Special Article – Ειδικό Άρθρο

# *Cancer Patients Holistic Needs Assessment Test. Why Patients Engage. A Study of Drivers in Scotland*

***Raptis Sotirios****

*PhD*

*School of Design and Informatics, Abertay University, Dundee, Scotland Macmillan Cancer Support, United Kingdom*

**Correspondance: e-mail: sotiris.raptis.n@gmail.com*



**ABSTRACT**

**Cancer Patients Holistic Needs Assessment Test. Why Patients Engage. A Study of Drivers in Scotland**

**Raptis Sotirios**

There are approximately 1,000 new cancer cases diagnosed in the UK every day. Macmillan Cancer Support UK (MCS) offers persons affected by cancer a tool the called Holistic Needs Assessment Test (eHNA) to access more than 200 support services. The paper seeks to (1) find detailed non-drivers that make participants not engage in full, (2) compare metrics used to capture reasons for not answering in full, and extend them to compute the cost of the service.

The paper processed administrative, managerial, and clinical characteristics from published thesis on data kept by Macmillan and their histograms as the participants (patients) progressed from registration, to expressing their unmet needs after treatment and then to reaching a medical recommedation from Macmillan experts at the end of the test. The data used were very detailed characteristics histograms kept per stage per characteristic that were common to all eHNA stages. The paper checked for drivers using: (a) the Cross Correlation of the histograms, (b) the $x^2$ signifinace test between pairs of histograms, (c) the mere difference between pairs of histograms, (d) the Bayesian dependence of reaching the end on the common characteristics kept.

This study of characteristics was carried out in order to identify drivers of engagement and classified participants for their engagement progress, using nonclinical questions (test setting) to predict engagement for cancer sufferers. The study can help to shape policies of cancer treatment and found that those with comorbidities, with breast cancer, generally collaborating, taking pretest at NHSS hospitals, and following a specific treatment-management schema would likely reach the

end. Age and the intensity of a need to consult proved not determinants of a full engagement. Finally, around 20% of those registering reached a recommendation by participants.

**Keywords:** people affected by cancer (pabcs), electronic Hollistic Need Assessment test (eHNA), engagement driver, differences, significance test, Bayesian dependence, cross-correlations

## INTRODUCTION

The present work is a part of systematic study on a Macmillan organized set of questionnaires, called *electronic Hollistic Need Assessment (eHNA) test*.

This work innovates in that it was found that the quality of the treatment depends on the hospital where the treatment took place. Also, the intensity or the age does not seem to determine the degree of engagement as it would have been expected. The goals of eHNA are:

(a) highlight the unmet needs of people affected by cancer, (b) enable healthcare professionals to focus on those needs in a structured way, (c) enable appropriate services to meet those needs, and (d) aid the development of an individualized care plan (National Cancer Action Team).

The work analyses participants' responses to the questionnaires and covers a period from 2017 to 2020, and it is conducted across the UK.

There are approximately 1,000 new cancer cases diagnosed in the UK every day with a new diagnosis every two minutes on average. However, many cancer sufferers have a poor quality of life during and after treatment, Macmillan cite{mcs}[1] offers access to more than 200 support services, ranging from information provision to arranging access to specialist care. The work on eHNA data revealed that the type of the facility does drive the full engagement or not to the test.

With the eHNA, the participants initially register with the system (personal and medical data) in stage '*Registration'*, and are supported in completing an eHNA concern checklist cite {mm5}[2], cite {mmehna}[3], cite {mmchecklist}[4] with a Macmillan member in the stage of *'Expressing concerns'*. A Macmillan staff member later in stage 'Reaching a recommendation' determines a set of follow-up actions (information provision, care services such as dog walking, referral to specialist care, etc.

A substantial fraction of participants complete the *'Registration'* stage but do not complete the pathway to access the information on follow-up actions. The work in this paper sought to identify characteristics that discriminate among participants who complete the pathway, those who only register and those who progress part-way through, *'Reaching a recommendation'*.

The work looked into a range of methods to find fully engaging participants such as (1) eHNA stage count differences, (2) age/score stratified analysis, (3) similarity of counts (cross-correlations) before and after these stages, (4) the Bayesian dependence on non-engagement on the characteristics registered.

## MATERIALS AND METHODS

The participants' data have a natural segmentation that comprises the 3 subsets above. This degree of engagement is also statistically linked to capturing the so-called *'High-Resource Individuals'* or patients who finally follow complex treatments when this was not predicted before from their initial diagnosis. Overall, the participants expressed 57 different concerns that were allocated to the above 5 concern types. The intensity of need (part of the test) ranged from 1 to 10.

### *Participants' subsets and High-Resource Individuals*

The goal of capturing the full engagement is also linked to capturing the risky patients in terms of reaching an action at the end that incurs a cost for NHSS *(National Health Services Scotland).* The participants can be further divided in those in stage *'Registration'* but not progressing to those *'Expressing concerns'*, those in *'Registration*' and progress to *'Expressing concerns'*, those who finally those *'Expressing concerns'* but not *'Reaching a recommendation'* (leave) the test beforehand. The work analysed around 2000 answers, overall, across these subsets and used the statistics similarity tests to find fully engaging participants. Those who are at the end do have characteristics but other than statistics there is no scientifically established way to confirm these characteristics are indeed drivers.

### *Capturing the size and complexity of the data*

Macmillan data is a mix of needs, contextual, and clinical information. Each participant had multiple concerns. Each was a separate entry in the data and all assigned a percentage (probability) to be observed in the finalists. The reduction in participants from *'Registration'* to *'Expressing concerns'* and then to *'Reaching a recommendation'* was approximately half at each stage.

The participants tend to engage with any of the 33 eHNA questions $\{q_i, i \in (1,33)\}$, and respond differently choosing out of the 2000 available answers, in total. Roughly speaking, for any pair of questions $\{i, j \in (1,33)\}$, the frequencies of choosing single answers from the $N_i$ or $N_j$ available for the questions $i$ (or $j$) is usually much higher than the joint frequency $N_{i,j}=1/N_i * N_j$. Also, if one studies the combined counts of individual pairs of answers then the lower one (the rare answer) is a better discriminator for full engagement compared to an answer that is often observed (has a high count) and makes the joint probability very low across the participants. It is also likely that many questions have the same counts for their individual answers.

### *Methods*

Briefly, one counted how often (using a histogram) each answer is observed in all in eHNA like in *'Registration'* then after *'Expressing concerns'* and compared the same characteristics counts between *'Expressing concerns'* , and *'Reaching a recommendation'*. Also, one ap-

plied the same checks for the above participants subsets namely, "*Registration and Expressing concerns*", "*Registration without Expressing concerns*", "*Expressing concerns and reaching a recommendation*", "*Expressing concerns without reaching a recommendation*", "*Registration and Expressing concerns without reaching a recommendation*", "*Registration through reaching a recommendation*", "*Registration without reaching a recommendation*", "*Doing the full path*". That is, one examined those who make a step forward and those who do not. Due to the nature of Macmillan tests the mere counts of participants do not identify real persons but therapy sessions {called session (ID)s}. Hence, one can see more data in *'Expressing concerns'* than the participants initially in *'Registration'*. The histogram of the answers (per question) were taken either for combined questions (thus, more possible answers and then combinations of more answers per question) or for single questions (thus, single-variable histograms). Then, the differences were taken as in Fig.1 and Fig.2

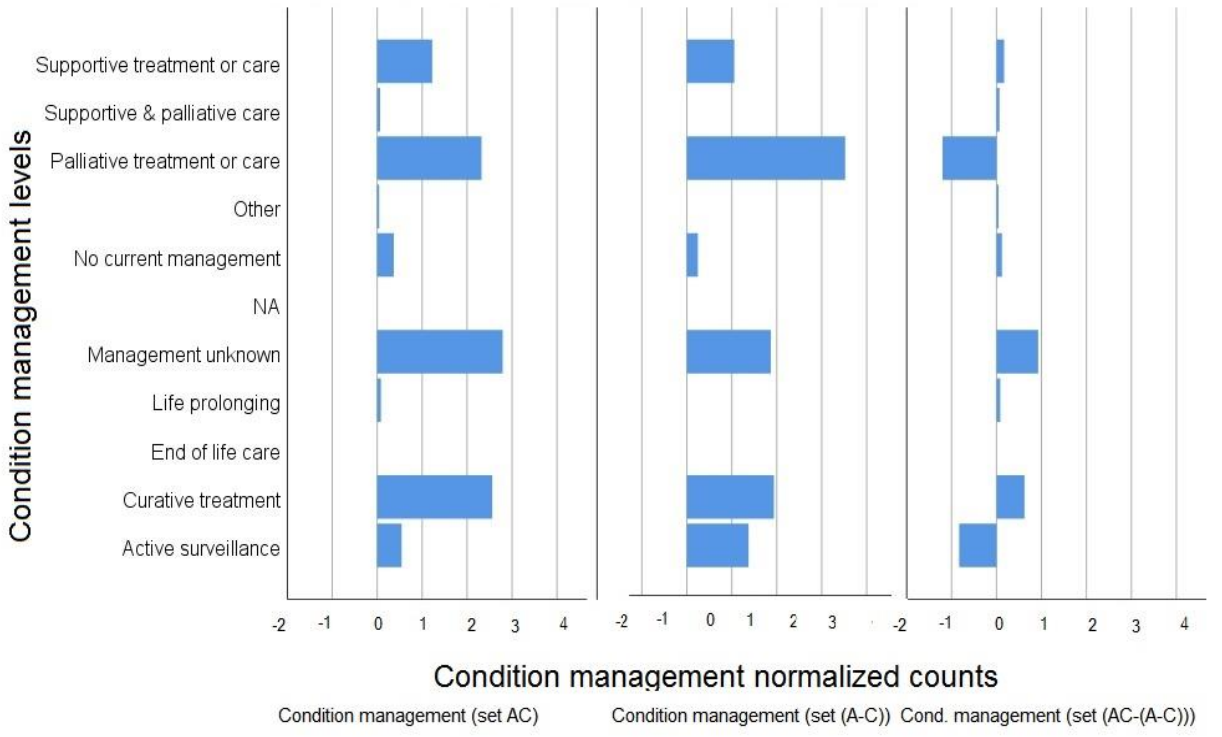


**Figure 1** The 3 plots are the histograms for those under a "*condition management"* schema (11 possible plans) for the groups *"Registration and Expressing concerns"* (1st plot), *"Registration without Expressing concerns"*(2nd plot) and the difference set (*"Registration and Expressing concerns"- "Registration without Expressing concerns"*). *The plot shows a nondriver since the last plot has many highly negative peaks indicating no full engagement.*

ISSN 1109-6888


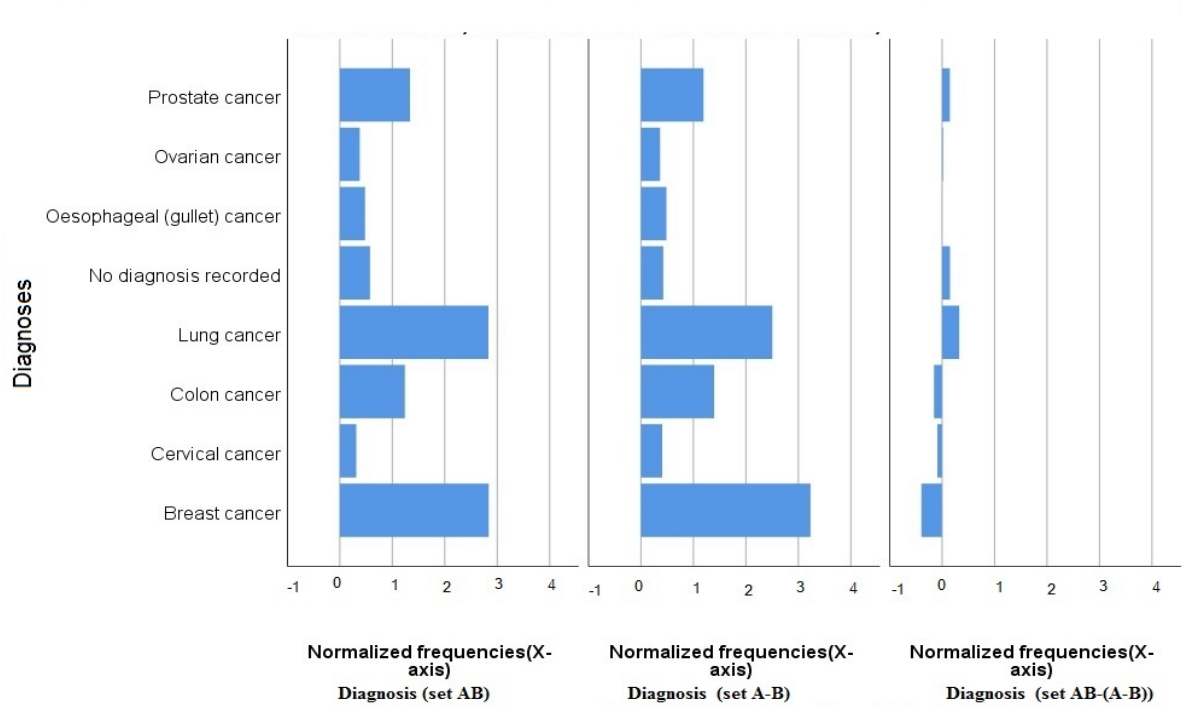


**Figure 2:** The 3 plots are the histograms for the first diagnosis for the groups *"Registration and Expressing concerns"* (1st plot), *"Registration without Expressing concerns"*(2nd plot) and the difference set (*"Registration and Expressing concerns"-"Registration without Expressing concerns"*). *The plot shows a driver since all plots are roughly even populated with no sharp differences indicating likely full engagement. The diagnosis cannot inform on who might answer all questions.*

Also, the cross-correlation was applied between single or combined histograms before and after reaching a certain eHNA subset. No results are given due to space limitations. The cross-correlations or the differences were further fed to significance tests. For example, the $x^2$ was applied to see if the histograms were really coming (sampled) from the same theoretical big sample (called initial participants population) before and after some eHNA stage. The full analysis conducted as well as the full results obtained cannot be presented for the economy of space.

The significance tests define a confidence interval to assume the NULL hypothesis holds, that is, two histograms came from the same distribution or not (the alternative hypothesis). If the so called *PV=p.value* that stems from these tests is *PV>0.05* then when considering a specific characteristic (question) with high PV the two populations (before and afte some eHNA stage) are alike and the participants do engage. Hence,

this characteristic is a candidate driver as a characteristic for becoming a *'High-Resource Individual'. If PV<0.05* it is the opposite and we have a non-driver. This would point to the fact that the specific question does not really capture participants' unmet needs.

The participants' age and score space was constructed by dividing the age and the score found in into 4 bands each. Thus we had 16 combinations (or $4^2$) as for example (*Age* $_{very\ low}$, *Age* $_{low}$, *Age* $_{high}$, *Age* $_{very\ high}$) and same for Score. Then one counted the participants or each of the sets (*Age* $_{some\ band}$, *Need* $_{some\ band}$)s who complete eHNA and then variably focused on a third different characteristic. This produced three-variable joint counts (age, need, third characteristic), that look like the 16 plots as the one in Fig. 3. This method showed that age and score do not play any specific role (either as drivers or as non-drivers).

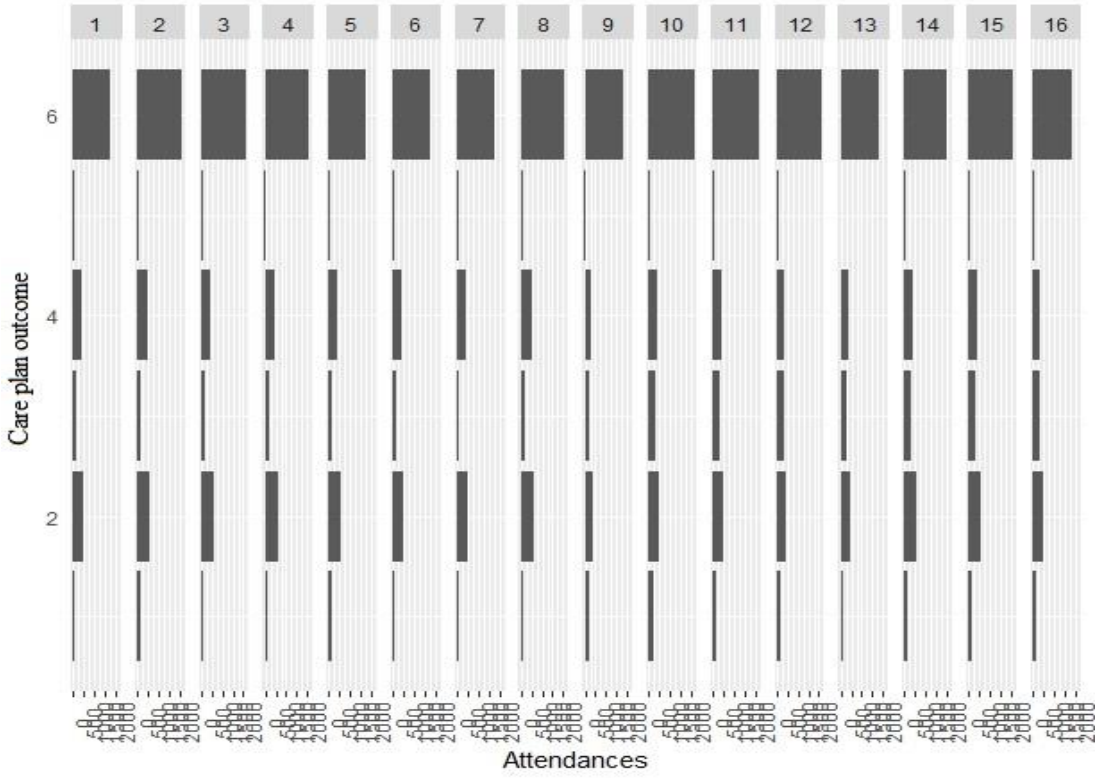


**Figure 3.** Age/score band subplots against the way the outome of the eHNA test is communicated to participants. *One can see 16 subplots each per single the 16 age/score bands or 4x4 combinations used. The plot shows that participants counts across all age/score band combinations do not change indicating that neither age nor score play a pivotal role in making engaging or not.*

***The engagement in a Bayesian context***

Bayesian approach was based on the idea of a *"characteristic's propagation"* and refers to answers of participants that are observed throughout the eHNA process. It was used to capture the questions *i* $\epsilon$*(1,33)/answers j* $\epsilon$*(1,* $N_{answer\ for\ i}$*)* or *answer* $_{i,\ j}$ that were drivers. The question that can arise is whether one can predict, for

newly registered participants in *'Registration'* and without knowing how they will progress, that some will quit the test. Normally, we cannot tell that from a single characteristic since the participants are not uniquely determined by some answer. The Bayesian method, though, can link in a probabilistic context *P {engage in full (or not)/ answer* $_{i,j}$*}* how a participant with a specific characteristic and when considered alone (age, region, gender, other) or in a specific eHNA setting defined for him (for example, emailed answers, taking the test from home, on-site test, other) will reach the end. Fig. 4 shows the separate Bayesian posteriors. These are finally weighted to yield the Bayesian belief for reaching the end. With these separate Bayesian posteriors the rest of the test is ignored and one links a single answer to the fact of '*Reaching a recommendation'*.

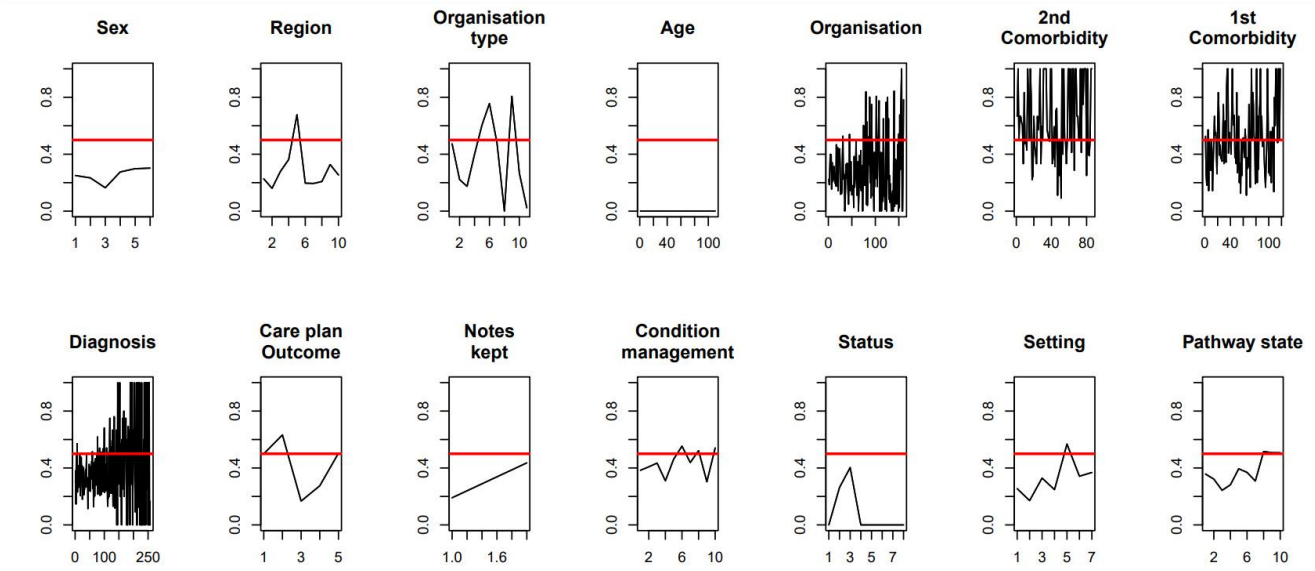


**Figure 4** The 14 subplots per major characteristic Bayesian posteriors P updated = P *(reaching the end/characteristic X* is observed) in completing participants. *The red line is the prior probability of quitting without knowledge of the characteristic (*$P_{prior}$*= 0.25) the x-axes are the possible levels (example gender has levels, "Male", "Female") per characteristic*.

## RESULTS AND DISCUSSION

The results from histogram differences and from age/score space studies are explained using graphs alone. The significance of the similarities *(p values)* of histograms cross-correlations are given in tabular format. Finally, the Bayesian approach is presented as single subplots of posteriors per characteristic.

No single characteristic was found to alone prevail among those in *'Reaching a recommendation'* or in differences while other characteristics

did not. One cannot define safely a single main driver that makes specific categories of participants engaging or not.

It is rather a combination (based on joint frequencies) that would likely point to fully engaging participants or to likely '*High-Resource Individuals'*. The method of differences provided better results comparing the above data sets in *"Registration and Expressing concerns"* and in *"Registration without Expressing concerns"* but not with the sets *"Expressing concerns then Reaching a recommendation"* and *"Expressing concerns without Reaching a recommendation"*.

The method considering the intensity of the need (called "score" by Macmillan) and the age proved less informative than expected and this informed that these characteristics are not crucial to determine full engagement. The method of cross-correlation and $x^2$ had the advantage of integrating significance tests. The cross-correlation was less granular than the differences. That is, as it sums up products of counts it cannot capture subtle peaks that make a difference prevail on an item-by-item basis (befores and afters).

| **Characteristics** | **p value** | **P:driver**<br>**NP: non-driver** |
|---|---|---|
| **Status** | 0.03 | NP |
| **Declined the test** | 0.36 | P |
| **Completed the test** | 1 | P |
| **Treatment follow-up model** | 0.47 | P |
| **Age** | 9.20E-011 | NP |
| **Talking to GP** | 0.1 | P |
| **Treatment management model** | 0.06 | P |
| **Hospital** | 2.20E-014 | NP |
| **Health condition** | 0.03 | NP |
| **Type of hospital** | 0.03 | NP |
| **Health condition management model** | 0.01 | P |
| **Notes kept** | 0.33 | P |
| **Communication of outcome** | 2.00E-003 | NP |
| **Interview setting** | 0.21 | P |
| **Region** | 0.01 | NP |

**Table 1.** The p.values are taken for those in set *"Registration and Expressing concerns" as compared to those in set "Registration without Expressing concerns" that is the first instance of non-engagement.*

The generic probability (prior) of engagement was roughly 25% {(participants in the end/participants registering) 100. The Bayesian posterior is the updated probability using the propagation rate of the characteristics for full engagement. It proved that it was 0.33. It was also observed that the comorbidities are trusted factors that drive participants to quit. One could also see that the primary diagnosis in registration could not lead with certainty to full engagement or one could not tell. One could also see that for those participants for whom the "Status" was known (that is, they were in a therapy protocol upon registration) were found almost certain they would quit. p. values are given in Table 1. For $PV \geq 0.05$ we have a driver and otherwise a non-driver. These are noted below as *NP (no-driver) and P (driver).*

## CONCLUSION

This study of characteristics was carried out in order to identify drivers of engagement and classified participants for their engagement progress using nonclinical questions (test setting) to predict engagement for cancer sufferers. The study can help to shape policies of cancer treatment and found that those with comorbidities, with breast cancer, generaly collaborating, taking pretest at NHSS hospitals, and following a specific treatment-management schema would likely reach the end. Age and the intensity of a need to consult proved not determinants of a full engagement. Finally, around 20% of those registering reached a recommendation by participants.

**Addittional materials:** No

**Acknowledgements:** The author is grateful to Macmillan Cancer Support, UK and especially to Mrs. Margaret Greer (Project Manager at Macmillan) for providing access to eHNA data and to Peter Gerry (Group Leader) for suggestions on how to work on these data. The author is also thankful to Dr. James Bown who devoted time and effort to supervise and revise the original thesis along with, Dr. Euan Dempster and Dr. Ann Savage, all from Abertay University, Dundee.

**Authors' contributions**: RS: manuscript preparation, primary case-management, literature review. The author read and approved the final manuscript.

**Funding:** Not applicable

**Availability of supporting data:** Not applicable.

**Ethical approval and consent to participate:** No IRB approval required.

**Competing interests:** The author declares no competing interests.

Received: November 2024, Accepted: November 2024, Published: December 2024.

## REFERENCES

1. Macmillan Cancer Support, UK, "Holistic Needs Assessments". Available at: https://www.macmillan.org.uk/healthcare

-professionals/innovation-in-cancer-care/holistic-needs-assessment.

2. Macmillan Cancer Support, UK, 'A guide for professionals providing holistic needs assessment care and support planning'. Available at: https://be.macmillan.org.uk/Downloads/CancerInformation/ResourcesForHSCP/InformationResources/MAC16266HNAHowtoguide2016A4DIGITALAW.pdf
3. Macmillan Cancer Support, UK, "Holistic Needs Assessments". Available at: https://www.macmillan.org.uk/healthcare-professionals/innovation-in-cancer-care/holistic-needs-assessment.
4. Macmillan Cancer Support, UK, "Concerns Checklist - identifying your concerns". Available at: https://www.macmillan.org.uk/dfsmedia/1a6f23537f7f4519bb0cf14c45b2a629/17780-10061/options/download/english-hna-concerns-checklist.

**Publisher's Note**

The publisher remains neutral with regard to jurisdictional claims in published maps and institutional afliations.

**Citation**: Raptis Sotirios. Cancer Patients Holistic Needs Assessment Test. Why Patients Engage. A Study of Drivers in Scotland. *Greek e j Perioper Med. 2024;23(d): 51-60.*