Volume-outcome relationship in colon cancer surgery: another biased logical short cut towards questionable centralization policies

Vincenzo Violi[1,2,3,4], Renato Costi[1,2], Massimo De Bernardinis[3], Clara Pavlidis[3,4], Luigi Roncoroni[1,2]

[1]Dipartimento di Scienze Chirurgiche, Università degli Studi di Parma; [2]Sezione di Clinica Chirurgica Generale e Terapia Chirurgica, Azienda Ospedaliero-Universitaria di Parma; [3]Azienda Unità Sanitaria Locale di Parma; [4]Unità Operativa Complessa di Chirurgia Generale, Ospedale di Fidenza

Abstract. The association between hospital high volumes and good outcomes after complex surgery has given rise to a worldwide controversial debate. Important and unsolved questions have followed, both theoretical and practical, which could have repercussions on health care and health economic policies, such as the centralization/regionalization of major surgical procedures. We read a recent study on the impact of surgery volumes on short-term outcomes after colon cancer resection in Emilia Romagna, Italy, the same geographic area where we operate. Ten issues were submitted to critical analysis and many sources of planning and methodology bias were identified, which, in our opinion, paradigmatically led to unreliable results, inadequate statistical analysis and deceptive conclusions. Despite the authors' admitted awareness of their study's limits, their conclusive message was, surprisingly, that centralization of colon cancer surgery should be substantially encouraged. Unrecognized, systemic biases may easily turn into cognitive biases, into logical short cuts which could confuse healthcare policymakers. The volume-outcome relationship, in which a direct causal link has never been demonstrated, should not be used as a reliable measure of quality, rather than less implementable process indicators, to address centralization policies. A disregarded negative consequence of centralization could be that non-high-volume centres, after a further progressive workload decrease and depletion in resources and surgical skills, will have to cope with patients in bad general condition and at high risk, who must be treated in emergency or cannot anyway afford the move for age, indigence or severe co-morbidities. Thus, centralization policies might disadvantage the weak segments of the population, thereby moving towards an iniquitous health service.

Key words: Colorectal cancer; cancer surgery; hospital volume; surgeon volume; surgery outcomes; volume-outcome relationship; centralization of care; healthcare quality control; healthcare policy

Introduction

An association between hospital high volumes and good outcomes after complex surgical procedures, which might support centralization/regionalization policies, has been witnessed since the late '70s (1) and has given rise to a worldwide controversial debate through a great number of articles: systematic reviews and meta-analyses (2-6), retrospective analyses on administrative databases (7-12), experts' commentaries and editorials (12-15). In addition, the volume-outcome relationship, in which a direct causal link between the two variables has never been demonstrated (8, 16), has generated a series of important sub-topics and questions which complicate and confuse, rather than clarify, the primary issue. 1) Assuming that there

[1] Editor's note: This section contains only the personal views of the Authors and acts as a stimulus for an open and constructive debate

is a direct causality between volume and outcome, which is the causation direction? That would mean whether to address healthcare policies to scale economies, if high volumes improve the outcomes, or to selective referral criteria, if good outcomes attract greater demand (17). 2) In compliance with these two opposite interpretations, how should knowledge, learning and experience themselves be considered and addressed (18)? 3) More simply but crucially, which is more determining, hospital or surgeon volume (4, 6, 7, 14, 19-23)? 4) How much is the volume-outcome relationship affected by organization, services and technology factors (8, 24)? 5) What social, demographic, geographic and healthcare system characteristics and settings are we dealing with (5, 25-27)? And, consequently, how does the centralization/regionalization concept in itself comply with the latter variables? 6) Could socio-demographic disparities and barriers to centralized healthcare access turn into an economic disadvantage for individual patients and their families (28-30)?

Another important question is in what surgical procedures a significant volume-outcome relationship was identified. Whenever various procedures were taken into account for volume-outcome relationship in a single study (7, 24, 31-33), only a selected proportion of them, varying from one article to another, showed a hospital or surgeon volume association with outcome, usually in low prevalence diseases which require major surgery, such as pancreatectomy, oesophagectomy, cystectomy, lung resection, gastrectomy, cardiovascular and major hepatic procedures, and also rectal but rarely colon cancer (CC) resections (9). By contrast, and curiously, CC procedures, without or with rectal resections, showed a volume-outcome relationship only in focused studies (6, 11, 20, 21, 34). However, in other focused studies on colorectal cancer surgery, there was no, or not convincing, evidence of a relationship (23, 35-39) or was it stronger for individual surgeons than for the hospital (6, 20, 21, 23, 40).

The scientific production concerning this relationship did not show convincing evidence, not so much as to the volume-outcome association in itself, but for the identification of a causal link and interpretation of its implications on health care and health economic policies. On the whole, the uncritical assumption has prevailed that "increasing volumes" means "decreasing mortality" and vice versa, and that surgical volumes are a reliable measure of quality, as an alternative to less implementable process indexes, even establishing, empirically or arbitrarily, volume thresholds as conditions enabling hospitals and surgeons to perform specific procedures (36, 37, 41, 42).

This summary approach has been repeatedly and authoritatively questioned and criticized (8, 16, 24, 38, 43-54), since most studies are compromised by too many biases, which should be a strong caveat for policy makers, since policies which appear at first sight reasonable may have unforeseen and uncontrollable consequences if they are prompted by misleading, unrecognized mechanisms.

Descending from a world context into our regional setting, we now present a typical example of how such mechanisms may lead to the drawing of wrong conclusions.

A paradigm of systematic biases

The present paper was conceived after reading, in a free on line journal, an article entitled *"Impact of procedure volumes and focused practice on short-term outcomes of elective and urgent colon cancer resection in Italy"*, by Lenzi and co-workers (11). The study analyzes data extracted over six years (2005–2010) from the Hospital Discharge Records (HDRs) database, collected from 86 General Surgery Units (GSUs) of 66 public and private hospitals in the Regione Emilia Romagna (RER), Italy. The authors operate at the University of Bologna (Department of Biomedical and Neuromotor Sciences) and at the Ospedale Maggiore of Bologna (General Surgery Unit, Department of Surgery). Thus, our concern in this study is threefold. 1) We work in the same geographic area as the authors' and the database includes the cases treated in the two GSUs of the institutions to which we are affiliated. 2) The article in question is a revealing paradigm of the biases which structurally affect the studies based on administrative data, but it could also have important specific implications, as it deals with a high prevalence disease, the treatment of which should be one of the main reasons why any general hospital provides

healthcare. 3) Since we have a longstanding and acknowledged interest, both clinical and scientific, in colorectal cancer (55-57), we mean to express our critical opinion on how this topic has been approached and developed.

The main criticisms emerging will be pointed out in the ten items reported below, but other questionable issues and statistical analysis methods will be focused and commented on later.

1. The authors define a GSU's volume as the mean annual number of CC procedures carried out over 6 years.

Although the article's title refers to resections, CC procedures are not only colectomies, either curative or palliative, but also non-resective ones (intestinal bypasses and enterostomies). The authors do not explain what procedures they took into account for GSU volumes definition, which should be based only on colectomies, as non-resective palliation increases the volumes but decreases their quality. This missed information causes a recurring misunderstanding, also because patient selection codes include any surgery (items 7, 8 and 9 below). However, such a definition is too summary: 360 may theoretically derive from 60+60+60+60+60+60, but also from 110+90+70+50+30+10, or 10+30+50+70+90+110. This is a paradoxical example, but a mere mean of a 6-year activity which does not use time series data does not take into account a GSU's workload trend and its underlying reasons, not excluding changes in human resources and individual competences.

2. The authors use a "tertile split" to classify GSUs into three volume categories: low-volume (<40 cases/year), intermediate-volume (40-64 cases/year) and high-volume (>64 cases/year).

As a result, the 86 GSUs were thus divided: 60 at low-, 17 at intermediate-, and 9 at high-volume. In a continuous variable, a "tertile split" to classify three categories means that the three groups are numerically equal and that the two cut-off points are identified such that one third of the observations are in each group. Hence, the authors should explain how and why they chose "those" two values which split the GSUs into "those" three different volume groups: are they justified by any evidences or at least authoritative sources? If any cut-offs can be arbitrarily chosen just to support one's own thesis, a modicum of common sense should be used, otherwise about 70% of GSUs in the RER result as having a low-volume and somehow an inadequate threshold of activity for a high-prevalence disease. To clarify this issue, each of the two GSUs in which we operate, according to the above partition criteria and assuming that only the colectomies have been accounted for, should be ascribed to the intermediate-volume group, whereas they are in the ninth decile (probably in the tenth decile, i.e. in the high volume group, if we were to include also non-resective procedures; these data, which would increase our procedure rate by 20-25% overall, are non reliably selectable, since their codes also include diverting enterostomies, i.e. synchronous to left colectomy). The partition criteria used for the study seem to be aimed at identifying the overall excellence of few very-high-volume GSUs, which does not imply the excellence of each of them, rather than the actual inadequacy of few at very low volume. It is possible that a correct tertile split would have given results which excluded the study's hypothesis. In any case, clustering criteria should be addressed and explained in such a way that the hypothesis could be more plausible.

3. The authors use a "median split" to classify GSUs as non-focused (<5% CC cases over total operations) or focused (≥5% CC cases).

"Median split" means that the percentage cut-off is chosen so that the two groups are numerically equal, whilst there were 64 non-focused and 22 focused GSUs. So, this cut-off was arbitrarily chosen too. In addition, any percentage cut-off over "total operations" is unreliable, since it does not take into account the organization of the single GSUs. The entirety of the operations coded in a GSU may vary greatly, for instance, according to the presence or not, inside it, of a high-volume day-surgery, or a breast surgery centre, or a 24-hour emergency surgery room, or a surgical endoscopy sub-unit, and to the individual GSUs' way of organizing and coding their activity. All these procedures in some hospitals might be recorded in the HDRs by autonomous units, but when all activity is carried out and coded within the same GSU, the rate

of any surgical "focused" practices decreases, although they could be performed at objectively high quanti-qualitative standards and even by dedicated teams.

4. The authors describe hospital characteristics by categorizing them as private or public and teaching or non-teaching.

The latter distinction is unclear and misleading, as we do not know what teaching type the authors are talking about. In the RER there are four University Medical Schools, which are all public institutions, but postgraduate surgical training is also done in non-university hospitals. In the university hospitals there are also non-teaching GSUs and in non-teaching hospitals there may be teaching GSUs affiliated to a University Medical School. In addition, the condition itself of teaching unit may be a determinant factor, in either a positive or negative sense, also on the GSUs' size (in terms of numbers of beds), depending on policies and relations between medical school and hospital management. This may affect the yearly admission potentiality of a GSU, independently of its quality standards. We do not call in question the GSUs' research activity, since this would move us far away. As regards private hospitals, which in the RER have healthcare standards lower on average than those of public institutions, a source of bias could derive from their great variability as regards organization, staff, provision contract with the public health service, professionals' mission and vision, and even the fact that a GSU may be a mere container of various, heterogeneous surgical specializations.

5. The authors analyze data extracted from an administrative, retrospective database.

The data from the HDRs are generally considered of poor quality, partly because of a well-known intrinsic inadequacy of the coding system, but mostly because there is a great variability in HDRs' recording accuracy and performance. This is likely due to the assigning of this task to individual surgeons without substantial administrative support, or to other imponderable factors, such as a GSUs' differing interest to demonstrate high performances. Non-surgical comorbidities, which strongly affect the case mix, are even more unreliably recorded. The authors acknowledge, in the Discussion, the potential biases of the administrative databases, but claim in their own support the good HDR performances reported, in comparison with those of cancer registries, totally different as to finalities, in a study not on colorectal but on breast cancer (58), which often has dedicated operators, besides a lesser case mix variability.

6. The authors use ICD-9-CM codes to identify patients with a primary diagnosis of carcinoma in situ (230.3) and malignant neoplasm (153.x) of the sole colon.

The exclusion of the codes 154.x (malignant neoplasms of the rectum) is very questionable. Actually, the article's title refers only to colon cancer, but colorectal cancer is intended, everywhere and under all viewpoints, as a single disease, although some procedures on the rectum may be technically more demanding and therefore qualifying. In any case, such a selection excludes, besides the subperitoneal rectum, even tumours of the recto-sigmoid junction (154.0), which are intraperitoneal and, like those of the sigmoid colon, require procedures not involving a total mesorectal excision and a low anastomosis

7. The authors include, among the independent variables for case mix-adjusted analyses, the urgent/elective admission status.

This distinction is a very strong source of bias: an urgent admission does not imply an urgent operation, since the real surgical timing and priority are not coded in the HDRs. The elective or urgent admission status is a merely administrative datum, which depends sometimes on real clinical needs, but in many cases only on organization or social reasons. From an administrative viewpoint, admissions are considered "urgent" whenever they are not preceded by an out-patient management (the so-called "pre-admission"). This is typical in frail patients or those living far away. Hence, there are no databases available which allow the authors to reliably explain their findings as to "urgent" procedures, since such data may result only in the clinical or surgical records, not in the HDRs. Thus, the fact that a relationship between "urgent" procedures and outcomes was or was not found, possible discrepancies with other studies and any other priority-related findings, are meaningless. However,

any reference to colorectal emergency surgery would require a better definition of the procedures carried out, since GSUs' strategies as regards one- or more-step surgical management have a wide variability which may also affect the re-intervention rates.

8. The authors identify as the primary procedure an operation in the digestive system (codes 42-54) to decrease the risk of excluding from the analyses patients undergoing multi-visceral resections for locally advanced tumours.

The primary identification of synchronous resective procedures other than colectomies implies that also a secondary procedure is anyway coded, otherwise they would be lost when, as usual, they are coded as secondary procedures. So, what does coding priority matter? However, the codes 42-to-54 include all the gastrointestinal procedures, both surgical and endoscopic ones, hepatic resections and even abdominal wall surgery, not to mention intestinal bypasses and enterostomies. As a consequence, the use of this long code sequence means that, rather than decreasing the risk of excluding the multivisceral resections, there is an increased risk of including synchronous procedures carried out for non-malignant diseases, such as cholecystectomy, hysterectomy, iatrogenic splenectomy, ventral hernia repair, and even diverting enterostomies, not to mention the possibility that diagnostic or therapeutic endoscopic resections of carcinomata in situ (230.3, see item 6 above) are also selected in the GSUs where operative endoscopy is carried out and codified. In an attempt to verify this methodological step at the Fidenza Hospital GSU, which had its own endoscopy service until 2010, a selection of all 42-54 codes as both primary and secondary procedures in association with the diagnosis codes 230.3 and 153.x, gave an increase, as compared to the number of colectomies (codes 45.7 and 45.8), of about 150%. It is evident that there was something wrong, either in the authors' methodology or in our comprehension of it. Moreover, with such a code selection, hepatic M1 resections (50.x) are also recorded and included besides those for T4 tumours. We therefore wonder, since the HDR coding system is very inaccurate as to liver surgery, how the major hepatic synchronous resections were distinguished from diagnostic metastasectomies, and what the overall R0 resection rate was. The authors say in the Discussion that the HDRs cannot report data about the radical nature of the resections nor data regarding recurrence rate and disease-free survival, but this information, albeit important for the long-term results, is useless for the study's end-points, whereas all stage- and treatment-related variables may affect the short-term outcomes.

9. The authors categorize the procedures as partial and total colectomies, and the remaining interventions as "other".

The Authors do not explain what they included in this third group. Three hypotheses are possible. The first is that they refer to non-resective surgery, the sole residual option after the partition into partial and total colectomies. This would reinforce our initial doubt that not only resective procedures were selected for volume definition (which means that even the article's title is inaccurate). Non-resective surgery is an important palliation-related variable, although there is also a hidden resective palliation rate, but it was not adequately focused on. However, if this interpretation is correct, we wonder what the destiny was of the multivisceral resections, which worried the authors as to their possible exclusion from the analyses. Is it possible that such a variable, which regards a more aggressive surgery, has not been specifically considered as to its possible relationship with short-term outcomes? A simple four-cell test could and should have matched single vs multivisceral colectomies for each no/yes alternative of the three outcomes, since total colectomies, besides being irrelevant for the tumour's stage, were numerically negligible. The second hypothesis is that the non-resective procedures were initially somehow excluded, so that the term "other" refers to the multivisceral resections. If so, the latter were not alternative, but associated with either partial or total colectomies. Provided this second interpretation is right, as it is unlikely, the non-resective procedures would be consequently lacking, with all implications on patients' case mix. The third possibility, very improbable, would be that the group "other" includes all the procedures which are not mere CC resections, that is, non-resective CC surgery, multivisceral resections and other procedures wrongly selected by codes 42-54. So, the biases of both the previous possible interpretations

would be confusingly mixed. "Other intervention", the rate of which is, in any case surprisingly, only 3.52%, is an issue which should have been much better explained and analyzed, since, apart from patients' age and the meaningless elective/urgent admission status, the type of procedure was the sole variable significantly related to all the three outcomes.

10. The authors use, as an outcome measure, the 30-day re-admission rate, which they consider "as a fairly good surrogate of surgical complications occurring after hospital discharge".

This outcome measure is very weak since readmissions within such an interval, which are recorded irrespective of which was the discharging or receiving unit and of the readmission cause, may vary in accordance with locally diversified hospital network organizations and healthcare standards. Older patients and/or patients living in peripheral areas – a more common situation where there are low-volume hospitals – are more likely to be readmitted to other non-surgical units of the same or another hospital, even for mild non-surgical residual problems. In addition, with regard to outcome measures, we wonder why the authors did not take into account the postoperative hospital stay, which is a better, albeit indirect, surrogate of surgical complications, since these usually occur before a patient's discharge, whereas more than two thirds of readmissions after a surgical discharge are authoritatively reported as due to medical sequelae (59). So, it is incredible, except for the effect of any hidden biases, that the overall comorbidity rates were not significantly related to 30-day readmissions.

Other questionable issues

Besides the above sources of bias, the authors also run into some totally omitted-variable biases which could have relevant implications for the results. The first was not evaluating how many GSUs had in their hospital a postoperative Intensive Care Unit (ICU), staffed by its own specialists, and whether this significantly affected the outcomes, since the presence of an ICU is a hospital quality high-validity indicator (60) or even a requisite needed to undertake major surgery. Such a hospital characteristic could really identify the few GSUs at very critical low volume. Another neglected issue was the diverting enterostomy rates, which are an important variable, indirectly and inversely indicating a GSU's quality and CC treatment strategy for any possible timing: before resection (decompression rather than immediate resection of an occluding CC); during resection (emergency conditions, intra-operative technical problems, trend to protect the anastomoses); after resection (anastomotic leaks). Hartmann's procedure rates are another important issue. We are aware that these kinds of data are difficult to manage, but we would expect that this issue was at least discussed as a possible source of bias. The authors also ignore the possible impact of laparoscopy on outcomes. This question, and the related issue of the learning curve, would open too complex a scenario, but we cannot deny that further bias can derive from omitting laparoscopy from an outcome analysis after CC surgery.

All the noted biases invalidate, by themselves, all the results preceding the statistical analysis which cannot consequently remedy the wrong methodological premises, the poor quality database and its management, despite the use of a sophisticated statistical analysis, such as the described two-step model, probably built with the appreciable intention of overcoming the distribution variability of patients' characteristics in the different GSUs. However, also forgetting for a moment the pre-analysis biases, such a method cannot be properly considered a case mix-adjusted analysis. The case mix has not the same distribution in the GSUs, either singly considered or anyhow clustered for volume and focused practice, and even such imponderable factors as patient selection criteria for surgery and differently aggressive surgical attitudes (resectability judgement, performing multivisceral or hepatic resections) might show a great variability in the different GSUs. In addition, and most of all, many variables cannot be assumed as being independent, as the authors' method does. How can independence be assumed where as much as 12 out of 20 selected specific comorbidities regard the sole cardiovascular system? How can the length of hospital stay be independent of comorbidities or age, the hepatic and multivisceral resections be independent of the presence of metastases, or the metastases be independent of non-resective surgery? Similarly to many other studies,

these aspects have not been properly accounted for and faced, since the multilevel logistic regression, which like all generalized linear models assumes the independence of the variables and the exponential probability distribution, does not enable the extraction of this kind of information. For this purpose, other types of multivariate analysis, such as the multivariate analysis of the variance and the principal components analysis, could be more suitable.

Besides being ineffective, such a model seems rather contrived, just as the previous criteria for GSUs' volume and focused practice clustering, to obtain preconceived results. For instance, an important statistical issues is what method the authors used, among the various pseudo R-squareds, to evaluate how much of the total variation was explained by the co-variates included in the model. When analyzing data with a logistic regression, since an equivalent statistic to R^2 does not exist, several pseudo R^2 models have been developed, which can arrive at very different values, especially when dealing with non-continuous variables, as in the case of this analysis. If the authors do not explain which method they used, the legitimate doubt arises that they have chosen the most suitable one to support their own thesis. In this case, the pseudo R^2 values reported in Table 4 could be, in reality, not as determinant as they seem. A similar remark could be made for the variance proportional change values resulting in Table 5 for the "hospital characteristics".

The issue of how much each co-variable (patient, disease, hospital and GSU characteristics, and even surgeons' skills) contributes to a given outcome is decisive, much more so than the statistical significance of each variable-outcome relationship, since the modifiable variables are very few: at present, addressing the patient to the GSU or hospital where – or to the surgeon by whom – he/she will be operated on; in future, addressing health policies towards centralized or diffuse hospitals' and surgeons' high quality performances. Therefore, a selection of all the clinical and environmental variables which could affect the outcomes and a measurement of the "specific weight" of each of them are mandatory to establish whether and how much the modifiable variables are worth modifying. Without such an evaluation we are only constructing "houses of cards".

Discussion and conclusions

It is very appreciable that Lenzi and co-workers, besides being sober and cautious, are aware and sometimes even critical of their own study's limits and possible biases. Our criticism are not exhaustive, indeed, since various other issues could be questioned, partly acknowledged and answered by the authors themselves. They also acknowledge the negative consequences on patient accessibility by referring a large number of cases to a limited number of centres. Nonetheless, since their results *"suggest a relationship between GSU volumes and outcomes in elective patients"* they argue that *"centralization may facilitate the quality of surgery for these patients, including for screen-selected ones, to avoid exposure of apparently healthy people to unnecessary harmful treatment"*. This would imply that higher-risk, critically-ill patients, who cannot move in order to receive electively the supposed best care, can be treated in low-volume – i.e. low-quality – GSUs, thereby feeding the vicious circle of the volume-outcome relationship.

The authors' conclusive message is that, albeit with some doubts as to the pros and cons of centralization on health service organization, *"clinicians, policy makers and hospital administrators should consider the opportunity of centralizing CC surgery"*. Our conclusive opinion, by contrast, is that these kinds of articles – and we refer to many other studies, also regarding other gastrointestinal and HPB major procedures – are at best useless and at worst harmful, since they supply an erroneous message. There are no simple solutions to complex problems, and the above systemic biases may easily turn into an unsettling cognitive bias, a misleading logical short cut, which could confuse clinicians, policy-makers, hospital administrators and, we add, insurance companies exper and even magistrates.

Policy makers and healthcare service managers have their right or questionable strategies, but should avoid supporting and justifying changes in healthcare access pathways and hospital network organization by relying on structurally-biased volume-outcome studies, especially when they are carried out in totally different geographic areas, contexts and/or healthcare systems. It is also unacceptable, in our opinion, that policy makers arbitrarily assign resources, address pa-

tient flows, and decide whether a single hospital or a surgeons' team is of high, intermediate or low quality, except on the basis of an objective and reliable quality control. Although a legitimate purpose of policy-makers is "better care at lower costs" – and we cannot exclude that a centralization policy may lead to lower costs, but neither can we rule out that diseconomies or diminishing returns may occur – "better care" cannot, in any case, be rated *a priori*.

However, even assuming that a hospital's high volume is, in itself, a good reason to address to that hospital individual patients who can afford to move, we are unable to look at the other side of the coin, at the hidden problem, at the disregarded negative consequence of a centralization policy: non-high-volume centres, after a progressive further volume decrease and depletion in local resources and surgical skills, will likely be coping with patients in bad general condition and at high risk, who must be treated in emergency or, for elective surgery, cannot afford the move for reasons such as age, indigence or severe co-morbidity requiring close family support. Thus, centralization involves a potential impairment of experience and skills, and diminishing chances of quality improvement, in hospitals and areas in which – and for patients for whom – such policies might be problematic or unfeasible. Ultimately, such policies disadvantage the weak segments of the population, thereby moving towards an iniquitous health service. A health service system which, in Italy, is prevalently public, charge-free for the patient, and should have equity as one of its governance's principles.

In conclusion, in our opinion policy-makers should favour programs aimed at empowering the healthcare quality in all hospital contexts, by evaluating the process elements which result in improved outcomes and by understanding how to transfer these improvements from centres already assessed as being of excellence to other centres. They should thereafter evaluate hospitals' and GSUs' quality and cost/effective standards by *ex-post* methodologies, rather than labelling the institutions on the basis of preconceived and deceptive criteria. Only based on these premises and fulfilments can they adequately intervene, even by restrictive measures, if necessary, on hospitals or GSUs which cannot really sustain designated standards of performance. Economical sustainability problems and related healthcare policies are legitimate and can be shareable, but they cannot be concealed behind pseudoscientific reasoning.

References

1. Luft HS, Bunker JO, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med 1979; 301:1364-9.
2. Gooiker GA, van Gijn W, Wouters MW, Post PN, van de Velde CJ, Tollenaar RA. Systematic review and meta-analysis of the volume-outcome relationship in pancreatic surgery. Br J Surg 2011; 98: 485-94.
3. Halm EA. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Int Med 2002;137:511–20.
4. Chowdhury M.M., Dagash H., Pierro A. A systematic review of the impact of volume of surgery and specialization on patient outcome. Br J Surg 2007; 97: 145-61.
5. Gruen RL, Pitt V, Green S, Parkhill A, Campbell D, Jolley D. The effect of provider case volume on cancer mortality: systematic review and meta-analysis. CA Cancer J Clin 2009; 59: 192-211.
6. Archampong D, Borowski D, Wille-Jorgensen P, Iversen LH. Workload and surgeon's specialty for outcome after colorectal cancer surgery. Cochrane Database Syst Rev 2012; 3: CD005391.
7. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003; 349: 2117-27.
8. Urbach DR, Baxter NN. Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. BMJ 2004; 328: 737-40.
9. Ho V, Heslin MJ, Yun H, Howard L. Trends in hospital and surgeon volume and operative mortality for cancer surgery. Ann Surg Oncol 2006; 13: 851-8.
10. Balzano G, Zerbi A, Capretti G, Rocchetti S, Capitanio V, Di Carlo V. Effect of hospital volume on out come of pancreaticoduodenectomy in Italy. Br J Surg 2008; 95: 357-62.
11. Lenzi J, Lombardi R, Gori D., Zanini N, Tedesco D, Masetti M., Jovine E, Fantini MP. Impact of procedure volumes and focused practice on short-term outcomes of elective and urgent colon cancer resection in Italy". PLoS ONE 2013; 8:e64245.
12. Birkmeyer JD. Should we regionalize major surgery? Potential benefits and policy considerations. J Am Coll Surg 2000; 190: 341-349.
13. Epstein AM. Volume and outcome – it is time to move ahead. N Engl J Med 2002; 346: 1161-4.
14. Eppsteiner RW, Csikesz NG, McPhee JT, Tseng JF, Shah SA. Surgeon volume impacts hospital mortality for pancreatic resection. Ann Surg 2009; 249: 635-40.

15. Hogan AM, Winter DC. Does practice make perfect? Ann Surg Oncol 2008; 15: 1267-70.
16. Harrison A. Assessing the relationship between volume and outcome in hospital services: implications for service centralization. Health Serv Manag Res 2012; 25:1-6.
17. Gaynor M, Seider H, Vogt WB. The volume–outcome effect, scale economies, and learning-by-doing. Am Econ Rev 2005; 95: 243-7.
18. Huesch MD, Sakakibara M. Forgetting the learning curve for a moment: how much performance is unrelated to own experience? Health Econ 2009; 18: 855–62.
19. Eppsteiner RW, Csikesz NG, Simons JP, Yseng JF, Shah SA. High volume and outcome after lever resection: surgeon or center? J Gastrointest. Surg 2008; 12: 1709-16.
20. Billingsley KG, Morris AM, Dominitz JA, Marrews B, Dobic S., Barlow W, Wright GE, Mac Baldwin L. Surgeon and hospital characteristics as predictors of major adverse oucomes follwing colon cancer surgery. Arch Surg 2007; 142: 23-31.
21. Karanicolas PJ, Dubois L, Colquhoun PHD, Swallow CJ, Walter SD, Guyatt GH. The more, the better?: the impact of surgeon and hospital volume on in-hospital mortality following colorectal resection. Ann Surg 2009: 249:954-9.
22. Schmidt CM, Turrini O, Parikh P, House MG, Zyromski NJ, Nakeeb A, Howard TJ, Pitt HA, Lillemoe KD. Effect of hospital volume, surgeon experience, and surgeon volume on patient outcomes after pancreaticoduodenectomy. A single-institution experience. Arch Surg 2010; 145: 634-40.
23. Drolet S, MacLean AR, Myers RP, Shaheen AA, Dixon E, Buie WD. Elective resection of colon cancer by high-volume surgeons is associated with decreased morbidity and mortality. J Gastrointest Surg 2011; 15: 541-50.
24. Finks JF, Nicholas H, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. N Engl J Med 2011; 364: 2128-37.
25. Gasper WJ, Glidden DV, Jin C, Way LW, Patti MG. Has recognition of the relationship between mortality rates and hospital volume for major cancer surgery in California made a difference?: A follow-up analysis of another decade. Ann Surg 2009; 250: 472-83.
26. Chappel AR, BA, Zuckerman RS, Samuel RG, Finlayson SRG. Small rural hospitals and high-risk operations: how would regionalization affect surgical volume and hospital revenue? J Am Coll Surg 2006; 203: 599-604.
27. Swan RZ, Lau KN, Sindram D, Iannitti DA, Martinie JB. Pancreatic resection in a large tertiary care community-based hospital: building a successful pancreatic surgery program. Surg Oncol Clin N Am 2011; 20: 487-500.
28. Finlayson SR, Birkmeyer JD, Tosteson AN, Nease RF Jr. Patient preferences for location of care: implications for regionalization. Med Care 1999; 37:204-9.
29. Stitzenberg KB, Sigurdson ER, Egleston BL, Starkey RB, Meropol NJ. Centralization of cancer surgery: implications for patient access to optimal care. J Clin Oncol 2009; 27: 4671-8.
30. Riall TS, Eschbach KA, Townsend CM Jr, Nealon WH, Freeman JL, Goodwin JS. Trends and disparities in regionalization of pancreatic resection. J Gastrointest Surg 2007; 11: 1242-51.
31. Birkemeyer JD Sun Y, Goldfaden A, Birkmeyer NJ, Stukel TA. Volume and process of care in high-risk cancer surgery. Cancer 2006; 106:2476-81.
32. Killeen SD, O'Sullivan MJ, Coffey JC, Kirwan WO, Redmond HP. Provider volume and outcomes for oncological procedures. Br J Surg 2005; 92: 389-402.
33. Nathan H, Cameron JL, Choti MA, Schulick RD, Pawlik TM. The volume-outcomes effect in hepato-pancreato-biliary surgery: hospital versus surgeon contributions and specificity of the relationship. J Am Coll Surg 2009; 208: 528-38.
34. Rogers SO Jr, Wolf RE, Zaslavsky AM, Wright WE, Ayanian JZ. Relation of surgeon and hospital volume to processes and outcomes of colorectal cancer surgery. Ann Surg 2006;244:1003-11.
35. Engel AF, OomenJL, Knol DL, Cuesta MA. Operative mortality after colorectal resection in The Netherlands. Br J Surg 2005; 92: 1526-32.
36. Burns EM, Bottle A, Almoudaris AM, Namidanna R, Aylin P, Darzi A, Nicholls RJ, Faiz OD. Hierarchical multilevel analysis of increased caseload volume and postoperative outcome after elective colorectal surgery. Brit J Surg 2013; 100: 1531-38.
37. Van Erning FN, van Steenbergen LN, van den Broek WT, Rutten HJT, Lemmens VEPP. No difference between lowest and highest volume hospitals in outcome after colorectal cancer surgery in the southern Netherlands. Eur J Surg Oncol (2013), ahead of print: http://dx.doi.org/10.1016/j.ejso.2013. 08.020.
38. Ko CY, Chang JT, Chaudhry S, Kominski G. Are high-volume surgeons and hospitals the most important predictors of inhospital outcome for colon cancer resection? Surgery 2002; 132: 268-73.
39. Kee F, Wilson RH, Harper C, Patterson CC, McCallion K, Houston RF, Moorehead RJ, Sloan JM, Rowlands BJ. Influence of hospital an clinician workload on survival from colorectal cancer: cohort study. BMJ 1999; 318: 1381-85.
40. Iversen LH, Harling H, Laurberg S, Wille-Jørgensen P. Influence of caseload and surgical speciality on outcome following surgery for colorectal cancer: a review of evidence. Part 1: short-term outcome. Colorectal Dis 2007; 9: 28-37.
41. Birkemeyer JD, Dimick JB. Potential of the new Leapfrog standards: effect of process and outcomes measures. Surgery 2004; 135:569-75.
42. Birkmeyer JD, Finlayson EV, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. Surgery 2001: 130: 415-22.
43. Sowden A, Aletras V, Place M, Rice N, Eastwood A, Grilli R, Ferguson B, Posnett J, Sheldon T. Volume of clinical activity in hospitals and healthcare outcomes, costs, and patient access. Qual Health Care 1997; 6: 109-14.
44. Wexler MJ. More procedures, better quality of care? Is there a case for regionalization of pancreatic resection for neoplasm? CMAJ 1999; 160; 671-3.

45. Teh SH, Diggs BS, Deveney CW, Sheppard BC. Patient and hospital characteristics on the variance of perioperative outcomes for pancreatic resection in the United States: a plea for outcome-based and not volume-based referral guidelines. Arch Surg 2009; 144: 713-21.
46. Luft HS. Better for whom? Policy implications of acting on the relation between volume and outcome in coronary artery bypass grafting. J Am Coll Cardiol 2001; 38: 1931-33.
47. Shaian DM, Normand SL. The volume-outcome relationship:from Luft to Leapfrog. Ann Thor Surg 2003;75:1048-58.
48. Khuri SF, Henderson WG. The case against volume as a measure of quality of surgical care. World J Surg 2005; 29: 1222-29.
49. Sfekas A. Learning, forgetting and hospital quality: an empirical analysis of cardiac procedures in Maryland and Arizona. Health Econ 2009;18: 697-711.
50. Lymam S. Higher hospital volume and specialization. BMJ 2010; 340 :160-1
51. Wright FC, Fitch M, Coates AJ, Simunovic M. A qualitative assessment of a provincial quality improvement strategy for pancreatic cancer surgery. Ann Surg Oncol 2011; 18: 629-35.
52. Urbach DR, Bell CM. The effect of patient selection on comorbidity-adjusted operative mortality risk: implications for outcome studies of surgical procedures. J Clin Epidemiol 2002; 55: 381-5.
53. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ The volume-outcome relationship: don't believe everything you see. World J Surg 2005; 29: 1241-4.
54. Pieper D, Mathes T, Neugebauer E, Eikermann M. State of evidence on the relationship between high-volume hospitals and outcomes in surgery: a systematic review of systematic reviews. J Am Coll Surg. 2013; 216: 1015-25.
55. Violi V, Pietra N, Grattarola M, Sarli L, Choua O, Roncoroni L, Peracchia A. Curative surgery for colorectal cancer: long-term results and life expectancy in the elderly. Dis Colon Rectum 1998; 41: 291-8.
56. Violi V, Boselli AS, De Bernardinis M, Costi R, Pietra N, Sarli L, Roncoroni L. Anorectal reconstruction by electrostimulated graciloplasty as part of abdominoperineal resection. Eur J Surg Oncol 2005; 31:250-8.
57. Costi R, Di Mauro D, Giordano P, Leonardi F, Veronesi L, Sarli L, Roncoroni L, Violi V. Impact of palliative chemotherapy and surgery on management of stage IV incurable colorectal cancer. Ann Surg Oncol 2010; 17: 432-40.
58. Yuen E, Louis D, Cisbani L, Rabinowitz C, De Palma R, Maio V, Leoni M, Grilli R Using administrative data to identify and stage breast cancer cases: implications for assessing quality of care. Tumori 2011; 97: 428-35.
59. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med 2009; 360: 1418-28.
60. Bilimoria KY, Bentrem DJ, Lillemoe KD, Talamonti MS, Ko CY, on behalf of American College of Surgeons' Pancreatic Cancer Quality Indicator Development Expert Panel. Assessment of pancreatic cancer care in the United States based on formally developed quality indicators. J Natl Cancer Inst 2009; 101: 848-59.

Accepted: 10 October 2013
Correspondence:
Vincenzo Violi, M.D.
Full Professor of Surgery
Department of Surgery,
Parma University Medical School
University Hospital, Parma, Italy
Tel. 0039-0521-033533
Fax: 0039-0521-940125
E-mail: vincenzo.violi@unipr.it