JBUON 2021; 26(1): 275-277
ISSN: 1107-0625, online ISSN: 2241-6293 • www.jbuon.com
Email: editorial_office@jbuon.com

SHORT COMMUNICATION

Decision support systems for the prediction of lymph node involvement in early breast cancer

Raffaella Massafra[1], Domenico Pomarico[1], Daniele La Forgia[2], Samantha Bove[3], Vittorio Didonna[1], Agnese Latorre[4], Anna Orsola Russo[5], Pasquale Tamborra[1], Vito Lorusso[4], Annarita Fanizzi[1]

[1]Struttura Semplice Dipartimentale di Fisica Sanitaria, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Viale Orazio Flacco 65, 70124 Bari, Italy. [2]Struttura Semplice Dipartimentale di Radiologia Senologica, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Viale Orazio Flacco 65, 70124 Bari, Italy. [3]Dipartimento di Matematica, Università degli Studi di Bari, Bari, Italy. [4]Unità Operativa Complessa di Oncologia Medica, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Viale Orazio Flacco 65, 70124 Bari, Italy. [5]Dipartimento di Oncologia Medica, Università degli Studi di Napoli, Napoli, Italy.

Summary

The prediction of lymph node involvement represents an important task which could reduce unnecessary surgery and improve the definition of oncological therapies. An artificial intelligence model able to predict it in pre-operative phase requires the interface among multiple data structures. The trade-off among time consuming, expensive and invasive methodologies is emerging in experimental setups exploited for the analysis of sentinel lymph nodes, where machine learning algorithms represent a key ingredient in recorded data elaboration. The accuracy required for clinical applications is obtainable matching different kind of data. Statistical associations of prognostic factors with symptoms and predictive models implemented also through on-line softwares represent useful diagnostic support tools which translate into patients quality of life improvement and costs reduction.

Key words: *bioinformatics, complex diseases, data science, machine learning, personalized medicine*

Early stage cancer detection represents a crucial step to prevent cells diffusion in metastatic processes. This status recognition is implemented clinically and radiologically.

An example of this kind is referred to lymph nodes in breast cancer, whose positive status shows a non-lethal metastatic diffusion, usually treated through dissection or biopsy [1]. A metastatic lymph node can present various characteristics such as dimensional increase, morphological variation, increase in cortical thickness and vascularization, lack of visualization or displacement of the hilum, characteristics often detectable by clinical or ultrasound examinations [2]. The sensitivity and specificity of this type of evaluation is not absolute. In fact, in the absence of lymph node abnormalities detectable on clinical examination or imaging, the guidelines suggest in any case of malignant breast lesions the removal of the first axillary draining lymph nodes. This is worked out via injecting a radionuclide that allows the lymph nodes localization and their removal before and during surgery with sentinel node and occult lesion localization (SNOLL) technique. These lymph nodes, named as

---

Corresponding author: Daniele La Forgia, MD. Via Orazio Flacco, 65, 70124, Bari, Italy,
Tel: +39 0805555111, Email: d.laforgia@oncologico.bari.it
Received: 28/09/2020; Accepted: 21/10/2020

sentinel lymph nodes, are inspected according to an intra-operative pathology exam, known as one-step nucleic acid amplification (OSNA), with range 87.5-100% for sensitivity and specificity in 90.5-100% [3], whose result may mismatch with the radiological one. Although the one-step nucleic acid amplification is time consuming and expensive, it is still the intra-operative exam endowed with the best performance, even though the elastic scattering spectroscopy achieves the same accuracy. Latest developments benefit from a two-stage partially supervised image classification model applied to recorded optical spectra [4]. These methods, due to the relatively high costs and the necessary multi-specialist commitment, although more reliable than the clinical and ultrasound examination, must be limited to actually negative cases: the clinical-ultrasound doubts should in fact be resolved before surgery by means of a withdrawal with a needle thus further selecting the cases to be sent directly to axillary dissection or to SNOLL. However, it is not always possible to arrive at surgery without diagnostic doubts and therefore the natural need for decision support emerges which could provide further assessments through artificial intelligence techniques [4]. This article aimed at providing a concise overview of current data science approaches in statistical and predictive models used for lymph node status detection in pre-operative stage.

The usefulness of a decision support system emerges therefore accordingly as a data elaboration system able to exploit information available from retrospective cases and then to compare learned structures with current cases. If it is endowed with a high sensitivity and specificity, it can alleviate both surgery activity and medical responsibility. In medical purposes such utilities can be broadcast more generally as patient safety, quality of care and healthcare delivery efficiency.

Joint analysis of different kind of data is the fingerprint of nowadays artificial intelligence. This is the data science approach to biological examples like lymphangiogenesis, whose mechanism includes the interaction between vessel remodelling, caused by the tumour mass, and biomarker indicators variation. The enlargement of lymph nodes is strictly related with such a condition, thus yielding the enhancement of growth factors, whose accurate recognition could lead to significant improvements in diagnosis and therapeutic approaches in different types of tumour [5]. Concerning lymph nodes, a clinical decision support system (CDSS) is required to avoid expansive and invasive procedures, such as OSNA: a CDSS with a high sensitivity allows a surgeon to directly eradicate axillary lymph nodes during the breast cancer dissection surgery.

A first class of algorithmic procedures aims at characterizing statistical associations among prognostic factors and symptoms. A set of features potentially endowed of information useful for the problem has to be identified to train a decision support system. In the scientific literature the status of lymph nodes associated with a surgical protocol as dissection or biopsy is related with molecular subtypes consisting of the combination of hormone receptors (HRs: ER and PgR), proliferation marker (Ki67) and human epidermal growth factor receptor 2 (HER2), as verified through statistical analysis based on univariate, multivariate logistic regression and regression trees [1]. Concerning the prediction of sentinel lymph node status both luminal A (positive HRs, low Ki67, negative HER2) and triple negative (negative HRs, negative HER2) show a significant association with nodal negativity [1]. Moreover there are much more features identified using both statistically significant associations [1] and predictive models [6,7] which are exploited through the calculation of multivariate functions, assuming values playing the role of scores, as nomograms [2,6]. Recent literature studies propose a wide variety of artificial intelligence methods aimed at prediction of lymph nodal involvement, established through a probability associated with the occurrence of an event (e.g. positive/negative) for each sample cases.

According to aforementioned algorithmic methods some prognostic factors play a dominant role, such as age, primary tumour features (size, grade, type, location) and a vascular one (lymphovascular invasion) [6], the latter representing a promising interface with genomics [5]. Unfortunately performances in the prediction of a positive lymph node, initially categorized as a false negative, are still not suffciently satisfying with regard to sensitivity and specificity. CancerMath [8] is a software on line which offers the possibility to estimate the probability of lymph node involvement as well as the others wider cancer survival analysis [9,10]. This score corresponds to the estimated probability to observe a positive sentinel lymph node. A primary importance is attributed to tumour size, weighted by prognostic factors such as age, HRs, tumour type and grade. Different databases were tested using on-line software to study their behaviour with respect to population variations [8-10]. The tumour survival probability estimation was less reliable in some subgroups, such as older patients [9,10], while the percentage of patients with positive lymph nodes (40.6%) in the southeast Asia sample (43.6%) was underestimated [8].

Relations among different kinds of data play a fundamental role to overcome these limitations.

These multivariate statistical techniques are based on graphical models, as a matter of fact their structure intrinsically includes the interaction among schemes of multiple data elaboration. A multivariate logistic regression analysis exploiting both biomarkers indices and data acquired by means of radiological or ultrasound images (lymph node diameter, cortex thickness) of clinically positive patients reaches a mean classification accuracy equal to 0.864 [2]. Genomic analyses of metastatic cancer reveal complex patterns of dissemination and understanding the complex interactions between tumour cells, the immune system and lymphatics could be essential to develop therapeutic and prognostic approaches to cancer [5].

Concluding, the common front of the scientific community is the development of support systems that make it possible to define models of personalized medicine: for each patient, multiple data types are taken into account, such as the ones provided by histology, radiology and genomics. Indeed, the treatment of multifactorial or complex diseases, the category of the vast majority of tumours, is intrinsically related with the management of a interplay effect among genetic, environmental and complex factors: their mutual optimization allows the definition of a targeted surgical and therapeutic care plan which translates into the reduction of health care costs and the improvement of the quality of life of patients. The increase of sensitivity to lymph node metastases in cases of clinical-radiological suspicion not supported by positive needle sampling could lead directly to axillary dissection, while reducing the use of SNOLL technique and related costs; moreover, the reduction of false positive cases could limit unnecessary axillary lymph node dissections, sources of physical and psychological discomfort due to lymphedema consequences.

Authors' contribution

R.M., A.L., A.F., and D.P. conceived the study; R.M., A.F., supervision; R.M., A.F., and D.P. writing–original draft; D.L.F., S.B., V.D., A.O.R., P.T., and V.L., writing–review and editing. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

This work was supported by funding from the Italian Ministry of Health "Ricerca Corrente 2018–2020".

Conflict of interests

The authors declare no conflict of interests.

References

1. Diotaiuti S, De Summa S, Altieri R et al. Biomarker phenotyping drives clinical management in axillary sentinel node: A retrospective study on women with primary breast cancer in 2002. Oncol Lett 2020;20:2469-76.
2. Qiu HC, Zeng F, Zhang C et al. A nomogram to predict the probability of axillary lymph node metastasis in early breast cancer patients with positive axillary ultrasound. Sci Rep 2016;6:21196.
3. Banerjee SM, Michalopoulos NV, Williams NR et al. Detailed evaluation of one step nucleic acid (OSNA) molecular assay for intra-operative diagnosis of sentinel lymph node metastasis and prediction of non-sentinel nodal involvement: Experience from a London Teaching Hospital. The Breast 2014;23:378-84.
4. Zhu Y, Fearn T, Chicken DW et al. Elastic scattering spectroscopy for early detection of breast cancer: partially supervised Bayesian image classification of scanned sentinel lymph nodes. J Biomed Optics 2018;23:085004.
5. Farnsworth RH, Achen MG, Stacker SA. The evolving role of lymphatics in cancer metastasis. Curr Opin Immunol 2018;53:64-73.
6. Chen J, Chen B, Yang Z et al. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: assessment of an existing nomogram and a new predictive nomogram, Breast Cancer Res Treat 2012;135:839-48.
7. Liu C, Zhao Z, Gu X et al. Establishment and Verification of a Bagged-Trees-Based Model for Prediction of Sentinel Lymph Node Metastasis for Early Breast Cancer Patients. Front Oncol 2019;9:282.
8. Miao H, Hartman M, Verkooijen HM et al. Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia. BMC Cancer 2016;16:820.
9. de Glas NA, van de Water W, Engelhardt EG et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. Lancet Oncol 2014;15:722-29.
10. de Glas NA, Bastiaannet CC, Engels AJM et al. Validity of the online PREDICT tool in older patients with breast cancer: a population-based study. Br J Cancer 2016;114:395-400.