Journal of Statistics on Health Decision

A1 Importance-Performance Analysis: a management tool on health decision making

Ricardo São João[1,2], Filipe Correia[3], Elsa Vieira[4]

[1]Department of Computer Science and Quantitative Methods, School of Management and Technology-Polytechnic Institute of Santarém, Complexo Andaluz, Apartado 295, 2001-904 Santarém, Portugal.
[2]Centro de Estatística e Aplicações Universidade de Lisboa (CEAUL), Faculty of Sciences of the University of Lisbon, Block C6 - Floor 4, 1749-016 Lisbon, Portugal.
[3]The Surgical Department, High-Resolution Scheduled Outpatient Clinic. Santarém District Hospital, Av. Bernardo Santareno, 2005-177, Santarém, Portugal.
[4]Sport Sciences School of Rio Maior - Polytechnic Institute of Santarém. NECE - Research Unit in Business Sciences. Av. Dr. Mário Soares -2040-413 Rio Maior – Portugal

Introduction

The High-Resolution Scheduled Outpatient Clinic (HRSOC) of the District Hospital of Santarém aims to improve the health service provided to the user, through a new concept of care: The Single Act Consultation (SAC). This new consultation allows the user to be seen by a health professional on the same day, perform the necessary complementary examinations that after which he/she will return home with a treatment or therapeutic proposal. Additionally, SAC offers its users the possibility of performing surgical procedures (minor surgeries) with discharge on the same day, reducing the number of trips to the hospital. This is a pioneering concept in Portugal: a new method of outpatient care with high clinical resolution. With the aim of measure the experience of the SAC users, to improve the quality of the health services provided in HRSOC, this research is being carried out. It is a cross-sectional epidemiological study focusing on the use of this health service. The information to be analysed was collected from a designed questionnaire structured for this purpose, addressed to 400 users, containing an evaluation scale for the health services integrated in the SAC in 26 items related to 5 domains which the Importance-Performance Analysis (IPA) model in the decision-making process in hospital settings will be implemented later on. The questionnaire also includes questions that allow a sociodemographic characterisation of its users.

Importance-Performance Analysis

The Importance-Performance Analysis-IPA was initially proposed by Martilla and James [1] and suffered several adjustments and derivations being applied to numerous sectors of activity where health is no exception [2-3]. In its essence, this analysis combines Importance (I) and Performance (P) in a two-dimensional graph, in which each of the quadrants represents the intersection of the level of I and P. The value of each attribute under analysis is identified in the IPA chart as an ordered pair (P, I). The set of all ordered pairs is called the I-P matrix. In the present context, the attributes are the 26 items that make up the SAC evaluation scale. Figure 1 depicts the IPA quadrants. On the x-axis are the values of the attributes relative to P, while the y-axis is assigned to I. For simplicity of reasoning, we define as barriers of the coordinate axes, the centre (three) of the five-point Likert scale. These quadrants (IQ,IIQ,IIIQ,IVQ) have specific meanings whose IPA will allow us to identify and evaluate which attributes experienced by users of the SAC should be: raise/need to concentrate (IQ), reduce/low priority (IIQ); Maintain/Keep Up the Good Work (IQ), or Eliminate/Possible Overkill (IV) in the formulation of a strategy by hospital decision-makers aimed at maximizing satisfaction from the perspective of these users. A key issue in IPA is what values to be set as barriers for the P and I axes. In most cases, the value to be attributed to each of the barriers lacks statistical justification and is left to the subjectivity of the decision-maker. As barriers of the co-ordinate axes P and I usually fall: the centre of the scale, the average value or the midpoint of the P and I attributes.

Motivation and aim

The absence of statistical methodology that supports the establishment of such barriers makes the predictive and discriminating power of IPA unfeasible [4]. The pertinence of adopting robust criteria in the setting of barriers therefore becomes an imperative. This research aims to fill this gap by making IPA a more robust and valuable tool, particularly for hospital decision-makers to maximize user satisfaction in the SAC. To this end, the authors propose the adoption of the logistic regression model to help determine IPA barriers, whose cut-off threshold will be guided by the maximization of the discriminant power assessed either by the area under the Receiver Operating Characteristic (ROC) curve [5] or by the adoption of

Keywords:
Health, Importance-Performance Analysis, Satisfaction, Single Act Consultation, User.

Corresponding author:
Ricardo Miguel Vieira de São João
ricardo.sjoao@esg.ipsantarem.pt

Supplementary material:
Archived online: Link

Conflict of interest:
The authors declare no conflict of interests

First published: 22.JUL.2021

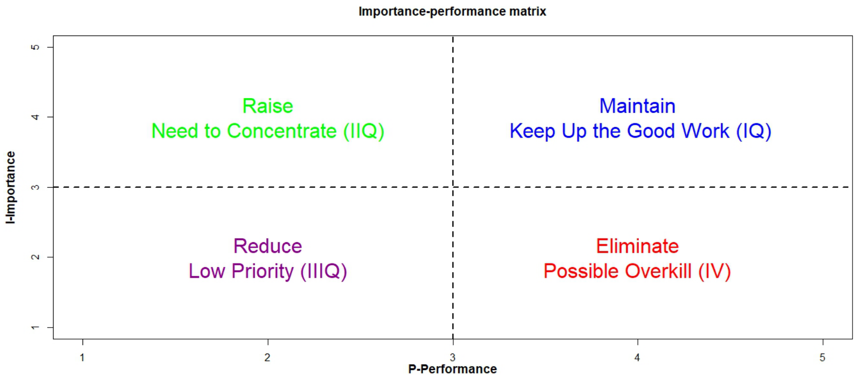

Figure 1 - Quadrants of the IPA chart and their reading regarding decision making for each attribute under analysis. Adapted from Martilla and James [1].

Cohen's k [6], Somers' correlation [7] and Matteus [8] coefficients. In addition, we intend to carry out a sociodemographic characterization of the users of this health service.

Methods:

The sample size determination and the choice of sampling criteria deserved particular attention. Based on the history of clinical records between January 2014 and October 2018, the medical specialties with the highest monthly inflow were identified. We chose to adopt a non-proportional stratified random sampling with these specialties as strata. The adoption of a non-proportional stratified random sample is justified because in this way the four medical specialties with the highest affluence, our strata, will have their respective representativeness assured, while in a proportional stratified random sample all specialties would have equal weight, i.e. 1/4 each. The strata considered and their size were: dermatology (115), ophthalmology (68), otorhinolaryngology (84), surgery (98). An additional 10% increase was also chosen to make up for possible losses in the collection of information, totalling 400 users to be sampled. The level of significance and the tolerable sampling error were set at 5%. Data collection took place between July to September 2019.

The questionnaire includes sociodemographic questions and 26 items distributed over 5 domains, namely: Access/Admission to health care (AcAd), Quality of Care (QuCa), Facilities (Fa), Satisfaction (Sa), Receptivity to the SAC (Re). The items are evaluated on a Likert-type scale of five points where the Importance (not at all important to very important) and Performance (not at all satisfied to very satisfied) experienced by the user is assessed in order to contribute to a greater and more detailed knowledge of the user's profile.

Exploratory and confirmatory factor analysis will be used to validate the domains present in the questionnaire as well as the use of Cronbach's Alpha coefficient [9] and Guttman's Lambda 6 [10] coefficient to measure the levels of internal consistency of the scale. In the descriptive study, the sample mean (± standard deviation) and median (interquartile range) were used in the sociodemographic characterization of the respondents. Having adopted the logistic regression model as a support for fixing the barriers in IPA, the need arises to define a new Gold Standard (GS) that allows for a correct classification of the attributes under analysis. The authors propose that this GS be based on the following criteria:

$$GSi = \begin{cases} 1; \bar{X}(P_i) - \bar{X}(I_i) \geq 0 \\ 0; \bar{X}(P_i) - \bar{X}(I_i) \geq 0 \end{cases}, i = 1, ..., 400 \qquad \text{(equation 1)}$$

where index i refers to the 400 sampled users, $\bar{X}(P_i)$ and $\bar{X}(I_i)$ represent respectively the sample mean of the attributes associated with P and I in the i users. This criterion only takes into account the classification given by each respondent and is not influenced by other measures [11]. The logistic regression model will have as response variable and as predictors the attributes referring to the five domains with regard to P and I. Table 1 summarises the frequencies and formulas that support the calculations intervening in the discriminant measures associated with the logistic regression model. The value L therein refers to the cut-off threshold that maximizes Cohen's κ coefficient [12]. An equally considered measure was Somers' D coefficient, determined on the basis of the Area Under the ROC curve-AUC by the expression D_{Somers}=2(AUC-0,5).

The Ethics Committee of the hospital approved this study in June 2019 (see supplementary material). All participants were asked to sign the informed consent form, after being previously informed about the study and its purpose, confidentiality and anonymity were ensured, as well as the possibility of withdrawing from the questionnaire at any moment during its application. All statistical analysis was performed in R (version 4.0.2).

Table 1 - Support for the calculation of discriminant measures associated with the logistic regression

	Gold Standard		
Estimated Probability	**1**	**0**	**Total**
$\hat{\pi}_i \geq L \Rightarrow 1$	A	B	A+B
$\hat{\pi}_i < L \Rightarrow 0$	C	D	C+D
Total	A+C	B+D	A+B+C+D=N

Legend: L refers to the cut-off threshold that maximizes Cohen's κ coefficient; $\hat{\pi}_i$ estimated probability for the ith user from the logistic regression model. Discriminant measures: Sensitivity= $\frac{A}{A+B}$; Specificity= $\frac{D}{B+D}$; Matheus= $\frac{AD+BC}{\sqrt{(A+C)^2(B+D)(C+D)}}$; Cohen's κ= $\frac{A+B-\frac{(B+D)(D+C)+(A+C)(A+B)}{N}}{N-\frac{(B+D)(D+C)+(A+C)(A+B)}{N}}$.

Results:

An exploratory and confirmatory factor analysis was performed, which corroborated the 5 domains present in the questionnaire to assess the delivery of health services integrated into the SAC, deserving equal attention the levels of internal consistency of the scale (Cronbach's Alpha=0,84; Guttman's Lambda 6=0,98).

The following results were obtained from the stratified sample consisting of 400 users, which allow tracing the profile of its users: a greater demand for the SAC was registered among users living in rural areas (58,25%), almost all of them from the district of Santarém (98,75%), female (53,75%), retired people and pensioners being the group with the highest adherence (57%) and the overwhelming majority through the social security/national health service (83%). The median age was 63 years (AIQ=29), being slightly higher in males (64 years; AIQ=26). Regarding marital status, 60,25% of the users are married; 17,71% are single; 15,5% are widowed and the remaining are divorced. As for literacy, the education with the greatest weight is represented by basic education (67,25%), followed respectively by professional education (14,74%) and higher education (10,75%). The users who are illiterate and/or do not attend school have a residual representation. Working users (self-employed or employed) represented approximately 29% of the consultations, with the unemployed and students having an equal quota (7%). When questioned about who they live with, a minority (3%) reported being institutionalized, 10,5% living alone and 86,25% living with relatives. In the SAC assistance plan the answers with the highest demand were: (i) first consultation with examinations/treatment (42,25%); (ii) first consultation with examinations and surgery (30%), (iii) only consultation (15%) and finally (iv) 12,25% first consultation, examinations and surgical proposal. The average time of stay at the SAC was 1h14m ± 41m, with different values per specialty: otorrino 1h10m ± 36m, surgery 1h35m ± 53m, dermatology 1h ± 27m and ophthalmology 1h10m ± 34m. 72,25% of the users found out about the SAC through other health institutions; 22,75% through the hospital itself; 2,5% through family and friends, and the remaining cases through the media and the Internet. Regarding the service provided they were asked to evaluate their overall satisfaction with the SAC, in which 99,75% of respondents said they were very satisfied or satisfied; all would choose the SAC again, recommend it to friends/family and think it should be replicated in other institutions.

The following frequencies were counted for the new Gold Standard (equation 1): 283 zeros and 117 ones. According to Prabhakaran [13], in a logistic regression model, the proportion of events in each category should be similar/balanced. In statistical modelling, to avoid class bias, it is suggested to subdivision of the information into two data sets: one for training (adjustment) and another for testing (validation and prediction). Table 2 presents summary information concerning the logistic regression model for the two data sets. The logistic regression model had in its structure the response variable GS (equation 1) and the following set of significant predictors related to SAC: ease of booking; flexible hours; well-signposted consultation office; parking for users and carers; reduced waiting list. Table 3, present in the supplementary material, contains the list of all the study attributes and domains, in particular the predictors of the regression model. The assumptions of the logistic regression model were considered in the modelling. The existence of multicollinearity was not detected in both models through the VIF (variance inflator factor) measure. The significance of the models was also tested (null p-values) based on the deviance and other measures presented in Table 2.

As the Test model was used to define the barriers in IPA, the Hosmer & Lemeshow test (goodness of fit) was used to corroborate its adequate adjustment (X^2_{Test}=12,467, df=8, p-value=0,132). This information was complemented by Nagelkerke's pseudo R^2 (0,746). The values of the discriminant measures associated to the Test model are: (i) L=0,41 is the cut-off threshold that maximizes Cohen's κ coefficient; (ii) the area under the ROC curve (0,969) points to an exceptional discriminant power (AUC≥0,9) [14]; (iii) the Sommers correlation coefficient (0,939) corresponds to a perfect discriminant power; (iv) Matteus' correlation coefficient (0,819) indicates a high association between the values predicted by the model and the

Table 2 - Summary of the fit of the logistic regression models in the two data sets.

Model	Sample Size	Null deviance (df)	Residual deviance (df)	AIC	BIC	IRLS
Training	162	224,58 (161)	101,92 (155)	115,92	137,53	7
Testing	238	202,247 (237)	69,726 (231)	83,726	108,03	8

df- degrees of freedom, AIC - Akaike Information Criterion, BIC -Bayesian Information Criterion, IRLS - Iteratively Reweighted Least Squares.

observed ones. The adoption of this coefficient is justified by the fact that it a robust measure even in classes with distinct dimensions [15];(v) Cohen's k coefficient of agreement (0,794) presents a strong level of agreement (0,61<k<0,80) [16]. Thus, the Test model seems to meet the conditions in supporting the definition of the barriers in the IPA.

The test model is defined by the following mathematical expression, which represents the logistic regression model, particular case of a generalized linear model (glm) with logit link function:

$$glm(formula = Gsi \sim AcAd\_1\_P + AcAd\_4\_P + Fa\_4\_P + Fa\_6\_P$$
$$+AcAd\_1\_I + Sa\_4\_I, family = binomial(link = "logit"),$$
$$data = testingData)$$

The significance of the predictors can be found in Table 3 (supplementary material). In terms of nomenclature in the previous expression, predictors ending with the letter I refer to importance while those ending with the letter P refer to performance. The structure of the two models is analogous, differing only in the data set used.

Figure 2 shows the location of the barriers according to the methodology, attributes and respective domains. Table 3 (see supplementary material) summarises, based on the IPA analysis, a strategy to support hospital decision-makers aimed at maximising SAC user satisfaction. As can be seen, the strategy may change depending on the location of the barriers. These quadrants have specific meanings whose IPA will allow us to identify and evaluate which attributes experienced by users of the SAC should be raised, lowered, maintained, or eliminated in the formulation of a strategy by hospital decision-makers aimed at maximizing satisfaction from the users' perspective.

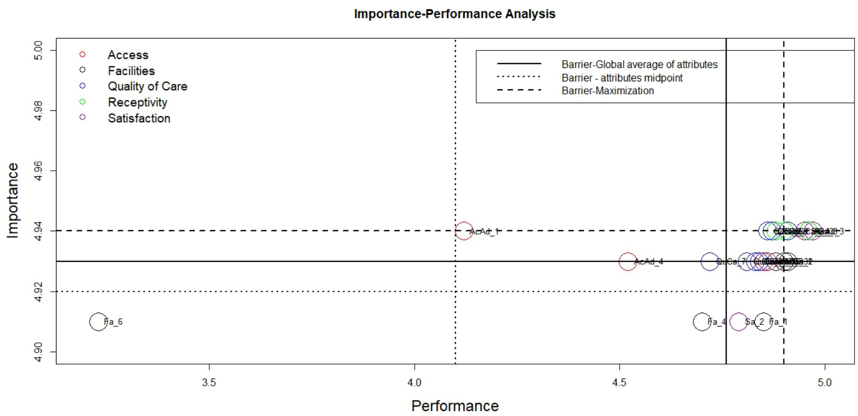

Figure 2 - Definition of the barriers in the IPA model according to the methodologies discussed and respective location of the quadrants aimed at supporting decision-making. The domain of each attribute can be identified by its colour.

Discussion:

Based on the IPA analysis we found that: (i) it is possible to define strategies aimed at maximising the satisfaction of users who have experienced the SAC based on the location of the barriers; (ii) the use of logistic regression supported by discriminate measures makes an additional contribution to the definition of the IPA barriers by challenging the hospital decision makers to a more interventive attitude; (iii) as we get closer to the maximization barrier, the attributes change more frequently in the decision quadrants; (iv) in the usual scenario, using the centre of class (P,I)=(3,3) nothing would be changed in the hospital management process (IQ: maintain/ keep up the good work); (v) it was also found that even in a scenario of excellence with regard to the positioning of the attributes on the P and I axes experienced by users of the SAC, there is always a "window of opportunity" for an increase in user satisfaction with this methodology. The authors believe that IPA is a valuable management tool in health decision making because it has a clear interpretation, manages in a two-dimensional graph and in a concise way to set barriers, barriers that will help the decision-maker. The absence of subjectivity in setting these barriers using a robust statistical methodology proves to be an added value. IPA also suggests that the quality of care should be maintained and prioritized. One of the strengths of this study is that this methodology has led to greater efficiency in the SAC translated by the release of vacancies and reduction of the waiting list with increased levels of user satisfaction.

As regards limitations, the authors are of the opinion that this study should be applied to more hospital units to gain scale and greater comparability.

Ethics committee and informed consent:

The current research was approved by an independent ethics committee and subjects gave their informed consent before they were enrolled in the study.

Acknowledgements:

This work is partially financed by national funds through FCT – Fundação para a Ciência e a Tecnologia under the project UIDB/00006/2020 & UIDB/04630/2020.

References:

1. Martilla J, James J. Importance-Performance Analysis. Journal of Marketing. 1977;41(1):77-79. https://doi.org/10.2307/1250495
2. La Fata C, Lupo T, Piazza T. Service quality benchmarking via a novel approach based on fuzzy ELECTRE III and IPA: an empirical case involving the Italian public healthcare context. Health Care Management Science. 2017;22(1):106-120. https://doi.org/10.1007/s10729-017-9424-4
3. Zarei E. Service quality of hospital outpatient departments: patients' perspective. International Journal of Health Care Quality Assurance. 2015;28(8):778-790. https://doi.org/10.1108/IJHCQA-09-2014-0097
4. Sever I. Importance-performance analysis: A valid management tool?. Tourism Management. 2015; 48:43-53. https://doi.org/10.1016/j.tourman.2014.10.022
5. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27(8):861-874. https://doi.org/10.1016/j.patrec.2005.10.010
6. Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 1960;20(1):37-46. https://doi.org/10.1177/001316446002000104
7. Somers R. Simple Approximations to Null Sampling Variances. Sociological Methods & Research. 1980;9(1):115-126. https://doi.org/10.1177/004912418000900107
8. Matthews B. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) - Protein Structure. 1975;405(2):442-451. https://doi.org/10.1016/0005-2795(75)90109-9
9. Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334. https://doi.org/10.1007/BF02310555
10. Guttman L. A basis for analyzing test-retest reliability. Psychometrika. 1945;10(4):255-282. https://doi.org/10.1007/BF02288892
11. Vieira E, Ferreira J, São João R. Creation of value for business from the importance-performance analysis: the case of health clubs. Measuring Business Excellence. 2019;23(2):199-215. https://doi.org/10.1108/MBE-09-2018-0076
12. São João R. Modelação da insuficiência cardíaca congestiva por modelos lineares generalizados com variável resposta binária. Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística: Estatística Arte de Explicar o Acaso. Vila Real: Sociedade Portuguesa de Estatística; 2008. p. 603-612.
13. Prabhakaran S. Logistic Regression With R [Internet]. R-statistics.co. 2021 [cited 2021 Apr 1]. Available from: https://r-statistics.co/Logistic-Regression-With-R.html
14. Hosmer D, Lemeshow S. Applied Logistic Regression. 2nd ed. New York: John Wiley & Sons; 2000.
15. Powers D. Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. Journal of Machine Learning Technologies [Internet]. 2011 [cited 2021 Apr 1];2(1):37-63. https://arxiv.org/abs/2010.16061
16. Landis J, Koch G. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159. https://doi.org/10.2307/2529310