



## Original Article

## Need differences by treatment phases between patients with colorectal cancer and their caregivers: A text mining analysis

Jaehee Yoon<sup>a</sup>, Heesook Son<sup>b,\*</sup><sup>a</sup> Wolchon Elementary School, Yangcheon-gu, Seoul, Republic of Korea<sup>b</sup> Red Cross College of Nursing, Chung-Ang University, Dongjak-gu, Seoul, Republic of Korea

## ARTICLE INFO

## Keywords:

Colorectal cancer  
Survivors  
Needs  
Caregivers  
Text-mining

## ABSTRACT

**Objective:** To improve the quality of life and health outcomes of an increasing number of colorectal cancer patients and their caregivers, the specific needs of both should be understood across the cancer treatment trajectory. This study attempted to identify the supportive needs of patients and caregivers during each treatment phase.

**Methods:** Using text mining, posts from a South Korean internet self-help support group for colorectal cancer were classified by assigning labels to the terms extracted according to the poster, treatment phase, and support type. The extracted terms were then classified based on term frequency, with higher term frequency indicating higher need.

**Results:** Both caregivers and patients showed the highest term frequencies during the treatment phase. Caregivers had more high-ranking terms that were ranked higher than patients, indicating high support needs. In addition, the terms regarding treatments and diseases were ranked high. In the pre-treatment phase, test-related needs were ranked higher for both caregivers and patients. Meanwhile, patients had more highly ranked terms in the post-treatment phase, showing their high need in this phase, especially terms related to post-treatment symptoms. Caregivers had higher term frequencies of food intake during the post-treatment phase.

**Conclusions:** This study demonstrated that the needs of colorectal cancer patients and their caregivers were not always consistent in each treatment phase. Therefore, the needs of both patients and caregivers should be considered separately in each treatment phase and supportive interventions should be provided accordingly.

## Introduction

The 5-year relative survival rate of colorectal cancer (CRC) has consistently increased in South Korea from 54.8% in the 1990s to 75.6% in 2012–2016<sup>1</sup> while in the US, more than 90% of CRC patients in the stable stage showed high 5-year survival rates.<sup>2</sup> As the survival rate has increased, the number of CRC survivors that survived from diagnosis to post-treatment has risen. Further, the definition of *cancer survivors* also includes caregivers since cancer has various effects on both patients and their caregivers.<sup>3</sup> As medical care has shifted from hospital-based to community-based treatments, caregivers have become the primary resources for patients' social and emotional support, playing a critical role by providing most of the patients' cancer management. Thus, the importance of caregivers in the management of CRC patients has become increasingly recognized.<sup>4–8</sup>

Caregivers meet healthcare providers on behalf of patients when patients are unable to report their symptoms or management to

healthcare providers because of their condition, treatment, and surgery.<sup>7</sup> However, when caregivers represent patients, the caregivers' needs may differ from patients' practical needs, which can present notable challenges when the needs of patients and caregivers are diverse<sup>9,10</sup> resulting in the unmet needs of patients.

CRC caregivers may experience various psychological problems, including anxiety, depression, and psychological distress comparable to that observed in patients and physical difficulties because of the burden of caregiving.<sup>10–13</sup> Such challenges can negatively affect the quality of life and health outcomes of both patients and caregivers.<sup>6,12–14</sup> Therefore, both patients' and caregivers' needs should be considered along with appropriate supportive help to support CRC survivors effectively.<sup>6,15–17</sup>

Most studies have evaluated the needs of CRC survivors without distinguishing between patients and caregivers or have focused solely on either patients' or caregivers' needs. Further, the investigations have been fragmented as they have only examined needs at certain treatment

\* Corresponding author.

E-mail address: [hson@cau.ac.kr](mailto:hson@cau.ac.kr) (H. Son).<https://doi.org/10.1016/j.apjon.2022.03.013>

Received 21 February 2022; Accepted 31 March 2022

2347-5625/© 2022 The Authors. Published by Elsevier Inc. on behalf of Asian Oncology Nursing Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

stages, particularly in the treatment phase<sup>18</sup> hindering a comprehensive understanding of the needs of patients with CRC and their caregivers throughout the entire cancer trajectory.

As a big data analysis method for unstructured data,<sup>19</sup> a text-mining approach has mostly been used to analyze perceptions and trends of individuals using previously generated large-scale unstructured data. Text mining analysis has been shown to acquire useful information in medical research<sup>20,21</sup> and posts on the internet can be a valuable source of data to understand patients and develop healthcare services. Specifically, many CRC survivors have utilized internet support groups to share experiences and information as a resource for social support.<sup>13</sup> Therefore, text mining can provide an opportunity to analyze communications among CRC survivors naturally formed on the web,<sup>22</sup> but only a few active studies have applied text mining to the CRC field, such using text mining for trend analysis<sup>23</sup> or CRC experience analysis using descriptive and responsive data.<sup>24</sup> In the current study, we analyzed posts freely written on a South Korean CRC internet self-help support group using a text-mining approach to identify the support needs of patients and caregivers in each phase of their cancer trajectory.

**Methods**

This descriptive study analyzed posts on an internet self-help support group using a text-mining approach to identify the support needs of patients with CRC and their caregivers. The data analysis involved a three-step process: preprocessing, operations, and postprocessing<sup>25,26</sup> of text. The KoNLP and R word cloud packages (version 4.0.2) of the statistical program R were employed (<http://cran.rproject.org>, accessed on December 15, 2020).

*Data analysis*

*Step 1. Text preprocessing*

The data collection process yielded 2825 posts written from 2011 to 2022 from the largest (approximately 15,000 members) and most active internet self-help support group for CRC survivors in Republic of Korea. Two research assistants read posts and classified them by assigning a label based on the poster (ie, patient or caregiver) and treatment phase (ie, pre-treatment, treatment, and post-treatment). According to House's social support definition,<sup>27</sup> the identified support needs on the posts were labeled as *Informational Support Need* (ie, asking or sharing information), *Emotional Support Need* (ie, sharing emotional challenges or expressed emotion), and *Appraisal Support Need* (ie, relating to information that is useful for self-evaluation). The two research assistants independently assigned labels after reading each post multiple times. Single labels were assigned to each post for the posters, treatment phases, and support needs, whereas multiple labels were assigned when a post corresponded to one or more labels in each category. The two research assistants independently classified 2825 posts, of which 2597 were consistently classified. For the 228 posts with inconsistencies, discussions and decisions were made by all authors on the final labels.

During the data cleaning process, 11,927 nouns and adjectives were extracted from the text of the posts. Following a cleaning process that comprised noise removal and term unification, 4882 terms were included in the final analysis. Noise was defined as a term that was incompletely extracted with a difficult to understand meaning and of low importance (eg, size, body part, quantity and number, age, country, region, hospital, greetings, title). For text transformation and feature selection, the terms extracted through the cleaning process were converted into a matrix based on the term frequency (TF) for each label.

*Step 2. Text operations*

Classification analysis was performed based on term frequency (TF). High TF indicated an important term,<sup>28</sup> which was interpreted as a high social support need in this study. For each label, the 20 most frequent terms were analyzed, and supervised labeled latent Dirichlet allocation

(SL-LDA) was applied to analyze data with single and multiple labels. The 20 most frequent terms used by the patients and caregivers were visualized using word clouds, where a more frequently appearing word is visualized using a larger font size.<sup>29</sup>

*Step 3. Postprocessing*

The results were interpreted to determine what they meant statistically, discussed, and compared to previous studies.

**Results**

A total of 13,695 terms written by caregivers were ranked in the top 20 terms on the internet self-help support group for CRC survivors compared to the 9035 terms written by patients (Table 1). Fig. 1 shows that both caregivers and patients frequently used terms related to treatments and tests, including “surgery,” “chemotherapy,” “treatment,” and “examination.” For patients, terms related to defecation and pain, such as “anus,” “stool,” “colostomy,” and “pain,” appeared with higher ranks. For caregivers, terms related to the disease itself and hospitals such as “colorectal,” “cancer,” “metastasis,” and “hospital” had higher ranks.

Table 2 shows TFs for caregivers and patients by the treatment phases. Both groups had the highest TFs in the treatment phase, followed by the post-treatment and pre-treatment phases. In the pre-treatment phase, patients and caregivers showed similar TFs. However, in the treatment phase, caregivers’ had greater TF being twice more than patients. In contrast, patients had greater TF in the post-treatment phase compared to caregivers.

In the pre-treatment phase, patients and caregivers posted test-related terms most frequently, including “level,” “white blood,” “endoscopy,” and “result.” The terms mostly used by patients were “radiotherapy” and “anus,” while “surgery” was the highest-ranked for caregivers. In the pre-treatment phase, the terms that were high ranks for patients were those for negative emotions such as “anxiety” and “stress,” which was not observed among caregivers.

The top 20 terms in the treatment phase were those related to CRC treatments and adverse effects, including “chemotherapy,” “surgery,” “treatment,” “side effect,” “take medicine,” and “injection.” The term “pain” was ranked high for both patients and caregivers; however, patients posted the term more than caregivers. Meanwhile, “concern” was posted more by caregivers than by patients. Finally, “recurrence” was highly ranked only for the patients.

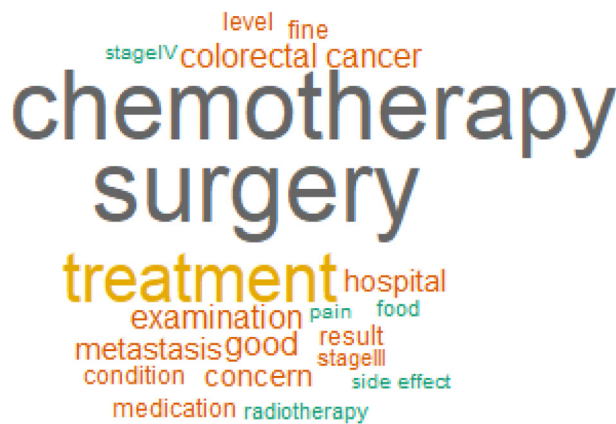
In the post-treatment phase, “surgery,” “examination,” and “recurrence” were ranked high for both patients and caregivers. For patients,

**Table 1**  
Term frequency ranks posted by patients and caregivers.

Rank	Patient (N = 9035)		Caregiver (N = 13,635)	
	Term	n	Term	n
1	Surgery	1563	Surgery	2545
2	Chemotherapy	752	Chemotherapy	2486
3	Examination	696	Treatment	1630
4	Treatment	668	Examination	629
5	Anus	547	Colorectal cancer	605
6	Stool	542	Good	577
7	Pain	398	Hospital	559
8	Colostomy	385	Metastasis	558
9	Good	379	Concern	540
10	Sick	364	Result	402
11	Endoscopy	303	Fine	383
12	Abdomen	303	Level	359
13	Result	295	Condition	349
14	Exercise	278	Medication	344
15	Colorectal cancer	268	Stage III	331
16	Concern	266	Radiotherapy	301
17	Medication	259	Food	278
18	Colon	258	Pain	259
19	Rectal cancer	256	Stage IV	251
20	Symptoms	255	Side effect	249



1-1. Patient



1-2. Caregiver

Fig. 1. Word cloud for patients (1-1) and caregivers (1-2).

terms related to pain and defecation, including “pain,” “symptoms,” “stool,” and “defecation,” were in higher ranks. While the term “exercise” was ranked high for patients, food intake-related terms such as “food,” “diet,” and “intake” intake were ranked higher for caregivers.

Table 3 presents the TF ranks according to social support. Both patients and caregivers used the high-frequency terms in reference to informational support the most, followed by appraisal and emotional support. For informational and emotional support, caregivers had greater TF than patients. In contrast, for appraisal support, patients had greater TF than caregivers. For both patients and caregivers, the highly ranked terms for informational support were related to CRC treatment, tests, and the disease itself, including “surgery,” “chemotherapy,” “examination,” “treatment,” and “colorectal cancer.” “Pain” was more highly ranked for patients than caregivers. The terms “symptoms” and “recurrence” were only identified in the higher ranks for patients.

Regarding emotional support, the highly ranked terms differed between patients and caregivers. For patients, highly ranked terms were related to recovery and health management, such as “exercise,” “health care,” “complete recovery,” “living,” “vitamin,” “food,” and “nutritional,” and a term related to emotion, “afraid,” was observed. In

contrast, for caregivers, highly ranked terms were related to treatments and adverse effects, such as “chemotherapy,” “surgery,” “colorectal cancer,” “vomit,” and “side effect,” and the emotion-related terms were “concern” and “stuffy.”

There were differences in the high-ranking terms used in relation to appraisal support between patients and caregivers. For patients, the highly ranked terms were those related to defecation, such as “stool,” “anus,” “colostomy,” “toilet,” “urge for defecation,” “gas,” and “bloody stool.” Caregiver’s highly ranked terms were those related to treatment, food intake, and test values, including “injection,” “take medicine,” “food,” “intake,” “eat,” “level,” and “white blood cell.”

Discussion

Differences between patients’ and caregivers’ desired and provided needs may lead to negative outcomes, making it important to provide tailored support resources during treatment.<sup>30</sup> This study provides insights into the support needs of patients and caregivers across different cancer trajectories.

We found that information support needs were ranked highest for both patients and caregivers, similar to the results of a study that analyzed an online community of cancer patients.<sup>22</sup> Terms related to CRC treatment, tests, and diseases were also highly ranked, similar to research showing a high need for treatment-related resources, treatment, unmet needs related to the cancer itself, and information.<sup>3,31</sup> Notably, there were some differences between caregivers and patients, with terms related to defecation and pain appearing in the higher ranks for patients. Such symptoms have significant implications for patients regarding the fear of recurrence and functional loss.<sup>32</sup> Thus, supportive interventions for patients’ symptoms should be actively provided.

While the term “recurrence” was posted by caregivers in the post-treatment phase, it was first used by patients in the treatment phase. Fear of recurrence does not persist for CRC patients but is induced by various events that are reminders of the possibility of cancer recurrence, such as symptom onset and hospital visits.<sup>33</sup> Treatment procedures involving various invasive treatments and symptoms may play a role in reminding patients of recurrence, which could be interpreted as patients’ recognition of treatments as active responses to reduce the risk of recurrence.<sup>33</sup> These findings suggest that supportive interventions to address fears related to cancer recurrence should be provided to patients in the early stages of their cancer treatment trajectory.

The need for cancer information has been shown to change across a patient’s cancer trajectory.<sup>34,35</sup> Similarly, we found a difference in the support needs of caregivers and patients across cancer trajectories.

In the pre-treatment phase, there were similar numbers of high-frequency terms for patients and caregivers, with test-related terms appearing mostly in high ranks. This may be because when patients undergo various tests in the diagnostic phase, both patients and caregivers seek additional information to understand the cancer diagnosis and make the best decision.<sup>34</sup>

Cancer patients manage the uncertainties that they commonly experience by obtaining information, helping them feel a sense of control over their lives.<sup>36,37</sup> Information-seeking behavior can be interpreted not only as a problem-focused coping strategy but also as an emotion-focused coping function.<sup>30</sup> In the diagnostic phase, cancer patients experience emotional challenges.<sup>10,36</sup> In this study, “anxiety” and “stress” were high-frequency terms used by patients, which might be interpreted as survivors’ emotion-focused coping to control their uncertainty. Thus, healthcare providers should provide sufficient information about various tests to caregivers and patients to mitigate the emotional challenges of CRC survivors.

It has been reported that patients experience many unmet needs in the treatment phase,<sup>38,39</sup> which was supported by the current finding that showed the highest term frequencies were found in the treatment phase.

**Table 2**  
Term frequency ranks by information providers and treatment phases.

Rank	Pre-treatment phase				Treatment phase				Post-treatment phase			
	Patient (n = 2142)		Caregiver (n = 2468)		Patient (n = 4478)		Caregiver (n = 10,281)		Patient (n = 3893)		Caregiver (n = 2638)	
	Term	n	Term	n	Term	n	Term	n	Term	n	Term	n
1	Radiotherapy	487	Colon	276	Chemotherapy	988	Chemotherapy	2171	Surgery	678	Surgery	452
2	Anus	430	Endoscopy	271	Surgery	817	Surgery	1977	Examination	516	Examination	258
3	Level	359	Examination	204	Treatment	519	Treatment	1379	Exercise	278	Concern	204
4	Normal	107	Tissue	190	Pain	192	metastasis	446	Treatment	191	Fine	196
5	White blood cell	92	Surgery	184	Good	172	Good	440	Pain	167	Good	186
6	Hemorrhoids	66	Result	133	Medication	150	Hospital	408	Symptoms	167	Treatment	136
7	Visit hospital	65	Colorectal cancer	128	Examination	148	Colorectal cancer	390	Food	163	Chemotherapy	126
8	Red ginseng	62	Abdominal pain	118	Colectomy	135	Concern	316	Abnormal	154	Colorectal cancer	125
9	Urine	54	Emergency room	118	Result	130	Colostomy	310	Health	152	Food	115
10	CT (Computer tomography)	53	Polyp	113	Take medicine	126	Examination	309	Concern	151	Diet	105
11	Strain	53	Removal	107	Colorectal cancer	121	Medication	305	Stool	151	Intake	100
12	Stage	48	Detection	93	Side effect	117	Condition	260	Rectal cancer	146	Result	88
13	Bleeding	43	Vomit	82	Severe	115	Result	244	Chemotherapy	145	Hospital	87
14	Permanent	40	Treatment	73	Concern	110	Pain	240	Endoscopy	145	Recurrence	84
15	Laparotomy	37	Advanced cancer	68	Rectal cancer	109	Stage III	193	Recurrence	123	Abnormal	69
16	Anxiety	32	Sleep	65	Injection	109	Take medicine	193	Defecation	120	Symptoms	69
17	Belly button	31	Digestion	64	Metastasis	108	Side effect	183	Hospital	113	Condition	61
18	Anastomotic site	29	Lymph node	63	Hospital	105	Severe	180	Medication	112	Visit hospital	60
19	Stage I	27	Hospital	59	Stage III	105	Possibility	170	Care	112	Problem	60
20	Stress	27	Sound	59	Recurrence	102	Stage IV	167	Colorectal cancer	109	Home	57

**Table 3**  
Term frequency ranks by information providers and types of social support.

Rank	Informational support				Emotional support				Appraisal support			
	Patient (n = 7612)		Caregiver (n = 11,880)		Patient (n = 1629)		Caregiver (n = 1978)		Patient (n = 3599)		Caregiver (n = 2073)	
	Term	n	Term	n	Term	n	Term	n	Term	n	Term	n
1	Surgery	1523	Surgery	2380	Exercise	237	Chemotherapy	515	Stool	542	Level	359
2	Chemotherapy	740	Chemotherapy	1971	Good	230	Surgery	165	Anus	456	Food	278
3	Examination	696	Treatment	1573	Heart	130	Colorectal cancer	131	Colostomy	385	Fine	277
4	Treatment	668	Examination	629	Health	124	Concern	107	Abdomen	303	Injection	213
5	Pain	398	Hospital	480	Care	112	Stage III	104	Toilet	195	Good	134
6	Endoscopy	303	Colorectal cancer	474	Body	102	Heart	98	Urge for defecation	178	Intake	100
7	Result	295	Metastasis	472	Colorectal cancer	68	Metastasis	86	Reconstruction	177	White blood Cell	92
8	Concern	266	Concern	433	Be diagnosed	67	Hospital	79	Blood	176	Eat	70
9	Medication	259	Good	409	Living	67	Stage IV	79	Defecation	176	Normal	62
10	Colon	258	Result	402	Complete recovery	66	Vomit	71	Sick	174	Suffering	59
11	Rectal cancer	256	Condition	349	Cold	53	Side effect	66	Gas	144	Low	58
12	Symptoms	255	medication	344	Vitamin	53	Doctor	63	Abdominal pain	118	Take medicine	57
13	Recurrence	241	radiotherapy	301	Prevention	48	Erbix	61	Constipation	99	Water	56
14	Abnormal	219	Pain	259	Diagnosis	44	Treatment	57	Bloody stool	86	Experience	42
15	Radiotherapy	216	Colon	247	Precaution	42	Cancer cells	54	Water	80	Immunity	41
16	Hospital	213	Possibility	238	Afraid	41	Ascites	52	Diet	68	Urine	38
17	Severe	209	Visit hospital	232	Surgery	40	Stuff	50	Hemorrhoids	66	Anticancer injection	36
18	Colorectal cancer	200	Rectal cancer	230	Home	37	Admission	49	Fine	62	Energy	34
19	Take medicine	199	Abnormal	230	Food	34	Body	47	Sound	59	Relieve	34
20	Rectal	198	Stage III	227	Nutritional supplements	34	Family	44	Sleep	55	Aspirin	33

Interestingly, caregivers had more than twice the number of terms in the treatment phase than patients, indicating that caregivers might have higher support needs. Similarly, another study revealed that 78% of caregivers showed behavior that searched for information besides that provided by healthcare providers.<sup>3</sup> Thus, one notable behavior patterns for CRC caregivers is information seeking. Furthermore, this finding indicates that caregivers recognize that information is critical to help

patients make the best decisions without compromising their autonomy. Caregivers recognize coping with patients' physical and emotional pain as a key challenge in the treatment phase. As caregiving burdens greatly increase in the treatment phase, they want sufficient information to effectively support patients.<sup>3,14,40,41</sup> However, despite this increase in their need for information, caregivers may not receive sufficient information from healthcare providers or understand the provided

information.<sup>3</sup> Caregivers prefer information provided by healthcare providers but tend to be passive in asking for desired information.<sup>3,42</sup>

In contrast to caregivers, patients have more support needs during the treatment process, and CRC treatment-related information needs are met within the healthcare system.<sup>43</sup> Caregivers may have easier access to an internet self-help support group than patients that face challenges related to hospitalization, treatment, and adverse treatment effects.<sup>10</sup> Furthermore, they may lack the opportunity to actively find information in internet self-help support groups. Therefore, strategies to improve information accessibility in the healthcare system should be developed to meet caregivers' informational needs since caregivers are the primary source of support for CRC patients and play a central role in helping patients integrate and digest information.<sup>41</sup>

The post-treatment phase was the only phase in which patients had more frequent terms than caregivers. These results differed from those of Ehemann et al that reported patients stopped actively seeking cancer information after treatment.<sup>34</sup> Since unmet needs were cancer-specific in the post-treatment phase,<sup>44</sup> this difference might be attributed to CRC-specific support needs. Unlike other cancer patients, CRC patients experience various adverse effects, such as changes in their bowel function after treatment and challenges adapting to daily life.<sup>45-47</sup> Such challenges can lead to unmet needs, resulting in behaviors like actively seeking information in online self-help support groups.

In addition, the terms related to delayed symptoms were highly frequent terms for patients, showing that patients with CRC have high support needs related to various post-treatment symptoms. This is consistent with research reporting that many CRC patients feel abandoned because they lacked satisfactory supportive care from healthcare providers in the post-treatment phase.<sup>33</sup> Moreover, most negative experiences of CRC patients are due to a lack of post-treatment care.<sup>24</sup> Thus, supportive interventions that manage delayed symptoms of CRC patients in the post-treatment phase should be actively provided to help patients adapt to their changed circumstances.

In the post-treatment phase, there were clear differences in the types of high-frequency terms used by patients and caregivers. Nutrition, diet, and bowel activity have been reported as major unmet information needs for CRC patients.<sup>43</sup> In this study, "bowel activity" was a high-frequency term for patients, while "nutrition" and "diet" were high-frequency terms for caregivers. For appraisal support, indicating the information needed to evaluate health conditions,<sup>48</sup> the highly ranked terms were related to bowel activity for patients and to nutrition and diet for caregivers. These results indicate that patients and caregivers have different support needs in the post-treatment phase. Patients may experience defecation symptoms, primarily in the post-treatment phase. They consider these symptoms as the standard for evaluating their health status, which increases patients' support needs to manage and cope with such symptoms in the post-treatment phase.<sup>43,45,47</sup> While CRC patients either lose appetite or feel uncomfortable with food intake due to intestinal stomas or changes in colonic function, caregivers also recognize higher appetite as a sign of recovery.<sup>33</sup> Therefore, increases in caregivers' support needs may be owed to their desire to provide healthy food to patients. These differences suggest that CRC patients and caregivers should have access to different supportive interventions in the post-treatment phase. Patients with CRC should be supported with programs that manage and adapt to their changes in bowel function, while CRC caregivers could benefit from programs that address preparing or providing healthy and easy-to-eat foods for patients with CRC.

#### Limitations

This study had some limitations. First, we focused on frequently used terms in the social media posts of CRC survivors to analyze the supportive needs of patients and caregivers, which cannot provide an in-depth understanding of the experiences of the patients and caregivers. In addition, interviewing CRC patients and caregivers about the results of text mining in the postprocessing step is needed to strengthen the validity.<sup>25</sup> Further

studies that used methods such as focus group interviews are needed to better understand the needs of patients with CRC and their caregivers. Second, the posts written by caregivers were not identified according to their relationship to the patient (eg, spouse, child, other family members), and there may be differences in needs between the patients' spouses/partners and other family members. For instance, the spouses of patients with breast cancer have been found to lack the necessary information about sexual attractiveness, but family members, such as daughters, were more concerned about their mother losing hair and breast loss during the treatment.<sup>49</sup> Identifying the differences between the spouses/partners of CRC patients and family members' needs would be valuable for providing more tailored interventions in the future. Third, this study could not identify the ages of the patients or their caregivers that posted in the self-help group. Given the finding that internet use and the age of cancer survivors are related,<sup>50</sup> additional research is needed to determine whether there are significant differences in supportive needs based on the age of the posters. Finally, selection bias may exist in the study's data because the individuals that posted on this internet self-help group may be more motivated than those not engaged in this group, limiting the generalizability of the results to other CRC patients and caregivers.

#### Conclusions

Supportive interventions based on the needs of CRC survivors can efficiently help improve their quality of life. However, few studies have examined the differences and similarities in patients' and caregivers' needs across cancer trajectories. The significance of this study is that it provided a comprehensive understanding of the specific needs of CRC patients and caregivers across cancer trajectories. In particular, this study uses a novel approach of using the internet posts written by CRC survivors to needs analysis. The posts were recently, naturally, and increasingly formed on the Internet and were considered important for investigation.

We found that the needs of patients and caregivers were not always similar, depending on the treatment phase, thereby, highlighting the importance of distinguishing the needs of both parties and providing supportive interventions accordingly. In particular, these data suggest the necessity of supportive interventions in which healthcare systems actively provide caregivers with satisfactory levels of information about disease treatment and the disease itself, as shown by the high support needs of the Internet self-help support group. Finally, programs that address the management of delayed symptoms for patients and help caregivers prepare food for patients in the post-treatment phase.

Moreover, these results can be used as fundamental data to provide supportive interventions tailored to the needs of patients and caregivers in each treatment phase. The similarities and differences in the caregivers' and patients' needs demonstrate the necessity of analyzing the specific needs of both groups across the cancer treatment trajectory. Future studies should investigate these needs using methods such as surveys and interviews and compare them with the results of this study.

#### Funding

This work was supported by a National Research Foundation of Korea Grant funded by the Korean Government (Grant No. NRF-2019R1A2C1006716).

#### Declaration of competing interest

None declared.

#### Authors' contributions

Conceived and designed the analysis: Jaehee Yoon and Heesook Son; Collected the data: Jaehee Yoon and Heesook Son; Contributed data or

analysis tools: Jaehee Yoon; Performed the analysis: Jaehee Yoon and Heesook Son; Wrote the paper: Jaehee Yoon and Heesook Son.

## References

- Jung K-W, Won Y-J, Kong H-J, Lee ES. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2016. *Cancer Res Treat*. 2019;51(2):417–430. <https://doi.org/10.4143/crt.2019.138>.
- Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA A Cancer J Clin*. 2019;69(1):7–34. <https://doi.org/10.3322/caac.21551>, 2019.
- Chua GP, Ng QS, Tan HK, Ong WS. Caregivers of cancer patients: what are their information-seeking behaviours and resource preferences? *Ecancermedicalscience*. 2020;14:1068. <https://doi.org/10.3332/ecancer.2020.1068>.
- Ferrell B, Wittenberg E. A review of family caregiving intervention trials in oncology. *CA A Cancer J Clin*. 2017;67(4):318–325. <https://doi.org/10.3322/caac.21396>.
- Girgis A, Lambert SD, McElduff P, et al. Some things change, some things stay the same: a longitudinal analysis of cancer caregivers' unmet supportive care needs. *Psycho Oncol*. 2013;22(7):1557–1564. <https://doi.org/10.1002/pon.3166>.
- Given BA, Given CW, Sherwood P. The challenge of quality care for family caregivers. *Semin Oncol Nurs*. 2012;28(4):205–212. <https://doi.org/10.1016/j.soncn.2012.09.002>.
- Havner RD, van Ryn M, Wilson PM, Bangerter LR, Griffin JM. Concordance of patient and caregiver reports on the quality of colorectal cancer care. *J Oncol Pract*. 2019; 15(11):e979–e988. <https://doi.org/10.1200/JOP.19.00136>.
- Maguire R, Hanly P, Hyland P, Sharp L. Understanding burden in caregivers of colorectal cancer survivors: what role do patient and caregiver factors play? *Eur J Cancer Care*. 2018;27(1), e12527. <https://doi.org/10.1111/ecc.12527>.
- McMullen CK, Schneider J, Altschuler A, et al. Caregivers as healthcare managers: health management activities, needs, and caregiving relationships for colorectal cancer survivors with ostomies. *Support Care Cancer*. 2014;22(9):2401–2408. <https://doi.org/10.1007/s00520-014-2194-3>.
- Mosher CE, Adams RN, Helft PR, et al. Family caregiving challenges in advanced colorectal cancer: patient and caregiver perspectives. *Support Care Cancer*. 2016; 24(5):2017–2024. <https://doi.org/10.1007/s00520-015-2995-z>.
- Graham J, Spiliopoulou P, Arbutckle R, Bridge JA, Cassidy J. A pilot study of subjective well-being in colorectal cancer patients and their caregivers. *Patient Relat Outcome Meas*. 2017;8:111–119. <https://doi.org/10.2147/PROM.S141815>.
- Makumi J. *Stress Reactions, Coping Strategies and Counselling Needs of Family Caregivers of Colorectal Cancer Survivors*. Aga Khan University Hospital Nairobi Kenya; 2017.
- Skalla KA, Smith EM, Li Z, Gates C. Multidimensional needs of caregivers for patients with cancer. *Clin J Oncol Nurs*. 2013;17(5):500–506. <https://doi.org/10.1188/13.CJON.17-05AP>.
- Rha SY, Park Y, Song SK, Lee CE, Lee J. Caregiving burden and the quality of life of family caregivers of cancer patients: the relationship and correlates. *Eur J Oncol Nurs*. 2015;19(4):376–382. <https://doi.org/10.1016/j.ejon.2015.01.004>.
- Ferrell B, Hanson J, Grant M. An overview and evaluation of the oncology family caregiver project: improving quality of life and quality of care for oncology family caregivers. *Psycho Oncol*. 2013;22(7):1645–1652. <https://doi.org/10.1002/pon.3198>.
- Luo X, Li J, Chen M, Gong J, Xu Y, Li Q. A literature review of post-treatment survivorship interventions for colorectal cancer survivors and/or their caregivers. *Psycho Oncol*. 2021;30(6):807–817. <https://doi.org/10.1002/pon.5657>.
- Santin O, Murray L, Prue G, Gavin A, Gormley G, Donnelly M. Self-reported psychosocial needs and health-related quality of life of colorectal cancer survivors. *Eur J Oncol Nurs*. 2015;19(4):336–342. <https://doi.org/10.1016/j.ejon.2015.01.009>.
- van Mossel C, Leitz L, Scott S, et al. Information needs across the colorectal cancer care continuum: scoping the literature. *Eur J Cancer Care*. 2012;21(3):296–320. <https://doi.org/10.1111/j.1365-2354.2012.01340.x>.
- Zhang Y, Chen M, Liu L. A Review on Text Mining. Paper Presented at: ICSESS 2015. 6th IEEE International Conference on Software Engineering and Service Science. 2015 Sep 23-25; Beijing, China. doi: 10.1109/ICSESS.2015.7339149.
- Lafferty NT, Manca A. Perspectives on social media in and as research: a synthetic review. *Int Rev Psychiatr*. 2015;27(2):85–96. <https://doi.org/10.3109/09540261.2015.1009419>.
- Lee YJ, Donovan H. Application of text mining in cancer symptom management. *Stud Health Technol Inf*. 2016;225:930–9311. <https://doi.org/10.3233/978-1-61499-658-3-930>.
- Verberne S, Batenburg A, Sanders R, van Eenbergen M, Das E, Lamboij MS. Analyzing empowerment processes among cancer patients in an online community: a text mining approach. *JMIR Cancer*. 2019;5(1):e9887. <https://doi.org/10.2196/cancer.9887>.
- Soffer S, Klang E, Tau N, et al. Evolution of colorectal cancer screening research in the past 25 years: text-mining analysis of publication trends and topics. *Therapeutic advances in gastroenterology*. 2020;13:1–12. <https://doi.org/10.1177/1756284820941153>, 2020.
- Wagland R, Recio-Saucedo A, Simon M, et al. Development and testing of a text-mining approach to analyse patients' comments on their experiences of colorectal cancer care. *BMJ Qual Saf*. 2016;25(8):604–614. <https://doi.org/10.1136/bmjqs-2015-004063>.
- Kobayashi VB, Mol ST, Berkers HA, Kismihók G, Den Hartog DN. Text mining in organizational research. *Organ Res Methods*. 2018;21(3):733–765. <https://doi.org/10.1177/1094428117722619>.
- Inzalkar S, Sharma J. A survey on text mining-techniques and application. *Int J Renew Sustain Energy*. 2015;24:1–14.
- House J. *Work Stress, and Social Support*. Reading, MA: Addison-Wesley; 1981.
- Salton G, Buckley C. Term-weighting approaches in automatic text retrieval. *Inf Process Manag*. 1988;24(5):513–523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0).
- Jockers ML, Thalken R. *Text Analysis with R for Students of Literature*. New York: Springer; 2014.
- Germeni E, Bianchi M, Valcarengi D, Schulz PJ. Longitudinal qualitative exploration of cancer information-seeking experiences across the disease trajectory: the INFO-SEEK protocol. *BMJ Open*. 2015;5(10), e008933. <https://doi.org/10.1136/bmjopen-2015-008933>.
- Watanabe K, Katayama K, Yoshioka T, Narimatsu H. Impact of individual background on the unmet needs of cancer survivors and caregivers—a mixed-methods analysis. *BMC Cancer*. 2020;20:263. <https://doi.org/10.1186/s12885-020-06732-5>.
- McGeechan GJ, McPherson KE, Roberts K. An interpretative phenomenological analysis of the experience of living with colorectal cancer as a chronic illness. *J Clin Nurs*. 2018;27(15-16):31483156. <https://doi.org/10.1111/jocn.14509>.
- McGeechan GJ, Byrnes K, Campbell M, et al. A systematic review and qualitative synthesis of the experience of living with colorectal cancer as a chronic illness. *Psychol Health*. 2021. <https://doi.org/10.1080/08870446.2020.1867137>. Epub 2021 Jan 26.
- Eheman CR, Berkowitz Z, Lee J, et al. Information-seeking styles among cancer patients before and after treatment by demographics and use of information sources. *J Health Commun*. 2009;14(5):487–502. <https://doi.org/10.1080/10810730903032945>.
- Vogel BA, Bengel J, Helmes AW. Information and decision making: patients' needs and experiences in the course of breast cancer treatment. *Patient Educ Counsel*. 2008; 71(1):79–85. <https://doi.org/10.1016/j.pec.2007.11.023>.
- Browne S, Dowie A, Mitchell E, et al. Patients' needs following colorectal cancer diagnosis: where does primary care fit in? *Br J Gen Pract*. 2011;61(592):e692–e699. <https://doi.org/10.3399/bjgp11X606582>.
- Miller LE. Uncertainty management and information seeking in cancer survivorship. *Health Commun*. 2014;29(3):233–243. <https://doi.org/10.1080/10410236.2012.739949>.
- Lim BT, Butow P, Mills J, Miller A, Goldstein D. Information needs of the Chinese community affected by cancer: a systematic review. *Psycho Oncol*. 2017;26(10): 1433–1443. <https://doi.org/10.1002/pon.4347>.
- Puts MT, Papoutsis A, Springall E, Tourangeau AE. A systematic review of unmet needs of newly diagnosed older cancer patients undergoing active cancer treatment. *Support Care Cancer*. 2012;20(7):1377–1394. <https://doi.org/10.1007/s00520-012-1450-7>.
- Kaynar GÖ, Vural F. Assessment of the caregiver burden of caregivers of colorectal cancer patients. *Turk J Colorectal Dis*. 2018;28(4):164–171. <https://doi.org/10.4274/tjcd.26121>.
- Tolbert E, Bowie J, Snyder C, Bantug E, Smith K. A qualitative exploration of the experiences, needs, and roles of caregivers during and after cancer treatment: "That's what I say. I'm a relative survivor. *J Cancer Surviv*. 2018;12(1):134–144. <https://doi.org/10.1007/s11764-017-0652-x>.
- Kav S, Tokdemir G, Tasdemir R, Yalili A, Dinc D. Patients with cancer and their relatives beliefs, information needs and information-seeking behavior about cancer and treatment. *Asian Pac J Cancer Prev APJCP*. 2012;13(12):6027–6032. <https://doi.org/10.7314/APJCP.2012.13.12.6027>.
- Dau H, Safari A, Saad El Din K, et al. Assessing how health information needs of individuals with colorectal cancer are met across the care continuum: an international cross-sectional survey. *BMC Cancer*. 2020;20(1):1031. <https://doi.org/10.1186/s12885-020-07539-0>.
- Harrison JD, Young JM, Price MA, Butow PN, Solomon MJ. What are the unmet supportive care needs of people with cancer? A systematic review. *Support Care Cancer*. 2009;17(8):1117–1128. <https://doi.org/10.1007/s00520-009-0615-5>.
- Brown S, Greenfield D, Thompson J. Knowledge and awareness of long-term and late treatment consequences amongst colorectal cancer survivors: a qualitative study. *Eur J Oncol Nurs*. 2016;20:191–198. <https://doi.org/10.1016/j.ejon.2015.08.005>.
- Han CJ, Yang GS, Syrjala K. Symptom experiences in colorectal cancer survivors after cancer treatments: a systematic review and meta-analysis. *Cancer Nurs*. 2020;43(3): E132–E158. <https://doi.org/10.1097/NCC.0000000000000785>.
- Ho MY, McBride ML, Gotay C, et al. A qualitative focus group study to identify the needs of survivors of stage II and III colorectal cancer. *Psycho Oncol*. 2016;25(12): 1470–1476. <https://doi.org/10.1002/pon.3988>.
- Glanz K, Rimer BK, Viswanath K. *Health Behavior and Health Education: Theory, Research, and Practice*. San Francisco, CA: John Wiley & Sons; 2008.
- Beaver K, Witham G. Information needs of the informal caregivers of women treated for breast cancer. *Eur J Oncol Nurs*. 2007;11(1):16–25. <https://doi.org/10.1016/j.ejon.2006.01.006>.
- Chou W-yS, Liu B, Post S, Hesse B. Health-related internet use among cancer survivors: data from the health information national trends survey, 2003–2008. *J Cancer Surviv*. 2011;5(3):263–270. <https://doi.org/10.1007/s11764-011-0179-5>.