The impact of a pathologist's personality on the interobserver variability and diagnostic accuracy of predictive PD-L1 immunohistochemistry in lung cancer

Butter, R.; Hondelink, L.M.; Elswijk, L. van; Blaauwgeers, J.L.G.; Bloemena, E.; Britstra, R.; ... ; Radonic, T.

Lung Cancer 166 (2022) 143–149

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

The impact of a pathologist's personality on the interobserver variability and diagnostic accuracy of predictive PD-L1 immunohistochemistry in lung cancer

Rogier Butter [a], Liesbeth M. Hondelink [b], Lisette van Elswijk [c], Johannes L.G. Blaauwgeers [d], Elisabeth Bloemena [e], Rieneke Britstra [e], Nicole Bulkmans [f], Anna Lena van Gulik [g], Kim Monkhorst [h], Mathilda J. de Rooij [i], Ivana Slavujevic-Letic [i], Vincent T.H.B.M. Smit [b], Ernst-Jan M. Speel [j], Erik Thunnissen [e], Jan H. von der Thüsen [k], Wim Timens [l], Marc J. van de Vijver [a], David C.Y. Yick [m], Aeilko H. Zwinderman [n], Danielle Cohen [b], Nils A. 't Hart [l,o], Teodora Radonic [a,e,*]

[a] Department of Pathology, Amsterdam University Medical Centers, Cancer Center Amsterdam, University of Amsterdam, The Netherlands
[b] Department of Pathology, Leiden University Medical Center, Leiden University, The Netherlands
[c] Independent Operating Psychologist, The Netherlands
[d] Department of Pathology, OLVG LAB BV, Amsterdam, The Netherlands
[e] Department of Pathology, Amsterdam University Medical Centers, Cancer Center Amsterdam, Vrije Universiteit Amsterdam, The Netherlands
[f] Department of Pathology, Spaarne Gasthuis, Haarlem, The Netherlands
[g] Department of Pathology, University Hospital Cologne, Germany
[h] Department of Pathology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands
[i] Department of Pathology, Symbiant Pathology Expert Center, Alkmaar, The Netherlands
[j] Department of Pathology, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, University of Maastricht, The Netherlands
[k] Department of Pathology, Erasmus University Medical Center, Erasmus University Rotterdam, The Netherlands
[l] Department of Pathology and Medical Biology, University Medical Center Groningen, University of Groningen, The Netherlands
[m] Department of Pathology, Amphia Hospital, Breda, The Netherlands
[n] Department of Clinical Epidemiology and Biostatistics, Amsterdam University Medical Centers, University of Amsterdam, The Netherlands
[o] Department of Pathology, Isala Hospitals, Zwolle, The Netherlands

ARTICLE INFO

Keywords:
Non-small cell lung cancer (NSCLC)
programmed death-ligand 1 (PD-L1)
Immunohistochemistry
Interobserver variability
Diagnostic reliability

ABSTRACT

Objectives: Programmed death-ligand 1 (PD-L1) is the only approved predictive biomarker for immunotherapy in non-small cell lung cancer (NSCLC). However, predictive PD-L1 immunohistochemistry is subject to interobserver variability. We hypothesized that a pathologist's personality influences the interobserver variability and diagnostic accuracy of PD-L1 immunoscoring.

Materials and Methods: Seventeen pathologists performed PD-L1 immunoscoring on 50 resected NSCLC tumors in three categories (<1%;1–49%;≥50%). Also, the pathologists completed a certified personality test (NEO-PI-r), assessing five personality traits: neuroticism, extraversion, openness, altruism and conscientiousness.

Results: The overall agreement among pathologists for a series of 47 tumors was substantial (kappa = 0.63). Of these, 23/47 (49%) tumors were entirely negative or largely positive, resulting in a kappa value of 0.93. The remaining 24/47 (51%) tumors had a PD-L1 score around the cutoff value, generating a kappa value of 0.32. Pathologists with high scores for conscientiousness (careful, diligent) had the least interobserver variability (r = 0.6, p = 0.009). Also, they showed a trend towards higher sensitivity (74% vs. 68%, p = 0.4), specificity (86% vs. 82%, p = 0.3) and percent agreement (83% vs. 79%, p = 0.3), although not significant. In contrast, pathologists with high scores for neuroticism (sensitive, anxious) had significantly lower specificity (80% vs. 87%, p = 0.03) and percent agreement (78% vs. 85%, p = 0.03). Also, a trend towards high interobserver variability (r = -0.3, p = 0.2) and lower sensitivity (68% vs. 74%, p = 0.3) was observed, although not significant.

* Corresponding author.
E-mail address: t.radonic@amsterdamumc.nl (T. Radonic).

https://doi.org/10.1016/j.lungcan.2022.03.002
Received 15 December 2021; Received in revised form 23 February 2022; Accepted 2 March 2022
Available online 5 March 2022

Pathologists with relatively high scores for conscientiousness scored fewer tumors PD-L1 positive at the ≥ 1% cut-off (r = −0.5, p = 0.03). In contrast, pathologists with relatively high scores for neuroticism score more tumors PD-L1 positive at ≥ 1% (r = 0.6, p = 0.017) and ≥ 50% cut-offs (r = 0.6, p = 0.009).
Conclusions: This study is the first to demonstrate the impact of a pathologist's personality on the interobserver variability and diagnostic accuracy of immunostaining, in the context of PD-L1 in NSCLC. Larger studies are needed for validation of these findings.

1. Introduction

The science of mobilizing the immune system to attack tumor cells, known as immunotherapy, has revolutionized cancer treatment. [1] In advanced non-small cell lung cancer (NSCLC) immunotherapy aimed at the interaction of tumor cells expressing programmed death-ligand 1 (PD-L1) and T-cell PD1 receptor has achieved spectacular clinical responses and has led to the expansion of the use of immunotherapy to benefit more patients. [2] Therefore, biomarkers which are capable to reliably predict the response to therapy are becoming increasingly important.

In NSCLC, a predictive biomarker for PD-L1 targeting immunotherapy is the expression of PD-L1 on tumor cells. [3] Up to date, PD-L1 expression is the only approved biomarker for treatment with pembrolizumab. [4] The assessment of PD-L1 expression is performed by using immunohistochemistry, and in NSCLC is scored as the percentage of PD-L1 expressing tumor cells out of all tumor cells in the sample. [4] NSCLC patients are stratified for immunotherapy either at a threshold of ≥ 1% PD-L1 positive tumor cells (in combination with chemotherapy) or ≥ 50% PD-L1 positive tumor cells (as monotherapy). [5–7] Patients with NSCLC tumors expressing <1% PD-L1 were until recently denied immunotherapy, but following the Keynote-189 trial, even PD-L1 negative patients were shown to benefit from combination chemotherapy/immunotherapy. [8].

As is the case in any laboratory test, PD-L1 immunohistochemistry has shown inter-assay and inter-laboratory variation in test interpretation. [9–11] Also, the protocol of estimating a percentage of PD-L1 expressing tumor cells by a pathologist inevitably results in variability among pathologists (i.e. interobserver variability). [12] Inter- and intraobserver variability remains a tenacious problem for the reliability and reproducibility of PD-L1 immunoscoring, as it may ultimately result in inconsistent patient selection for immunotherapy. [13] Measures to decrease interobserver variability have been implemented, such as mandatory training for the interpretation of PD-L1 staining. [14] As a correlation between personality and decision making has been recognized in many different settings, [15] we hypothesized that a pathologist's personality may influence their PD-L1 scoring profile, most likely in tumors around the 1% or 50% cutoff for PD-L1 positivity. Therefore, the objective of the present study was to investigate the impact of a pathologists' personality on the interobserver variability and diagnostic accuracy of PD-L1 immunoscoring in NSCLC.

2. Materials and methods

2.1. Study design

Seventeen pathologists who were trained and certified in PD-L1 immunoscoring and score PD-L1 on a daily basis in routine clinical practice, were asked to participate in the study. The pathologists were employed at teaching hospitals (n = 11) and community hospitals (n = 6) in the Netherlands (n = 16) and Germany (n = 1).

Participating pathologists were asked to score PD-L1 immunohistochemistry in 50 NSCLC tumors. For that purpose, a tissue microarray (TMA) was used, as described in a previously published study. [13].

Simultaneously with scoring the TMA, pathologists were asked to complete an accredited personality test (NEO-PI-r test) and to score the enjoyment they experienced in PD-L1 scoring (scale 0–10).

2.2. Case selection and TMA construction and PD-L1 immunohistochemistry

Histopathologically confirmed resected NSCLC samples were collected from the archives of the University Medical Center Groningen, the Netherlands. Tumors from 50 formalin-fixed and paraffin embedded (PPFE) tissue blocks were included in the TMA, each tumor represented by three 0.6 mm cores. One 3 μm thick FFPE section was stained using hematoxylin and eosin (HE) stain and one 3 μm thick section was used for PD-L1 immunohistochemistry.

Immunohistochemistry for PD-L1 was performed, using the monoclonal antibody 22C3 (Dako) as a validated laboratory-developed test (LDT) on a single tissue slide of the TMA. The HE stained slide and PD-L1 immunohistochemistry were digitized and uploaded using Philips PathXL software (PathXL, Ireland). The pathologists had access to the digitalized TMA slides using the corresponding software.

2.3. Scoring protocol PD-L1

The percentage of tumor cells with any linear membranous staining at any intensity was reported following the interpretation instruction of the 22C3 Dako pharmDx assay. The percentage of tumor cells expressing PD-L1 relative to all tumor cells contained in three cores per tumor was estimated. Scores were assigned in three categories: <1% PD-L1 expressing tumor cells (negative), 1–49% PD-L1 expressing tumor cells (intermediate) and ≥ 50% PD-L1 expressing tumor cells (high). The minimal cell count required for the PD-L1 score was 100 tumor cells, in agreement with the commercial package insert.

2.4. Personality test

The NEO-PI-r Big Five personality test, consisting of 240 multiple choice questions, was distributed to all pathologists. [16] For every question the pathologists had five answer options: 'totally disagree', 'disagree', 'neutral', 'agree' or 'totally agree'.

The 240 questions scored five personality domains: Neuroticism, Extraversion, Openness, Altruism and Conscientiousness (Table 1), with 48 questions per domain. Missing answers were regarded as 'neutral' and questionnaires were considered invalid if >40 questions were missing, according to the test's instructions.

Pathologists were also asked to score the enjoyment they experience while scoring PD-L1 (range 1–10).

2.5. Outcome measures and statistical analysis

The interobserver agreement was expressed as Cohen's kappa, which ranges from 0 to 1. A kappa value of 0–0.2 indicates slight agreement (i.

Table 1
The Big Five personality domains, which are scored using the NEO-PI-r Big Five personality test.

Personality domain	Characteristic low score	Characteristic high score
Neuroticism	Resilient, confident	Sensitive, anxious
Extraversion	Solitary, reserved	Outgoing, energetic
Openness	Consistent, cautious	Inventive, curious
Altruism	Challenging, callous	Friendly, compassionate
Conscientiousness	Extravagant, careless	Systematic, diligent

e. close to flipping a coin), 0.2–0.4 fair agreement, 0.4–0.6 moderate agreement, 0.6–0.8 substantial agreement and 0.8–1 indicates almost perfect agreement, according to the methodology proposed by Landis and Koch. [17].

The diagnostic accuracy of the pathologists was expressed in sensitivity, specificity and percent agreement, according to a method proposed by Marchevsky et al. [18] To calculate the diagnostic accuracy, a majority score for each tumor was determined, which was defined as the category (<1%, 1–49% or ≥ 50%) in which the majority of the pathologists score a tumor.

Pearson's correlation coefficient, multivariate linear regression, and an independent *t*-test were used for statistical analysis of the normally distributed variables in this study. All calculations were performed in IBM SPSS Statistics 26. Differences were considered statistically significant at a p-value of < 0.05.

3. Results

3.1. Participating pathologists and data collection

All 17 pathologists completed the NEO-PI-r personality test, which met the criteria for a valid test (Table 2). In total, three (out of 4080) questions were left unanswered, by two pathologists.

PD-L1 immunoscoring was completed by all pathologists. In three tumors, the tumor cell count was considered insufficient (⟨100) for PD-L1 immunoscoring by the majority (≥9). These cases were excluded, resulting in the inclusion of 47 tumors for analysis.

3.2. PD-L1 immunoscores by the pathologists

Based on the majority score, PD-L1 was negative in 24/47 (51%) tumors, intermediate (1–49%) in 11/47 (23%) tumors and high (≥50%) in 12/47 (26%) tumors (Fig. 1).

In 24/47 tumors (51%), the PD-L1 immunoscore was scored differently by ≥ 2 pathologists, and these were deemed "informative tumors" for interobserver variability. In the remaining 23/47 (49%) tumors, most (16/17 or 17/17) pathologists rendered an identical PD-L1 score, and these are referred to as "non-informative" hereafter. "Non-informative" tumors were either completely PD-L1 negative or largely positive.

3.3. Overall interobserver variability

The overall agreement between the pathologists in all tumors (n = 47) was substantial, with a kappa value of 0.63 (95% CI 0.61–0.65). In the "informative tumors" (n = 24) the interobserver agreement was fair, with a kappa value of 0.32 (95% CI 0.29–0.35) and in the "non-informative" tumors (n = 23) excellent, with a kappa value of 0.93 (95% CI 0.90–0.97).

3.4. Conscientious pathologists have the least interobserver variability

Between each pair of pathologists, the kappa value was calculated from the scores of the "informative tumors" (n = 24), resulting in 16 kappa values per pathologist. The mean kappa value per pathologist was calculated from the 16 kappa values.

Table 2
Basic characteristics of participating pathologists.

Total (n)	17
Female (n)	7
Median Age (y)	49 (34–68)
Median Joy in immunoscoring (1–10)	6.6 (3–9)
Teaching hospital (n)	11
Community hospital (n)	6

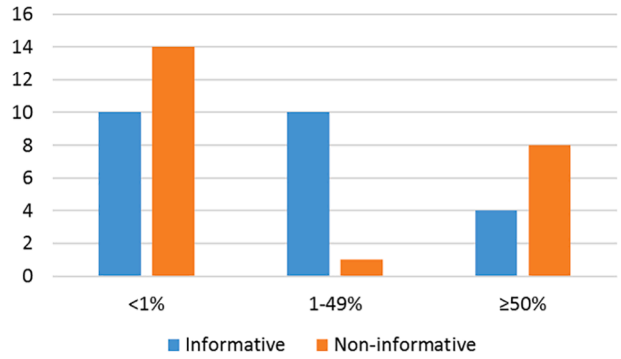

Fig. 1. The distribution of "informative tumors" ($n = 24$) and "non-informative tumors" ($n = 23$) over the three PD-L1 scoring categories (<1%, 1–49% and ≥ 50%).

A high score for conscientiousness correlated with higher mean kappa values of the pathologists (i.e. higher interobserver agreement), r = 0.61, p = 0.009 (Fig. 2). In contrast, higher scores for neuroticism showed a trend towards lower mean kappa values (i.e. lower interobserver agreement), although not statistically significant (r = -0.35, p = 0.17). For other personality traits no correlation with kappa values was found.

3.5. Diagnostic accuracy is highest in more conscientious pathologists

The diagnostic accuracy in informative tumors (n = 24) was highest in pathologists who have a higher than average score for conscientiousness (n = 8) and in pathologists with a lower than average score for neuroticism (n = 9) (Table 3). Despite the observed trend for differences in diagnostic accuracy between pathologists, few outcomes were statistically significant: the specificity and percent agreement of pathologists with high scores for neuroticism was significantly lower than that of pathologists with low scores for neuroticism (independent of the PD-L1 scoring category: sensitivity 74% ±11 vs. 68% ±15, p = 0.3; specificity 87% ±5 vs. 80% ±7, p = 0.03; percent agreement 85 ± 6 vs. 78% ±9, p = 0.03).

Abbreviations: SENS = Sensitivity, SPEC = Specificity, PA = Percent Agreement and SD = Standard Deviation.

3.6. More PD-L1 positive tumors in pathologists with high neuroticism scores

A positive correlation was found between the score of neuroticism and the number of PD-L1 positive tumors, for both the cut-off value of ≥ 1% (r = 0.6, p = 0.017) and ≥ 50% PD-L1 positive tumors (r = 0.6, p = 0.009) (Fig. 3). Thus, pathologists with high scores for neuroticism scored more tumors PD-L1 positive. Conversely, a negative correlation was found between the number of PD-L1 positive tumors and the score of conscientiousness, statistically significant at the cut-off value of ≥ 1% PD-L1 positive tumor cells (r = −0.5, p = 0.03), but not significant at the ≥ 50% cut-off value (r = −0.4, p = 0.1). Thus, pathologists with high scores for conscientiousness seem to score fewer tumors PD-L1 positive. The remaining personality traits were not correlated to the number of PD-L1 positive tumors.

Including the variables neuroticism, extraversion, openness, altruism, conscientiousness, enjoyment of scoring and age in a multivariate linear regression model, neuroticism remained significantly correlated to the number of PD-L1 positive tumors at both cut-off values.

3.7. Scores for neuroticism and conscientiousness are correlated

Pathologists with high scores for neuroticism have significantly

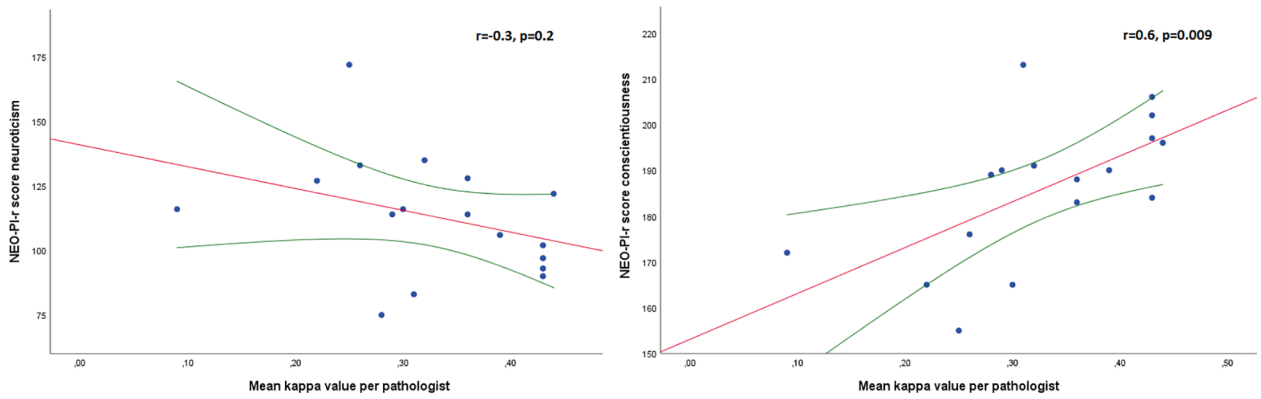

Fig. 2. Correlation between the NEO-PI-r score for conscientiousness (left) and neuroticism (right) and the mean kappa value per pathologist.

Table 3
The diagnostic accuracy of pathologists (n = 17) who score below or above the median value for neuroticism and conscientiousness. The analysis was performed in the "informative tumors" (n = 24).

Diagnostic Accuracy	Conscientiousness			Neuroticism		
	Below average (n = 9)	Above average (n = 8)	p-value	Below average (n = 9)	Above average (n = 8)	p-value
<1%						
SENS	64% ±27	81% ±12	0.12	83% ±10	60% ±27	0.026
SPEC	85% ±13	81% ±17	0.58	82% ±12	84% ±17	0.85
PA	76% ±10	80% ±10	0.42	83% ±7	73% ±10	0.044
1–49%						
SENS	73% ±14	75% ±22	0.81	75% ±19	73% ±17	0.77
SPEC	66% ±19	79% ±15	0.15	81% ±15	62% ±17	0.03
PA	69% ±15	77% ±11	0.23	79% ±10	66% ±14	0.05
≥50%						
SENS	67% ±32	66% ±32	0.95	63% ±32	70 ± 31	0.66
SPEC	96% ±6	97% ±6	0.67	98% ±4	94% ±7	0.23
PA	91 ± 4	92% ±5	0.67	93% ±4	91% ±6	0.34
All categories						
SENS	68% ±15	74% ±11	0.38	74% ±11	68% ±15	0.34
SPEC	82% ±7	86% ±6	0.28	87% ±5	80% ±7	0.025
PA	79% ±9	83% ±7	0.27	85% ±6	78% ±9	0.03

lower scores for conscientiousness (r = -0.7, p = 0.002). In addition, pathologists with high scores for extraversion also had high scores for openness (r = 0.68, p = 0.003). For the other personality domains, no correlations were found. The enjoyment of immunoscoring for PD-L1 or the pathologist's age were not correlated to the personality traits, interobserver variability, or diagnostic accuracy (Supplementary Table 3).

4. Discussion

This study demonstrated that a pathologist's personality is related to the interobserver variability, diagnostic accuracy and tendency of PD-L1 scoring in NSCLC. The least interobserver variability and the highest diagnostic accuracy, two important diagnostic features in pathology, were observed in pathologists who have high scores for conscientiousness. In contrast, more interobserver variability and lower diagnostic accuracy were observed in pathologists who have high scores for neuroticism. In addition, pathologists with high scores for conscientiousness assign fewer tumors as PD-L1 positive, while pathologists who have high scores for neuroticism assign more tumors as PD-L1 positive.

No literature, to our knowledge, is available about personality assessment and the performance of pathologists. However, in other fields of medicine some research on personality styles and medical decision making has been performed. In a Chinese study, elderly patients with acute myeloid leukemia (AML) were more likely to receive relatively intensive chemotherapy by attending physicians with high scores of conscientiousness or extraversion. [19] Also in elderly AML patients, French hematologists who have a relative aversion towards risk tend to choose more intensive chemotherapy. [20] A retrospective study at an intensive care and high dependency unit found that end-of-life decisions are more frequently made by medical or surgical consultants who score more towards a 'judging' (i.e. structured/decided) than a 'perceiving' (i.e. flexible/adaptable) personality. [21] A systematic review conducted in 2020 sought to include studies in which the personality of surgeons was assessed, and found two studies reporting on perioperative decision making and/or postoperative outcomes. [22] High scores for conscientiousness and openness affected decision making with respect to the choice between making an anastomosis or a stoma. [23] In a real-life setting, surgeons with high levels of 'constructive' and 'passive/defensive' personality styles had fewer adverse events in bariatric surgery than surgeons with other styles. [24] Lastly, a review from 2016 covering a wide range of medical professionals and different settings of decision making, found that some cognitive biases (i.e. the processing of observations in an irrational way or having a 'subjective reality') are associated with diagnostic inaccuracies, and 'therapeutic and management errors': anchoring-, availability-, confirmation-, information-, overconfidence-, premature closure- and representativeness bias. [25] All abovementioned studies have in common with our results that

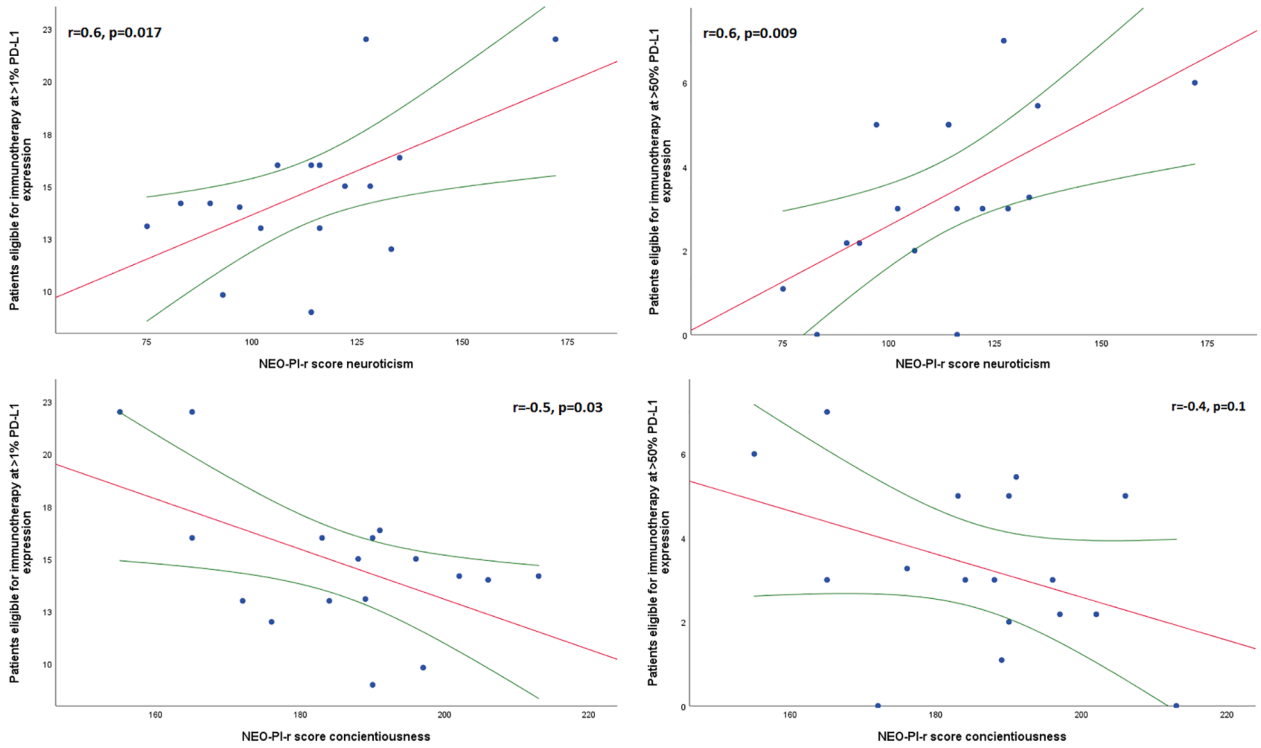

Fig. 3. Correlation between the NEO-PI-r score of neuroticism and the number of patients eligible for immunotherapy at a cut-off of ≥ 1% (top left) and ≥ 50%. (top right), and the correlation between the NEO-PI-r score of conscientiousness and the number of patients eligible for immunotherapy at a cut-off of ≥ 1% (bottom left) and ≥ 50% (bottom right).

personality styles indeed affect medical decision making. And more specifically, a conscientious/judging/constructive personality style was identified to affect medical decision making in multiple studies in a similar pattern as in our study: consistent decision making.

High scores for conscientiousness are related to characteristics such as being disciplined, systematic, diligent and well-organized. [26] Individuals who have high scores for conscientiousness tend to make more weighted decisions. [27] Moreover, multiple psychological studies have shown the relation between high scores of conscientiousness and a rational decision making style. [28–30] In particular, a large study by Weller et al. found that high scores for conscientiousness in 804 participants from an Italian community, are associated with improved decision making competence, which is a measure to quantify rational decision making. [31] This might be explained by a maximizing decision making style in individuals with high conscientiousness, which aims to make the most optimal decision after careful evaluation of all alternatives. [32] In the context of immunoscoring for PD-L1, the properties associated with conscientiousness clarify the results of our study. Pathologists who have high scores for conscientiousness possibly have a more diligent and systematic approach for scoring PD-L1, and therefore are more likely to score in similar categories, also resulting in a higher diagnostic accuracy. The diagnostic accuracy of PD-L1 scoring in this study is based on the majority score of all pathologists. The favorable diagnostic accuracy in pathologists who have higher scores for conscientiousness reflects the agreeableness among all the pathologists participating in this study.

Neuroticism, in contrast, is associated with a less rational style of decision making. [15,30] A study by Denburg et al., associated neurotic personality traits with poor decision making in older adults (≥65 years) from a random population, possibly due to the influence of strong emotions on cognitive performance. [33] Individuals with high scores for neuroticism tend to be more impulsive and anxious. In our study, tumors were scored in higher PD-L1 categories by pathologists with higher scores for neuroticism, while the diagnostic accuracy was lower. Hypothetically, a pathologist with higher scores for neuroticism might not be comfortable withholding immunotherapy in patients with a tumor around the threshold of PD-L1 positivity. However, it is difficult to separate the influence of neuroticism and conscientiousness on PD-L1 immunoscoring, as these two are known to be negatively correlated [34], which was also found in this study (supplementary material). This issue may be elucidated in larger future studies using other decision-making assessment tools.

Extraversion, openness and agreeableness were described to be associated with minimizing, maximizing, and satisfying decision making styles, [32] but were not associated with PD-L1 scoring in this study. No correlation was observed between age and joy of immunoscoring on one hand, and personality traits, interobserver variability or diagnostic accuracy on the other. One might expect that more joyful pathologists put more effort into scoring for PD-L1, but this is ultimately not related to personality style and the reliability of immunoscoring.

The participating pathologists in this study had an overall interobserver variability (kappa 0.63), independent of personality traits, which is comparable with results available in the literature. Three studies reported overall kappa values between 0.58 and 0.77, using different numbers of scoring categories and observers. [9,12,35] The largest validation study to date found interobserver kappa values among 18 pathologists ranging from 0.6 to 0.9 in 81 NSCLC tumors. [36] To the best of our knowledge, this current study is the first to evaluate PD-L1 immunoscoring using sensitivity and specificity as proposed by Marchevsky and colleagues. [18].

Immunotherapy is being applied in an increasingly wide range of NSCLC patients; based on recent clinical trials, nearly all patients are eligible for immunotherapy irrespective of PD-L1 status, in combination with chemotherapy or as monotherapy. [37–39] Nonetheless, reliable and consistent PD-L1 immunoscoring remains valuable as predictive biomarker for clinical response to immunotherapy (although not decisive for its application) and for the fundamental scientific understanding of immunotherapy efficacy. The results of this study indicate that the

general diagnostic accuracy and diagnostic style might also be affected by the personality style of pathologists. Therefore, the results of this study can serve as a starting point for further, broader studies on this topic.

TMAs are used in surgical pathology research as surrogates for resection specimens, and it has been previously described that PD-L1 immunoscores assessed in TMAs might not be representative for the whole resection [40,41]. However, this does not seem to affect our results, as no correlation with clinical outcomes was made and all pathologists assessed the same TMA section. A further limitation of our study is the relatively small sample size of pathologists for correlation with personality characteristics. It cannot be ruled out that other personality traits might affect the interobserver variability and diagnostic accuracy in a larger sample size. Also, selected tumors for the study consisted of a relatively higher number of PD-L1 negative tumors than reported in the literature [42]. Uniformly negative tumors provide little information about the scoring behavior of pathologists, as demonstrated in our study. In addition, the inclusion of five personality traits to investigate the correlation with PD-L1 immunoscores may lead to a degree of multiple testing bias. Finally, the most relevant outcomes in predictive testing are patient response to therapy and quality of life, both of which are not investigated in this study.

The results of our study may suggest that tumors with a PD-L1 value around the cut-off for positivity should be examined independently by two pathologists, and should be discussed if discordant. In reality, few centers will select or screen their pathologists for personality style. However, a pathologist who scores a tumor around the cut-off value of PD-L1, should be aware of his or her intrinsic motivation and argumentation for the final decision. Our study is the first on this topic in pathology, but it is likely that once validated, personality traits may also prove to have influence in other fields of pathology.

5. Conclusions

This study is the first to investigate the impact of a pathologist's personality style on interobserver variability, diagnostic accuracy and trend of immunoscoring for PD-L1. This topic has been left largely untouched in the medical literature so far and may serve as a starting point for similar studies in various fields of diagnostic or predictive testing performed by humans.

NSCLC tumors demonstrate PD-L1 expression around the clinically/scientific relevant thresholds in approximately 50% of the tumors, resulting in inconsistent immunoscores among pathologists. Personality styles of pathologists influence how these tumors are ultimately scored, and this may affect the predictive value of the clinical response to PD-L1 immunoscoring for immunotherapy in NSCLC, both in clinical practice and in research.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research. Research was performed following a waiver of Medical ethical approval, in line with the Code of Good Conduct that is followed in the Netherlands for retrospective studies using archival tissue samples. All participants (i.e. pathologists) agreed to participate in the current study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Karen Hilverda, medical psychologist at the Amsterdam University Medical Centers, for support in using the NEO-PI-r personality test.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lungcan.2022.03.002.

References

[1] M.F. Sanmamed, L. Chen, A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization, Cell. 175 (2) (2018) 313–326.
[2] D.B. Doroshow, M.G. Sanmamed, K. Hastings, K. Politi, D.L. Rimm, L. Chen, et al., Immunotherapy in non-small cell lung cancer: Facts and hopes, Clin. Cancer Res. 25 (15) (2019) 4592–4602.
[3] R. Brody, Y. Zhang, M. Ballas, M.K. Siddiqui, P. Gupta, C. Barker, et al., PD-L1 expression in advanced NSCLC: Insights into risk stratification and treatment selection from a systematic literature review, Lung Cancer. 112 (May) (2017) 200–215.
[4] C. Roach, N. Zhang, E. Corigliano, M. Jansson, G. Toland, G. Ponto, et al., Development of a Companion Diagnostic PD-L1 Immunohistochemistry Assay for Pembrolizumab Therapy in Non-Small-cell Lung Cancer, Appl Immunohistochem Mol Morphol. 24 (6) (2016) 392–397.
[5] R.S. Herbst, P. Baas, D.W. Kim, E. Felip, J.L. Pérez-Gracia, J.Y. Han, et al., Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): A randomised controlled trial, Lancet. 387 (10027) (2016) 1540–1550.
[6] M. Reck, D. Rodríguez-Abreu, A.G. Robinson, R. Hui, T. Csőszi, A. Fülöp, et al., Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer, N Engl J Med. 375 (19) (2016) 1823–1833.
[7] T.S.K. Mok, Y.L. Wu, I. Kudaba, D.M. Kowalski, B.C. Cho, H.Z. Turna, et al., Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial, Lancet. 393 (10183) (2019) 1819–1830.
[8] S. Gadgeel, D. Rodr, rapid communications abstract Updated Analysis From KEYNOTE-189: Pembrolizumab or Placebo Plus Pemetrexed and Platinum for Previously Untreated Metastatic Nonsquamous Non – Small-Cell, Lung Cancer. 38 (14) (2020).
[9] A.H. Scheel, M. Dietel, L.C. Heukamp, K. Jöhrens, T. Kirchner, S. Reu, et al., Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas, Mod Pathol. 29 (2016) 1165–1172.
[10] D.L. Rimm, G. Han, J.M. Taube, E.S. Yi, J.A. Bridge, D.B. Flieder, et al., A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non–small cell lung cancer, JAMA Oncol. 3 (8) (2017) 1051–1058.
[11] A.H. Scheel, G. Baenfer, G. Baretton, M. Dietel, R. Diezko, T. Henkel, et al., Interlaboratory concordance of PD-L1 immunohistochemistry for non-small-cell lung cancer, Histopathology. 72 (2017) 449–459.
[12] W.A. Cooper, P.A. Russell, M. Cherian, E.E. Duhig, D. Godbolt, P.J. Jessup, et al., Intra- and Interobserver Reproducibility Assessment of PD-L1 Biomarker in Non – Small Cell Lung Cancer, Clin Cancer Res. 23 (16) (2017) 4569–4577.
[13] R. Butter, Hart't NA, Hooijer GKJ, Monkhorst K, Speel EJ, Theunissen P,, et al., Multicentre study on the consistency of PD-L1 immunohistochemistry as predictive test for immunotherapy in non-small cell lung cancer, J Clin Pathol. 73 (7) (2020) 423–430.
[14] S. Lantuejoul, M. Sound-Tsao, W.A. Cooper, N. Girard, F.R. Hirsch, A.C. Roden, et al., PD-L1 Testing for Lung Cancer in 2019: Perspective From the IASLC Pathology Committee, J Thorac Oncol. 15 (4) (2020) 499–519.
[15] D.S. Jalajas, R. Pullaro, The effect of personality on decision making, J Organ Psychol. 18 (5) (2018) 66–72.
[16] P.T. Costa, R.R. McCrae, Revised NEO Personality lnventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual, Psychological Assessment Resources, Inc., Odessa, FL, 1992.
[17] J.R. Landis, G.G. Koch, The Measurement of Observer Agreement for Categorical Data, Biometrics. 33 (1) (1977) 159.
[18] Marchevsky AM, Walts AE, Lissenberg-Witte BI, Thunnissen E. Pathologists should probably forget about kappa. Percent agreement, diagnostic specificity and related metrics provide more clinically applicable measures of interobserver variability. Ann Diagn Pathol. 2020;47:151561.
[19] X. Wu, Y. Jiang, nan, Zhang Y lun, Chen J, Mao Y ying, Zhang L,, et al., Impact of Physicians' Personalities and Behavioral Traits on Treatment-Related Decision-making for Elderly Acute Myeloid Leukemia, J Gen Intern Med. 36 (10) (2021) 3023–3030.
[20] P. Bories, S. Lamy, C. Simand, S. Bertoli, C. Delpierre, S. Malak, et al., Physician uncertainty aversion impacts medical decision making for older patients with acute myeloid leukemia: Results of a national survey, Haematologica. 103 (12) (2018) 2040–2048.
[21] B. Poulton, S. Ridley, R. Mackenzie-Ross, S. Rizvi, Variation in end-of-life decision making between critical care consultants, Anaestesia. 60 (2005) 1101–1105.
[22] C.N. Bisset, T. McKee, M. Cawley, E. Tilling, S.J. Moug, A systematic review of the abdominal surgeon's personality: Exploring common traits in western populations, Behav Sci (Basel) 11 (2) (2021).
[23] S.J. Moug, N. Henderson, J. Tiernan, C.N. Bisset, E. Ferguson, D. Harji, et al., The colorectal surgeon's personality may influence the rectal anastomotic decision, Color Dis. 20 (11) (2018) 970–980.

[24] S.P. Shubeck, A.E. Kanters, J.B. Dimick, Surgeon leadership style and risk-adjusted patient outcomes, Surg Endosc. 33 (2) (2019) 471–474.
[25] G. Saposnik, D. Redelmeier, C.C. Ruff, P.N. Tobler, Cognitive biases associated with medical decisions: a systematic review, BMC Med Inform Decis Mak. 16 (138) (2016).
[26] B.W. Roberts, J.J. Jackson, J.V. Fayard, G. Edmonds, J. Meints, Conscientiousness, in: M.R. Leary, R.H. Hoyle (Eds.), Handbook of individual differences in social behavior, The Guilford Press, 2009, pp. 369–381.
[27] Widiger TA. Neuroticism. In: Handbook of individual differences in social behavior. 2009. p. 129–46.
[28] R.S. Bajwa, I. Batool, M. Asma, H. Ali, A. Ajmal, Personality Traits and Decision Making Styles among University Students Pakistan Journal of Life and Social Sciences Personality Traits and Decision Making Styles among University Students (Pakistan), Pakistan J Life Soc Sci. 14 (1) (2016) 38–41.
[29] N. Bayram, M. Aydemir, Decision-Making Styles and Personality Traits, Int J Recent Adv Organ Behav Decis Sci 3 (1) (2017) 905–915.
[30] R. El Othman, R. El Othman, R. Hallit, S. Obeid, S. Hallit, Personality traits, emotional intelligence and decision-making styles in Lebanese universities medical students, BMC Psychol. 8 (46) (2020).
[31] Weller J, Ceschi A, Hirsch L, Sartori R, Costantini A. Accounting for individual differences in decision-making competence: Personality and gender differences. Front Psychol. 2018;9(Nov 23):1–13.
[32] S. Miceli, V. de Palo, L. Monacis, S. Di Nuovo, M. Sinatra, Do personality traits and self-regulatory processes affect decision-making tendencies? Aust J Psychol. 70 (3) (2018) 284–293.
[33] N.L. Denburg, J.A. Weller, T.H. Yamada, D.M. Shivapour, A.R. Kaup, A. Laloggia, et al., Poor decision making among older adults is related to elevated levels of neuroticism, Ann Behav Med. 37 (2) (2009) 164–172.
[34] N. Beckmann, R.E. Wood, A. Minbashian, It depends how you look at it: On the relationship between neuroticism and conscientiousness at the within- and the between-person levels of analysis, J Res Pers. 44 (5) (2010) 593–601.
[35] E. Munari, G. Rossi, G. Zamboni, G. Lunardi, M. Marconi, M. Sommaggio, et al., PD-L1 Assays 22C3 and SP263 are Not Interchangeable in Non – Small Cell Lung Cancer When Considering Clinically Relevant Cutoffs, Am J Surg Pathol. 42 (10) (2018) 1384–1389.
[36] M.S. Tsao, K.M. Kerr, M. Kockx, M.-B. Beasly, A.C. Borczuk, J. Botling, et al., PD-L1 Immunohistochemistry Comparibility Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project, J Thorac Oncol. 13 (9) (2018) 1302–1311.
[37] L. Gandhi, D. Rodríguez-Abreu, S. Gadgeel, E. Esteban, E. Felip, F. De Angelis, et al., Pembrolizumab plus Chemotherapy in Metastatic Non–Small-Cell Lung Cancer, N Engl J Med. 378 (22) (2018) 2078–2092.
[38] L. Paz-Ares, A. Luft, D. Vicente, A. Tafreshi, M. Gümüş, J. Mazières, et al., Pembrolizumab plus Chemotherapy for Squamous Non–Small-Cell Lung Cancer, N Engl J Med. 379 (21) (2018) 2040–2051.
[39] M.A. Socinski, R.M. Jotte, F. Cappuzzo, F. Orlandi, D. Stroyakovskiy, N. Nogami, et al., Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC, N Engl J Med. 378 (24) (2018) 2288–2301.
[40] E. Thunnissen, K.M. Kerr, U. Dafni, L. Bubendorf, S.P. Finn, A. Soltermann, et al., Programmed death-ligand 1 expression influenced by tissue sample size. Scoring based on tissue microarrays' and cross-validation with resections, in patients with, stage I-III, non-small cell lung carcinoma of the European Thoracic Oncology Platform Lungs, Mod Pathol. 33 (2020) 792–801.
[41] H. Elfving, J.S.M. Mattsson, C. Lindskog, M. Backman, U. Menzel, P. Micke, Programmed Cell Death Ligand 1 Immunohistochemistry: A Concordance Study Between Surgical Specimen, Biopsy, and Tissue Microarray, Clin Lung Cancer. 20 (4) (2019) 258–262.e1.
[42] H. Yu, T.A. Boyle, C. Zhou, D.L. Rimm, F.R. Hirsch, PD-L1 expression in lung cancer, J Thorac Oncol. 11 (7) (2016) 964–975.