PrediCare: a New Diagnostic Tool Predicting Imminent Disease Progression in Advanced NSCLC Patients by Machine-Learning Integration of Three Serum Tumor Biomarkers

Yuri Kogan[1], Marina Kleiman[1], Shmuel Shannon[1], Moran Elishmereni[1], Eldad Taub[1], Larisa Aptekar[2], Ronen Brenner[3], Raanan Berger[4], Hovav Nechushtan[2], Zvia Agur[1]

[1] Optimata Ltd. Israel, [2] Hadassah Medical Center, Israel, [3] Wolfson Medical Center, Israel, [4] Sheba Medical Center, Israel

Background

In advanced non-small cell lung cancer (NSCLC), most patients deteriorate rapidly and die within 1 year of diagnosis[1]. Forecasting disease progression just prior to its clinical manifestation would allow an earlier switch to the next treatment line, thus preventing major deterioration in the patient's stature and potentially improving response to therapy. However, present serum tumor markers, e.g. carcinoembryonic antigen (CEA), lack the power to signal progression[2,3].

We developed PrediCare, an innovative predictive tool for continuous monitoring and alerting to forthcoming disease progression in late-stage NSCLC.

Methods & Results

Data of late-stage NSCLC patients under 1st-line standard-of-care therapeutics were collected within a retrospective observational trial (NCT identifier NCT02577627);
A total of 167 patients were collected, the median follow-up time being 101 days; **Figure 1, Table 1**.

Table 1. Population characteristics and administered treatments.

Patients in the study dataset	n= 167
Age	
Median (range)	69 (38-87) yrs
Gender	
Male	103
Female	64
Smoking status	
Non smoker	53
Ex Smoker	71
Smoker	43
Treatments	
Pemetrexed platinum doublet	118
Taxane platinum doublet	13
Other regimens	36
Median follow-up time	101 days

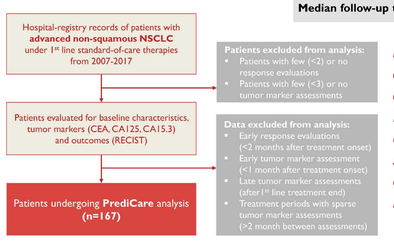

Figure 1. Overview of data collection and analysis. *Patient data were analyzed for inclusion of sufficient response evaluations and tumor marker assessments.*

Basic Marker (Value) Classifiers

To assess basic predictive potential of the tumor markers, **receiver operating characteristic (ROC) analysis** was applied to marker levels monitored during treatment.

ROC-derived classification trees were also tested for predictive accuracy. Cross-validation of classification trees was done using 80%/20% of the patients for training/testing; the threshold for the marker value was calibrated to reach 90% specificity for predicting progression. This was repeated 500 times (random sampling).

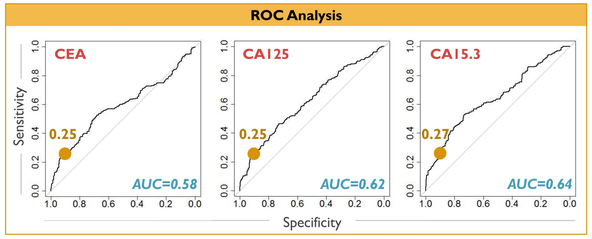

Figure 2. Weak association between tumor markers and progression. *ROC curves indicate the correlation between marker values and progressive disease (PD by Response Evaluation Criteria In Solid Tumors; RECIST 1.1) in NSCLC patients. Points corresponding to 90% specificity and respective sensitivity values are marked. AUC- area under curve.*

Each of the tumor markers had weak predictive ability on its own, and was incapable of predicting imminent progression, as shown by ROC curves **(Figure 2)** and simple classification trees **(Table 2)**.

	Classification trees			Reference
Markers	CEA	CA125	CA15.3	*No information*
Sensitivity	25.4%	25.9%	26.4%	22.1%
Specificity	89.9%	90.1%	90.1%	77.9%
Cohen's Kappa	0.177	0.185	0.192	0

Table 2. Performance of simple classifiers using tumor marker values. *Each classifier was derived from ROC-analysis using one marker. The results are shown in comparison to the reference index, specifying the performance of a classifier that uses no information.*

Single & Multiple-Marker PrediCare

PrediCare was designed to input, at any given time, one to three tumor markers and output a prediction of progression expected to occur within the following 1-2 months.

- ***Model:*** The algorithm was constructed by machine learning modeling on an R programming platform.
- ***Training:*** Features extracted from the tumor marker time course were tested as input, where output was progressive disease (PD) outcomes (y/n) by Response Evaluation Criteria In Solid Tumors (RECIST 1.1). After feature selection, training of trees/ regressions and stacking of best-performing models was done.
- ***Testing***: Cross-validation was performed (500 iterations; random sampling), where in each iteration 80%/20% patients were used for training/testing. Accuracy was evaluated in pooled test sets.

Tool	Single-marker PrediCare			**Multiple-marker PrediCare**
Markers	CEA	CA125	CA15.3	**CEA + CA125 + CA15.3**
Sensitivity	33.4%	34.3%	48.9%	**52.6%**
Specificity	90.1%	91.6%	91.3%	**91.1%**
Positive Predictive Value	80.3%	80.5%	83.8%	67.5%
Negative Predictive Value	53.1%	58.2%	66.1%	84.5%
Accuracy	76.0%	77.2%	80.4%	81.1%
Cohen's Kappa	0.269	0.301	0.440	**0.471**

Table 3. PrediCare performance metrics *in NSCLC patients using one, or three, tumor marker(s).*

The combination of all three tumor markers (CEA / CA125 / CA15.3) accurately predicted over half of the progression events, 1-2 months before they were clinically observed by RECIST (87/165 events; 52.6% sensitivity), with few false positives (15/165 events; 91.1% specificity); **Table 3**.

Merging features from three markers, PrediCare showed much better prediction capacity than any parallel single-marker algorithm **(Table 3)** or double-marker algorithms (data not shown).

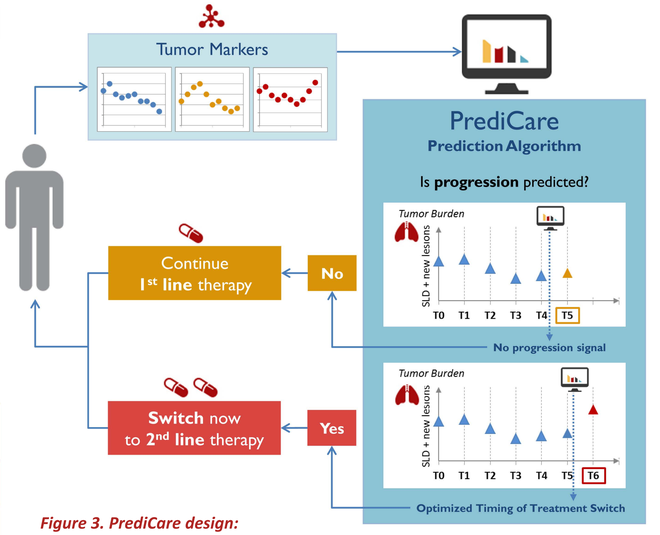

Figure 3. PrediCare design: *Optimizing the moment of therapy switch by predicting imminent progression.*

Conclusions

We used machine learning methods to create a new **individualized medicine software tool , which integrates weak signals from three tumor markers, monitored during treatment, into a combined strong signal for approaching disease progression**. This offers an improvement over the present clinical practice in advanced NSCLC, enabling to personally **optimize the timing of the switch** from 1st line treatment to 2nd line treatment, just prior to potential disease deterioration. Our PrediCare technology (Figure 3) uses easy-to-obtain tumor markers in a unique way that offers superiority over their current clinical interpretation. Testing of PrediCare under a larger marker panel is underway.

Website: www.optimata.com
Correspondence: moran@optimata.com

References: [1] Cetin K et al. Clin Epidemiol. 2011; 3: 139–148. [2] Accordino MK et al. J Oncol Pract 2016;12(1):65-6, e36-43. [3] Holdenrieder S et al. Biomed Res Int. 2016; 9795269.