

Do diagnostic criteria for ME matter to patient experience with services and interventions? Key results from an online RDS survey targeting fatigue patients in Norway

Journal of Health Psychology
2023, Vol. 28(13) 1189–1203

© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13591053231169191

journals.sagepub.com/home/hpq



Anne Kielland¹ , Jing Liu¹
and Leonard A Jason²

Abstract

Public health and welfare systems request documentation on approaches to diagnose, treat, and manage myalgic encephalomyelitis and assess disability-benefit conditions. Our objective is to document ME patients' experiences with services/interventions and assess differences between those meeting different diagnostic criteria, importantly the impact of post-exertional malaise. We surveyed 660 fatigue patients in Norway using respondent-driven sampling and applied validated DePaul University algorithms to estimate Canadian and Fukuda criteria proxies. Patients on average perceived most interventions as having low-to-negative health effects. Responses differed significantly between sub-groups for some key interventions. The PEM score was strongly associated with the experience of most interventions. Better designed and targeted interventions are needed to prevent harm to the patient group. The PEM score appears to be a strong determinant and adequate tool for assessing patient tolerance for certain interventions. There is no known treatment for ME, and "do-no-harm" should be a guiding principle in all practice.

Keywords

chronic fatigue syndrome, DePaul Symptoms Questionnaire, health and welfare services, myalgic encephalomyelitis, respondent-driven sampling

Introduction

In the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems (ICD)*, the World Health Organization (2019) has classified *myalgic encephalomyelitis* (ME), under code G93.3 as a disorder of the nervous system. The United States and United Kingdom health authorities have in their revised guidelines described the illness as a debilitating and complex medical

condition (Centers for Disease Control and Prevention, 2021; National Institute for Health and Care Excellence (NICE, 2021). The illness's exact mechanisms are unclear, and disagreement

¹Fafo, Norway

²DePaul University, USA

Corresponding author:

Anne Kielland, Fafo, Borggata 2b, Oslo 0608, Norway.

Email: aki@fafo.no

over the validity of case definition criteria prevails. Especially, there is disagreement on the role of post-exertional malaise (PEM), and whether it is meaningful to promote the Canadian Consensus Criteria (CCC) as the golden standard for diagnosis and research, assuming these case criteria can distinguish one illnesses from resembling condition.

Due to this uncertainty and the lack of well-documented treatment options, public health and welfare systems meet people with ME with few available tools in their toolboxes. A major challenge is for the welfare services to determine when a person with ME should be granted disability status. Currently, suspected ME cases undergo lengthy diagnostic processes that involve doctors, psychologists, and physiotherapists (see Owe et al., 2016 for more detail on this). After diagnosis, welfare authorities may require the patient to try several interventions to assess their potential for recovery and remaining labor capability. Such interventions commonly comprise seeing psychiatrists, cognitive therapists, and physiotherapists and attending symptom-management classes, rehabilitation, and work-assessment programs (NAV, 2022). A recent review of chronic fatigue syndrome (CFS)/ME patients under the public work-assessment-allowance program in Norway underscored that “the health and welfare systems lack knowledge about who will respond well to the various alternatives” (PROBA Samfunnsanalyse (PROBA, 2022: 6).

The NICE 2021 assessment of the research literature on the most common ME interventions spurred a heated debate over the guidelines’ research-quality grading (Flottorp et al., 2022; NICE, 2021; Vink and Vink-Niese, 2022). Some CFS/ME treatment trials previously considered positive were graded low quality. A common cause for the change was those trials’ broad inclusion criteria for research participants. If the trial could not document correctly using PEM as a trial-inclusion criterion, the NICE review did not consider the trial’s positive results as generally applicable to the ME population.

One of the most heated debates concerns the representativity of certain patient experiences shared in media. One side of the debate has been described as the “bio-psycho-social side,” focusing on cognitive techniques to help redirect illness-conserving thought patterns. The other side has been called the “bio-medical side,” calling for more medical research into an illness they perceive as somatically rooted. In the media, voices on both sides claim, respectively, to have been helped or harmed by some frequently discussed interventions.

This article shares statistical results on patient experiences with the most common interventions for ME affected and the public and private services that provided them. It derives from a 5-year social science research project which broader objectives were to increase knowledge of prevalence, identify the sociodemographic characteristics of persons affected, map patients’ use of public services, and capture variations and trends in patient experiences with and perspectives on such public services within the health and welfare sectors (SINTEF and Fafo, n.d.). In the project, public health register data were also obtained to map the overall use of services and interventions, and initial exploratory qualitative interviews with patients and their families were used to learn more about the variation in experiences with the different services and interventions. Our survey questions on interventions and services were informed by the qualitative interviews and specified to help complement the register data.

In this article, we used the term “ME”, not “CFS/ME”. Guided by APA, 7th Edition (APA, 2020), Section 5.4. we asked about patient identity in the survey and found that almost none of the participants identified as only “CFS”, but most identified with “ME”. One third accepted a combination of both terms (twice as many preferring ME put first). Two thirds found only ME—and not CFS—acceptable. In the qualitative interviews, participants explained that they found “CFS” trivializing of their more complex illness experience.

Objectives and hypotheses

This article has two main objectives. First, we considered that Norwegian Health and Welfare Authorities offer or impose standard interventions on ME affected who apply for disability status. Our first objective became to document how helpful or unhelpful persons with ME perceive the most common services and interventions to be for their symptoms. We also wanted to learn whether patients considered a service or intervention “good,” regardless of its usefulness. In the qualitative interviews, the target group was asked to define a “good” service; many said that would be a service where they felt “seen and understood.” This operationalization was used in the questionnaire.

The second objective was to assess similarities and differences in experiences between subgroups of fatigue patients. This could help inform policymakers of “who will respond well to the various alternatives” (PROBA, 2022). We compared patients meeting proxies for the CCC to those meeting *only* the less strict Fukuda criteria (herein, “Fukuda cases”) and estimated the relative impact of PEM on their service experiences.

Frequent observations made during the qualitative interviews and actualized by the ongoing debate over the new NICE guidelines’ scientific premises informed the tested hypotheses. We first tested the claims that many patients (a) experience public services and interventions as unhelpful to their symptoms and (b) do not feel seen or understood by service providers. Second, we hypothesized that meeting the CCC affected the intervention experiences differently than meeting only the Fukuda criteria. Third, we tested the PEM score as a potential indicator of whether a patient would experience an intervention as useful or harmful.

Widespread negative health experiences with services and interventions would suggest a need for refocusing the public services protocol vis-à-vis ME patients. Statistically significant differences between patient groups would support the hypothesis that diagnostic criteria matter to patient experiences with services and

interventions. It also would suggest that case definition criteria and the PEM score could be used to determine whether a patient should be recommended for certain interventions. Similarities in intervention responses between subgroups, on the other side, would lend support to the claim that the case definition and PEM score are of limited relevance to assessing intervention adequacy. Indirectly, it would support the counterhypothesis that research results from studies that recruited participants who did not expressly meet the stricter CCC may be considered generalizable to patients who did.

Method

In identifying ME affected according to the CCC, the sampling strategy had to be selected carefully. Some U.S. studies indicated that as many as 84%–91% of ME patients might go undiagnosed (Jason et al., 1999, 2006, 2020; Solomon and Reeves, 2004), and we did not want to miss undiagnosed cases. Given the diagnostic process’s on average 6-year duration in Norway, the choice of sampling approach had to consider that underdiagnosing ME with ICD register code G93.3 was nontrivial (Bragée et al., 2020; Kielland et al., forthcoming; ME Foreningen, 2021; Owe et al., 2016). Moreover, drawing a sample from lists derived from patient registers was increasingly challenging due to the new person-protection rules and declining response rates recently experienced with this sampling approach.

Faced with a target population with no clear sampling frame from which to draw a representative sample, we selected a respondent-driven sampling (RDS) approach (Heckathorn, 1997, 2002). On the surface, RDS may appear an ordinary snowballing method where each respondent is asked to recruit new participants from their social networks. Two features, however, distinguish RDS from less sophisticated snowball-sampling methods. First, the RDS software algorithms use network and recruitment information to develop specific variable estimators (Johnston and Luthra, 2014). The estimators are designed to remove recruitment-related bias

(from ingroup recruitment and network size variations); therefore, the numbers presented in this article resemble a weighted distribution. Second, specialized software for RDS analysis allows researchers to monitor how each variable's value distribution develops as the recruitment proceeds. Recruitment equilibrium is a key analytical concept in RDS: At some point in the survey, the distribution of values on key variables stops fluctuating and remains unchanged when new participant cohorts are added. When equilibrium is reached, adding new respondents to the sample is no longer likely to change the distribution for each variable (see examples in Supplemental Annex II). These two features determine when a large enough sample is obtained to support a claim of representativeness in a specified target group. (For more on how RDS produces unbiased estimates, see Johnston, 2014.)

To identify CCC cases in a population with fatigue illness, we targeted wider than our group of interest. In that way, we reduced the likelihood of exclusion errors and could use standardized approaches to identify inclusion errors. The targeted population was, therefore, persons experiencing fatigue-related illness to the extent that they could no longer work or study full-time. We selected four seed respondents (the surveys' starting points) who differed in age, gender, geography, and affiliation to the main ME association (Kubal et al., 2014). Each seed and each respondent recruited was given four new survey links. By using social acquaintances to help recruit others in the target group, RDS restricts access to—and, therefore, bias from—the most eager survey responders (who often spend much time online) more than in an open online survey. Instead, the RDS stimulates participation from those who feel socially committed to the acquaintance who invited them.

To proxy what cases may meet the CCC and Fukuda criteria, respectively, the survey used algorithms based on the two health questionnaires: the DePaul Symptoms Questionnaire (DSQ; DePaul University, n.d.) and the Medical Outcomes Study Questionnaire Short Form 36 (SF-36; Jason and Sunnquist, 2018; RAND

Corp., n.d.). The algorithms to assess the likelihood that a patient would meet the CCC or Fukuda criteria had been developed and validated by DePaul University (n.d.) researchers (Bedree et al., 2019; Jason et al., 2015a, 2015b, 2017; Jason and Sunnquist, 2018; Murdock et al., 2016).

An essential distinction between the CCC and Fukuda criteria is that PEM is required to meet the CCC but is a polythetic criteria in the Fukuda manual. This and the threshold for the five PEM indicators in the DSQ are among the elements that differentiate the two algorithms. In DSQ, PEM is assessed by severity and frequency scores for the five key fatigue symptoms (DSQ14–18; DePaul University, n.d.).

Analysis and results

The survey, conducted in the Fall of 2020, covered 660 respondents, representing an 89% response rate of the 741 who opened the recruitment link. Because RDS recruits from participant networks, there is a risk that respondents cluster on shared experiences and beliefs. An important RDS feature is the possibility of studying recruitment patterns through visual representations of recruitment trees and revealing potential bias-producing clustering. In the media and research debates, the ME field may appear ideologically polarized. A divide is often drawn between ME association members and non-members: The leading associations tend to support a biomedical approach to ME, and the others more often a biopsychosocial approach. There were no strong tendencies of in-group recruitment in the survey's recruitment pattern. Two-thirds of the respondents were association members, and members and non-members appeared to have recruited each other back and forth indiscriminately. That pattern suggested less polarization among patients than the impression from the media and academia. (See Supplemental Annex I for sample recruitment trees.)

Equilibrium on key variables was obtained after a maximum of 250 participants, validated by convergence plots. Adding more participants

to the sample of 660 was therefore unlikely to have changed the distributions in any significant way on these variables. (See examples of convergence plots and a further explanation for this claim in Supplemental Annex II.) We confirmed that the seeds' characteristics did not determine the sample's characteristics. With no bottlenecks experienced, the results should according to the theory be robust and thus representative of the target group when estimators are applied. (More on the method and its application in this study in Kielland et al., forthcoming).

Applying the DePaul algorithms to the sample, 584 respondents met the Fukuda criteria; among them, 474 also met the stricter CCC. In the analysis, the 474 participants who met the CCC were compared to the 110 who met *only* the Fukuda criteria. The mean PEM score in the survey population was 30 (scale of 0–40), with a median score of 32. For the assessment of the impact of PEM score on intervention experience, all 660 respondents were used.

In the main analysis following, we first looked at the most common encounters between patients and the health services in Norway, starting with their general practitioner (GP) and specialist health care services in charge of determining the correct diagnosis for their fatigue illness. Experiences with public and private specialists are described separately. (See Supplemental Annex III for conditional distribution.)

Second, we studied interventions intended to alleviate and manage fatigue patients' symptoms by psychologists, psychiatrists, physiotherapists, symptom-management training, CBT, and rehabilitation institutions. The selection of interventions presented correspond to those currently involved in the standard diagnostic procedure and those the welfare authorities, in agreement with the patient's GP, are instructed to test before granting disability benefits (NAV, 2022). Physiotherapy was the most common intervention, followed by psychologist consultations and symptom-management training. Rehabilitation programs are offered

regularly to those not too sick. (See Supplemental Annex IV for the number of interventions by intervention type.)

Third, we included in the analysis experience with some off-label interventions: the opioid receptor blocking medicine Naltrexone in low doses (LDN); the mental training program, the Lightening Process (LP); and the immunoglobulin Gammanorm. Despite a lower number of participants having tried the latter two, the results are relevant at a time when different patient groups and professionals suggest such interventions be more formally included in the public repertoire. (See Supplemental Annex V for the conditional distribution.)

Finally, we looked at experiences with the welfare services' work-assessment programs often imposed on fatigue affected during the disability-assessment period. We also addressed the overall experiences with the coordination of health and welfare services and between the health services.

For each service and intervention studied, we used three statistical approaches to address variations. First, we compared the *satisfaction rates* of CCC cases to Fukuda cases as the share of participants stating they were very or somewhat satisfied or a service was useful or improved their symptoms. Second, we estimated the mean *satisfaction score* with, or health benefit experienced from, each service on a scale of 2 (*very satisfied/useful*), 1 (*somewhat satisfied/ useful*), 0 (*neither/nor*), -1 (*somewhat dissatisfied/unhelpful*), and -2 (*very dissatisfied/unhelpful*).

Whereas the ratings expressed the share of participants helped by a service or intervention, the scores added information about the intensity/degree to which they experienced an intervention as helpful or unhelpful/harmful. We applied the RDS-specific estimators to the rates and scores presented here, and they can therefore be considered representative. We used *t* tests to identify statistically significant differences between the CCC and Fukuda cases in the unweighted sample because the RDS software does not allow for regular *t* testing. Although

other approaches to assessing standard errors are available in RDS, these typically produce very large standard errors, partly because of widely differing social network sizes, here ranging from 1–100. Although some differences between the CCC and the Fukuda groups remained statistically significant also using RDS estimators, important information on group differences would get lost if one restricted the testing to standard errors resulting from using such tools. Thus, the *t*-test scores are reported here based on unweighted rates and scores (all estimates are available in the attached logfile).

Third, Pearson's correlations addressed the statistical significance of the PEM scores' effects on the patients' service or intervention experiences.

The following analysis section covers experiences with services and interventions in four main groups (see Supplemental Annex VI for a visual overview):

- A. Services that manage the diagnostic process
- B. Interventions intended to help the patient manage or improve symptoms
- C. Experiences with off-label intervention approaches
- D. Experiences with welfare sector work-assessment programs and coordination between health and welfare services.

Table 1 summarizes the findings described in the article.

Getting the diagnosis

General practitioner. Almost all ($n=644$) participants reported having used a GP. The GP recommends the patient to specialist health care services and is responsible for the patient's follow-up. In the qualitative interviews, patients often raised concerns about the GP's knowledge about the fatigue illness, and again, how they felt seen and understood. The survey asked participants to evaluate their experience with their current or last GP on these two dimensions.

Table 1, section A, first shows that a *rate* of only around 45% of CCC cases were satisfied with the knowledge level of their GP, but 61% felt seen and understood. The mean satisfaction *scores* were also lower for the GP's knowledge of the fatigue illness than for feeling seen and understood. Rates and scores were generally lower for patients meeting the CCC, but there were no statistically significant differences between CCC and Fukuda patients. The PEM score was strongly associated with patient satisfaction levels, as dissatisfaction increased with higher PEM scores.

Specialist health care services. After referral from the GP, patients undergo further examination by specialized health care services. Most start the diagnostic process within the public system; some change to a private specialist in fatigue illnesses during the diagnostic course or to get a second opinion; others go directly to a private specialist because waiting lists in the public systems may be long. In our data, 549 participants had experiences with public specialist health care services and 299 participants had used private specialist health care services.

Table 1, section A, shows that only slightly more than half the patients were happy with the public specialists' knowledge level, but more appreciated the private specialists' knowledge. More respondents felt seen and understood by the private specialists than the public ones. A higher share of Fukuda cases felt seen and understood by both public and private providers; the satisfaction share was particularly high with private providers. Whereas some CCC patients had strongly negative experiences with being seen and understood by the public services, which dragged down the group's average score, many more appeared to experience being acknowledged in the private services. The differences between CCC and Fukuda cases thus far were not statistically significant, except Fukuda patients felt more seen and understood by private providers. The PEM scores were negatively associated with the patients' experiences with public services at a statistically significant level.

Table 1. Patient satisfaction with services and interventions.

	Satisfaction rate			Mean satisfaction score			PEM score	
	CCC	Fukuda	T test	CCC	Fukuda	T test	Pearson	Sig
A								
General practitioner (n = 656)								
Knowledge about illness	0.45	0.52	0.418	0.28	0.40	0.418	-0.132**	0.001
Feel seen and understood	0.61	0.66	0.156	0.75	0.89	0.156	-0.155***	0.000
Public specialist (n = 549)								
Knowledge about illness	0.52	0.59	0.082	0.41	0.71	0.083	-0.090**	0.035
Feel seen and understood	0.54	0.60	0.211	0.40	0.74	0.105	-0.116**	0.007
Private specialist (n = 299)								
Knowledge about illness	0.67	0.74	0.172	0.84	1.12	0.179	0.017	0.764
Feel seen and understood	0.72	0.96	0.000	0.96	1.55	0.000	-0.032	0.584
B								
Psychologist (n = 371)								
Usefulness/effect on symptoms	0.39	0.49	0.190	0.09	0.54	0.123	-0.151**	0.004
Feel seen and understood	0.57	0.72	0.036	0.50	0.96	0.100	-0.114**	0.028
Psychiatrist (n = 165)								
Usefulness/effect on symptoms	0.30	0.46	0.347	-0.25	0.27	0.478	-0.156*	0.046
Feel seen and understood	0.39	0.47	0.491	-0.01	0.46	0.526	-0.186*	0.017
Physiotherapist (n = 436)								
Usefulness/effect on symptoms	0.28	0.37	0.360	-0.24	0.20	0.102	-0.222***	0.000
Feel seen and understood	0.52	0.78	0.000	0.47	1.13	0.000	-0.140**	0.003
Cognitive-behavioral therapy (n = 241)								
Usefulness/effect on symptoms	0.16	0.54	0.002	-0.62	0.02	0.068	-0.268***	0.000
Feel seen and understood	0.50	0.58	0.234	0.12	0.45	0.389	-0.201**	0.002
Symptom management (n = 360)								
Usefulness/effect on symptoms	0.46	0.48	0.226	0.06	0.07	0.399	-0.136**	0.010
Feel seen and understood	0.77	0.69	0.356	1.01	0.99	0.686	-0.071	0.177
Rehabilitation (n = 342)								
Wanted to go	0.81	0.92	0.082	1.23	1.60	0.186	-0.072	0.187
Comfortable at the location	0.54	0.83	0.003	0.34	1.23	0.015	-0.124*	0.022
Enjoyed the social	0.80	0.91	0.290	1.13	1.50	0.167	-0.047	0.390
Program suited for illness	0.48	0.72	0.010	-0.01	0.75	0.008	-0.180**	0.001
Usefulness/effect on symptoms	0.20	0.40	0.030	-0.93	-0.20	0.014	-0.243***	0.000
C								
Low-dose naltrexone (n = 278)								
Usefulness/effect on symptoms	0.33	0.35	0.166	-0.21	-0.09	0.465	-0.117	0.051
Feel seen and understood	0.40	0.70	0.372	0.28	0.78	0.254	-0.147*	0.014
Lightening process (n = 59)								
Usefulness/effect on symptoms	0.09	0.00	0.044	-1.20	-1.58	0.486	-0.159	0.229
Feel seen and understood	0.02	0.31	0.174	-1.16	-0.68	0.759	-0.225	0.087
Gammanorm (n = 45)								
Usefulness/effect on symptoms	0.60	0.48	0.607	0.76	0.48	0.778	-0.253	0.094
Feel seen and understood	0.66	1.00	0.003	0.99	1.07	0.316	-0.262	0.083

(Continued)

Table 1. (continued)

	Satisfaction rate			Mean satisfaction score			PEM score	
	CCC	Fukuda	T test	CCC	Fukuda	T test	Pearson	Sig
D								
Work-assessment program (n = 292)								
Looked forward to	0.50	0.66	0.026	0.22	0.74	0.025	-0.093	0.110
Workplace suitable	0.31	0.53	0.007	-0.5	0.03	0.007	-0.151**	0.009
Tasks suitable	0.37	0.58	0.003	-0.41	0.30	0.001	-0.139*	0.017
Enjoyed the social	0.61	0.59	0.529	0.41	0.59	0.627	-0.102	0.082
Made me sicker	0.94	0.79	0.054	1.70	1.03	0.000	0.342***	0.000
Made me more ready for work	0.01	0.00	0.471	-1.86	-1.85	0.940	-0.169**	0.004
Service coordination (n = 660)								
Coordination among health services	0.14	0.23	0.119	-0.67	-0.34	0.136	-0.133**	0.001
Coordination health and welfare	0.10	0.16	0.222	-0.63	-0.41	0.315	-0.106**	0.006

Satisfaction rated 0 to 1; satisfaction scored -2 to 2.

CCC: Canadian consensus criteria; PEM: post-exertional malaise (range 0–40).

*** $p < .001$, ** $p < .01$, * $p < .05$.

Interventions to alleviate and manage symptoms

Concerning the five interventions aimed at treating and alleviating symptoms, we focused on (a) the questions of the *perceived usefulness* of the interventions for symptom relief and (b) the experience of being seen and understood, corresponding to patients' overall definition of a "good" intervention.

First, fatigue patients today generally are required to see a psychologist as part of the diagnostic process before granting disability status. A total of 371 participants in the sample had seen a psychologist about their fatigue illness. Second, 165 patients had seen a psychiatrist related to their fatigue illness. Third, physiotherapists are part of the diagnostic process and are also used for pain prevention, alleviation, and training schemes. In the survey, 436 participants had seen a physiotherapist: 238 received mainly help with relaxation and pain management, 120 mainly movement training, and 86 guidance on pacing. Fourth, CBT can be intended for both treatment and symptom management in fatigue and pain patients. A total of 241 survey participants had been to CBT. Finally, symptom-management training, which

360 participants attended, is another standard component of the current public plan for ME patients in Norway.

Table 1, section B, shows satisfaction rates and scores for the five interventions. Looking at the perceived helpfulness of interventions for the patients' symptoms in CCC and Fukuda cases, respectively, 39% and 49% of patients perceived the psychologist as helpful, 30% and 46% the psychiatrists, and 28% and 37% the physiotherapists. Symptom management was perceived as the most helpful intervention by 46% and 48% of patients, respectively. Only 16% of CCC cases found CBT helpful for their symptoms compared with over half the Fukuda cases, and this difference was statistically significant. For average scores, the CCC group reported that psychiatrists, physiotherapists, and CBT had a negative impact on symptoms. The positive and negative experiences of symptom-management training balanced each other. On average, the Fukuda cases had slightly positive experiences with psychologists, psychiatrists, and physiotherapists, and their positive and negative experiences with CBT and symptom management again balanced each other.

Were these seen as "good" services, operationalized as providers who see and understand

the patients? Most patients appeared to appreciate most service providers, but less than half felt seen and understood by their psychiatrist. Psychologists, physiotherapists, and providers of symptom-management training scored relatively high on acknowledging the patients despite relatively low ratings and scores for the usefulness of their interventions. In almost all estimates, the PEM score was negatively associated with satisfaction with these two aspects of these interventions and, most decisively, with the physiotherapy and CBT benefits.

Rehabilitation. Rehabilitation occurs at specialized establishments across Norway, and 342 respondents had been to rehabilitation at different locations. Among the 247 who had not been to rehabilitation, 113 stated they had been too ill to go. Probing into the degree of illness for patients reporting having been too ill for rehabilitation programs, we found that mean score for Physical Functioning in SF-36 for this group at the time of the survey was of only 23 compared to a sample mean of 36 (on a scale from 0 to 100). Therefore, we assumed that the experiences with rehabilitation reported in this section did not reflect what *would have been* the experiences of some of the most severely ill patients.

The rehabilitation programs varied but commonly comprised combinations of interventions such as pacing, GET, symptom-management guidance, mindfulness, psychologist and physician appointments, diet, sleep hygiene, and CBT. Other studies have indicated a possible negative association between these rehabilitation programs and indicators such as health-related quality of life in this patient group (Similä et al., 2020). In addition, no participant in the qualitative interviews preceding this survey reported positive health benefits from the rehabilitation centers but rather the contrary. Many, however, valued the programs' social aspects, often after suffering from illness-related isolation. The survey subsequently asked about both.

Most participants stated that they were motivated (Table 1, section B), and were comfortable at the rehabilitation center, Fukuda patients significantly more often than CCC patients. A

clear majority in both groups were happy with the placement's social aspects. Although Fukuda patients more often than CCC patients agreed that the rehabilitation center program suited their illness, half the CCC patients agreed. Critically, however, very few—only 20% of CCC patients and 40% of Fukuda patients—experienced the programs as helping their symptoms. Many *strongly disagreed* that the programs had helped their health, leaving the mean satisfaction scores strongly negative, especially in the CCC group. The PEM score was negatively associated with experiencing comfort at the rehabilitation center, program suitability, and program benefits to health.

Off-label approaches

The survey asked about five common alternative or off-label approaches often advocated for and tried by ME affected. Relatively few participants had tried each type, but 278 patients had tried LDN, 59 had tried the LP program, and 45 had tried the immunosuppressant Gammanorm.

Table 1, section C, shows a relatively low satisfaction rate with the effect of LDN treatment among both CCC and Fukuda cases, although around one third of all patients felt somewhat helped. The mean satisfaction score was slightly negative in both groups. More Fukuda than CCC cases felt seen and understood by the providers of this treatment. Satisfaction rates and scores with the health benefits of LP were very low: only five patients were very or somewhat satisfied with the effects of the training on their symptoms, all five in the CCC group. Satisfaction rates were higher for Gammanorm; this time, they were highest among the CCC patients. Gammanorm patients also much more often experienced being seen and understood. The PEM score was associated only with the experience of being seen and understood by LDN providers.

Welfare authorities and service coordination

Unless a patient's health state is clearly too serious, they are required in Norway to participate

in a work-assessment program as part of the evaluation for disability status. The arrangement is intended to help welfare authorities better evaluate an individual's work potential. During the qualitative interviews, patients' stories of such work-assessment programs were concerning from a health perspective. Some described drastic worsening of their symptoms and additional loss of functions during and following the assessment periods.

In the survey, experiences with work assessment were mapped by six indicators, with 292 respondents sharing their experiences. Again, on the background of the qualitative interviews, the survey tested common claims by asking whether participants looked forward to the program, whether the workplace and tasks offered were suitable for their fatigue illness, whether they enjoyed the program's social aspects, whether participation worsened their symptoms, and whether they felt more ready to return to work after the program ended.

Table 1, section D, shows that more Fukuda patients than CCC patients looked forward to the program and, more often, found the workplace and tasks suitable for their fatigue illness. In short, they appeared to tolerate the program content better. Almost three of five in both groups enjoyed the social aspects of the assessment. The most alarming results concerned how 94% of the CCC cases and 79% of the Fukuda cases stated that the participation worsened their illness. The CCC patients more often strongly agreed to this being the case, again yielding statistically significant differences between the two patient-group scores. Almost none of the participants stated that the program had made them more ready to return to work. Arguably, this survey did not target persons who today are back working or studying full time, but it does include those working or studying part time. Again, the PEM score was systematically associated with poorer outcomes.

Service coordination. All 660 patients responded to the questions on how well they thought the health services coordinated their efforts among themselves and how well health and welfare

services coordinated with each other. Satisfaction rates with the two types of coordination efforts were overall very low for both CCC and Fukuda patients, and the satisfaction scores were strongly negative (Table 1, section D). The higher the PEM score, the more difficulty patients reported in their experiences with service coordination.

Discussion

The data presented raise four issues. First, the scores for the experienced health benefits of some of the most common health and welfare interventions are overall low and, at times, negative in both groups of participants. Second, the experiences of patients meeting the proxy for the CCC and Fukuda criteria are similar concerning some services but differ significantly for some of the most debated interventions in the area. Third, patients' PEM scores are negatively associated at a statistically significant level with their experience with most services and interventions studied. Finally, despite poor experiences with many interventions, patients more often feel seen and understood by their public, private, and off-label intervention providers. This section presents a graphic figure to highlight some overall tendencies in the data and discusses some possible limitations to the method applied.

Figure 1 compares the average health benefits or disadvantages fatigue patients reported with the most common interventions, to their perception of being seen and understood by service and intervention providers. Group means are shown for patients meeting the CCC and Fukuda criteria in the DSQ algorithms.

Looking at the upper part of the figure, it is evident that few public interventions are experienced as very helpful, and, for the CCC cases, almost all experiences are marginal to negative. Fukuda patients rate psychologists, psychiatrists, and physiotherapists as slightly beneficial to their symptoms and CBT and symptom management as, on average, neither helpful nor unhelpful. The CCC cases report CBT as having, on average, a negative impact on their health, and

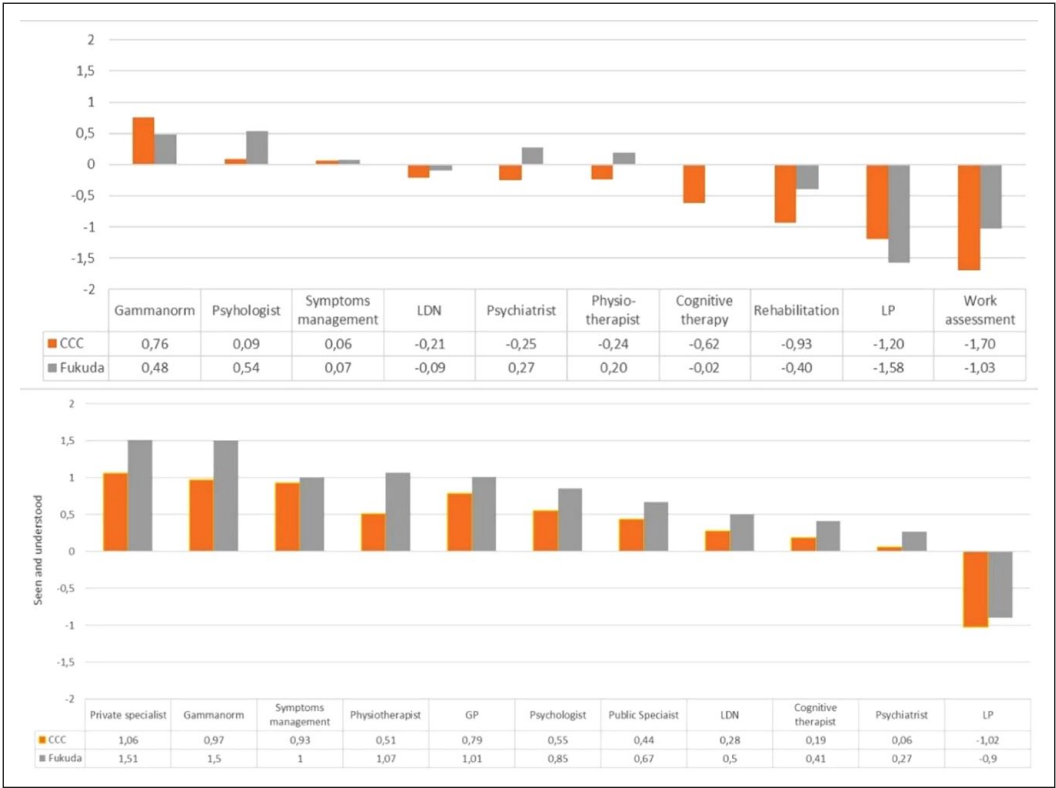


Figure 1. Satisfaction/helpfulness of common interventions by mean patient score: Canadian consensus criteria (DSQ) versus Fukuda criteria (RDS estimates) cases, compared to extent patients felt seen and understood by public and private service providers. DSQ: DePaul Symptoms Questionnaire; RDS: respondent-driven sampling.

both groups report a considerable and negative impact from rehabilitation, LP, and work-assessment programs. The figure demonstrates how marginal is the perceived overall utility of the most common interventions the Fukuda patients experienced and how the CCC cases experienced the interventions overall as a complex of stressors on their symptoms and health.

Especially in the CCC group, relatively few patients experience health benefits from rehabilitation and CBT, and the diagnostic criteria differentiate most concerning CBT. The share of patients who find physiotherapy helpful is also low. The share of Fukuda cases who experience a service as somewhat or very helpful is moderate to low. (See Supplemental Annex VIII for more detail).

In the qualitative interviews, most informants responded that a “good service” is where they feel seen and understood and, by that, acknowledged by the service provider. The question on being seen and understood was not given for rehabilitation and work-assessment programs because a complex set of providers are involved in both. Instead, the lower part of Figure 1 includes experiences with the GP, public and private specialist health care services, in addition to the services and interventions from the upper part of the figure.

This part of the figure shows that private health care specialists are well liked, more so than public ones. Among the three alternative approaches, patients feel more acknowledged by Gammanorm providers, less by LDN

providers, and generally not by LP providers. Physiotherapists score relatively high on seeing and understanding the patients despite physiotherapy being among the interventions the patients found least helpful. The GPs, followed by psychologists, are also generally seen as understanding. The professions experienced as least acknowledging are cognitive therapists and psychiatrists. For all services, the CCC patients report experiencing acknowledgment markedly less often than those scoring positive for the Fukuda criteria only.

Can some indicators help health and welfare authorities determine whether a patient is likely to benefit from specific interventions? From the results here, it seems that patients meeting the CCC respond differently but that the degree of PEM matters the most. It should be noted that we in Table 1 chose a standard cut-off point for statistical significance of $p < 0.5$, and borderline results always are at a higher risk of being spurious. Comparing the correlation coefficients, p values (two-tailed), and degrees of freedom from all areas, the PEM scores are systematically negatively associated with the experience of services and the experienced health benefits of the interventions. Tendencies are particularly clear for physiotherapy, CBT, rehabilitation, and work-assessment programs but also statistically significant at a level less than $p < 0.05$ for psychologists, psychiatrists, and symptom-management training. When comparing CCC and Fukuda patients, the CCC are negatively related to health outcomes from rehabilitation and work-assessment programs. For CBT, the negative effects of meeting the CCC are only statistically significant for the overall satisfaction *rate*. (Results are summarized in Supplemental Annex VII).

How accurate are the algorithms? A study comparing the DSQ results on the CCC to clinical examinations found the DSQ sensitivity to be high (98%) and specificity moderate (38%), suggesting that using the DSQ as a diagnostic tool should imply limited inclusion errors (Strand et al., 2016). However, the group defined here as only Fukuda cases is more

likely to have included false negatives to the CCC algorithm. A questionnaire will never be as good as a full diagnostic procedure for ME but should provide a reasonable proxy for an article like this.

Despite the qualities of the RDS approach in producing unbiased population estimates, some limitations to the survey representativeness should be addressed. First, the RDS estimators can correct for underrepresentation but not for the characteristics of unreachable subgroups. By targeting only those unable to work or study full-time, we lost the experience of former patients who were back in full-time jobs. Evidence, however, suggests this group to be small (Cairns Hotopf, 2005; Vink and Vink-Niese, 2019). Second, there are population-specific limitations, mainly concerning those who are too sick to respond and have nobody to help them. The survey allowed caregivers to respond for a consenting patient, but not all patients have such help. Third, there are method-specific limitations in RDS because someone with no acquaintances in the target group, even online, would have a zero-recruitment probability. We do not know how many that may be, but Norwegian ME patients are known to be very active in online support groups, and internet coverage is as good as universal (Ytre-Arne, 2016).

Conclusion

The RDS sampling approach in the Norwegian fatigue-patient population worked well. Analyses showed that values on most key variables stabilized at population equilibrium. This indicated that, after applying estimators, we could obtain representative estimates for the recruitable population. A main conclusion is that fatigue patients, on average, benefit relatively little from most interventions offered to or imposed on the patient group today, and many reported that the interventions negatively affected their health. Comparing patients who met the CCC to patients meeting only the Fukuda criteria, we identified statistically significant differences in the experienced

outcomes for the most contested among the interventions studied: CBT, rehabilitation programs, and work-assessment programs. This particular result indicates that the two groups may be qualitatively different, and that this difference is captured by the DSQ algorithms. Patients meeting the CCC consistently experienced interventions as more harmful to their illness than did the Fukuda patients.

The PEM score proved to be systematically associated with negative experiences, especially for physiotherapy, CBT, rehabilitation, and work-assessment programs. Among the CCC patients, experiences with the three latter interventions were very negative overall. This finding suggests that considerable public investments are going into interventions that are experienced as useful to the health of only a few but are reported as harmful to many. This is especially the case for patients meeting the CCC and having higher PEM scores.

In considering health-outcome experiences compared with experiences of being seen and understood across intervention types, patients more often feel seen and understood by service providers than they feel helped with their symptoms. Among the providers considered to most acknowledge the patients are private specialists, Gammanorm providers, and physiotherapists. At the low end are providers of LDN, LP, CBT, and psychiatrists.

There is no known treatment for ME, but, awaiting new research, it seems pertinent to remind decision-makers of the “do-no-harm” principle for this fragile patient group. Most patients reported being sicker after rehabilitation and work-assessment programs. This result should be taken seriously because such interventions are currently imposed on many patients as part of the procedure to qualify for disability benefits. Potentially harmful interventions that risk additional functional loss and further deterioration of life quality should be offered with caution. In considering a meaningful process toward potential disability status for ME patients, health and welfare authorities may tentatively consider the CCC and PEM scores

before suggesting or imposing interventions that might help a only few in this group regain function and work capacity but potentially harm others.

Acknowledgements

We would like to thank the Research Council of Norway for generous funding under the program BEHOV-ME, our person-protection ombudsman, the Norwegian Center for Research Data and the Regional Committees for Medical and Health Research Ethics, for advice and support. We also thank Dr. Guri Tyldum for her expert advice on the respondent-driven sampling approach, Dr. Karl E. Conroy for quality checking all DSQ algorithms estimates, and Donna Magnani for editing.

Author contributions

Anne Kielland, lead; Jing Liu, RDS specialist; Leonard Jason, DSQ algorithm specialist.

Data sharing statement

The current article is accompanied by the relevant raw data generated during and/or analyzed during the study, including files detailing the analyses and either the complete database or other relevant raw data. These files are available in the Figshare repository and accessible as Supplemental Material via the Sage Journals platform. Ethical approval, participant permissions, and all other relevant approvals were granted for this data sharing.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Research Council of Norway (NFR# 272674).

Ethics approval

Person-protection ombudsman The Norwegian Center for Research Data (NSD#391132) and The Regional Committees for Medical and Health Research Ethics (REK #7345).

ORCID iDs

Anne Kielland  <https://orcid.org/0000-0001-8321-9427>

Leonard A Jason  <https://orcid.org/0000-0002-9972-4425>

References

- APA. (2020) *Publication Manual of the American Psychological Association*, (7th ed.n). APA.
- Bedree H, Sunnquist M and Jason LA (2019) The DePaul Symptom Questionnaire-2: A validation study. *Fatigue: Biomedicine, Health & Behavior* 7(3): 166–179.
- Bragée B, Michos A, Drum B, et al. (2020) Signs of intracranial hypertension, hypermobility, and craniocervical obstructions in patients with myalgic encephalomyelitis/chronic fatigue syndrome. *Frontiers in Neurology* 11: 828.
- Cairns R and Hotopf M (2005) A systematic review describing the prognosis of chronic fatigue syndrome. *Occupational Medicine* 55(1): 20–31.
- Centers for Disease Control and Prevention (2021) Myalgic encephalomyelitis/chronic fatigue syndrome. Available at: <https://www.cdc.gov/me-cfs/treatment/index.html> (accessed 9 November 2022)
- DePaul University (n.d.) DePaul Symptom Questionnaire, including SF-36 and case-definition criteria algorithms. Available at: <https://csh.depaul.edu/about/centers-and-institutes/ccr/myalgic-encephalomyelitis-cfs/Documents/DSQ-1.pdf> (accessed 9 November 2022).
- Flottorp SA, Brurberg KG, Fink P, et al. (2022) New NICE guideline on chronic fatigue syndrome: More ideology than science? *The Lancet* 399: 611–613.
- Heckathorn DD (1997) Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems* 44(2): 174–199.
- Heckathorn DD (2002) Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems* 49(1): 11–34.
- Jason LA, Katz BZ, Sunnquist M, et al. (2020) The prevalence of pediatric myalgic encephalomyelitis/chronic fatigue syndrome in a community-based sample. *Child & Youth Care Forum* 49: 563–579.
- Jason LA, McManimen S, Sunnquist M, et al. (2017) Clinical criteria versus a possible research case definition in chronic fatigue syndrome/myalgic encephalomyelitis. *Fatigue: Biomedicine, Health & Behavior* 5(2): 89–102.
- Jason LA, Richman JA, Rademaker AW, et al. (1999) A community-based study of chronic fatigue syndrome. *Archives of Internal Medicine* 159(18): 2129–2137.
- Jason LA, So S, Brown AA, et al. (2015a) Test-retest reliability of the DePaul Symptom Questionnaire. *Fatigue: Biomedicine, Health & Behavior* 3(1): 16–32.
- Jason LA and Sunnquist M (2018) The development of the DePaul Symptom Questionnaire: Original, expanded, brief and pediatric versions. *Frontiers in Pediatrics* 6: 330.
- Jason LA, Sunnquist M, Brown A, et al. (2015b) Factor analysis of the DePaul Symptom Questionnaire: Identifying core domains. *Journal of Neurology and Neurobiology* 1(4): 1–9.
- Jason LA, Torres-Harding S and Njok M (2006) The face of CFS in the U.S. *CFIDS Chronicle* 2005–2006: 16–21. http://www.researchgate.net/profile/Leonard_Jason/publication/236995875_The_Face_of_CFS_in_the_U.S/links/00b7d51acf6823bccb000000.pdf (accessed 9 November 2022)
- Johnston LG (2014) Sampling migrants: How respondent-driven sampling works. In: Tyldum G and Johnston LG (eds) *Applying Respondent Driven Sampling to Migrant Populations: Lessons from the Field*. London: Palgrave Macmillan, pp.9–16.
- Johnston LG and Luthra R (2014) Analyzing data in RDS. In: Tyldum G and Johnston LG (eds) *Applying Respondent Driven Sampling to Migrant Populations: Lessons from the Field*. London: Palgrave Macmillan, pp.84–100.
- Kielland A, Liu J, Tyldum G, et al. (Forthcoming) Bias in diagnosing myalgic encephalomyelitis as G93.3: Applying a respondent-driven sampling design to a fatigue population.
- Kubal AI, Shvab I and Wojtynska A (2014) Initiation of the RDS recruitment process: Seed selection and role. In Tyldum G and Johnston LG (eds) *Applying Respondent Driven Sampling to Migrant Populations: Lessons from the Field*. London: Palgrave Macmillan, pp.37–48.
- ME Foreningen (2021) Pan-European survey for ME-patients: Results and conclusions. Available at: <https://www.youtube.com/watch?>

- v=jL5StmKCP5w (accessed 15 November 2022).
- Murdock KW, Wang XS, Shi Q, et al. (2016) The utility of patient-reported outcome measures among patients with myalgic encephalomyelitis/chronic fatigue syndrome. *Quality of Life Research* 26: 913–921.
- National Institute for Health and Care Excellence (2021) Myalgic encephalomyelitis (or encephalopathy)/chronic fatigue syndrome: Diagnosis and management, NICE Guideline (NICE Guideline 206). Available at: <https://www.nice.org.uk/guidance/ng206/resources/myalgic-encephalomyelitis-or-encephalopathy-chronic-fatigue-syndrome-diagnosis-and-management-pdf-66143718094021> (accessed 9 November 2022).
- NAV (2022) Rundskriv til Folketrygdlovens artikkel 12. Available at: <https://lovdata.no/nav/rundskriv/r12-00> (accessed 9 November 2022).
- Owe JF, Næss H, Gjerde IO, et al. (2016) Investigation of suspected chronic fatigue syndrome/myalgic encephalopathy. *Tidsskriftet*. Available at: <https://tidsskriftet.no/en/2016/02/investigation-suspected-chronic-fatigue-syndromemyalgic-encephalopathy> (accessed 15 November 2022).
- PROBA Samfunnsanalyse (2022) AAP-mottakeres brukerløp i helsetjenester og NAV (Rapport 2022-3 Prosjekt nr. 18016 I). Available at: <https://www.nav.no/no/nav-og-samfunn/kunnskap/forskningsrapporter-og-evalueringer-finansielt-av-nav/sykefravaer-og-arbeidsmiljo-rapportarkiv2/aap-mottakeres-brukerlop-i-helsetjenester-og-nav> (accessed 15 November 2022).
- RAND Corp (n.d.) SF-36. Available at: https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html (accessed 15 November 2022).
- Similä WA, Halsteinli V, Helland IB, et al. (2020) Health-related quality of life in Norwegian adolescents living with chronic fatigue syndrome. *BMC* 18: 170.
- SINTEF and Fafo (n.d.). Tjenesten OG MEG: et forskningsprosjekt [Individuals and families affected by CFS/ME: Population characteristics, service use and needs]. Available at <https://www.tjenestenogmeg.com/>
- Solomon L and Reeves WC (2004) Factors influencing the diagnosis of chronic fatigue syndrome. *Archives of Internal Medicine* 164(20): 2241–2245.
- Strand EB, Lillestøl K, Jason LA, et al. (2016) Comparing the DePaul Symptom Questionnaire with physician assessments: A preliminary study. *Fatigue: Biomedicine, Health & Behavior* 4(1): 52–62.
- Vink M and Vink-Niese A (2022) The updated NICE guidance exposed the serious flaws in CBT and graded exercise therapy trials for ME/CFS. *Healthcare* 10(5): 898.
- Vink M and Vink-Niese F (2019) Work rehabilitation and medical retirement for myalgic encephalomyelitis/chronic fatigue syndrome patients: A review and appraisal of diagnostic strategies. *Diagnostics* 9(4): 124.
- World Health Organization (2019) ICD-10: Postviral fatigue syndrome. Available at: <https://icd.who.int/browse10/2019/en#/G93.3> (accessed 9 November 2022).
- Ytre-Arne B (2016) The social media experiences of long-term patients: Illness, identity, and participation. *Nordicom Review* 37: 57–70.