Towards identifying cancer patients at risk to miss out on psycho-oncological treatment via machine learning

Moritz Philipp Günther1 | Johannes Kirchebner2 | Jan Ben Schulze1 | Roland von Känel1 | Sebastian Euler1

1Department of Consultation-Liaison-Psychiatry and Psychosomatic Medicine, University Hospital Zurich, University of Zurich, Zurich, Switzerland
2Department of Forensic Psychiatry, University Hospital of Psychiatry Zurich, Zurich, Switzerland

Correspondence
Moritz Philipp Günther, Department of Consultation-Liaison-Psychiatry and Psychosomatic Medicine, University Hospital Zurich, Rämistrasse 100, Zurich 8091, Switzerland.
Email: moritzphilipp.guenther@usz.ch

Abstract
Objective: In routine oncological treatment settings, psychological distress, including mental disorders, is overlooked in 30% to 50% of patients. High workload and a constant need to optimise time and costs require a quick and easy method to identify patients likely to miss out on psychological support.

Methods: Using machine learning, factors associated with no consultation with a clinical psychologist or psychiatrist were identified between 2011 and 2019 in 7,318 oncological patients in a large cancer treatment centre. Parameters were hierarchically ordered based on statistical relevance. Nested resampling and cross validation were performed to avoid overfitting.

Results: Patients were least likely to receive psycho-oncological (i.e., psychiatric/psychotherapeutic) treatment when they were not formally screened for distress, had inpatient treatment for less than 28 days, had no psychiatric diagnosis, were aged 65 or older, had skin cancer or were not being discussed in a tumour board. The final validated model was optimised to maximise sensitivity at 85.9% and achieved an area under the curve (AUC) of 0.75, a balanced accuracy of 68.5% and specificity of 51.2%.

Conclusion: Beyond conventional screening tools, results might contribute to identify patients at risk to be neglected in terms of referral to psycho-oncology within routine oncological care.

KEYWORDS
Cancer, machine learning, mental disorders, psychological support, psycho-oncology

1 INTRODUCTION

Between 30% and 50% of patients with cancer suffer from significant psychological distress impeding their private, social and work life (Mehnert et al., 2018). In 30–40% of patients, depression, anxiety and other mood or adjustment disorders are to be diagnosed (Mitchell et al., 2011). Because such symptoms often go unnoticed in primarily somatic treatment settings (Fallowfield et al., 2001; Hallet et al., 2020; Passik et al., 1998; Sharpe et al., 2004; Söllner et al., 2001), consensus-based treatment guidelines and certification requirements of cancer treatment centres require formal screening for psychological symptoms using a validated instrument (such as the...
distress thermometer, DT) in order to provide psycho-oncological (i.e., psychiatric/psychotherapeutic) treatment where necessary (American College of Surgeons, 2019; AWMF, 2014). Psycho-oncological interventions have shown to reduce distress significantly (Faller et al., 2013; Meijer et al., 2013), increase treatment adherence (Kennard et al., 2004) and satisfaction (Bui et al., 2005; von Essen et al., 2002), improve course and prognosis of cancer (Carlson & Bultz, 2003; Geerse et al., 2019; McCarter et al., 2018; Sanjida et al., 2016), reduce the length of hospitalizations (Prieto et al., 2002), reduce the financial and emotional burden of health care providers (McCarter et al., 2018) and maximise the quality of life in general (Sstarstein et al., 2000).

Despite all, 25% to 80% of patients with cancer fail to receive adequate professional psycho-oncological treatment (Hollingworth et al., 2013; Mitchell, 2013). Nineteen per cent of patients do not even know about the availability of psycho-oncological support (Dilworth et al., 2014). Reasons include deficits in organisational structure (e.g., treatment team resources) (McCarter et al., 2018) and patient related factors, such as lower education (Eakin & Strycker, 2001; Mehnert & Koch, 2005; Nekolaichuk et al., 2011; Waller et al., 2013), older age (Ellis et al., 2009; Faller et al., 2016; Merckartz et al., 2010; Waller et al., 2013) and male gender (Curry et al., 2002; Merckartz et al., 2010; Nekolaichuk et al., 2011).

The current study aimed to identify the most important factors for clinical oncologists to avoid overlooking patients when considering psycho-oncological treatment. In addition, barriers to treatment should be easily identifiable for clinicians by reviewing the patient files before the next consultation and regardless of any formal screening for distress. Therefore, the aim of this study is to develop a prediction model answering the question ‘what patients with cancer are least likely to consult with a psychiatrist/psychologist?’ To this end, a relatively large number of variables need to be explored in terms of their significance. Machine learning (ML) is a relatively new and promising statistical tool to sort variables in terms of relevance. Computer algorithms (such as logistic regression, support vector machines [SVM], decision trees or k-nearest neighbour [KNN] depending on the data structure) can be employed using all available data from the electronic files of patients (so-called predictor variables) to identify those patient-related variables which are most significantly correlated with the respective outcome variable (no psychological/psychiatric consultation). In contrast to contemporary statistical approaches, ML can deal with larger data sets (data mining) and uncover previously hidden (linear and non-linear) relationships between variables without being limited by pre-devised hypotheses. The study was conducted in adherence to recommendations for Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD), and the TRIPOD-checklist can be found in the supporting information (Moons et al., 2015).

2 | METHOD

2.1 | Source of data and data preparation

The present study was approved by the Ethics Committee of the State of Zurich (Ref. No. BASEC-NR 2020-01949).

The current study analysed electronic case file data of 7,318 patients, who were initially diagnosed with (and treated for) cancer between 2011 and 2019 at the Comprehensive Cancer Centre Zurich (C3Z), which is a subunit of the University Hospital of Zurich, Switzerland. About 15,000 patients with cancer are treated every year. Treatment occurs on different specialised wards and outpatient clinics (e.g., gynaecology, haematology and dermatology); however, according to institutional guidelines, all cases should be discussed in an interdisciplinary tumour board consisting of radiologists, oncologists, pathologists, other medical specialists and a psychiatrist to optimise treatment. In addition, nurses are instructed to screen for distress using an ultrashort standardised screening instrument, the distress thermometer and problem list (Fulcher & Gosselin-Acomb, 2007; Roth et al., 1998), as recommended in international guidelines. See Table 1 for study group characteristics. In a first step, all routinely collected data in the electronic health records via the clinical management software (®KISIM, Cistec AG) used at the C3Z and via an institutional cancer register (®OncoStar, IT-Choice) were reviewed by the authors. Data were reviewed for its clinical and data value in terms of predicting the event of consulting with a psychologist/psychiatrist at the C3Z at any point of time during treatment (outcome variable). The almost 800 predictor variables were thus reduced to a final set of 47 (see the supporting information) based on availability of data and clinical relevance (author judgement based on literature cited here) to avoid overfitting in the final model (see below). All 47 variables were dichotomized, but this should not impact predictive power. For brief definitions, dichotomizations and the number of missing values of these variables, see the supporting information. Of this set of 47 variables, three variables had to be eliminated from further analysis due to conservative use of imputation, which only permitted imputation of variables with less than one third of missing values. See Figure 1 for data selection.

2.2 | Statistical procedures—ML

A preliminary description of ML is provided in Günther et al. (2020) and was partially adopted and extended here. An overview of the statistical procedure can be seen in Figure 2. All steps were performed using R version 3.6.3 and the MLR package v2.171 (Bischl et al., 2016). Confidence intervals of the balanced accuracy were computed using MATLAB R2019a (MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) with the add-on ‘computing the posterior balanced accuracy’ v1.0 (Brodersen et al., 2010).

After initial data preparation, 44 dichotomous predictor variables and the dichotomous outcome variable (psychological or psychiatric consultation occurred/no consultation occurred) remained. A total of 6,222 patients (85%) did not consult with a psychiatrist/psychologist and were defined to be the positive class, whereas 1,096 patients (15%) did and were defined to be the negative class. Next, the data set was divided into a training subset (70%, 5,123 patients) and a validation subset (30%, 2,195 patients). This was done to separate
model building from model validation and reduce the risk of model overfitting.

### 2.3 Initial ML modelling—Nested resampling

For model building, only the training subset was used. To further reduce overfitting and to avoid final model selection to be influenced by data processing, nested resampling (Moons et al., 2014; Studerus et al., 2017) was employed for initial modelling. This means, in an inner loop, data processing and model training are performed imbedded in fivefold-cross-validation and then, in an outer loop, the performance of these models is tested—also embedded in fivefold-cross-validation. Cross-validation is a technique to artificially create different subsamples of a data set (Browne, 2000).
2.4 Initial model construction—Data processing and model building

This step was performed within the inner loop of the nested resampling of the training subset. Because some ML algorithms need complete data sets, missing values were imputed via random forest algorithms in the MLR package, using randomForestSRC ad on (Ishwaran & Kogalur, 2020). Weights were stored and used in the validation subset later on. Due to the imbalance in the distribution of the outcome variable (85% without vs. 15% with consultation of a psychiatrist/psychologist), the less frequent state of ‘consultation occurred’ was randomly upsampled at a rate of 5.6, thus balancing the data subset, as is recommended for optimal model building (Wei & Dunbrack, 2013). Since the extraction of the most predictive variables without overfitting was a key objective of the current study, a random forest algorithm (randomForestSRC package; Ishwaran & Kogalur, 2020) was used to filter the initial 44 variables. Hyperparameter tuning was used to adjust the default functioning of algorithms in order to identify the most efficient model (see the supporting information for final hyperparameters). Finally, discriminative model building was applied with logistic regression, trees, random forest, gradient boosting, KNN, SVM and naive Bayes, as an easily applicable generative model (for a more detailed description, see James et al., 2013).
2.4.1 | Initial model construction—Model selection

In the outer loop of the nested resampling procedure, the final model was selected by assessing performance of each model. Model performance was evaluated in terms of balanced accuracy (i.e., the average of true positive and true negative prediction rate, which is suggested for imbalanced data; Brodersen et al., 2010) and goodness of fit (measured with the receiver operating characteristic, balanced curve area under the curve method and ROC balanced AUC) (Campbell, 1994). Moreover, specificity, sensitivity, positive predictive value (PPV) and negative predictive value (NPV) were evaluated. As our training dataset was artificially balanced, the model with the highest AUC was chosen for final model validation in the validation subset of the data (Campbell, 1994).

The final set of identified predictor variables was tested for multicollinearity to avoid dependencies between the variables.

2.5 | Final model evaluation

The validation subset of the original data set was not manipulated, except for the imputation of missing values via the stored weights from initial model building (see above). The best performing final model (with set hyperparameters) was used, and performance measures were reassessed. The predictor variables of the outcome variable in this final model were sorted by indicative power through means of a sensitivity analysis using the gbm package (Cortez & Embrechts, 2013).

3 | RESULTS

An overview of the performance parameters of the different calculated models during the nested resampling procedure is shown in Table 2 (see the supporting information for specific hyperparameters used and the corresponding confidence intervals). With a balanced accuracy of 81% and an AUC of 0.77, gradient boosting outperformed all other ML algorithms.

The absolute and relative distribution of the six most influential variables identified during nested resampling, which were subsequently used for model building, is shown in Table 3. The variables most predictive of no psychological/psychiatric consultation were no prior screening for distress (see above), length of longest inpatient stay less than 28 days, absence of a psychiatric diagnosis according to ICD-10 (World Health Organization, 2016), age 65 or older, presence of a skin cancer and no inclusion in a tumour board. Testing for multicollinearity showed no dependencies of concern between the variables (for detailed results, see the supporting information).

The quality of the final model was assessed in a validation step with results provided in Table 4. As expected, the balanced accuracy of 68.5 and AUC of 0.75 was less than the results of the initial training model but still meaningful. With a sensitivity of 85.9%, most patients

| TABLE 2 | Machine learning models and performance in nested cross-validation on training dataset |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Statistical procedure | Balanced Accuracy (%) | AUC | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
| Logistic regression | 77 | 0.76 | 83.8 | 70.1 | 97.1 | 26.3 |
| Tree | 80 | 0.74 | 80.6 | 79.4 | 97.9 | 25.3 |
| Random Forest | 80.6 | 0.74 | 80.1 | 81.2 | 98.1 | 25.2 |
| Gradient boosting | 80.8 | 0.77 | 80.4 | 81.2 | 98.1 | 25.5 |
| KNN | 76.5 | 0.54 | 54.3 | 98.7 | 99.8 | 15.2 |
| SVM | 79.5 | 0.74 | 80.9 | 78.1 | 97.8 | 25.3 |
| Naive Bayes | 79.6 | 0.74 | 78.7 | 80.5 | 98.1 | 23.2 |

Abbreviations: AUC, area under the curve (level of discrimination); KNN, k-nearest neighbours; NPV, negative predictive value; PPV, positive predictive value; SVM, support vector machines.

| TABLE 3 | Absolute and relative distribution of indicative variables on the complete dataset |
|-----------------|-----------------|-----------------|-----------------|
| Variable code | Variable description | No psychiatric/psychotherapeutic treatment (%) | Psychiatric/psychotherapeutic treatment (%) |
| DTSREE | Distress screening applied | 1507/6222 (24.2) | 480/1096 (43.8) |
| HOSDUR | Longest inpatient treatment 28 days or more | 774/6,222 (12.4) | 410/1096 (37.4) |
| FDX | Mental disorder present | 538/6,222 (8.6) | 256/1096 (23.4) |
| Alt/Jung | Patient age 65 or older | 3,287/6,222 (52.8) | 409/1096 (37.3) |
| Haut | Skin cancer present | 1,975/6,222 (31.7) | 147/1096 (13.4) |
| Tumorboard | Tumour board held | 2,616/4,832 (54.1) | 622/732 (85) |
who did not consult with a clinical psychologist or psychiatrist were identified in the final model. With a specificity of 51.2%, more than half of patients who did consult with a psychologist/psychiatrist were detected correctly.

A one-sided tornado graph comparing the relative importance of the identified variables during model validation is presented in Figure 3. It shows the effect on the output variable by varying each predictor variable at a time, keeping all the other predictor variables at their initial values. Consequently, the predictor variables were ranked from the most influential to the least influential.

### Discussion

Patients not formally screened for distress, staying for less than 28 days in inpatient treatment, who do not have a psychiatric diagnosis, are aged 65 or older, have skin cancer or were not discussed in a tumour board meeting were least likely to receive psycho-oncological treatment.

With a sensitivity of 86%, the final model allows oncologists to quickly review patient files for six variables in order to identify those patients least likely to receive psycho-oncological treatment. The lower specificity of 51% seems less clinically relevant, because it does not impede on the objective of the study of identifying patients likely to miss out on a psycho-oncological consultation. Although between 30% and 40% of patients with cancer are dealing with a mood, anxiety or adjustment disorder according to the international classification of diseases (Mitchell et al., 2011), the rate of referral to a psycho-oncological consultation in the C3Z was a low 15% (1,096 of 6,222 patients; see Table 3). Based on this, it can be assumed that a high number of patients with cancer in need do not receive psycho-oncological treatment. Results presented here should help to easily identify patients who are least likely to consult a psychiatrist to increase the number of referrals. Prior research found that 19% of patients did not know about psycho-oncological support for patients with cancer (Dilworth et al., 2014). Oncological nursing teams may only find the time to screen an average of 40% of inpatients for distress due to workload (Götz et al., 2019, 2020). This is why results of the present study are needed to allow prioritisation for oncological treatment teams. For example, the list of variables presented here could allow oncological treatment teams on ward rounds to be alert for patients frequently missing out on psycho-oncological treatment.

As a major limitation, the final model provided is only useful in health care systems providing sufficient resources to allow all patients to receive professional psycho-oncological support. In less privileged health care systems, a triaging of patients may be necessary with formal psychological support only available to the most severely distressed patients, whereas support services run by specialist nurses or support workers would be available for less distressed patients.

Further, the present study confirms the preponderance of prior research and international consensus-based guidelines proposing formal distress screening in patients with cancer to identify those needing psycho-oncological treatment (American College of Surgeons, 2019; AWMF, 2014). Possible ‘informal’ distress screening by oncologists does not comply with these standards of evidence based medicine. Results also support prior research showing that oncological patients aged 65 and over are likely to miss out on psycho-oncological treatment (Ellis et al., 2009; Faller et al., 2016; Merckaelert et al., 2010; Waller et al., 2013). Similarly, prolonged length of hospitalisation has been linked to barriers to psychiatric consultation in general hospitals in a recent systematic review (Oldham et al., 2019), regardless of the somatic condition being treated for.

New findings from the current study indicate that patients with skin cancer, those who are not discussed at a tumour board and those without a psychiatric diagnosis are unlikely to receive psycho-oncological treatment. It is quite conceivable that patients not being discussed at a tumour board would receive less overall medical attention, which might include psycho-oncological treatment. This highlights the importance of tumour boards beyond identification of the best somatic treatment. In other words, tumour board discussions may not only be valuable in optimising survival rates or efficacy of oncological treatment, but also for the identification of psychological needs. An interpretation of our finding that patients with skin cancer are unlikely to receive psycho-oncological treatment is not straightforward. One reason may be that only about 16% of skin cancers are melanoma (Guy et al., 2015), which is still difficult to treat and likely to cause distress due to its poor prognosis on survival (Robert et al., 2019). In contrast, all other skin cancers have an above average prognosis in survival in comparison to other cancers overall (Leiter et al., 2014). Future research should examine in greater depth to what extent the psychological needs of this subgroup are currently being met. Some may argue that patients without a psychiatric diagnosis do not need psycho-oncological treatment (Ullrich, 2020), while the majority of researchers found it to be beneficial to all patients with cancer even if no mental disorder is to be diagnosed (American College of Surgeons, 2019; AWMF, 2014; Waller et al., 2013).

We identified a number of factors, which were not predictive of no psycho-oncological treatment. Among these were other types of cancer, sex, nationality, insurance status, occupation, relationship status, prescription of psychopharmacology, number of comorbidities,
palliative treatment perspectives, surgical interventions and religious confession. We may conclude that there is no evidence for discrimination of patients based on economic factors (i.e., insurance status), their beliefs, nationality, sex or medical circumstances. This is despite that some medical circumstances can be assumed to be emotionally challenging for psychotherapists (i.e., palliative care settings and many comorbidities).

5 | LIMITATIONS

Our study has several limitations, including the generalizability of results derived from data collected from electronic files at only one cancer treatment centre. Nonetheless, data were collected in a large cancer treatment centre in Switzerland, treating all known entities of cancer. Given the strict internal regulations to ensure evidence-based treatment for international accreditation as a leading cancer treatment centre, human error should be minimal when assessing and documenting data in the electronic health record.

Further, data from all types of cancers were included, which may lead to factors specific to certain (rare) entities of cancer to be overlooked, while increasing broad applicability of findings presented to oncologists working with patients with all sorts of cancers.

Another limitation is that there was no information available on whether or not patients who were not screened for distress were informally offered psycho-oncological treatment and how they reacted to that. Instead, the current study aimed to help oncologists increase efficacy in offering psycho-oncological treatment in extremely busy workload settings.

As ML was used with this set and type of data for the first time (to the authors knowledge), there is a residual risk of overfitting although precautionary measures were taken (see Section 2). Therefore, further (prospective) research, testing the usefulness of results presented here in clinical practice (and ideally in a different cancer treatment centre), is necessary for thorough performance evaluation of the final model. Such research might also evaluate further variables, currently not available in routine documentation of the treatment centre studied here.

6 | CONCLUSION

In conclusion, the results presented here might contribute to screen patients efficiently for their risk to be neglected in referral to psycho-oncological treatment within routine oncological care in health care systems aiming to offer such treatment to a majority of patients with cancer. The identified factors could serve as ‘yellow flags’ to directly ask patients during the next consultation. This approach might help to ensure that no patient misses the opportunity to receive psycho-oncological treatment. It may be a golden middle path to other health
care systems in which such treatment is offered to those in absolute need or to all patients with cancer regardless of their need. Similarly, it may be a compromise between perspectives viewing psychosocial oncology as the sixth vital sign in cancer treatment (Waller et al., 2013) and those suspecting a mere business model for psychologists (Ullrich, 2020).

ETHICS STATEMENT
The study was reviewed and approved by the Ethics committee of Zurich, Switzerland (Ref.-No. BASEC-NR 2020–01949). This is a retrospective study. For this type of study, formal consent is not required. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

ACKNOWLEDGEMENTS
Open Access Funding provided by Universitat Zurich.

CONFLICT OF INTEREST
The authors declare that they have no competing interest.

AUTHOR CONTRIBUTIONS
MG, JS and SE conceived and designed the study. Data collection was performed by JS and MG. Material preparation and analysis were performed by JK. The first draft of the manuscript was done by MG. RvK provided senior advice for the study. All authors edited multiple drafts and supervised the statistical analyses. All authors read and approved the final manuscript.

FUNDING INFORMATION
This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Moritz Philipp Günther https://orcid.org/0000-0002-7707-5532
Johannes Kirchebner https://orcid.org/0000-0002-6072-9958
Jan Ben Schulze https://orcid.org/0000-0002-5252-3976
Roland von Känel https://orcid.org/0000-0002-8929-5129
Sebastian Euler https://orcid.org/0000-0002-5009-8355

REFERENCES
American College of Surgeons. (2019). Optimal resources for cancer care 2020 standards. https://www.facs.org/-/media/files/quality-programs/cancer/coc/optimal_resources_for_cancer_care_2020_standards.aspx
AWMF, (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften), Deutsche Krebsgesellschaft, Deutsche Krebshilfe, (2014). Guideline for psychooncological diagnostics, counseling and treatment of adult patients with cancer [S3-Leitlinie Psychoonkologische Diagnostik, Beratung und Behandlung von erwachsenen Krebspatienten]. 1.1. https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Psychoonkologieleitlinie_11.1_LL_PSO_Langversion_1.1.pdf
Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., & Jones, Z. M. (2016). mlr: Machine learning in R. The Journal of Machine Learning Research, 17(1), 5938–5942.
Brodersen, KH, Ong, CS, Stephan, KE, & Buhmann, JM. (2010). The balanced accuracy and its posterior distribution. Paper presented at the 2010 20th international conference on pattern recognition.
Browne, M. W. (2000). Cross-validation methods. Journal of Mathematical Psychology, 44(1), 108–132. https://doi.org/10.1006/jmps.1999.1279
Bui, Q. U. T., Ostir, G. V., Kuo, Y. F., Freeman, J., & Goodwin, J. S. (2005). Relationship of depression to patient satisfaction: Findings from the barriers to breast cancer study. Breast Cancer Research and Treatment, 89(1), 23–28. https://doi.org/10.1007/s10549-004-1005-9
Campbell, G. (1994). Advances in statistical methodology for the evaluation of diagnostic and laboratory tests. Statistics in Medicine, 13(5–7), 499–508. https://doi.org/10.1002/sim.4780130513
Carlson, L. E., & Bultz, B. D. (2003). Cancer distress screening. Needs, models, and methods. Journal of Psychosomatic Research, 55(5), 403–409. https://doi.org/10.1016/s0022-3999(03)00514-2
Cortez, P., & Embrechts, M. J. (2013). Using sensitivity analysis and visualization techniques to open black box data mining models. Information Sciences, 225, 1–17. https://doi.org/10.1016/j.ins.2012.10.039
Curry, C., Cossich, T., Matthews, J., Beresford, J., & McLachlan, S. (2002). Uptake of psychosocial referrals in an outpatient cancer setting: Improving service accessibility via the referral process. Supportive Care in Cancer, 10(7), 549–555. https://doi.org/10.1007/s00520-002-0371-2
Dilworth, S., Higgins, I., Parker, V., Kelly, B., & Turner, J. (2014). Patient and health professional’s perceived barriers to the delivery of psychosocial care to adults with cancer: A systematic review. Psycho-Oncology, 23(6), 601–612. https://doi.org/10.1002/pon.3474
Eakin, E. G., & Strycker, L. A. (2001). Awareness and barriers to use of cancer support and information resources by HMO patients with breast, prostate, or colon cancer: Patient and provider perspectives. Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer, 10(2), 103–113. https://doi.org/10.1002/pon.500
Ellis, J., Lin, J., Walsh, A., Lo, C., Shepherd, F. A., Moore, M., Li, M., Gagliese, L., Zimmermann, C., & Rodin, G. (2009). Predictors of referral for specialized psychosocial oncology care in patients with metastatic cancer: The contributions of age, distress, and marital status. Journal of Clinical Oncology, 27(5), 699–705. https://doi.org/10.1200/JCO.2007.15.4864
Faller, H., Schuler, M., Richard, M., Heckl, U., Weiss, J., & Küffner, R. (2013). Effects of psycho-oncologic interventions on emotional distress and quality of life in adult patients with cancer: Systematic review and meta-analysis. Journal of Clinical Oncology, 31(6), 782–793. https://doi.org/10.1200/jco.2011.40.8922
Faller, H., Weis, J., Koch, U., Brähler, E., Härter, M., Keller, M., Schulz, H., Wegscheider, K., Boehncke, A., Hund, B., & Reuter, K. (2016). Perceived need for psychosocial support depending on emotional distress and mental comorbidity in men and women with cancer. Journal of Psychosomatic Research, 81, 24–30. https://doi.org/10.1016/j.jpsychres.2015.12.004
Fallowfield, L., Ratcliffe, D., Jenkins, V., & Saul, J. (2001). Psychiatric morbidity and its recognition by doctors in patients with cancer. British Journal of Cancer, 84(8), 1011–1015. https://doi.org/10.1054/bjc.2001.1724
Fulcher, C. D., & Gosselin-Acomb, T. K. (2007). Distress assessment: Practice change through guideline implementation. Clinical Journal of Oncology Nursing, 11(6), 817–821. https://doi.org/10.1188/07.CJON.817-821
Skarstein, J., Aass, N., Fosså, S. D., Skovlund, E., & Dahl, A. A. (2000). Anxiety and depression in cancer patients: Relation between the hospital anxiety and depression scale and the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire. *Journal of Psychosomatic Research*, 49(1), 27–34. https://doi.org/10.1016/s0022-3999(00)00080-5

Söllner, W., DeVries, A., Steixner, E., Lukas, P., Sprinzl, G., Rumpold, G., & Maislinger, S. (2001). How successful are oncologists in identifying patient distress, perceived social support, and need for psychosocial counselling? *British Journal of Cancer*, 84(2), 179–185. https://doi.org/10.1054/bjoc.2000.1545

Studerus, E., Ramyead, A., & Riecher-Rössler, A. (2017). Prediction of transition to psychosis in patients with a clinical high risk for psychosis: A systematic review of methodology and reporting. *Psychological Medicine*, 47(7), 1163–1178. https://doi.org/10.1017/S0033291716003494

Ullrich G. (2020). Management bias in psycho-oncology? A critical view by a career changer. Retrieved from https://figshare.com/articles/preprint/Management_bias_in_der_Psychoonkologie_pdf/12059442

von Essen, L., Larsson, G., Öberg, K., & Sjödén, P.-O. (2002). ‘Satisfaction with care’: Associations with health-related quality of life and psychosocial function among Swedish patients with endocrine gastrointestinal tumours. *European Journal of Cancer Care*, 11(2), 91–99. https://doi.org/10.1046/j.1365-2354.2002.00293.x

Waller, A., Williams, A., Groff, S. L., Bultz, B. D., & Carlson, L. E. (2013). Screening for distress, the sixth vital sign: Examining self-referral in people with cancer over a one-year period. *Psycho-Oncology*, 22(2), 388–395. https://doi.org/10.1002/pon.2102

Wei, Q., & Dunbrack, R. L. Jr. (2013). The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS ONE*, 8(7), e67863. https://doi.org/10.1371/journal.pone.0067863

World Health Organization (WHO). (2016). *ICD-10: International statistical classification of diseases and related health problems: Tenth revision* (5th ed.). World Health Organization.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Günther, M. P., Kirchebner, J., Schulze, J. B., von Känel, R., & Euler, S. (2022). Towards identifying cancer patients at risk to miss out on psycho-oncological treatment via machine learning. *European Journal of Cancer Care*, 31(2), e13555. https://doi.org/10.1111/ecc.13555

https://doi.org/10.1111/ecc.13555