk. Inappropriate medication use has been documented in about 20% of elderly patients [11] and can have sizable adverse effects on patient health outcomes [12]. Moreover preventable adverse drug events (ADEs) result frequently in urgent medical attention or hospitalizations [13]. Considering the relationship between antimicrobial consumption and antimicrobial resistance [14], overutilization of antibiotics may contribute to the emergence of resistances. Inaccurate test results represent true risks to the patient and costs to both patients and the health care system. Indeed it is necessary to remember that imaging tests may observe anomalies or artifacts erroneously diagnosed as pathological, as well as anomalies that have no pathological meaning (“incidentalomas”). In this contest “control tests” required to reassure the doctor and patient can sometimes give unclear answers and activate a more or less long journey through imaging laboratories in order to “control” the “checks” performed (“Ulysses syndrome “) [15–17]. Finally the risks associated with ionizing radiation should be not underestimated and X-ray imaging exams should be performed only after careful consideration of the patient's health needs and radiation doses should be “As Low as Reasonably Achievable” to minimize radiation exposure to the patient.

3. Economic problems

Waste in health care is a major component of health costs [18,19]. Certainly many factors contribute to these unnecessary costs, such as: failure of care coordination; administrative inefficiency; fraud and abuse. However, overtreatment (and sometimes also failure of care delivery), as an expression of inappropriate choices, seems to be the most important factor [19]. Nevertheless, the potential saving due to the elimination of waste is not clear. An exact quantification of the problem is difficult or impossible and the published data probably lead to conclusions that are over-optimistic. Brody [18] affirms in an editorial that “waste in U.S. health care…actually amounts to…at least 30% of the budget – and that this waste is a major driver of cost increases”. This statement is based on the study of Berwick and Hackbarth [19] who analysed many complex data regarding all types of waste, leading to the conclusion that, at a prudential estimate, the total amount of waste in USA exceeds 20% of health costs and “overtreatments” should vary between 150 and 226 billions dollars in the year 2011. Also Fisher et al. [20], using a different approach, calculated that “about” 30% of health expenditure had to be considered waste of non-beneficial measures. They studied the differences per capita in Medicare spending from 1992 through to 2006 in five US hospital referral regions and noticed important variations between regions. For example, per capita spending in Miami grew at an annual rate of 5.0% as compared with 2.3% in Salem (Oregon) or 2.4% in San Francisco. These differences did not appear to be due to differences in health or to other factors (payment system for example) but rather to the propensity to intervene in grey areas of decision making (for example more frequent referrals to subspecialties, hospital and intensive care admissions). The Authors calculated that the reduction of the annual growth in per capita spending from 3.5% (national average) to 2.4% (San Francisco) would leave Medicare with a healthy estimated balance of 758 billions dollars, a cumulative savings of 1.42 trillion dollars over about 15 years. In a more sectorial study focused on frequency and costs of five clinical activities (suggested by The Good Stewardship Working Group) of common use in primary care but considered of little benefit to patients, Kale et al. [21] calculated that the approximate annual cost (national estimate) was 6.76 billion dollars. Finally, Mello et al. [22] tried to calculate the costs of “defensive medicine”, strictly connected with the problem of inappropriateness. They estimated that the total amount of hospital and physician spending was 45.6 billion dollars in the year 2008.

It is difficult and probably inappropriate to try to make comparisons between these data or try to reach reliable or exact conclusions as to the amount of waste, costs and potential savings. Indeed, the different studies analysed different and variably selected populations (hospital versus outpatients, different interventions etc.). Also debatable are the extrapolations of data, expression of different geographical, temporal and clinical situations for a national or international estimate. The most mentioned and accepted data (as in the Lancet [23] and in the Organisation for Economic Co-operation and Development [24]) are based on the study of Berwick and Hackbarth [19] but all these researches have a common denominator i.e. that waste (and in particular inappropriate choices) has a relevant and sometimes an extraordinary impact in the economic balance of our health care systems and that all communities have a pool of wasted resources which could be more rationally used. As we mentioned it is challenging to translate low-value services into meaningful metrics also because available data sources often lack the clinical detail necessary to distinguish appropriate from inappropriate use. There is a broad interest in systems that use big data medical systems providing upstream information which could allow greater discretion in identifying inappropriateness and further improve the quality of cost effectiveness analysis. Moreover professional societies, regulators, and payers need to standardize definitions that can be applied uniformly and fairly.

4. The choosing wisely campaign

The Choosing Wisely campaign represents a physician-driven effort to create conversations between physicians and patients regarding the increasing awareness that tests, procedures and medications are frequently overused. The purpose was to disseminate the recommendations of the scientific societies and to promote a process of appropriateness within a relation of dialogue and decision sharing with the patients and the public. Building on the work of Dr. Brody [18] and the National Physicians Alliance (NPA), the Board of Internal Medicine (ABIM) Foundation lunched the campaign with the release of the top 5 lists [25]. The Top Five list consists of five diagnostic tests or treatments commonly ordered, that are among the most expensive services provided, and that have been shown not to provide any meaningful benefit to at least some major categories of patients [18]. Choosing Wisely campaign begun > 7 years ago [18,25] and has now involved 19 countries, > 80 specialty societies which have produced > 500 recommendations. The Choosing Wisely campaign is important and innovative but some open problems and points of criticism have to be considered.

Many professional societies have published “top-five” lists, but most have not detailed the methods by which the list was created, without much input from frontline practitioners, using a process without clear criteria for inclusion on the list [26]. Some scientific societies rarely

included, in their suggestions, procedures that are characteristic of their specialty (otorhinolaryngology, orthopaedics for example) and preferred suggestions regarding imaging or use of antibiotics [27].

Moreover, it is advisable to re-evaluate the criteria of choice for some of the top five suggestions which are very rarely used in medical practice, such as, for example, a stress-test before low risk surgical intervention (0.76–2.4%), an indication of limited importance in a campaign pointed toward an extensive improvement of appropriateness [28]. Other data indicated problems in the implementations of the Choosing Wisely suggestions into clinical practice [29]. For example, a group of primary care physicians considered some recommendations regarding tests or treatments for symptomatic conditions difficult to follow (9.8–32%) with a high level of reported difficulties in acceptance from patients (35.7–87.1%) [29]. All these data indicate the importance of a solid methodology to develop the list of advance services. A methodology including broad college discussion, practitioner surveys, transparency, buy in, modified Delphi technique, and solid scientific foundations can be used to identify actionable targets of overuse [30].

Probably the most common and still unsolved issue is the lack of a rigorous and of a systematic methodology for the evaluation of the proposed initiatives. In this absence, it is impossible to know the extent and the characteristics of the obtained results but also to affect the future programming. In fact, according to some recent reviews [31,32] most published studies have to be considered “low quality” and many used weak quasi-experimental methods such as simple “pre-post design”, which make it difficult to separate the intervention effects from unrelated effects [31].

The extent to which the concepts and the principles of the Choosing Wisely campaign are part of the cultural background of the medical community is practically unknown. In a super-selective sample, 80% of members of the academic chairs or division chiefs of emergency medicine in USA were aware of the Choosing Wisely campaign (although about half did not recall any recommendation) but this sample certainly cannot be considered representative of the “real world” [33]. Probably more realistic are the data reported by Colla et al. [34] indicating a limited knowledge in a group of primary care physicians (43.2%) and a much lower rate in medical specialists (37.4%) and surgical specialists (27%).

Moreover, there is the risk that some scientific societies may enthusiastically follow the Choosing Wisely program (“fad effect”) but limited to the identification and publication of the top 5 recommendations without an active participation in the following steps (educational activities, meetings, evaluation of results, behavioural changes etc.). In fact, the knowledge of the problem, per se, is insufficient for its solution. Indeed the evaluation of Choosing Wisely recommendations showed only limited results [35] and, as reported by Rosenberg et al. [36], the decrease of some services (from 14.9% to 13.4% for example) may not represent clinically significant changes. Conversely some services showed an increased frequency and the most highly used services such as preoperative chest x-rays and low back pain imaging, remained stable [36].

All these data indicate that till now the Choosing Wisely recommendations have not substantially impacted care and further efforts are needed to translate the many recommendations into behavioural changes. In conclusion the Choosing Wisely campaign has certainly to be considered significant and innovative but, after more than five years' experience, some observations have to be made:

- the criteria for the formulation of some of the “top 5” recommendations are not always evident and they should be clearly defined. Moreover, the > 500 recommendations should be critically evaluated and, also considering the overlap between different disciplines, a stricter cooperation among the scientific societies should be favoured;
- the behavioural changes in the medical community in different countries induced by the Choosing Wisely recommendations are largely unknown but, till now, they seem to be modest. The evidence of a significant impact is absent and certainly the implementation of a rigorous robust methodology of evaluation of results is needed;
- It is essential to translate the Choosing Wisely recommendations into behavioural changes, involving not only the medical community but also the general public;
- Changes in physician behaviour seem to be slow and more active multidisciplinary coordinated interventions are needed.

We are living in a historical period characterised by an apparent paradox i.e. we have the possibility to improve the quality of our work and at the same time to achieve cost savings and waste reduction. However, the evolution of these changes is very slow and probably due to conscious or unconscious conservativeness of the medical community, which cannot be an indifferent bystander but must play a more active role in the process of change.

Acknowledgments

We would like to thank Gino Tosolini and Marco Bobbio who have constantly drawn attention to the concepts of appropriateness and sustainability and for their meaningful analyses on the topic.

References

[1] Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. Int J Technol Assess Health Care 1986;2:53–63.
[2] Matulevicius SA, Rohatgi A, Das SR, Price AL, deLuna A, Reimold SC. Appropriate use and clinical impact of transthoracic echocardiography. JAMA Intern Med 2013;173:1600. https://doi.org/10.1001/jamainternmed.2013.8972.
[3] Ioannidis JPA. Appropriate vs clinically useful diagnostic tests. JAMA Intern Med 2013;173:1607–9. https://doi.org/10.1001/jamainternmed.2013.6582.
[4] Gibbons RJ, Miller TD, Hodge D, Urban L, Araoz PA, Pellikka P, et al. Application of appropriateness criteria to stress single-photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. J Am Coll Cardiol 2008;51:1283–9. https://doi.org/10.1016/j.jacc.2007.10.064.
[5] Gimelli A, Rovai I, Liga R, Pasanisi EM, Marzullo P. Appropriate use criteria in clinical routine practice: implications in a nuclear cardiology lab. Int J Card Imaging 2016;32:1003–9. https://doi.org/10.1007/s10554-016-0864-6.
[6] Desai NR, Bradley SM, Parzynski CS, Nallamothu BK, Chan PS, Spertus JA, et al. Appropriate use criteria for coronary revascularization and trends in utilization, patient selection, and appropriateness of percutaneous coronary intervention. JAMA 2015;314:2045. https://doi.org/10.1001/jama.2015.13764.
[7] Thomas MP, Parzynski CS, Curtis JP, Seth M, Nallamothu BK, Chan PS, et al. Percutaneous coronary intervention utilization and appropriateness across the United States. PLoS One 2015;10:e0138251https://doi.org/10.1371/journal.pone.0138251.
[8] Chan PS, Patel MR, Klein LW, Krone RJ, Dehmer GJ, Kennedy K, et al. Appropriateness of percutaneous coronary intervention. JAMA 2011;306:53–61. https://doi.org/10.1001/jama.2011.916.
[9] Ayyad A-EM, Cole J, Syed A, Desai MY, Halliburton S, Schoenhagen P, et al. Temporal trends in utilization of cardiac computed tomography. J Cardiovasc Comput Tomogr 2009;3:16–21. https://doi.org/10.1016/j.jcct.2008.10.009.
[10] Guadagnoli E, Landrum MB, Peterson EA, Gahart MT, Ryan TJ, McNeil BJ. Appropriateness of coronary angiography after myocardial infarction among Medicare beneficiaries. Managed care versus fee for service. N Engl J Med 2000;343:1460–6. https://doi.org/10.1056/NEJM200011163432006.
[11] Fialová D, Topinková E, Gambassi G, Finne-Soveri H, Jónsson PV, Carpenter I, et al. Potentially inappropriate medication use among elderly home care patients in Europe. JAMA 2005;293:1348. https://doi.org/10.1001/jama.293.11.1348.
[12] Fu AZ, Liu GG, Christensen DB. Inappropriate medication use and health outcomes in the elderly. J Am Geriatr Soc 2004;52:1934–9. https://doi.org/10.1111/j.1532-5415.2004.52522.x.
[13] Hanlon JT, Schmader KE, Koronkowski MJ, Weinberger M, Landsman PB, Samsa GP, et al. Adverse drug events in high risk older outpatients. J Am Geriatr Soc 1997;45:945–8.
[14] Seppälä H, Klaukka T, Lehtonen R, Nenonen E, Huovinen P. Outpatient use of erythromycin: link to increased erythromycin resistance in group A streptococci. Clin Infect Dis 1995;21:1378–85.
[15] Picano E. The risks of inappropriateness in cardiac imaging. Int J Environ Res Public Health 2009;6:1649–64. https://doi.org/10.3390/ijerph6051649.
[16] Wardlaw JM, Davies H, Booth TC, Laurie G, Compston A, Freeman C, et al. Acting on incidental findings in research imaging. BMJ 2015;351:h5190. https://doi.org/10.1136/bmj.h5190.
[17] Rang M. The Ulysses syndrome. Can Med Assoc J 1972;106:122–3.
[18] Brody H. Medicine's ethical responsibility for health care reform — the top five list. N Engl J Med 2010;362:283–5. https://doi.org/10.1056/NEJMp0911423.

[19] Berwick DM, Hackbarth AD. Eliminating waste in US health care. JAMA 2012;307:1513. https://doi.org/10.1001/jama.2012.362.
[20] Fisher ES, Bynum JP, Skinner JS. Slowing the growth of health care costs–lessons from regional variation. N Engl J Med 2009;360:849–52. https://doi.org/10.1056/NEJMp0809794.
[21] Kale MS, Bishop TF, Federman AD, Keyhani S. "Top 5" lists top $5 billion. Arch Intern Med 2011;171:1856–8. https://doi.org/10.1001/archinternmed.2011.501.
[22] Mello MM, Chandra A, Gawande AA, Studdert DM. National costs of the medical liability system. Health Aff (Millwood) 2010;29:1569–77. https://doi.org/10.1377/hlthaff.2009.0807.
[23] Saini V, Brownlee S, Elshaug AG, Glasziou P, Heath I. Addressing overuse and underuse around the world. Lancet 2017;390:105–7. https://doi.org/10.1016/S0140-6736(16)32573-9.
[24] Tackling Wasteful Spending on Health. OECD Publishing; 2017. https://doi.org/10.1787/9789264266414-en.
[25] Good Stewardship Working Group. The "top 5" lists in primary care. Arch Intern Med 2011;171:1385–90. https://doi.org/10.1001/archinternmed.2011.231.
[26] Grady D, Redberg RF, Mallon WK. How should top-five lists be developed? JAMA Intern Med 2014;174:498. https://doi.org/10.1001/jamainternmed.2013.8272.
[27] Morden NE, Colla CH, Sequist TD, Rosenthal MB. Choosing wisely — the politics and economics of labeling low-value services. N Engl J Med 2014;370:589–92. https://doi.org/10.1056/NEJMp1314965.
[28] Kerr EA, Chen J, Sussman JB, Klamerus ML, Nallamothu BK. Stress testing before low-risk surgery. JAMA Intern Med 2015;175:645. https://doi.org/10.1001/jamainternmed.2014.7877.
[29] Zikmund-Fisher BJ, Kullgren JT, Fagerlin A, Klamerus ML, Bernstein SJ, Kerr EA. Perceived barriers to implementing individual choosing wisely(®) recommendations in two national surveys of primary care providers. J Gen Intern Med 2017;32:210–7. https://doi.org/10.1007/s11606-016-3853-5.
[30] Schuur JD, Carney DP, Lyn ET, Raja AS, Michael JA, Ross NG, et al. A top-five list for emergency medicine. JAMA Intern Med 2014;174:509. https://doi.org/10.1001/jamainternmed.2013.12688.
[31] Kerr EA, Kullgren JT, Saini SD. Choosing wisely: how to fulfill the promise in the next 5 years. Health Aff (Millwood) 2017;36:2012–8. https://doi.org/10.1377/hlthaff.2017.0953.
[32] Colla CH, Mainor AJ, Hargreaves C, Sequist T, Morden N. Interventions aimed at reducing use of low-value health services: a systematic review. Med Care Res Rev 2017;74:507–50. https://doi.org/10.1177/1077558716656970.
[33] Maughan BC, Baren JM, Shea JA, Merchant RM. Choosing wisely in emergency medicine: a national survey of emergency medicine academic chairs and division chiefs. Acad Emerg Med 2015;22:1506–10. https://doi.org/10.1111/acem.12821.
[34] Colla CH, Kinsella EA, Morden NE, Meyers DJ, Rosenthal MB, Sequist TD. Physician perceptions of choosing wisely and drivers of overuse. Am J Manag Care 2016;22:337–43.
[35] McCarthy M. US choosing wisely campaign has had only modest success, study finds. BMJ 2015:h5437https://doi.org/10.1136/bmj.h5437.
[36] Rosenberg A, Agiro A, Gottlieb M, Barron J, Brady P, Liu Y, et al. Early trends among seven recommendations from the choosing wisely campaign. JAMA Intern Med 2015;175:1913. https://doi.org/10.1001/jamainternmed.2015.5441.

Fulvio Camerini, Enrico Fabris, Gianfranco Sinagra*
Cardiovascular Department, University of Trieste, Trieste, Italy
E-mail address: gianfranco.sinagra@aots.sanita.fvg.it (G. Sinagra).

* Corresponding author at: Polo Cardiologico-Ospedale di Cattinara, Via Valdoni 7 – 34129 Trieste, Italy.