*Original Article*

# Intelligent Palliative Care Based on Patient-Reported Outcome Measures

Margaret H. Sandham, BHSc (Nursing), DClinPsych, Emma A. Hedgecock, BHSc (Nursing), MHSc, Richard J. Siegert, PhD, Ajit Narayanan, PhD, Mevhibe B. Hocaoglu, PhD, and Irene J. Higginson, PhD
*School of Clinical Sciences (M.S., R.S.), Auckland University of Technology, Auckland, New Zealand; Specialty Medicine and Health of Older People, Waitemata District Health Board, Private Bag (E.A.H.), Takapuna, New Zealand; School of Engineering, Computer and Mathematical Sciences (A.N.), Auckland University of Technology, Auckland, New Zealand; Cicely Saunders Institute of Palliative Care, Florence Nightingale Faculty of Nursing, Midwifery and Palliative Care (M.B.H., I.J.H.), King's College London, London, UK*

## Abstract

**Context.** The growth of patient reported outcome measures data in palliative care provides an opportunity for machine learning to identify patterns in patient responses signifying different phases of illness.

**Objectives.** The study will explore if machine learning and network analysis can identify phases in patient palliative status through symptoms reported on the Integrated Palliative Care Outcome Scale (IPOS).

**Methods.** A partly cross-sectional and partially longitudinal observational study was undertaken using the Australasian Karnofsky Performance Scale (AKPS); Integrated Palliative Care Outcome Scale (IPOS); Phase of Illness (POI). Patient palliative records ($n = 1507$, 65% stable, 20% unstable, 9% deteriorating, 2% terminal) from 804 adult patients enrolled in a New Zealand palliative care service were analysed using a combination of statistical, machine learning and network analysis techniques.

**Results.** Data from IPOS showed considerable variation with phase. Also, network analysis showed clear associations between items by phase. Six machine learning techniques identified the most important variables for predicting possible transition between phases of illness. Network analysis for all patients showed that Poor Appetite and Loss of Energy were central IPOS items, with Loss of Energy linked to Drowsiness, Shortness of Breath and Lack of Mobility on the one hand, and Poor Appetite linked to Nausea, Vomiting, Constipation and Sore and Dry Mouth on the other.

**Conclusion.** These preliminary results, when coupled with the latest technological developments in mobile apps and wearable technology, could point the way to increased use of digital therapeutics in continuous palliative care monitoring. J Pain Symptom Manage 2022;63:747−757. *© 2021 The Authors. Published by Elsevier Inc. on behalf of American Academy of Hospice and Palliative Medicine. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

***Key Words***
*Machine learning, network analysis, palliative care, psychometrics, Integrated Palliative Outcome Scale IPOS, Australasian Karnofsky Performance Scale AKPS, Phase of Illness POI, wearable electronic devices*

**EDITORIAL NOTE** ***Debra Parker Oliver, PhD, MSW***
This paper may be challenging to read but is worth the thought it provokes. It challenges us to think of ways technology can improve our care.

## *Key Message*

Preliminary risk prediction using ML and identification of key symptom networks, when coupled with the latest technological developments in mobile apps and wearable technology, could pave the way to increased use of digital therapeutics in continuous palliative care monitoring.

*Address correspondence to:* Margaret Sandham, BHSc (Nursing), DClinPsych, School of Clinical Sciences, Auckland University of Technology (AUT), 90 Akoranga Drive, Northcote, North Shore, Auckland, 0627, New Zealand. E-mail: Margaret.sandham@aut.ac.nz

*Accepted for publication: 16 November 2021.*



0885-3924/$ - see front matter
https://doi.org/10.1016/j.jpainsymman.2021.11.008

## Background

Measurement based palliative care in general and patient reported outcome measures (PROMs) in particular constitute key strategies for ensuring timely appropriate delivery of quality of health services.[1] Measurement based care involves the use of regularly measured symptoms to inform treatment decisions, whereas evidence-based care relies on evidence from research studies or practice. The use of measurement based care in mental health service delivery and decision making, for instance, has resulted in superior outcomes compared to treatment as usual and these outcomes transcend patient diagnosis and clinician factors.[2] Growing evidence suggests that in comparison to complex and burdensome evidence-based practices, measurement based care is a simple intervention that could produce substantial impact on patient outcomes.[3-5]

Of particular interest is whether patient-driven methods for measurement based care can be merged with appropriate analytic methods for identifying patterns in PROMs data that may indicate a possible change in level and type of support care required to improve quality of life and survival rates. That is, while clinical measurements provide one source of valuable data, there is a growing realization that most palliative and end-of-life care in future will be provided in hospices or at home rather than in clinics and hospitals.[6] Access to specialist facilities is likely to be limited, with increasing emphasis on PROMs by care-givers to provide appropriate care and support. This foreseeable trend towards increasing use of PROMs is not matched by progress in the application of methods and techniques from artificial intelligence (AI) and machine learning (ML) for improving the quality of care through non-specialist decision support. Measurement-based care is a data-driven approach suited for the application of methods from ML that 'learn' from the data to identify underlying patterns and make classifications and predictions based on those patterns. AI is involved if such learning produces models that resemble human reasoning.[7] Also, the vision of digital therapeutics (DTx) in palliative care, whereby digital sensors use internet technology for measuring and storing healthcare data to improve therapeutic intervention,[8] requires the merging of everyday technologies such as mobile apps and wearable devices with ML to augment care delivery, especially in non-clinical settings such as private homes and social care settings.[9] A number of applications of ML in healthcare are now documented, including automated classification and prediction of cardiovascular risk, eye disease, and personalized cancer risk from genomic information[10]

So far, papers in AI and ML dealing with palliative care have typically followed statistical practice in trying to predict mortality rather than changes in palliative care status, and applying random forests, feature selection and logistic regression to general medical records rather than palliative-specific measurement outcomes (see[11] for a review). Deep learning (DL), an area of ML that uses mathematical and statistical models inspired by brain processing,[12] has also been tried[13] to predict mortality and whether patients would benefit from palliative care by using a combination of clinical features including disease diagnosis and patient demographics. That study used the previous year's electronic health records of patients to predict mortality over the next 12 months, with the ML model learning to separate positive patients (those who have a recorded date of death in their electronic health records) from negative cases (patients without a recorded date of death). Due to imbalances in the data set (prevalence of death at 7%), traditional accuracy figures were difficult to apply but the study reported an average precision (the ratio of correctly predicted positive cases to the total number of positive cases) of 0.69. Another key initiative involving a DL model[14] used electronic health records, clinical notes as well as demographic data of dementia patients to predict mortality within six months, one year or two years. Receiver Operating Characteristic (ROC) curves are used in ML classification and prediction models to summarize the relationship between true positive rate and false positive rate. A ROC figure of 0.5 denotes no model discrimination, above 0.7 as acceptable and above 0.8 as very good. The DL model achieved ROC AUC figures of 0.978, 0.956 and 0.943, respectively. Topics extracted from health notes identified possible predictive topics such as cognitive function and laboratory testing in the absence of interpretability of DL models, which are considered 'black box' because of a lack of transparency and interpretability concerning how input data are transformed to output.[15] More interpretable ML decision trees (gradient-boosted trees) used electronic health records for chemotherapy patients, including palliative care patients, to predict 30-day mortality after start of a new regimen. The model used over 5000 predictors and achieved AUC figures of 0.94 for all patients and 0.924 for palliative care patients across all cancers. Predictors were identified through proportion of variance explained by each variable, with cancer of the brain and other nervous systems being the top predictor, followed by age and specific prior diagnoses.[16] A combination approach involving logistic regression, gradient boosting and random forests using electronic health records and demographic data for oncology patients predicted six-month mortality with AUC figures between 0.86 and 0.88. Variable importance measures identified albumin and alkaline phosphate levels as well as diagnostic codes for solid tumour with and without metastasis.[17] Also, there is wide variability in the accuracy of clinicians' predictions of prognosis and survival (around 78% for cancer and less for non-cancer), indicating that predicting survival is a challenging task for both clinicians and ML.[18]

None of the palliative care mortality studies has dealt with predicting level or type of palliative care required, or how to predict possible transition from one level to another. Also, while EHRs have been used in previous studies, there is currently very little understanding of how to use PROMs palliative care data for identifying level of palliative care and change of level of care given differences in such data after admission. In particular, the association between PROMs measures and phase of illness, which is critical for identifying the most appropriate care strategies, is still not fully understood and is gaining more attention.[19] There is growing acceptance that definitions of 'phase of illness' need to be standardized internationally as part of the strategy to improve quality of palliative care.[20] The definitions used in this study are as follows[21]:

**Stable**: When the patients problems are being managed by the current plan of care and family/carer situation stable.
**Unstable**: When there is a need for a new plan of care to manage new symptoms, escalating current symptoms or change in family situation.
**Deteriorating**: Patient's plan of care requires constant review and / or family circumstances are changing.
**Terminal**: Death is likely within days.

Another challenge will be converting any understanding of reasons for change of phase into a practical monitoring solution in all palliative care settings (clinical and non-clinical) without requiring constant and possibly intrusive measurement.

The present study will explore the feasibility of applying ML and network analysis to palliative care PROMs data, in order to provide useful information to support future timely decision-making. The study will identify if ML can model differences in phase of illness by identifying different association patterns in palliative-specific PROMs. Identifying key symptom changes on PROMs through ML may enable prompt early outreach in non-clinical settings to provide early interventions and help ease the transition or possibly prevent rapid deterioration.

## Methods

### Study Setting

[blinded for review] is a charitable trust that provides palliative care for approximately 230 community dwelling patients living in a mixed urban and rural geographical area of [blinded], New Zealand. Palliative health and social care are provided predominantly as outpatients, in a patient's home environment with some specialist medical and nursing services provided in Hospice House (inpatient). Patients are accepted if they have life-limiting conditions with a life-expectancy of less than 12 months and complex or high palliative care needs, and normally remain in the service until death. In contrast to other parts of New Zealand, these patients are more likely to be indigenous Māori or of Pacific ethnicity and have lower socioeconomic status.

### Design

The present data-based study is an observational psychometric study implementing measurement based palliative care, assessed to be low-risk. Exemption from further ethics approval was granted by the New Zealand Health and Disability Ethics Committee since the data being used in this study had already been gathered and stored for routine monitoring purposes. Also, patients whose anonymised data were used in the research had signed a consent to have their de-identified data used for research or audit purposes.

### Participants

Patients under the care of Hospice during the data collection period were included if they were aged 18 or above and able to complete the English language study measures independently or with support. Patients were excluded if the clinical team judged them as being too unwell or distressed to participate, or had very limited English language or moderate to severe cognitive impairment. The data were collected between January 2018 and October 2019.

### Study Measures

The Palliative care Outcome Scale (POS)[22] was a psychometrically reliable tool[23-24] that was refined and became the Integrated Palliative care Outcome Scale (IPOS;[25]) with evidence of reliability and validity.[26-27] The IPOS measures symptom burden through ten items on a zero to four Likert scale (ranging from 'not at all' (0) to 'overwhelmingly' (4) and a summative total score is calculated. Patients self-assess or, if they are not able, complete the items with help from family members or caregivers. A description of the items and labels can be found in (Table 1) and associated caption.

The clinician-scored Australasian Karnofsky Performance Scale (AKPS;[28]) is based on observations of a patient's performance in the domains of work, activity, and self-care. Scores range from 100% (normal, no complaints, no evidence of disease) down to 0% when the patient has died. Criteria are given for increments of 10% (see caption to Table 1 for more details). In the present study the AKPS provides a measure of corroboration between self-administered IPOS items and expert-provided clinical judgement.

*Table 1*
**Overview of Patient Demographics and Diagnosis**

| | Total Patients = 751 | *n* | % |
|---|---|---|---|
| Gender | Female | 374 | 49.8 |
| | Male | 376 | 50.1 |
| | Not specified | 1 | 0.1 |
| Ethnicity | Caucasian | 405 | 53.9 |
| | Maori | 96 | 12.8 |
| | Pacific Island | 82 | 10.9 |
| | Asian | 49 | 6.5 |
| | Other | 119 | 15.8 |
| Diagnosis | Oncological cancer | 548 | 73.0 |
| | Organ failure | 127 | 16.9 |
| | Haemotological cancer | 37 | 4.9 |
| | Dementia | 11 | 1.5 |
| | Other | 28 | 3.7 |

The Phase of Illness (POI; [19]) classifies patients according to their care needs into one of five, distinct stages of illness (stable, unstable, deteriorating, terminal/dying, deceased). Patients can move back and forward through the first four phases, and a change of phase signals a requirement to undertake a holistic reassessment. The POI has validity in capturing information related to clinical need and has acceptable interrater reliability.[21] Its use here is to group patients for sub-group analysis.

### *Analysis*

First, descriptive statistics were used including percentages and correlations ($r$) to check on questionnaire properties and reliability for subsequent analysis. Since the purpose here is exploratory with no specific research hypotheses, no adjustments were made to $P$ values to take into account multiple testing. Imputation of missing IPOS values was also undertaken to maximise data entry (details below). Second, analysis of variance (ANOVA) and psychometric networks[29] were used to explore group comparisons using phase, where nodes represent IPOS items and edges the associations between items.

Psychometric network analysis is a rapidly growing area in health science aimed at identifying structural relations among psychological and other health-related variables through graphical representations, where nodes represent variables such as attitudes, cognitive states, symptoms and behaviours, and links (edges) represent a statistical relationship, usually correlation or other forms of association measures, that convey information about the strength or direction of the relationship.[30] The aim is to identify mutually-influencing relationships among variables to explain a phenomenon rather than extract dominant, latent factors.[31] Partial correlations were used for edge calculation, where correlations between two variables are calculated after removing the effects of all other variables.[32] Spring layout (a force-directed graph drawing algorithm[33]) was adopted for network visualisation, where edges are regarded as springs pulling nodes together. Larger edge weights (partial correlations) imply a stronger inter-node attractive force. Nodes, however, have a repulsive force and the aim is to produce a graph that reaches equilibrium between edge attraction and node repulsion. Nodes with the strongest links tend to be placed at the centre of the graph using this method. Centrality measures provide further information on network nodes, such as 'betweenness' (how well a node acts as a connecting point based on the number of paths through that node to other nodes), 'closeness' (how close a node is to other nodes using the average weight of the paths from that node) and 'degree' (the sum of all weights from that node, or strength[34]). Centrality measures are presented as standardized z scores to allow comparison. Third, classification by phase was achieved through supervised ML techniques. Six techniques were used given that no previous ML approach to modelling IPOS data appears to have taken place: neural networks, naïve Bayes, decision forest, ID3, logistic model tree, random forest. These techniques cover a range of black box, statistical, rule-induction and tree-based methods. ROC AUC figures are provided for comparison between ML techniques, where phase is the class of interest.

### *Patient Details*

Patients enrolled in the Hospice between December 2017 and September 2019 had their routinely collected assessment data included in the study. A total of 1434 separate assessments were made of 751 patients almost equally split between females and males, with ages ranging from 23 to 101 years (mean = 71.9, median = 73.0, SD = 13.7). The majority of patients were Caucasian (54%) and the main cause of disease was oncological cancer (73%). Table 1 provides an overview of the sample demographics and diagnosis. The average number of assessments per patient was 1.91, with 61.1% of patients assessed once, 16.2% assessed twice, 9.3% assessed three times, and 5.5% assessed four times. The remaining 4.6% of patients had a range between 7 (three patients) and 12 (one patient) assessments.

## Results

### *Descriptive*

Table 2 provides an overview of all 10 IPOS items, summed IPOS score and AKPS score. The average rate of missing values per IPOS item was 2.1%, leading to 178 cases (12.4%) not being included in any analysis that required a full set of scores due to listwise deletion. Little's missing completely at random value test[35] was $P$ = .018, indicating that data were not missing completely at random and therefore could be imputed

*Table 2*

**Overview of 10 IPOS Items (0 = not at all; 1 = Slightly; 2 = Moderately; 3 = Severely; 4 = Overwhelmingly), Total IPOS SCORE (Summed IPOS Items) and AKPS Score (0 = Dead; 10 = Comatose or Barely Rousable; 20 = Almost Totally Bed-bound, Requiring Extensive Nursing Care by Professionals/Family; 30 = Almost Completely Bed-bound; 40 = In Bed More Than Fifty Percent of the Time; 50 = Considerable Assistance and Frequent Medical Care Required; 60 = Able to Care for Most Needs, but Requires Occasional Assistance; 70 = Able to Care for Self but Unable to Work or Carry on Other Normal Activities; 80 = Normal Activity Requires Effort, Signs and Symptoms of Disease More Prominent; 90 = Minor Signs of Illness Present; 100 = No Complaints or Evidence of Disease). The Final Two Columns Describe the IPOS Items and Revised IPOS SCORE After Imputation**

| Item | Original | | | | | Imputed $n = 1434$ | |
|---|---|---|---|---|---|---|---|
| | | | | Missing | | | |
| | *n.* | Mean | Standard Deviation | Count | Percent | Mean | Std. Deviation |
| Pain | 1410 | 1.40 | 1.13 | 24 | 1.7 | 1.40 | 1.13 |
| Shortness of breath | 1408 | 1.19 | 1.14 | 26 | 1.8 | 1.19 | 1.13 |
| Weakness or lack of energy | 1407 | 1.96 | 1.09 | 27 | 1.9 | 1.96 | 1.09 |
| Nausea | 1405 | 0.53 | 0.88 | 29 | 2.0 | 0.53 | 0.88 |
| Vomiting | 1413 | 0.23 | 0.62 | 21 | 1.5 | 0.24 | 0.64 |
| Poor appetite | 1409 | 1.20 | 1.20 | 25 | 1.7 | 1.20 | 1.19 |
| Constipation | 1408 | 0.84 | 1.10 | 26 | 1.8 | 0.84 | 1.09 |
| Sore or dry mouth | 1387 | 0.90 | 1.02 | 47 | 3.3 | 0.91 | 1.02 |
| Drowsiness | 1394 | 1.00 | 1.05 | 40 | 2.8 | 1.02 | 1.05 |
| Poor mobility | 1405 | 1.45 | 1.22 | 29 | 2.0 | 1.46 | 1.22 |
| *SCORE* | *1256* | *10.52* | *6.06* | *178* | *12.4* | *10.75* | *6.09* |
| *AKPS* | *1370* | *60.82* | *16.53* | *64* | *4.5* | - | - |

using linear regression from samples with complete data. The range of missing values was constrained to between zero and four in line with the original IPOS item allowable values, with rounding to the nearest integer.

Reliability analysis (one way random, single measure) of the 10 IPOS items produced a Cronbach's $\alpha$ of 0.78, indicating good internal consistency, with an average inter-item correlation of $r$=.26. Factor analysis confirmed that the 10 IPOS items could not be reduced to a smaller set of underlying components, with each item contributing on average 10% of the variance in the entire data using varimax rotation.

The average AKPS score was 60.82% (median = 60, st.dev. 16.53, st.err 0.45, $n$ = 1370), with 64 missing values due to AKPS not being measured at the same time as IPOS items. AKPS score was significantly and negatively correlated with IPOS score ($r = -.34$, $P \leq .001$), as expected. That is, low AKPS scores signify worse states than high scores, and low IPOS item scores signify better states than high scores. Significant negative correlations also persisted for every item of IPOS with AKPS except for Nausea.

### *Subgroup Analysis and Network Analysis*

947 patient records indicated Phase one (Stable), 291 Phase two (Unstable), 124 Phase three (Deteriorating) and 20 Phase four (Terminal), with 51 missing values due to Phase not being measured at the same time as IPOS items. Table 2 provides a breakdown of IPOS total score and AKPS, with both measures varying significantly by phase overall ($P \leq .001$) and indicating that at least one group differs from other groups. When scores were compared by phase pairwise (post-hoc analysis), only IPOS scores for the first group (Stable) could be significantly distinguished from other groups. AKPS scores were significantly different between all phases ($P \leq 0.001$), however. The 51 missing phase records were removed from further analysis. The ML and network analyses were conducted on the remaining 1383 patient records with full IPOS and phase values.

Fig. 1 provides the partial correlation network for the 10 IPOS items, irrespective of phase, and can be considered the 'reference' network for IPOS where correlations between pairs of items are generated with variances from all other variables controlled for. Similarly, Fig. 2 provides the partial correlation networks for each of Stable, Unstable, Deteriorating and Terminal. Fig. 3 provides the standardized centrality plots for each node in terms of betweenness, closeness and degree by phase.

Network analysis for all patients showed that Poor Appetite and Loss of Energy were central IPOS items, with Loss of Energy linked to Drowsiness, Shortness of Breath and Lack of Mobility on the one hand, and Poor Appetite linked to Nausea, Vomiting, Constipation and Sore and Dry Mouth on the other (Fig. 1). Loss of Energy and Poor Appetite are central to all phases, according to network analysis (Fig. 2), indicating their importance and influence on other symptoms.[36] Centrality measure of betweenness, closeness and degree all indicate that these two IPOS symptoms have the highest standardized values (just under +2.0) and form a coherent and cohesive substructure (Fig. 3). Network links show increasingly negative

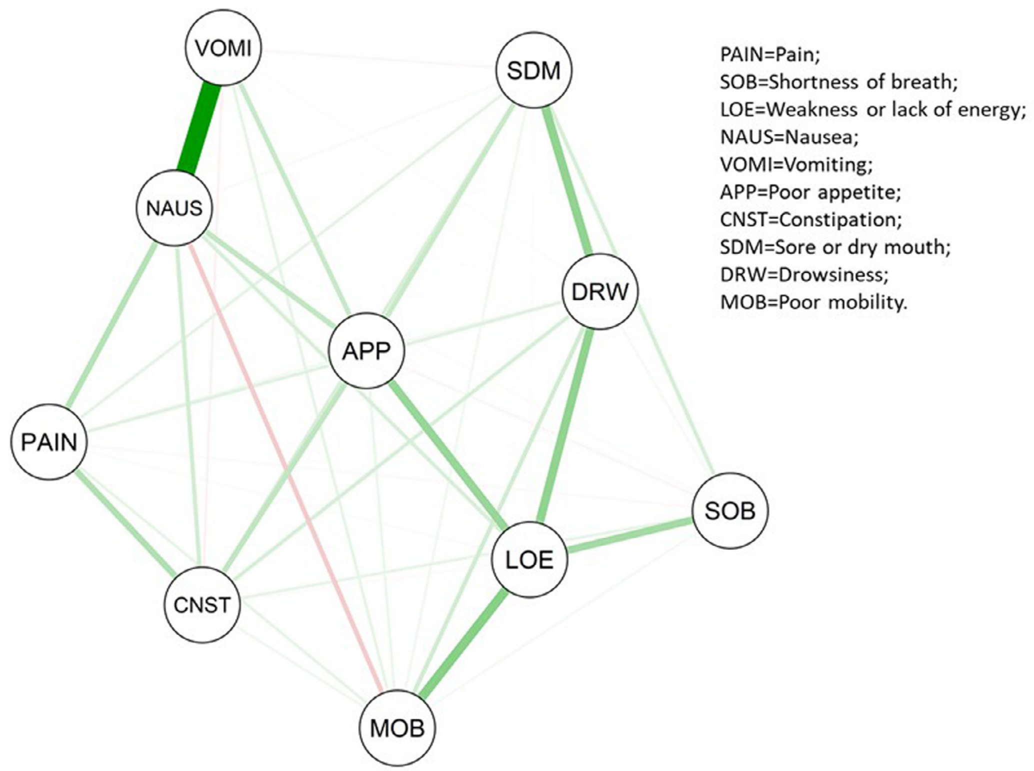


Fig. 1. Network for all 10 IPOS items using partial correlation (every pair of associations having the effects of all other items removed) and spring layout ($n = 1383$). Green edges signify positive associations and red edges negative. Thickness of line signifies strength of correlation.

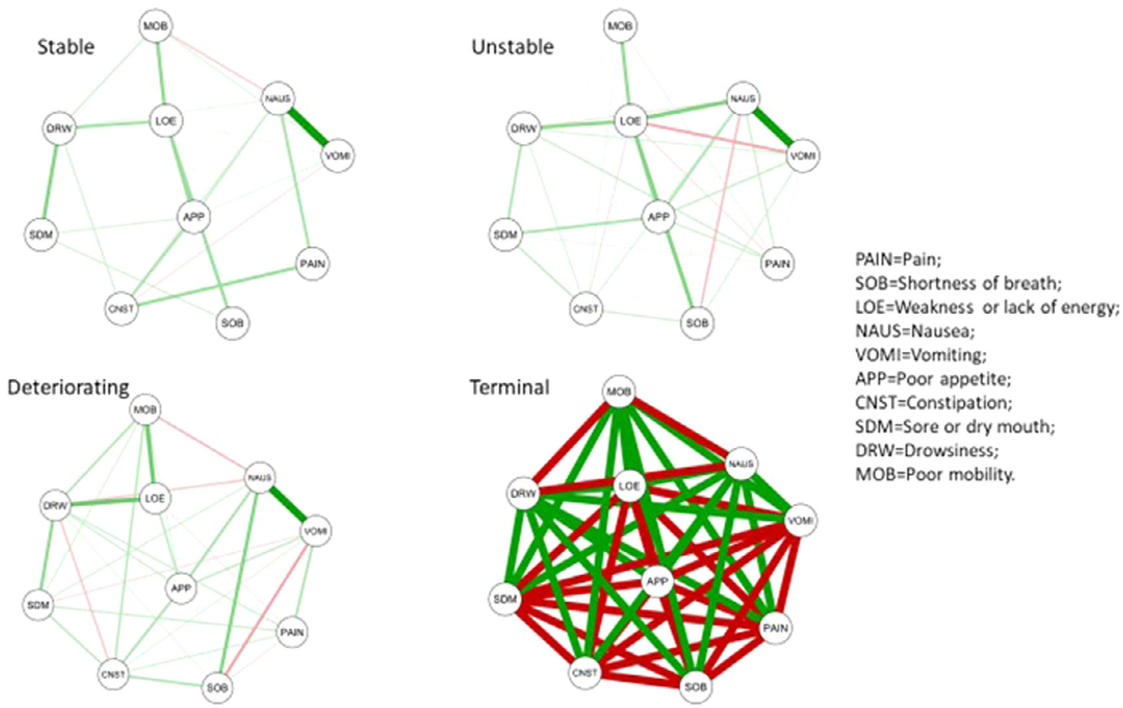


Fig. 2. Partial correlation network visualisations for Stable ($n = 948$), Unstable ($n - 291$), Deteriorating ($n = 124$) and Terminal ($n = 20$), using spring layout.

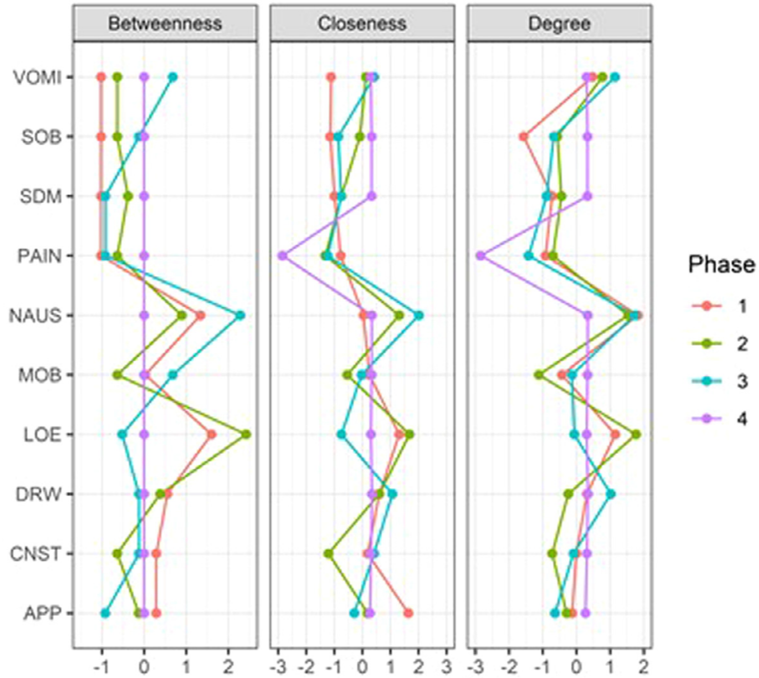


Fig. 3. Centrality plots for Betweenness, Closeness and Degree by phase for IPOS nodes in Fig. 2 (see Fig. 1 for item descriptions).

associations through phases, indicating that an increase in one symptom leads to a decrease in another symptom (Fig. 2).

*Machine Learning (ML)*

The 10 IPOS items and AKPS were used to predict phase through ML. After a series of initial experiments to determine suitable parameters for neural network learning, a two hidden-layer perceptron (10 input nodes, two layers of 50 hidden nodes, output layer of four nodes with one node per phase) was used for classifying and predicting phases (10-fold cross-validation) using the 11 variables. Subsequent experiments identified five other ML techniques for comparison (naïve Bayes, decision forest, ID3, logistic model tree, random forest). Given the imbalance of cases in the four phases, precision-recall curves (PRCs) were calculated in addition to ROC area under curve (AUC). Table 3 provides an overview of the final AUC and PRC figures using six ML techniques. A deeper ML exploration focusing on separating Stable from Unstable provided no significant improvement in the results.

## Discussion

Machine learning with the six applied techniques was moderately successful in being able to correctly predict cases falling within phases, with average ROC AUC figures of 0.639, 0.60, 0.627 and 0.724 for the four

*Table 3*
**Breakdown of IPOS Score and AKPS by Phase**

| Measure | Phase | *n* | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| *IPOS* | Stable | 948 | 9.38 | 5.51 | 0.18 | 9.03 | 9.74 |
| | Unstable | 291 | 12.99 | 6.09 | 0.36 | 12.29 | 13.69 |
| | Deteriorating | 124 | 14.18 | 6.52 | 0.59 | 13.02 | 15.34 |
| | Terminal | 20 | 16.35 | 5.74 | 1.28 | 13.66 | 19.04 |
| | Total | 1383 | 10.67 | 6.06 | 0.16 | 10.35 | 10.99 |
| *AKPS* | Stable | 941 | 64.06 | 14.64 | 0.48 | 63.12 | 65.00 |
| | Unstable | 287 | 59.34 | 14.21 | 0.84 | 57.69 | 60.99 |
| | Deteriorating | 122 | 47.38 | 16 | 1.47 | 44.46 | 50.29 |
| | Terminal | 20 | 12.00 | 11.52 | 2.58 | 6.61 | 17.39 |
| | *Total* | *1370* | *60.82* | *16.53* | *0.45* | *59.95* | *61.70* |

*Table 4*
**Outcomes for Six Machine Learning Methods for Predicting Phase Using ROC AUC (Area Under Curve) and PRC (Precision-recall Figures) Using 10-Fold Cross-Validation (total $n$ =1383)**

| ML Type | | ROC AUC | | | | PRC | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Stable ($n$ = 948) | Unstable ($n$ = 291) | Deteriorating ($n$ = 124) | Terminal ($n$ = 20) | Stable | Unstable | Deteriorating | Terminal |
| Perceptron | 10 input; 2 × 50 hidden; 4 output | 0.632 | 0.582 | 0.641 | 0.944 | 0.768 | 0.582 | 0.197 | 0.300 |
| Naïve Bayes | | 0.695 | 0.649 | 0.680 | 0.789 | 0.812 | 0.318 | 0.178 | 0.102 |
| Decision Forest PA | 10 trees with penalized attributes | 0.657 | 0.631 | 0.639 | 0.705 | 0.781 | 0.309 | 0.189 | 0.058 |
| ID3 | Unpruned decision tree | 0.554 | 0.503 | 0.519 | 0.515 | 0.706 | 0.215 | 0.096 | 0.016 |
| LMT | Logistic model trees | 0.645 | 0.621 | 0.634 | 0.656 | 0.790 | 0.291 | 0.157 | 0.030 |
| Random Forest | Forest of random trees | 0.651 | 0.614 | 0.648 | 0.732 | 0.781 | 0.276 | 0.207 | 0.039 |
| | | *0.639* | *0.600* | *0.627* | *0.724* | *0.773* | *0.332* | *0.171* | *0.091* |

stages of Stable, Unstable, Deteriorating and Terminal, respectively (Table 3). These figures provide preliminary support for the feasibility of ML in predicting stages in palliative care. However, PRC figures decreased from Stable to Terminal (from 0.724 to less than 0.1), leading to reduced relevance of the model for the later stages due to greater proportions of patients being in earlier palliative stages. Repeating the analysis with larger datasets with equal representation across phases may help to increase the relevance of ML for patients in these later stages.

The lack of underlying latent factors or components in IPOS is demonstrated by each item contributing equally and separately to the overall variance. The average and median AKPS scores of 61% − 60% are in line with previously reported figures for this measure.[28] ANOVA showed that IPOS total score and AKPS varied significantly by phase, with post-hoc analysis showing that this was confined to the Stable and Unstable groups. AKPS scores were significantly and negatively correlated with IPOS items (except Nausea), as expected, providing further assurance concerning the reliability of the data and their use in machine learning and network analysis.

While the AUC figures are not as high as previous ML for mortality prediction, this could be due to the non-homogeneity of the patient population in terms of disease. Nevertheless, the study has identified ML benchmarks for further research on phase prediction using IPOS items and AKPS. Experienced clinical researchers and caregivers also have difficulty in palliative care prognosis[37] indicating that there is still much to learn about how to combine the most appropriate measures and judgements into a reliable prediction framework concerning possible transition of phase.

The identification of clusters of central IPOS items using network analysis provides a direction for key symptoms to detect in future digital therapeutics in palliative care. Mobile apps and wearable technologies could focus on sensors dealing with the central symptoms found in this study of pain, weakness, mobility, nausea and shortness of breath. Regular (e.g., hourly) readings from these sensors can constitute automatic and non-intrusive methods for detecting possible change in level in advance of patient-reported outcomes or periodic clinical assessment. For instance, for patients considered stable, increases in associations between sensor readings for pain, weakness, nausea, poor mobility and loss of appetite could signal a possible change of level to unstable. For patients considered unstable, increases in association between shortness of breath, pain and weakness, as well as drowsiness and nausea, may be early indicators of transition to deteriorating. For patients considered deteriorating, further increases in association between weakness, poor mobility and drowsiness, and between vomiting and nausea, could be early indicators of transition to terminal.

Mobile and wearable sensors and systems already exist for mobility (steps taken and distance travelled), weight, calories consumed, blood pressure, sleep levels, balance and other range of motion analysis. Increased use of virtual reality headsets has led to the need to identify possible 'cybersickness' (nausea, vomiting) through physiological measurements of heart rate variability, blood pressure, galvanic skin reaction and breathing.[38] Heart rate variability in older adults has also been shown to correlate with physical health status.[39] Decreased movement and increased pain in the present study was associated with deterioration in phase. Decrease in movement could be detected early through wearable sensors and medical teams notified to allow early preventative intervention. Data gathered from these sensors and systems can feed into a continuously updated patient profile for use by both specialists and non-specialists in addition to their periodic monitoring activities. Such is the progress in these systems and technologies in digital therapeutics that the question for palliative care is not whether they will be used but when and how. How and if this technology should enter palliative care is an emerging discussion[9] and is a consideration based on philosophical stance and models of care. The present paper is intended to offer one of many potential tools for future researchers and clinicians to consider.

### *Limitations*

The data set is heterogeneous with regard to disease, with varying numbers in each category and a preponderance (over 70%) in oncological cancer. The generalizability of the modelling results to patients with other diseases and performance scores needs further exploration. Also, patients in the deteriorating and terminal phase constitute a relatively small 11.5% of the cohort, leading to possible bias in results towards patients in the other two phases limiting the ability to make predictions in the later phases. This drop-off in assessments in the later phases reflects the emphasis of the service focus on implementing measurement at initial intake assessments.[40] While some of the data is longitudinal in that some patients were measured more than once, the measurement intervals varied as did the number of measurements per patient. It was not possible to conduct an analysis of phase change over time in a consistent and reliable manner. Finally, the relatively small number of samples made any attempt to produce cross-validated models using ML difficult.

## Conclusion

The present study demonstrated a novel application of network analysis to understand changes of phase in patient conditions. Previous research had focused on survival prediction and mortality rates, and have analysed general patient health records rather than specific palliative care data. As far as we are aware, this is the first time that palliative-specific data have been exclusively used in network analysis and ML. Future research using larger samples is needed to explore the phase changing by diagnostic groups and the application of wearable devices in the palliative care context. Together with ML and network analysis, digital therapeutics has the potential to enhance clinical decision making to improve the quality of care across settings where palliative care is delivered.

## Author's Contribution

M.S. (Corresponding author). Conceptualised the project, oversaw data collection, supported data analysis and interpretation, led on drafting the manuscript.

E.H. Contributed to the implementation of the project, acquisition of data, and interpretation of the data.

A.N. Undertook the data analysis and interpretation, assisted drafting of the article and provided critical conceptual discussion.

M.H. Contributed to planning the project and interpretation of findings, input into drafts, provided a critical review of the article and approved the final version.

I.J.H. Contributed to planning and interpretation of the project and findings, critical analysis of the drafts of the paper and approved the final version.

R.S. Provided mentorship to [author one], supported the conceptualisation of the project, contributed to the critical analysis of the manuscript.

All authors provided intellectual content and approved the version for publication.

## Research Ethics and Patient Consent

The present research did not require specific ethical approval due to (1) patients signing a consent form that enabled the use of their data for quality improvement purposes, (2) data was anonymised prior to collection for the research, and (3) the Health and Disability Ethics Committee (Northern) provided a letter confirming that the research conducted was out of scope.

## Data Management and Data Sharing

Data may be obtained through contacting the corresponding author.

## Disclosures and Acknowledgments

The authors have no conflicts of interest to disclose.

This research received no specific grant from any funding agency in the public, commercial, or not for profit sectors.

We wish to acknowledge the support of [blinded] Community Palliative Care team and patients, in facilitating this research to be undertaken. We wish to thank [blinded] for efforts in supporting the implementation of PROMs into [blinded] and the broader study.

## References


1. Dudgeon D. The impact of measuring patient-reported outcome measures on quality of and access to palliative care. J Palliat Med 2018;21:S76–s80.

2. Peterson K, Anderson J, Bourne D. Evidence brief: use of patient reported outcome measures for measurement based care in mental health shared decision-making. VA evidence synthesis program evidence briefs. Washington (DC): Department of Veterans Affairs (US); 2011.

3. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. Cogn Behav Pract 2015;22:49–59.

4. Martin-Cook K, Palmer L, Thornton L, et al. Setting measurement-based care in motion: practical lessons in the implementation and integration of measurement-based care in psychiatry clinical practice. Neuropsychiatr Dis Treat 2021;17:1621–1631.

5. Gual-Montolio P, Martínez-Borba V, Bretón-López JM, Osma J, Suso-Ribera C. How are information and communication technologies supporting routine outcome monitoring and measurement-based care in psychotherapy? A systematic review. Int J Environ Res Public Health 2020;17:3170. https://doi.org/10.3390/ijerph17093170.

6. Higginson IJ, Brooks D, Barclay S. Dying at home during the pandemic. BMJ 2021;373:n1437.

7. Tohka J, van Gils M. Evaluation of machine learning algorithms for health and wellness applications: a tutorial. Comput Biol Med 2021;132:104324.

8. Capobianco E. On digital therapeutics. Front Digit Humaniti 2015;2:1–3. https://doi.org/10.3389/fdigh.2015.00006.

9. Kamdar M, Lakin J, Zhang H. Artificial intelligence, machine learning, and digital therapeutics in palliative care and hospice: the future of compassionate care or rise of the robots? (TH363). J Pain Symptom Manage 2020;59:434–435.

10. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019;380:1347–1358.

11. Storick V, O'Herlihy A, Abdelhafeez S, Ahmed R, May P. Improving palliative care with machine learning and routine data: a rapid review. HRB Open Res 2019;2:13. https://doi.org/10.12688/hrbopenres.12923.2.

12. Williams D, Hornung H, Nadimpalli A, Peery A. Deep learning and its application for healthcare delivery in low and middle income countries. Front Artif Intell 2021;4:553987. https://doi.org/10.3389/frai.2021.553987.

13. Avati A, Jung K, Harman S, et al. Improving palliative care with deep learning. BMC Med Inform Decis Mak 2018;18:122.

14. Wang L, Sha L, Lakin JR, et al. Development and validation of a deep learning algorithm for mortality prediction in selecting patients with dementia for earlier palliative care interventions. JAMA Network Open 2019;2. e196972-e196972.

15. Sheu Yh. Illuminating the black box: interpreting deep neural network models for psychiatric research. Front Psychiatry 2020;11:551299. https://doi.org/10.3389/fpsyt.2020.551299.

16. Elfiky AA, Pany MJ, Parikh RB, Obermeyer Z. Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy. JAMA Network Open 2018;1. e180926-e180926.

17. Parikh RB, Manz C, Chivers C, et al. Machine learning approaches to predict 6-month mortality among patients with cancer. JAMA Network Open 2019;2. e1915997-e1915997.

18. White N, Reid F, Harris A, Harries P, Stone P. A systematic review of predictions of survival in palliative care: how accurate are clinicians and who are the experts? PLoS One 2016;11. e0161407.

19. Mather H, Guo P, Firth A, et al. Phase of illness in palliative care: cross-sectional analysis of clinical data from community, hospital and hospice patients. Palliat Med 2018;32:404–412.

20. de Wolf-Linder S, Dawkins M, Wicks F, et al. Which outcome domains are important in palliative care and when? An international expert consensus workshop, using the nominal group technique. Palliat Med 2019;33:1058–1068.

21. Masso M, Allingham SF, Banfield M, et al. Palliative care phase: inter-rater reliability and acceptability in a national study. Palliat Med 2015;29:22–30.

22. Hearn J, Higginson IJ. Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. Palliative care core audit project advisory group. Qual Health Care 1999;8:219–227.

23. Siegert RJ, Gao W, Walkey FH, Higginson IJ. Psychological well-being and quality of care: a factor-analytic examination of the palliative care outcome scale. J Pain Symptom Manage 2010;40:67–74.

24. Collins ES, Witt J, Bausewein C, et al. A systematic review of the use of the palliative care outcome scale and the support team assessment schedule in palliative care. J Pain Symptom Manage 2015;50:842–853.

25. Schildmann EK, Groeneveld EI, Denzel J, et al. Discovering the hidden benefits of cognitive interviewing in two languages: the first phase of a validation study of the integrated palliative care outcome scale. Palliat Med 2016;30:599–610.

26. Murtagh F, Ramsenthaler C, Firth A, et al. A brief, patient- and proxy-reported outcome measure in advanced illness: validity, reliability and responsiveness of the integrated palliative care outcome scale (IPOS). Palliat Med 2019;33:1045–1057.

27. Sandham MH, Medvedev ON, Hedgecock E, Higginson IJ, Siegert RJ. A rasch analysis of the integrated palliative care outcome scale. J Pain Symptom Manage 2019;57:290–296.

28. Abernethy AP, Shelby-James T, Fazekas BS, Woods D, Currow DC. The Australia-modified karnofsky performance status (AKPS) scale: a revised scale for contemporary palliative care clinical practice. BMC Palliat Care 2005;4. 7-7.

29. Epskamp S, Rhemtulla M, Borsboom D. Generalized network psychometrics: combining network and latent variable models. Psychometrika 2017;82:904–927.

30. Hevey D. Network analysis: a brief overview and tutorial. Health Psychol Behav Med 2018;6:301–328.

31. De Schryver M, Vindevogel S, Rasmussen AE, Cramer AO. Unpacking constructs: a network approach for studying war exposure, daily stressors and post-traumatic stress disorder. Front Psychol 2015;6:1896.

32. Epskamp S, Fried EI. A tutorial on regularized partial correlation networks. Psychol Methods 2018;23:617–634.

33. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. Softw Pract Exp 1991;21:1129–1164.

34. Montazeri F, de Bildt A, Dekker V, Anderson GM. Network analysis of behaviors in the depression and autism realms: inter-relationships and clinical implications. J Autism Dev Disord 2020;50:1580–1595.

35. Roderick JAL. A test of missing completely at random for multivariate data with missing values. J Am Statist Assoc 1988;83:1198–1202.

36. Salavaty A, Ramialison M, Currie PD. Integrated value of influence: an integrative method for the identification of the most influential nodes within networks. Patterns (N Y) 2020;1:100052.

37. Chu C, White N, Stone P. Prognostication in palliative care. Clin Med (Lond) 2019;19:306–310.

38. Wang Y, Chardonnet JR, Merienne F. VR sickness prediction for navigation in immersive virtual environments using a deep long short term memory model. In: IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 2019. p. Osaka, Japan1874–1881.

39. Graham SA, Jeste DV, Lee EE, et al. Associations between heart rate variability measured with a wrist-worn sensor and older adults' physical function: observational study. JMIR mHealth and uHealth 2019;7. e13757-e13757.

40. Sandham MH, Hedgecock E, Palmer C, Siegert RJ, Narayanan A. Strengthening community service delivery through measurement based palliative care: a mixed methods implementation study. Contemp Nurse.