Acta Scientific MEDICAL SCIENCES (ISSN: 2582-0931)

Volume 7 Issue 9 September 2023

Review Article

Implications of the EGF +61G>A SNP and platelet count on the serum EGF profile: A Critical Review of the proposed predictors for the success of CIMAvax-EGF

Idania González-Pérez[1]*, Elizabeth Cuétara Lugo[1], Elizabeth Reyes Reyes[1] and Enrique Iglesias[2]

[1]Department of Pharmacology, National Institute of Oncology and Radiobiology, Havana, Cuba

[2]Department of Vaccines, Center for Genetic Engineering and Biotechnology, Havana, Cuba

***Corresponding Author:** Idania González-Pérez, Department of Pharmacology, National Institute of Oncology and Radiobiology, Havana, Cuba.

Received: July 31, 2023

Published: August 10, 2023

Abstract

Epidermal growth factor (EGF) is an important ligand for the epidermal growth factor receptor (EGFR), the overexpression of which is directly related to the genesis of some epithelial tumors, including non-small cell lung cancer (NSCLC). Therefore, EGF is considered both a marker of cancerous disease and a target for its treatment. However, some methodological problems related to the quantification of this molecule have been affecting its use as a biomarker for disease diagnosis, prognosis, response prediction, and therapy monitoring. CIMAvax-EGF® is a Cuban vaccine for the treatment of NSCLC, targeting EGF. In the last five years, several biomarkers have been proposed for the personalized indication of this vaccine, which have shown limitations in predicting and explaining the response to therapy in some patients. Here we comparatively review these potential predictors of vaccine success and hypothesize that the performance of pretreatment serum EGF concentration, the efficacy biomarker currently in use, is limited by variability in serum EGF quantitation due to: the progressive release of EGF by platelets during serum collection, interindividual differences in platelet count, and different EGF expression caused by the functional single nucleotide polymorphism (SNP) EGF +61G>A. In addition, we discuss the impacts of standardizing serum collection and normalizing serum EGF concentration, on the discriminatory capacity (diagnostic accuracy) of several EGF-related biomarkers described in a more recently conducted prospective exploratory study. The greater discriminatory accuracy of the new variables suggests their better performance in predicting EGF tumor dependence and the response to anti-EGF/EGFR therapies such as the CIMAvax-EGF® vaccine.

Keywords: NSCLC; EGF; +61G>A SNP; EGFR; Diagnostic Accuracy; Predictive Biomarker; CIMAvax-EGF®

Introduction

Cancer is a worldwide health concern. According to the Global Cancer Observatory [1], in 2020 lung cancer (LC) was the leading cause of death from malignant tumors globally (18%) and the second in incidence (11.4%), only surpassed by breast cancer (11.7%). In Cuba, carcinomas of the trachea, bronchi and lung are among the most frequent malignant tumors, and they accounted for the highest cancer mortality rate in both genders at the end of 2019 [2]. The early diagnosis of LC continues to be one of the most important unmet needs [3], which along with the personalization of therapies allows for a better response and survival of patients [4]. Therefore, it is considered critical for better management

of the disease. Currently, approved diagnostic biomarkers lack the required sensitivity, specificity, and reproducibility, even for diagnosis in advanced stages [5]. Consequently, and due to the absence of obvious symptoms, the majority of patients with non-small cell lung cancer (NSCLC), the main LC histological subtype [6], are diagnosed in locally advanced or metastatic stages. This late diagnosis makes unfeasible surgery, which is closely linked to the concept of cure, and determines an overall survival (OS) rate in patients at five years between 4%-12%; which could rise to 50%-80% if the diagnosis occurred early in stage I [4,5,7].

In recent years, low-dose computed tomography (LDCT), the current standard of LC screening, has substantially reduced the likelihood of death from LC in high-risk populations [8,9]. Nevertheless, the test has several limitations, such as high cost and the occurrence of false positives results to be confirmed by occasionally more invasive procedures. Additionally, it is known that annual checkups with procedures based on ionizing radiations can cause cancer in healthy people [10].

The effective application of targeted therapies and immunotherapies, which depends on the evaluation of accurate biomarkers, has also increased OS in patients with LC. The Cuban vaccine CIMAvax-EGF®, immunotherapy directed at the epidermal growth factor (EGF), has several potential biomarkers evaluated for its personalized indication. However, all candidates have limitations in predicting/explaining response to therapy in some patients. In this manuscript, we critically review these potential predictors of vaccine success and hypothesize that the performance of pretreatment serum absolute EGF concentration, the current biomarker of vaccine efficacy, is limited by variability in EGF quantitation due to: the progressive release of EGF by platelets during serum collection, interindividual differences in platelet count, and different EGF expression levels caused by the functional EGF +61G>A single nucleotide polymorphism (SNP). Furthermore, we discuss the implications of standardizing serum collection [11] and normalizing the absolute serum EGF concentration [12], on the ability of alternative EGF-related biomarkers to discriminate/diagnose NSCLC patients (diagnostic accuracy) and to stratify them in order to predict response to CIMAvax-EGF® or other EGF/EGFR-targeted therapies.

Cuban vaccine CIMAvax-EGF®

Active immunotherapy or vaccination [13], aimed at inducing a cellular and/or humoral response against specific or tumor-associated and differentially-expressed antigens, represents a promissory alternative for cancer treatment, despite its modest benefits in terms of survival [14]. The EGF shows differentiated levels in sera from patients with NSCLC and other epithelial tumors [15-22]; is involved in tissue invasion and metastasis [23,24]; and its concentration in serum correlates with the stage of the disease, its clinical course and prognosis [25]. That is why it represents an attractive target for the rational design of antitumor vaccines. The Cuban therapeutic vaccine CIMAvax-EGF® is an EGF-targeted immunotherapy with very low toxicity and proven efficacy in NSCLC treatment [26,27]. The vaccine is indicated in advanced disease, as maintenance therapy with drug switching after first-line treatment, in patients who attained an objective response or disease stabilization and have acceptable general conditions to respond satisfactorily to treatment [28]. The vaccine induces anti-EGF antibodies that recognize the EGF in the tumor microenvironment and blood circulation, preventing signaling through the EGFR pathway. In this way, it reduces cell proliferation, angiogenesis, and metastasis, with arrest or deceleration of tumor progression [29,30]. The efficacy and safety profile of the vaccine allows its long-term use, either as monotherapy or in combination [30,31], with a significant impact on the survival of patients having EGF-dependent tumors [26,32].

Diagnostic capacity and biomarker value of serum EGF concentration

The accuracy of serum EGF levels to discriminate between NSCLC patients and healthy controls (diagnostic value) determines its performance as a biomarker [33]. However, in NSCLC, the discriminatory capacity of serum EGF concentrations has been poorly studied and the results are inconclusive [16,34]. Similarly, there are conflicting reports on this issue in ovarian and colon cancers [35-38], and in conditions such as Alzheimer's disease, arteriosclerosis, and mood disorders; in which the involvement of EGF has been reported [39-43]. The main differences between the studies could be explained by the variability in the estimates of serum EGF concentrations, due to the lack of harmonization and even standardization of blood collection and sera separation procedures. In that sense, it is important to consider that platelets are the main reservoir and source of serum EGF [44]. Consequently, there is a correlation between serum EGF concentrations and platelet count [11,45], which is dependent on serum separation time [44,46].

The study by González-Pérez., *et al*. [11] identified one and four hours as the times relevant for serum separation after phlebotomy. The EGF concentration measured in serum collected 1h after phlebotomy ($[EGF]_{1h}$) is interpreted as a good estimator of the actual concentration of free EGF in the bloodstream. On the other hand, the EGF concentration measured at 4h ($[EGF]_{4h}$) is very close to the plateau achieved by the progressive release of EGF by platelets during their aggregation in the coagulation process, which is why it is considered the average total pool of EGF in the blood of an individual. These assumptions are consistent with a previous report by Savage., *et al*. [46]. Other factors, including the type of tube for blood collection and its coagulation, as well as the clotting temperature [47], also affect the kinetics of EGF release by platelets and the estimated levels of EGF in serum [11,12]. The lack of control over these factors, which greatly influence the serum concentration of EGF, has prevented drawing conclusions about its diagnostic capacity in several pathologies including NSCLC; which has been an obstacle to the broader consolidation of serum EGF as a biomarker. Thus, reliable assessment of the diagnostic value of serum concentration of EGF depends on a standardized quantification methodology that controls for the variety of factors that determine serum EGF values.

Another source of variability in EGF estimates is the SNP +61G>A located in the promoter region of the EGF gene [48], which affects EGF expression levels causing natural differences between individuals [49]. This polymorphism has been associated with cancer risk [48-50], its severity [51], prognosis [52], and response to therapy [53]. However, to the best of our knowledge, no previous studies have addressed its impact on the diagnostic capacity of serum EGF levels in NSCLC and its performance as a biomarker.

Potential predictors of CIMAvax-EGF® success

The application of targeted therapies and immunotherapies requires the assessment of accurate biomarkers. The search for such biomarkers for the indication of tailored treatments and better patient management is currently an area of active investigation in NSCLC. In the last five years, several potential biomarkers have been evaluated for the personalized indication of the CIMAvax-EGF® vaccine [12,27,54-56]. Table 1 compares them, taking as reference the absolute serum EGF concentration before treatment, the biomarker currently approved for the indication of the vaccine [26,27].

Reference	Study Type	Study subjects	Biomarker/ (Cut-off value)/ MOS advantage	Cut-off selection	Statistical significance of the prediction	Biomarker assay/ Analytical Complexity/ Affordability	EGF evaluation	Advantages	Disadvantages
Rodríguez., *et al*. [27]	Retrospective, exploratory	BM+ Patients 70 vaccinated/ 25 non-vaccinated	Serum EGF/ (≥ 870 pg/ mL)/ 6 months	Arbitrarily established (median EGF concentration of study population after 1st-line CRT, n = 188)	HR[a] = 0.41 p[a] = 0.0001 p[b]< 0.0001	ELISA Quantikine EGF[c] (R&D Systems Inc., USA)/ Low/Relatively cheap	After 1st-line CRT	The biomarker is related to EGF, the main target of CIMAvax-EGF® Potential biases due to the retrospective nature of the study Non-standardization of blood collection and serum separation The biomarker does not take into account natural interindividual differences in serum EGF levels Dichotomization of serum EGF values, with loss of information, reduction in power, and uncertainty in defining the cut-off point Estimation of survival advantage by *post-hoc* analysis, a procedure with known limitations Absence of serum EGF measurements in naïve patients

Saavedra., *et al*. [55]	Prospective, exploratory	BM+ Patients 13vaccinated/ 3 non-vaccinated	CD4+ T cells/ (> 40%)/ 34.1 months	Selected by *post-hoc* Cox´s regression	HR^a = 0.17 p^a = 0.01	FACS Flow cytometry/ High/Relatively expensive	None	The prospectiveness of the study Immunocompetence parameters are relevant to the individual´s immunity and response to immunotherapy	The biomarkers are not related to EGF, the main target of CIMAvax-EGF® vaccine. The bias inherent to *post-hoc* analysis Dichotomization of the biomarkers, with loss of information, reduction in power, and uncertainty in the definition of the cut-off point
		9 vaccinated/ 6 non-vaccinated	CD8+CD28- T cells/ (< 24%)/ 22.9 months		HR^a = 0.2 p^a = 0.03				
		7 vaccinated/ 4 non-vaccinated	CD4+/CD8+/ T cells ratio (>2)/ 36.1 months		HR^a = 0.16 p^a = 0.04				
González-Pérez., *et al*. [12]	Prospective, exploratory	Patients 25 (treatment-naïve) Healthy Controls 18	$[EGF]_{1h}$/ platelets/L/ (>0.63)/ ND ($[EGF]_{4h}$-$[EGF]_{1h}$)/ platelets/L/ (<1.80)/ ND	Selected by ROC analysis (Youden Index maximized by Se(%) and Sp(%))	AUC = 0.7389 p = 0.01308 AUC = 0.8875 p<0.0001	UMELISA EGF® (SUMA platform, Cuba)/ Low/Cheap	Basal (in treatment-naïve patients)	The prospectiveness of the study The relationship of biomarkers with EGF, the main target of CIMAvax-EGF® Standardization of sample collection and estimation of EGF and the use of an *IVD* kit for quantitation The evaluation of biomarkers in healthy controls, thus allowing the identification of confounding factors The evaluation of biomarkers in naïve patients, contributing to the knowledge of tumor biology The combination of variables and/ or their normalization, which produces biomarkers with improved diagnostic accuracy compared to absolute serum EGF concentrations The use of ROC analysis for the evaluation of the diagnostic accuracy of biomarkers and their comparisons The use of multiple criteria to select the best potential biomarker to predict the efficacy of CIMAvax-EGF®	The biomarkers were not evaluated in patients vaccinated with CIMAvax-EGF®, preventing their correlation with survival data associated with the use of this immunotherapy
		Patients 18/25 (after 1st-line CRT) Healthy Controls 18	$[EGF]_{1h}$/ platelets/L/ (>1.62)/ ND ($[EGF]_{4h}$-$[EGF]_{1h}$)/ platelets/L/ (<1.84)/ ND		AUC = 0.7431 p = 0.0071 AUC = 0.7059 p = 0.0349		After 1st-line CRT		

Luaces., *et al*. [54]	Retrospective, exploratory (data from [27])	Patients Good responders[d] 11 vaccinated/ 5 non-vaccinated Bad responders[e] 12 vaccinated/ 5 non-vaccinated	Multivariate predictors[f]/ (-)/ ND Serum EGF NLR Monocytes Neutrophils Proportion of CD4+ T cells	None	PCI = 0.74 $p^a = 0.033$ $p^a = 0.97$	FACS Flowcytometry/ High/Relatively Expensive ELISA Quantikine EGF[c]/Low/Relatively cheap	After 1st-line CRT	The use of the multivariate causal inference approach, which makes it possible to analyze the role of several biomarkers jointly, taking advantage of the continuous scale of measurements to identify a good predictive combination of vaccination success.	Potential biases due to the retrospective nature of the study Non-standardization of blood draw and serum separation for EGF measurements The measurement of multiple predictors, which could be a limitation for its application in clinical practice
Santos Morales., *et al*. [56]	Retrospective, exploratory (data from [27,60])	BM+ Patients 83 vaccinated/ 37 non-vaccinated	Serum EGF/ (≥ 870 pg/mL)/ 6.1 months	According to the phase III trial by Rodriguez., *et al*. [54]	$HR^a = 0.44$ $p^a = 0.001$ $p^b = 0.001$	ELISA Quantikine EGF[c]/ Medium/Relatively cheap	After 1st-line CRT	The biomarker is related to EGF, the main target of CIMAvax-EGF®	The retrospective nature of the study The disadvantages listed in this Table for the Rodriguez study The bias caused by minor differences in the designs of pooled phase II and III trials

Table 1

The predictive value of the serum EGF concentration for CIMAvax-EGF® was first described by Rodríguez., *et al*. [27,57]. The authors found a survival advantage in patients with pretreatment serum EGF concentration ≥870 pg/ml, suggesting different levels of dependence on serum EGF availability among the patients/tumors. However, the application of this classification threshold does not explain the benefits of vaccination in some patients, nor some treatment failures [27,57]. Such limitations of the biomarker can be partly explained by biases due to the retrospective selection of its cut-off point in the context of nonstandardized measurements [27]. Certainly, it has been acknowledged that the biomarker´s deficiencies could be related to its identification from serum samples not optimally separated [26], a very critical issue in precision medicine [58]. The selection of the biomarker by *post-hoc* survival analysis, and dichotomization, is according to Luaces., *et al*. [54] another source of bias that may affect its efficacy (see Table 1). In fact, *post-hoc* data analyses are considered a concern in statistical inference [59]. Nevertheless, possibly the most relevant aspect contributing to the biomarker failure is that the serum separation time, although not controlled in [27], was probably around four hours on average, given the similarity between the median concentration reported and that corresponding to $[EGF]_{4h}$ according to González-Pérez., *et al*. (873 pg/ml and 829 pg/ml, respectively) [12]. The closeness between these medians could partially explain the limitations in the efficacy of the biomarker, considering that in the last-mentioned study [12], performed under standardized conditions, healthy controls and treatment-naive patients had the same total EGF stock on average ($[EGF]_{4h}$) (Figure 1A). This result not only compromises the diagnostic capacity of 4h-concentrations and its performance as a biomarker [33] but also the relevance of the total stock of EGF in the pathophysiology of NSCLC. Another limiting aspect for the biomarker is that it does not take into account the interindividual variability in EGF concentrations, associated with different platelet counts, nor the functional SNP +61G>A, which undermines the classification of the absolute levels of EGF in serum as "low" or "high".

Recently, Santos., *et al*. [56] analyzed the combined data from phase II [60] and III [27] clinical trials with the CIMAvax-EGF® vaccine, to further assess whether pretreatment serum

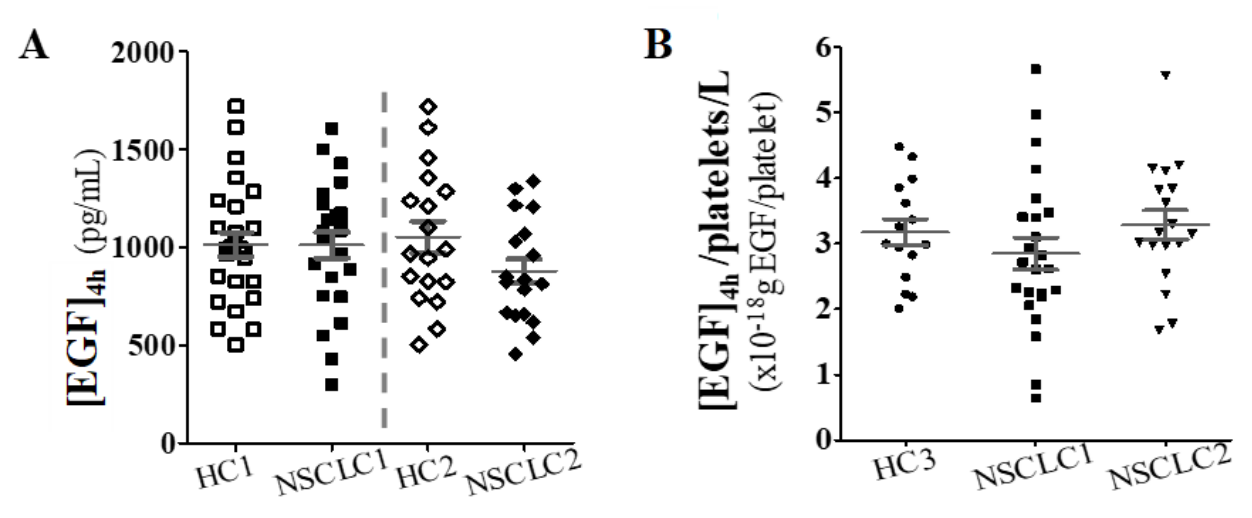

Figure 1: Serum EGF concentrations measured at 4h as absolute values or normalized by platelet count: patients vs. healthy controls. Panel (A) shows the absolute serum EGF concentrations of patients at diagnosis (NSCLC1) and after first-line chemoradiotherapy (NSCLC2), compared with those in age- and gender-matched healthy controls (HC1 and HC2, respectively). Panel (B) compares the same serum EGF concentrations, normalized by platelet count. The control group HC3 was selected based on the availability of platelet count.

EGF concentration is a predictive biomarker of a survival benefit in advanced NSCLC patients treated with the vaccine. As in Rodriguez´s phase III trial, patients were classified as biomarker positive (≥870 pg/mL) or negative (<870 pg/mL) based on pretreatment serum EGF concentration, and survival was analyzed using the Kaplan-Meier and log-rank tests. Median overall survival for vaccinated patients was 13.03 months compared with 6.93 months for non-vaccinated (hazard ratio HR=0.44, p=0.001), results quite similar to those of the Rodriguez study (14.66 months versus 8.63 months for controls, HR=0.41, p=0.0001) that included 25 fewer patients with a positive biomarker (13 vaccinated and 12 in the control arm). Therefore, with the same limitations as the phase III trial (Table 1) and limited scope due to bias resulting from minor differences in phase II and III trial designs, as acknowledged by the author [56], the analysis of pooled data did not provide stronger evidence in favor of pretreatment serum EGF concentration as a predictor of vaccine success.

Also looking for alternative predictors of CIMAvax-EGF® success, Saavedra., *et al*. studied several immunocompetence parameters in patients with advanced NSCLC and evaluated them as potential biomarkers of vaccine efficacy [55]. *Post-hoc Cox* regression analysis revealed that patients who had pre-vaccination frequencies of $CD8^+CD28^-$ T cells < 24%, CD4 T cells > 40%, and a CD4/CD8 ratio > 2, achieved at least a 20-month increase in the median OS compared to vaccinated controls who did not meet these criteria (Table 1). Previous studies in cancer patients have associated the presence of $CD8^+CD28^+$ T cells and $CD8^+CD28^-$ T cells with prolonged or shorter progression-free survival, respectively [61]. Recently, Liu., *et al*. published a similar result, but limited to specific LC histological types [62]. The authors showed that high levels of peripheral $CD8^+CD28^+$ T cells (considered those above the mean value of the analyzed cell subgroup) are a factor of favorable response in patients with advanced lung adenocarcinomas treated with chemoradiotherapy, while high levels of $CD8^+CD28^-$ T cells predict an unfavorable survival in patients with squamous cell carcinomas. In Saavedra´s work, although the peripheral lymphocytes evaluated must be relevant to the individual´s anti-tumor immunity and their cellular and humoral responses to CIMAvax-EGF®, their values are only associated with the state of immunosenescence of the patients, and the chronic inflammation

characteristic of the oncological disease, never to the dependence/independence of the tumor with respect to EGF. Therefore, patients with positive immunocompetence biomarkers and non-EGF-dependent tumors might not benefit– or even receive the most benefit– from the use of CIMAvax-EGF® or other EGF/EGFR-targeted therapies. On the other hand, in our opinion, the positivity of the biomarkers (immunocompetence parameters) before vaccination, instead of predicting the efficacy of the vaccine, could explain a good antibody response to vaccination (high antibody titers), whose magnitude is critical to target inhibition but insufficient for the success of CIMAvax-EGF®. As Popa and Crombet previously concluded [63], the magnitude of the anti-EGF antibody response alone does not determine treatment response or clinical benefit. In this sense, the quality of the interaction between these antibodies and circulating EGF, measured as EGFR phosphorylation inhibition, is predominant and is not associated with antibody titers. Neither has a correlation been reported between the antibody titers obtained and the inhibitory effect (response to treatment) for CIMAvax-EGF® in combination with nivolumab [31]. However, the weakest aspect of this study is probably that the serum EGF concentration was not estimated in the patients before treatment, although it is the proposed predictor of vaccine response and therefore a factor to control for when other possible independent predictors of response to CIMAvax-EGF® are univariately evaluated. The scope of the study is also limited by the known weaknesses of *post-hoc* analysis and dichotomization, previously acknowledged by Luaces., *et al*. [54].

In the Luaces., *et al*. [54] study, an interesting approach aimed at identifying biomarkers predictive of CIMAvax-EGF® efficacy, the authors retrospectively evaluated several potential biomarkers following the causal inference approach described by Abad., *et al*. in 2015 [64]. The analysis included the concentrations of serum EGF from Rodríguez´s phase III clinical trial [27], which were further evaluated in 2021 by Santos Morales., *et al*. [56]; some peripheral blood parameters such as monocytes, neutrophils, and eosinophils, among others; and the immunocompetence parameters previously studied by Saavedra., *et al*. in 2016 [55]. The study evaluated all possible combinations of predictors, revealing that peripheral blood parameters as the neutrophil-to-lymphocyte ratio, the monocytes and neutrophils counts, the percent of $CD4^+$ T cells, and the pretreatment serum EGF concentration are, together, a good predictor of CIMAvax-EGF® success in patients with advanced NSCLC. Mean predictive causal information, a measure of prediction accuracy defined as the correlation between treatment effect and predictors, increased from 0.486 with one predictor to 0.740 with all five predictors selected by the multivariate approach. Kaplan Meier survival estimation was carried out in good and poor responder patients. The log-rank-test analysis showed a significant survival advantage within biomarker-selected vaccinated patients (good responders), as shown in Table 1. The greatest achievement of this study is its multivariate approach compared to univariate techniques for biomarker validation. However, as accepted by the authors themselves, measuring and collecting data from multiple predictors increases the costs and burden of clinical investigations and routine follow-up, which could be a limitation for the clinical applicability of this approach. The lack of standardization of blood collection and separation of sera for EGF measurements [26,27] and possible biases due to the retrospective nature of the study could also be noted as limitations (Table 1).

Alternative EGF-related biomarkers

In González-Pérez´ proposal [12], serum EGF concentrations and platelet count were considered primary variables. EGF levels were measured in sera separated at 1h and 4h after phlebotomy ($[EGF]_{1h}$ and $[EGF]_{4h}$), using the UMELISA EGF® kit validated for *in vitro* diagnosis (IVD) [65]. In addition to these main variables, five others were constructed from their combination and/or normalization by platelet count. The ratio $r=[EGF]_{1h}/[EGF]_{4h}$ is interpreted as the fraction that represents the EGF in the bloodstream regarding its total stock, which also includes the EGF retained in platelets $d=[EGF]_{4h}-[EGF]_{1h}$ (not available to circulation). The variables normalized by platelets: $[EGF]_{1h}$/platelets/L, $[EGF]_{4h}$/platelets/L and d/platelets/L, are interpreted, respectively, as the average EGF contributed to circulation per platelet, the average total stock of EGF per platelet, and the average EGF retained per platelet (that is, not in circulation). Consistently with the lack of discriminatory capacity for $[EGF]_{4h}$ [12] (Figure 1A), Figure 1B shows that the variable cannot discriminate patients even when normalized by platelet count ($[EGF]_{4h}$/platelets/L), although this normalization corrects for variability in serum EGF concentration due to different platelet counts. This result suggests, once again, that the total EGF stock does not reflect the physiology of the tumor, a fact that largely justifies the limitations and failures in clinical practice of the serum EGF concentration measured at 4h ($[EGF]_{4h}$), the current predictor

of vaccine success. At the same time, González-Pérez´s approach evidenced the key role of circulating EGF in NSCLC disease and the potential of the new EGF-related variables to predict the success of therapies targeted to EGF/EGFR, including the CIMAvax-EGF® vaccine, whose mechanism of action directly involves EGF. This is relevant because, although biological understanding is not an absolute prerequisite for biomarker testing development, knowledge of the underlying biological pathways provides the rationale to support them [66]. The study did evidence, in this case, an increase in free/accessible EGF in the blood circulation of patients with NSCLC compared to healthy controls. Thereby, patients have more EGF in circulation ($[EGF]_{1h}$), a higher fraction of circulating EGF regarding its total stock ($r=[EGF]_{1h}/[EGF]_{4h}$), and consequently, a lower amount of EGF retained in platelets ($d=[EGF]_{4h}-[EGF]_{1h}$) (see [12]). Similarly, the EGF retained per platelet (d/platelets/L) is lower in these patients, while the EGF circulating per platelet ($[EGF]_{1h}$/platelets/L) is significantly higher, even after chemoradiotherapy, although treatment reduces circulating levels ($[EGF]_{1h}$) to the normal values present in healthy controls [12]. In contrast to the total EGF stock ($[EGF]_{4h}$), the increase of free EGF circulating in patients' blood supports its relevant role in the biology of NSCLC, most likely because it reflects the increased accessibility of this growth factor to tumor cells. The study also showed that the normalized variables $[EGF]_{1h}$/platelets/L and d/platelets/L provide similar selections of patients for treatment with CIMAvax-EGF® (see [12]), but that these selections could be different from those obtained using the median of the absolute concentrations of EGF in serum as a cut-off value, especially when using 4h measurements ($[EGF]_{4h}$), which could not discriminate cases from controls (Figure 1A). Therefore, although the serum EGF measured at 4h could predict, to some extent, the overall survival of patients after chemoradiotherapy and their responses to CIMAvax-EGF®, as Rodríguez., *et al*. evidenced [27], the normalized variables, discriminatory in this scenario, could be more valuable for these purposes than 4h concentrations $[EGF]_{4h}$.

The rationality of combining and normalizing serum EGF concentrations

Despite the advances in metabolomics, the discovery of differentially expressed tumor-associated proteins capable of reliably discriminating patients from healthy individuals remains a challenge today [66]. Until now, only a few proteins have been included in FDA-approved cancer diagnostics tests, most of which have low clinical sensitivity and specificity [3]. Furthermore, there is a growing scientific consensus on the superiority of marker panels to warrant the specificity and sensitivity that individual markers lack [67]. Particularly in multifactorial diseases such as cancer, often the approach that provides the expected discrimination and confidence is the combination of multiple "weak" markers in a single "strong" multivariate model [66]. Another approach with a similar purpose is the construction of biomarkers that combine different variables [55,62,68].

In González-Pérez´s study, the constructed markers combine two to three variables (circulating EGF, total EGF, and platelet count) into potential biomarkers, whose biological rationality contributes to their better performance in discriminating patients, compared with absolute serum EGF concentrations. This is most remarkable in the case of variables normalized by platelet count, which achieved the best discrimination accuracy in the study [12]. The constructed biomarkers allow estimating the amount of EGF retained in platelets ($d=[EGF]_{4h}-[EGF]_{1h}$) or on average per platelet ($([EGF]_{4h}-[EGF]_{1h})$/platelets/L), as well as how much EGF is mobilized regarding its total ($r=[EGF]_{1h}/[EGF]_{4h}$) and per platelet ($[EGF]_{1h}$/platelets/L). Thus, they reflect the kinetics of EGF in the individual, especially the normalized ones that could be useful to stratify NSCLC patients, infer the EGF dependence on the tumor, and predict the response to EGF/EGFR-targeted therapies, including CIMAvax-EGF®.

The better performance of the normalized variables is associated with the control they achieve over natural interindividual differences in serum EGF concentrations [48,49]. Normalization of circulating EGF by platelets takes into account the interindividual variability due to variations in the platelet count, while the EGF retained per platel*et al*so ponders the differences in the expressed EGF (total EGF) due to the +61G>A SNP. As a result, these variables are able to discriminate patients even after first-line chemoradiotherapy, although most responded to treatment by modifying their circulating levels to those typical of healthy individuals [12]. In patients with NSCLC, the impact of normalization by platelets is also associated with the apparently aberrant relationship between serum EGF concentrations and platelet count, as evidenced in the study by González-Pérez., *et al*. [12]. The correlation between these

variables, previously described in healthy individuals [11], is lost in naïve NSCLC patients probably due to the modification exerted by thrombocytosis, which significantly reduces circulating EGF levels per platelet [12]. The better discrimination achieved by the variables normalized by platelets proves the altered relationship between EGF and platelets in patients with NSCLC.

Methodological strengths and limitations of the new approach

The study by González-Pérez., *et al*. [12], with a sample size equal to 25, classifies as a pilot study for biomarker selection [66]. These exploratory studies are usually based on small representative samples that include 25 to 200 subjects [69,70] and very often only 25 [71-76]. The sample size necessary to achieve representativeness of the patient population in the sample is not a trivial matter. In biomarker discovery, it depends, among other factors, on the scope envisioned for the study, the variability of the sample, and the precision of the biomarker test [77]. Despite the small size of the sample evaluated in the study by González-Pérez., *et al*. [12], the minimum exclusion criteria imposed on the selection of the subjects must guarantee the representation of the characteristics of the population in the sample, in terms of biological variability [78]. The evaluated sample seems to be representative of the Cuban population with NSCLC when analyzed by its main demographic and clinicopathological characteristics. The mean age of the patients was 63 years and the male gender (80%) predominated, in agreement with other Cuban reports that show a mean age at diagnosis of lung cancer of 64 years [79] and a higher incidence in men [2,79]. Furthermore, 60% of the patients were diagnosed with stage IV disease, which is also consistent with the data from the Cuban National Cancer Registry, that reports a predominance of stage IV at diagnosis [79]. Moreover, the sample studied compares well with that of Rodríguez., *et al*. [27], made up of 400 patients, from 19 clinical centers and several provinces of the country. The only difference lies in the predominant stage of the disease at the time of diagnosis: stage III in the Rodríguez sample (62%) and IV in the González-Pérez study (60%).

The aforementioned biological rationality of the alternative biomarkers, together with some methodological advantages in their evaluation and selection, determined the best discrimination achieved. Advantages include standardization of blood collection and serum separation and the evaluation of a panel of healthy individuals [11]. This allowed the identification of the confounding factors that influence EGF measurements, as well as the choice of healthy controls, matched to patients by gender and age, thus contributing to the reliability of the data collected and the results obtained. The prospective character of the exploratory study and the evaluation of the study variables in treatment-naïve patients that show the biology of the disease in the untreated tumor [12], as well as the quantification of EGF using an IVD UMELISA [65], also add soundness and reliability to the new approach.

Additionally, the diagnostic accuracy of the alternative biomarkers was estimated using univariate nonparametric analysis of receiver operating characteristic (ROC) [80], widely considered the most objective and statistically valid test for this purpose [66,81] and for the comparison of biomarkers contrasting their respective ROC curves [82]. Furthermore, since in a ROC curve each point represents, in percentage values, the sensitivity (Se(%)) and specificity (Sp(%)) at a different cutoff point, it is possible to optimize the cut-off value by maximizing Se(%) and Sp(%) simultaneously [66,83], in contrast to the selection of thresholds from the dichotomization of variables. Therefore, the inclusion of the ROC analysis can be considered another methodological strength of the new approach.

Finally, the use of multiple criteria to select the best potential predictor of response [70] also represents a strong methodological advantage of the study by González-Pérez., *et al*. Only normalized variables discriminate patients before and after therapy, indicating their more direct relationship with tumor biology compared to absolute serum EGF levels and, consequently, their better performance in predicting EGF dependence and the response to CIMAvax-EGF®. Of these variables, d/platelets/L shows the highest partial area under the ROC curve (pAUC) [66,81] in the clinically relevant range--Se(%) and Sp(%) greater than 70--, both at diagnosis and after chemoradiotherapy. This performance, together with the robustness (absence of bias) of the variable d/platelets/L in the presence of thrombocytosis (see [12]), determined its selection as the best potential biomarker to predict the efficacy of CIMAvax-EGF®.

However, the proposed alternative biomarkers were not evaluated in patients vaccinated with CIMAvax-EGF®, preventing their correlation with survival data associated with the use of this immunotherapy (see Table 1). Their accuracies in predicting

response to CIMAvax-EGF® are therefore not estimated, but inferred from the respective diagnostic accuracies and other selection criteria discussed above. Although of note, the study provided preliminary evidence of an association between the proposed alternative biomarkers and response to chemoradiotherapy, also suggesting their potential utility in predicting vaccine efficacy (unpublished data). Nevertheless, the clinical utility of these alternative biomarkers in predicting response to therapies targeting the EGF/EGFR system, prognosis, therapy monitoring, and evaluation of treatment response needs to be demonstrated in clinical trials in patients treated with these therapies.

Scope of the new approach

The proposal of potential serum predictors of response to CIMAvax-EGF® is clinically relevant. Most of the efficacy biomarkers introduced into oncology clinical practice are genetic alterations that require biopsy material for their evaluation. Unlike clinical pathology techniques for the analysis of tissue biopsies, the UMELISA EGF® test [65] for the evaluation of the proposed biomarkers is simple, minimally invasive, and inexpensive. In our opinion, the new normalized variables, if clinically validated and based on their positivity as occurs with other targeted therapies, could support the indication and use of the CIMAvax-EGF® vaccine in the first line of NSCLC treatment. This is very significant as well. Immunization with CIMAvax-EGF® contrasts with treatments such as chemotherapy and radiotherapy, which are non-specific and aggressive and can compromise the immune status of patients. The use of this vaccine in the first line of treatment could enhance its effect, which depends on an effective anti-EGF immune response that could be affected in the second line setting by prior administration of chemoradiotherapy.

On the other hand, these alternative biomarkers could help to identify other EGF-dependent tumors, in which CIMAvax-EGF® or other EGF/EGFR-targeted therapies could be effective. Further studies should also focus on the value of proposed biomarkers in predicting response to chemoradiation or other neoadjuvant therapies in patients with tumors amenable to surgery. Likewise, the new approach could contribute to the study of the role of EGF in other pathologies in which this ligand has a key pathophysiological function. It could also contribute to the study of other growth factors such as VEGF and TGF beta that are also stored and released by platelets, with a similar impact in clinical practice. However, the greatest impact of the new variables could be associated with earlier detection of the malignant process. The simplicity of both, the UMELISA EGF® test for the quantification of biomarkers and the collection of blood samples, as well as their low costs, would allow the search for lung cancer and its detection in populations at risk. In other words, if clinically validated, the new variables could complement traditional radiology and other methods in use for NSCLC diagnosis and detection of disease progression, including LDCT, which remains the current gold standard for LC screening.

Conclusion

The standardization of sample collection and the normalization of serum EGF levels, made it possible to elucidate the diagnostic value of the serum EGF concentration in NSCLC, an issue that is not yet conclusive in ovarian and colon cancers, among others. The higher discriminatory capacity of the normalized biomarkers, compared to absolute serum EGF concentrations, suggests their closer relation to tumor biology and their better performance as biomarkers of the disease.

Acknowledgements

We thank Dr. Ernesto Galbán Rodríguez, from the Center for Genetic Engineering and Biotechnology, for his careful review of the first draft, and for his valuable comments and suggestions to improve the quality of our manuscript.

Conflict of Interest

The author IGP worked at the Center for Molecular Immunology (CIM, Havana, Cuba) from 2012 to 2017. The authors declare that the manuscript was conceived and written in the absence of commercial or financial relationships that could be interpreted as a potential conflict of interest.

Bibliography

1. Global Cancer Observatory. Global Cancer Observatory, (2020).

2. Statistician, Yearbook. “Cuba. Ministry of Public Health”. National Directorate of Medical Records and Health Statistics. Havana 2019. Havana: MINSAP; (2020).

3. Latini G., *et al.* “Clinical biomarkers for cancer recognition and prevention: a novel approach with optical measurements”. *Cancer Biomarkers* 22.2 (2018): 179-198.

4. Miller KD., *et al*. "Cancer treatment and survivorship statistics". *CA: A Cancer Journal for Clinicians* 72.5 (2022): 409-436.

5. Calabrese F., *et al*. "Are there new biomarkers in tissue and liquid biopsies for the early detection of non-small cell lung cancer?". *Journal of Clinical Medicine* 8.3 (2019): 414.

6. Rosell R and Karachaliou N. "Large-scale screening for somatic mutations in lung cancer". *The Lancet* 387.10026 (2016): 1354-1356.

7. Liang B., *et al*. "Predicting diagnostic gene biomarkers for non-small-cell lung cancer". *BioMed Research International* 2016 (2016).

8. Lam S. and Tammemagi M. "Contemporary issues in the implementation of lung cancer screening". *European Respiratory Review* 30.161 (2021).

9. Seijo LM., *et al*. "Cribado en cáncer de pulmón: últimas evidencias". *Archivos de bronconeumología* 56.1 (2020): 7-8.

10. Smith RA., *et al*. "Cancer screening in the United States, 2018: a review of current American Cancer Society guidelines and current issues in cancer screening". *CA: A Cancer Journal for Clinicians* 68.4 (2018): 297-316.

11. Idania GP., *et al*. "Measurement of serum EGF levels, a methodological approach: learning what means low-/high-concentration of EGF in serum. Some clinical implications". *Journal of Molecular Biomarkers and Diagnosis* 8.335 (2017): 2.

12. Idania GP., *et al*. "Normalized serum EGF levels as a potential biomarker in non-small cell lung cancer: the role of platelets". *Journal of Molecular Biomarkers and Diagnosis* 9.402 (2018): 2.

13. Chodon T., *et al*. "Active immunotherapy of cancer". *Immunological investigations* 44.8 (2015): 817-836.

14. Oliveres H., *et al*. "Vaccine and immune cell therapy in non-small cell lung cancer". *Journal of Thoracic Disease* 10.13 (2018): S1602.

15. Erol-Demirbilek M., *et al*. "EGF and TNF-α levels and oxidative/nitrosative stress in breast and non-small cell lung cancer patients". *Hacettepe Journal of Biology and Chemistry* 41.4 (2013): 357-363.

16. Blanco-Prieto S., *et al*. "Serum calprotectin, CD26 and EGF to establish a panel for the diagnosis of lung cancer". *PLoS One* 10.5 (2015): e0127318.

17. Huang P., *et al*. "The role of EGF-EGFR signalling pathway in hepatocellular carcinoma inflammatory microenvironment". *Journal of Cellular and Molecular Medicine* 18.2 (2014): 218230.

18. Goodison S., *et al*. "Bladder cancer detection and monitoring: assessment of urine-and blood-based marker tests". *Molecular Diagnosis and Therapy* 17.2 (2013): 71-84.

19. Li H., *et al*. "Aspirin prevents colorectal cancer by normalizing EGFR expression". *EBioMedicine* 2.5 (2015): 447-455.

20. Bracher A., *et al*. "Epidermal growth factor facilitates melanoma lymph node metastasis by influencing tumor lymphangiogenesis". *Journal of Investigative Dermatology* 133.1 (2013): 230238.

21. Grieco MJ., *et al*. "Minimally invasive colorectal resection is associated with a rapid and sustained decrease in plasma levels of epidermal growth factor (EGF) in the colon cancer setting". *Surgical Endoscopy* 24.10 (2010): 2617-2622.

22. Soulitzis N., *et al*. "Expression analysis of peptide growth factors VEGF, FGF2, TGFB1, EGF and IGF1 in prostate cancer and benign prostatic hyperplasia". *International Journal of Oncology* 29.2 (2006): 305-314.

23. Pudełek M., *et al*. "Epidermal Growth Factor (EGF) Augments the Invasive Potential of Human Glioblastoma Multiforme Cells via the Activation of Collaborative EGFR/ROS-Dependent Signaling". *International Journal of Molecular Sciences* 21.10 (2020): 3605.

24. Schelch K., *et al*. "EGF Induces Migration Independent of EMT or Invasion in A549 Lung Adenocarcinoma Cells". *Frontiers in Cell and Developmental Biology* 9 (2021): 634371.

25. Kuroi K and Toi M. "Circulating angiogenesis regulators in cancer patients". *The International Journal of Biological Markers* 16.1 (2001): 5-26.

26. Crombet T., *et al*. "CIMAvax EGF (EGF-P64K) vaccine for the treatment of non-small-cell lung cancer". *Expert Review of Vaccines* 14.10 (2015): 1303-1311.

27. Rodriguez PC., *et al*. "A phase III clinical trial of the epidermal growth factor vaccine CIMAvax-EGF as switch maintenance therapy in advanced non–small cell lung cancer patients". *Clinical Cancer Research* 22.15 (2016): 3782-3790.

28. Crombet T., *et al*. "The Position of EGF Deprivation in the Management of Advanced Non-Small Cell Lung Cancer". *Frontiers in Oncology* 11 (2021): 639745.

29. Garcí a B., *et al*. "Effective Inhibition of the Epidermal Growth Factor/Epidermal Growth Factor Receptor Binding by Anti–Epidermal Growth Factor Antibodies Is Related to Better Survival in Advanced Non–Small-Cell Lung Cancer Patients Treated with the Epidermal Growth Factor Cancer Vaccine". *Clinical Cancer Research* 14.3 (2008): 840-846.

30. Neninger E., *et al*. "Combining an EGF-based cancer vaccine with chemotherapy in advanced nonsmall cell lung cancer". *Journal of immunotherapy* 32.1 (2009): 92-99.

31. Evans R., *et al*. "Augmenting antibody response to EGF-depleting immunotherapy: Findings from a phase I trial of CIMAvax-EGF in combination with nivolumab in advanced stage NSCLC". *Frontiers in Oncology* 12 (2022): 958043.

32. Saavedra D., *et al*. "CIMAvax-EGF: Toward long-term survival of advanced NSCLC". *Seminars in Oncology* 45.1-2 (2018): 34-40.

33. Duffy MJ., *et al*. "Validation of new cancer biomarkers: a position statement from the European group on tumor markers". *Clinical Chemistry* 61.6 (2015): 809-820.

34. Lemos-Gonzalez Y., *et al*. "Alteration of the serum levels of the epidermal growth factor receptor and its ligands in patients with non-small cell lung cancer and head and neck carcinoma". *British Journal of Cancer* 96.10 (2007): 1569.

35. Baron AT., *et al*. "Serum sErbB1 and epidermal growth factor levels as tumor biomarkers in women with stage III or IV epithelial ovarian cancer". *Cancer Epidemiology, Biomarkers and Prevention: A publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 8.2 (1999): 129-137.

36. Balcan E., *et al*. "Serum levels of epidermal growth factor, transforming growth factor, and c-erbB2 in ovarian cancer". *International Journal of Gynecologic Cancer* 22.7 (2012): 1138-1142.

37. Masiak W., *et al*. "Evaluation of endostatin and EGF serum levels in patients with gastric cancer". *Polish Journal of Surgery* 83.1 (2011): 42-47.

38. Scho lch S., *et al*. "Serum PlGF and EGF are independent prognostic markers in non-metastatic colorectal cancer". *Scientific Reports* 9.1 (2019): 10921.

39. Lim NS., *et al*. "Plasma EGF and cognitive decline in Parkinson's disease and Alzheimer's disease". *Annals of Clinical and Translational Neurology* 3.5 (2016): 346-355.

40. Humpel C and Hochstrasser T. "Cerebrospinal fluid and blood biomarkers in Alzheimer's disease". *World Journal of Psychiatry* 1.1 (2011): 8.

41. Lundstam U., *et al*. "Epidermal growth factor levels are related to diastolic blood pressure and carotid artery stiffness". *Scandinavian Cardiovascular Journal* 41.5 (2007): 308-312.

42. Berrahmoune H., *et al*. "Association between EGF and lipid concentrations: A benefit role in the atherosclerotic process?". *Clinica Chimica Acta* 402.1-2 (2009): 196-198.

43. Skibinska M., *et al*. "Elevated Epidermal Growth Factor (EGF) as Candidate Biomarker of Mood Disorders—Longitudinal Study in Adolescent and Young Adult Patients". *Journal of Clinical Medicine* 10.18 (2021): 4064.

44. Oka Y and Orth D. "Human plasma epidermal growth factor/beta-urogastrone is associated with blood platelets". *The Journal of Clinical Investigation* 72.1 (1983): 249-259.

45. Lev-Ran A., *et al*. "Human serum and plasma have different sources of epidermal growth factor". *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 259.3 (1990): R545-R548.

46. Savage AP., *et al*. "Epidermal growth factor in blood". *Regulatory Peptides* 16.3-4 (1986): 199-206.

47. Yucel A., *et al*. "Effect of blood collection tube types on the measurement of human epidermal growth factor". *Journal of Immunoassay and Immunochemistry* 28.1 (2007): 47-60.

48. Shahbazi M., *et al*. "Association between functional polymorphism in EGF gene and malignant melanoma". *The Lancet* 359.9304 (2002): 397-401.

49. Pantsulaia I., *et al*. "Heritability of circulating growth factors involved in the angiogenesis in healthy human population". *Cytokine* 27.6 (2004): 152-158.

50. de Mello RA. "EGF+ 61 A> G "Polymorphisms and lung cancer risk: Future directions". *Journal of Pulmonary Medicine and Respiratory Research* 5 (2012): 2.

51. Kohla MAS., *et al*. "Association of serum levels of epidermal growth factor with disease severity in patients with unresectable hepatocellular carcinoma". *Hepatoma Research* 2 (2016): 18-25.

52. Yoshiya S., *et al*. "Impact of epidermal growth factor single-nucleotide polymorphism on recurrence of hepatocellular carcinoma after hepatectomy in patients with chronic hepatitis C virus infection". *Cancer Science* 105.6 (2014): 646-650.

53. Hu-Lieskovan S., *et al*. "EGF61 polymorphism predicts complete pathologic response to cetuximab-based chemoradiation independent of KRAS status in locally advanced rectal cancer patients". *Clinical Cancer Research* 17.15 (2011): 5161-5169.

54. Lorenzo-Luaces P., *et al*. "Identifying predictive biomarkers of CIMAvaxEGF success in non-small cell lung cancer patients". *BMC cancer* 20 (2020): 1-8.

55. Saavedra D., *et al*. "Biomarkers related to immunosenescence: relationships with therapy and survival in lung cancer patients". *Cancer Immunology, Immunotherapy* 65.1 (2016): 37-45.

56. Santos O., *et al*. "High Levels of Serum EGF Concentration Predict Response to Cimavax-EGF Treatment in Advanced NSCLC". *Journal of Cancer Science and Clinical Therapeutics* 5.1 (2021): 1-10.

57. Crombet T., *et al*. "EGF-based cancer vaccine: Optimizing predictive and surrogate biomarkers". *Journal of Clinical Oncology* 31.15 (2013): 3013-3013.

58. Seijo LM., *et al*. "Biomarkers in lung cancer screening: achievements, promises, and challenges". *Journal of Thoracic Oncology* 14.3 (2019): 343-357.

59. Curran-Everett D and Milgrom H. "Post-hoc data analysis: benefits and limitations". *Current opinion in Allergy and Clinical Immunology* 13.3 (2013): 223-224.

60. Vinageras EN., *et al*. "Phase II randomized controlled trial of an epidermal growth factor vaccine in advanced non-small-cell lung cancer". *Journal of Clinical Oncology* 26.9 (2008): 1452-1458.

61. Song G., *et al*. "Elevated level of peripheral CD8+ CD28− T lymphocytes are an independent predictor of progression-free survival in patients with metastatic breast cancer during the course of chemotherapy". *Cancer Immunology, Immunotherapy* 62 (2013): 1123-1130.

62. Liu C., *et al*. "Prognostic significance of peripheral CD8+ CD28+ and CD8+ CD28− T cells in advanced non-small cell lung cancer patients treated with chemo (radio) therapy". *Journal of Translational Medicine* 17 (2019): 1-11.

63. Popa X., *et al*. "Anti-EGF antibodies as surrogate biomarkers of clinical efficacy in stage IIIB/IV non-small-cell lung cancer patients treated with an optimized CIMAvax-EGF vaccination schedule". *Oncoimmunology* 9.1 (2020): 1762465.

64. Abad AA., Elst WVD and Molenberghs G. "Validating predictors of therapeutic success: a causal inference approach". *Statistical Modelling* 15.6 (2015): 619-636.

65. Castells EM., *et al*. "An enzyme immunoassay for determining epidermal growth factor (EGF) in human serum samples using an ultramicroanalytical system". *Journal of Immunoassay and Immunochemistry* 38.2 (2017): 190-201.

66. Xia J., *et al*. "Translational biomarker discovery in clinical metabolomics: an introductory tutorial". *Metabolomics* 9 (2013): 280-299.

67. Conrads TP., *et al*. "Cancer diagnosis using proteomic patterns". *Expert Review of Molecular Diagnostics* 3.4 (2003): 411-420.

68. Ramí rez JS., *et al*. "Evaluation of methodologies to determine the effect of specific active immunotherapy on VEGF levels in phase I clinical trial patients with advanced solid tumors". *Heliyon* 4.11 (2018).

69. Perez-Gracia JL., *et al*. "Strategies to design clinical studies to identify predictive biomarkers in cancer research". *Cancer Treatment Reviews* 53 (2017): 79-97.

70. Wallstrom G., *et al*. "Biomarker discovery for heterogeneous diseases". *Cancer epidemiology, biomarkers and prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 22.5 (2013): 747-755.

71. Marcis auskas S., *et al*. "Univariate and classification analysis reveals potential diagnostic biomarkers for early stage ovarian cancer Type 1 and Type 2". *Journal of Proteomics* 196 (2019): 57-68.

72. de Siqueira LT., *et al*. "A screening study of potential carcinogen biomarkers after surgical treatment of obesity". *Obesity Surgery* 28 (2018): 2487-2493.

73. Thanarajasingam G., *et al*. "Low Plasma Omega-3 Fatty Acid Levels May Predict Inferior Prognosis in Untreated Diffuse Large B-Cell Lymphoma: A New Modifiable Dietary Biomarker?". *Nutrition and Cancer* 70.7 (2018): 1088-1090.

74. Dunphy MP., *et al*. "In vivo PET assay of tumor glutamine flux and metabolism: in-human trial of 18F- (2 S, 4 R)-4-fluoroglutamine". *Radiology* 287.2 (2018): 667.

75. Yilmaz E., *et al*. "Neural precursor cell-expressed developmentally down-regulated 4-like: a new biomarker in the pathophysiology of endometrial cancer". *The Journal of International Medical Research* 46.9 (2018): 3709-3716.

76. McAnena P., *et al*. "Circulating microRNAs miR-331 and miR-195 differentiate local luminal a from metastatic breast cancer". *BMC Cancer* 19.1 (2019): 1-10.

77. Forshed J. "Experimental design in clinical 'omics biomarker discovery". *Journal of Proteome Research* 16.11 (2017): 3954-3960.

78. Oberg AL and Vitek O. "Statistical design of quantitative mass spectrometry-based proteomic experiments". *Journal of proteome research* 8.5 (2009): 2144-2156.

79. Gala n Y and Torres P. "Lung Cancer in Cuba". II Symposium on Lung Cancer, Dr. Miguel Betancourt Rodríguez in Memoriam" (2017).

80. DeLong ER., *et al*. "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach". *Biometrics* 44.3 (1988): 837-845.

81. Obuchowski NA. "ROC analysis". *American Journal of Roentgenology* 184.2 (2005): 364-372.

82. Obuchowski NA., *et al*. "ROC curves in clinical chemistry: uses, misuses, and possible solutions". *Clinical Chemistry* 50.7 (2004): 1118-1125.

83. Youden WJ. "Index for rating diagnostic tests". *Cancer* 3.1 (1950): 32-35.