Check for updates

RESEARCH ARTICLE

REVISED **Typing myalgic encephalomyelitis by infection at onset: A DecodeME study [version 2; peer review: 1 approved, 1 approved with reservations]**

Andrew D. Bretherick [1,2], Simon J. McGrath[3], Andy Devereux-Cooke[3], Sian Leary[3], Emma Northwood[3], Anna Redshaw[3], Pippa Stacey[3], Claire Tripp[3], Jim Wilson[3], Sonya Chowdhury[4], Isabel Lewis[4], Øyvind Almelid [1], Sumy V. Baby[1], Tom Baker[1], Hannes Becher[1], Thibaud Boutin[1], Malgorzata Clyde[1], Diana Garcia[1], John Ireland[1], Shona M. Kerr [1], Ewan McDowall[1], David Perry[1], Gemma L. Samms[1], Veronique Vitart[1], Jareth C. Wolfe[1], Chris P. Ponting [1]

[1]MRC Human Genetics Unit, The University of Edinburgh, Edinburgh, Scotland, EH4 2XU, UK
[2]Pain Service, Ninewells Hospital, NHS Tayside, Dundee, Scotland, DD1 9SY, UK
[3]c/o DecodeME, MRC Human Genetics Unit, The University of Edinburgh, Edinburgh, Scotland, EH4 2XU, UK
[4]42 Temple Street, Keynsham, Action For ME, Bristol, England, BS31 1EH, UK

Abstract

Background:

People with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) experience core symptoms of post-exertional malaise, unrefreshing sleep, and cognitive impairment. Despite numbering 0.2-0.4% of the population, no laboratory test is available for their diagnosis, no effective therapy exists for their treatment, and no scientific breakthrough regarding pathogenesis has been made. It remains unknown, despite decades of small-scale studies, whether individuals experience different types of ME/CFS separated by onset-type, sex or age.

Methods:

DecodeME is a large population-based study of ME/CFS that recruited

Open Peer Review

Approval Status ✔✔

	1	2
version 4 (revision) 21 Aug 2023	✔ view	
version 3 (revision) 20 Jul 2023	? view	
version 2 (revision) 04 Jul 2023	? view	✔ view
version 1 24 Apr 2023	✘ view	? view

1. **Sarah Tyson**, The University of Manchester, Manchester, UK

2. **Suzanne Vernon** , The Bateman Horne

17,074 participants in the first 3 months following full launch. Detailed questionnaire responses from UK-based participants who all reported being diagnosed with ME/CFS by a health professional provided an unparalleled opportunity to investigate, using logistic regression, whether ME/CFS severity or onset type is significantly associated with sex, age, illness duration, comorbid conditions or symptoms.

Results:

The well-established sex-bias among ME/CFS patients is evident in the initial DecodeME cohort: 83.5% of participants were females. What was not known previously was that females tend to have more comorbidities than males. Moreover, being female, being older and being over 10 years from ME/CFS onset are significantly associated with greater severity. Five different ME/CFS onset types were examined in the self-reported data: those with ME/CFS onset (i) after glandular fever (infectious mononucleosis); (ii) after COVID-19 infection; (iii) after other infections; (iv) without an infection at onset; and, (v) where the occurrence of an infection at or preceding onset is not known. Among other findings, ME/CFS onset with unknown infection status was significantly associated with active fibromyalgia

Conclusions:

DecodeME participants differ in symptoms, comorbid conditions and/or illness severity when stratified by their sex-at-birth and/or infection around the time of ME/CFS onset.

Plain English summary

Myalgic Encephalomyelitis / Chronic Fatigue Syndrome (ME/CFS) is a chronic disease that affects an estimated 250,000 people in the UK. Its defining symptom is post-exertional malaise, an excessive delayed worsening of symptoms following even minor physical or mental exertion. For those with it, ME/CFS means disability and poor quality of life.

DecodeME is a research study which is looking for DNA differences between people with ME/CFS and people without any health problems. People with ME/CFS who take part in DecodeME complete a questionnaire that assesses their symptoms and whether they will then be invited to donate a DNA sample. This paper analyses the answers to this questionnaire; we will publish results of the DNA analysis separately.

So far, more than 17 thousand people with ME/CFS have completed the DecodeME questionnaire. Their answers help us to address the question: "Are there different types of ME/CFS linked to different causes and how severe it becomes?"

Results show that people with ME/CFS do not form a single group

Center of Excellence, Salt Lake City, USA

Any reports and responses or comments on the article can be found at the end of the article.

reporting similar symptoms and additional medical conditions. Instead, participants who had an infection at the start of their ME/CFS reported a different pattern of symptoms and conditions compared to those without an infection.

It is well known that most people with ME/CFS are females. What was not clear previously was that females tend to have more additional health conditions. Also, being female, being older and being over 10 years from ME/CFS onset all make it more likely that someone is more severely affected by their ME/CFS.

These findings could indicate that by studying people with different ME/CFS onset-types separately – rather than analysing all people with ME/CFS together – it will be easier to understand what is going wrong.

Corresponding authors: Andy Devereux-Cooke (andy@s4me.info), Sonya Chowdhury (sonya.chowdhury@actionforME.org.uk), Chris P. Ponting (chris.ponting@ed.ac.uk)

Amendments from Version 1

The revised manuscript has benefited greatly from the reviewers' constructive suggestions.

In response, the study's main aim has been made explicit in the Abstract; the plain English summary has been rewritten; the DecodeME questionnaire is described in greater detail; a Methods paragraph has been completely rewritten to state the questions asked explicitly and to connect these to the logistic regression models used; comparisons of results to the existing literature are made in the Discussion section rather than in the Results; numbers and percentages of participants are shown together, and median and interquartile range values for ages are given; the principal finding of each Results section is stated in its heading, to improve comprehension; the opening paragraph of the Discussion has been rewritten to summarise the study's main findings; a fourth limitation of the study (on collinearity) has been included; text that was duplicated in error was removed; and, we clarify that the DecodeME questionnaire is not intended to be used as a clinical tool. In two places, results were revised due to previous errors (detailed in our response to Reviewers).

Results previously shown in Figure 3 are now discussed first, so this now becomes Figure 1; previous Figure 1 and Figure 2 are now Figure 2 and Figure 3. Acknowledgements have been updated. Results shown in the Figures, and the authorship list remain unchanged.

Extended Data has now been made available from the Open Science Framework website: https://osf.io/rgqs3/. This Project site contains CC-BY license DecodeME Study Documents: (i) the DecodeME Questionnaire (version 6) annotated by question identifier (Qid) which were used in the logistic regression analyses detailed in (ii) the regressionResults.txt file.

Any further responses from the reviewers can be found at the end of the article

Plain English summary

Myalgic Encephalomyelitis / Chronic Fatigue Syndrome (ME/CFS) is a chronic disease that affects an estimated 250,000 people in the UK. Its defining symptom is post-exertional malaise, an excessive delayed worsening of symptoms following even minor physical or mental exertion. For those with it, ME/CFS means disability and poor quality of life.

DecodeME is a research study which is looking for DNA differences between people with ME/CFS and people without any health problems. People with ME/CFS who take part in DecodeME complete a questionnaire that assesses their symptoms and whether they will then be invited to donate a DNA sample. This paper analyses the answers to this questionnaire; we will publish results of the DNA analysis separately.

So far, more than 17 thousand people with ME/CFS have completed the DecodeME questionnaire. Their answers help us to address the question: "Are there different types of ME/CFS linked to different causes and how severe it becomes?"

Results show that people with ME/CFS do not form a single group reporting similar symptoms and additional medical conditions. Instead, participants who had an infection at the start of their ME/CFS reported a different pattern of symptoms and conditions compared to those without an infection.

It is well known that most people with ME/CFS are females. What was not clear previously was that females tend to have more additional health conditions. Also, being female, being older and being over 10 years from ME/CFS onset all make it more likely that someone is more severely affected by their ME/CFS.

These findings could indicate that by studying people with different ME/CFS onset-types separately – rather than analysing all people with ME/CFS together – it will be easier to understand what is going wrong.

Introduction

Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a chronic multisystem disorder that affects an estimated 0.2–0.4% of the UK population[1,2]. Its core symptoms are post-exertional malaise, pain, fatigue, unrefreshing sleep, cognitive impairment and/or orthostatic intolerance that may each change across the life-course[3]. Many people with ME/CFS report an infectious episode prior to their initial symptoms. Up to 10% of people with glandular fever (also known as infectious mononucleosis) are eventually diagnosed with ME/CFS[4,5], with similar fractions of people with Ross River virus or *Coxiella burnetii* infections also developing ME/CFS[4]. Long COVID, whose symptoms can overlap those of ME/CFS, appears to arise at a similar rate after infection with severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2)[6,7]. Onset of ME/CFS can also occur without report of infection[8]. Pathogenesis is unknown, and effective treatment is not available. In one study, the health-related quality of life for people with ME/CFS was worse than 20 other conditions compared, including breast, prostate, colon or lung cancer, type I or II diabetes, stroke, multiple sclerosis and schizophrenia[9].

One priority from a 2022 priority setting exercise facilitated by the James Lind Alliance[10] was "Are there different types of ME/CFS linked to different causes and how severe it becomes? Do different types of ME/CFS need different treatments or have different chances of recovery?" To address this question, we took advantage of questionnaire data from DecodeME, a study launched in the UK in September 2022. Before the end of the year, over 17,000 people with a ME/CFS diagnosis from a health professional, and at least 16 years (y) old, had been recruited and completed the study questionnaire.

Over many decades, ME/CFS studies have addressed similar questions using symptom data for tens or hundreds of participants recruited using various inclusion and exclusion criteria[8,11,12]. However, they remain inconclusive on whether different ME/CFS types exist and whether symptoms are sex-biased. The DecodeME project provided a unique opportunity to perform adequately-powered analyses for detecting differences within a single large ME/CFS cohort, under an assumption that ME/CFS type is delineated by onset type.

Methods

Patient and Public Involvement

The DecodeME project grew out of the UK ME Research Collaborative (MERC), formerly known as the CFS/M.E. Research Collaborative or CMRC, which was first established in 2013. The MERC includes people with ME/CFS and carers within a Patient Advisory Group (PAG). As the project evolved in 2018–19, Patient and Public Involvement (PPI) was embedded in every discussion and workshop, resulting in the project becoming a co-production with its grant proposal, aims and outcomes being decided by researchers and PPI equally. In 2020, PPI Steering Group members were selected from across diverse charities and organisations, and for their breadth of experience. The project's name was suggested and decided by PPI Steering Group members. In DecodeME, PPI representatives serve on each of its delivery groups, lead on marketing and communication (including social media), and contribute the majority (two of three) members of the decision-making body, the Management Group. People with lived experience of ME/CFS led the co-creation of a new DecodeME questionnaire. This resulted in substantial improvements in comprehension and accuracy compared to initial drafts and reduced the burden on participants, thereby boosting recruitment.

DecodeME's genetics question ("What, if any, significant genetic differences are there between people with — and those without — ME/CFS?") was identified as a priority first by the MERC and its PAG, before being confirmed as a priority by a wider section of the patient community in the results of the Priority Setting Partnership for ME/CFS[10]. Established participant selection criteria were further refined with PPI throughout. PPI members, through their profound understanding of ME/CFS phenotypes, triggers, severity, symptom range, comorbidities and more, have improved the study's adherence to our chosen case definition and thus further assured the relevance of genetic associations to ME/CFS lived experience.

A substantial minority (16 of 41; 39%) of volunteer participants who trialled an initial paper questionnaire experienced difficulties when answering its questions, missing out questions, marking too many answers or adding their own responses. We then created a substantially revised version with which fewer participants found difficulties (88 of 470; 19%). This revised version again implemented Canadian Consensus criteria (CCC) and IOM/NAM criteria[3,13] as well as criteria introduced in response to peer reviewers' comments on the grant application. A total of 14,789 (86.6%) participants met CCC and/or IOM/NAM criteria. The final DecodeME questionnaire captures participants' age and sex at birth, their ME/CFS illness severity, duration, course, associated symptoms and co-occurring conditions, and whether they experienced an infection around the time of their first ME/CFS symptoms occurring. The questionnaire contains 10 questions on personal information and 29 questions on symptoms; it additionally allows participants to indicate whether a health professional has diagnosed them with any of 34 conditions.

Symptom questions allowed multiple choice, others require single answers. This questionnaire is freely available from the DecodeME website. As a co-production, PPI members advised and helped to create both our recruitment strategy and recruitment materials. Further description of DecodeME's recruitment methods and PPI aspects can be found elsewhere[14]. Before study launch, public awareness of DecodeME was enhanced using regular podcasts, webinars, blog posts and media interviews. These media channels will be used by PPI members and scientists to disseminate results to the international ME/CFS community. PPI team members maintain extensive input into reporting of the results of the questionnaire (including in this article), providing greater understanding and context, and ensuring accessibility. Our genome-wide association study and analysis plans were co-created by researchers and PPI members.

The DecodeME study was reviewed and given a favourable opinion by the North West – Liverpool Central Research Ethics Committee (21/NW/0169). Potential participation bias due to internet use was mitigated by providing a paper questionnaire and providing participants with assistance in completing their online questionnaires. Team members were available to answer phone calls and emails during working hours.

Cohort

Diverse methods used to identify potential participants are detailed in the open access DecodeME Study Protocol publication[14]. Between its full launch date of September 12, 2022 and a data freeze performed on December 19, 2022, DecodeME recruited 17,074 female or male participants who self-reported a diagnosis of ME, CFS, ME/CFS or CFS/ME by a health professional and consented to participate. It is this cohort that we analyse here. All participants were aged 16y or older and completed a questionnaire either online (98.1%) or with a paper version (1.9%). Participants were asked for their sex assigned at birth, how long they had experienced ME/CFS symptoms, and information about 34 conditions: "If a health professional has ever told you that you had any of the conditions below, please select all that apply. If the conditions don't apply to you, please do not select any box." Participants indicated whether each condition was Active ("If the condition has given you symptoms in the past 6 months") or not active ("If the condition has not given you symptoms in the past 6 months, either because it has died down or treatment has controlled it"). They were also asked about 9 fatigue- and 73 non-fatigue symptoms: "In the last 6 months, have you had any of the symptoms below often, repeatedly, or constantly? Please mark any that apply. If none apply, leave all the boxes blank." Respondents were asked: "How severe is your illness?" with answer options matching severity definitions from the UK's National Institute for Health and Care Excellence (NICE) guidelines (2021). Severity categories were consistent with participants' reports of their comorbidities and symptoms (**Results**). Participants indicated the duration of their ME/CFS illness by selecting from a set of predefined ranges, for example between 5 and 10 years, or over

10 years, since onset of symptoms. Questionnaire responses from participants who both consented to participate and self-reported being given a diagnosis of ME, CFS, ME/CFS or CFS/ME by a health professional (as of 19 December 2022) were analysed. Only those whose sex assigned at birth was male or female were analysed due to insufficient numbers of other identities. Participant ages were as of 19 December 2022. Further analyses of questionnaire and genotype data will be undertaken for the full DecodeME cohort once the recruitment phase of the project is completed.

Significance testing. Logistic regression analyses were used to evaluate the relationship between various predictor variables (e.g. age or sex or comorbidities) and a binary outcome (e.g. symptom or onset type). For this we used the glm function in R version 4.2.2. Only p-values surviving Bonferroni correction for multiple tests (nominal p-value, here 0.05, divided by the number of tests per analysis) are shown. To address the question "What self-reported symptoms are associated with sex and/or age?" for each of 80 symptoms we used the linear model: Symptom ~ age + sex + intercept – Figure 4. To address "What symptoms are associated with severity?" for 80 symptoms we used the model: Severity ~ age + sex + symptoms + intercept – Figure 5. To address the questions "What onset types are associated with each of 8 fatigue symptoms (Figure 6A) or 72 non-fatigue symptoms (Figure 6B) or 5 illness courses (Figure 6C)?" we used the

model: OnsetType ~ age + sex + symptoms + intercept. To address "What onset types are associated with 34 comorbidities (active and inactive)?" we used the model: OnsetType ~ age + sex + comorbidities + intercept – Figure 7. For the relevant analyses, severity was coded as mild versus others (i.e. moderate or severe or very severe) – Figure 5; 5 illness courses were compared with 'Fluctuating', the majority response – Figure 6C.

Results

This initial DecodeME cohort contained 17,074 participants (83.5% females) whose median age was 49y (interquartile range [IQR] 37y-59y). Male participants tended to be older than females (median 52y [IQR=40y-63y] and 48y [IQR=37y-59y] respectively; p< 2.2×10⁻¹⁶, Wilcoxon rank sum test). Only 3.3% (n=557) of 17,074 participants did not self-report their ethnicity as White, far fewer than the 18.3% in England and Wales who identify as non-White (https://www.ethnicity-facts-figures.service.gov.uk/). Most DecodeME participants' severity levels are categorised as Mild or Moderate, but Severe and Very Severe individuals are also represented (Figure 1).

Two-thirds (n=10,853; 63.6%) reported an infectious onset to their symptoms (Figure 1), such as glandular fever (n=2,936; 17.2%), COVID-19 (n=380, 2.2%) or another infection (n=7,537; 44.1%). However, only 68% (n=2,009), 51% (n=192) and
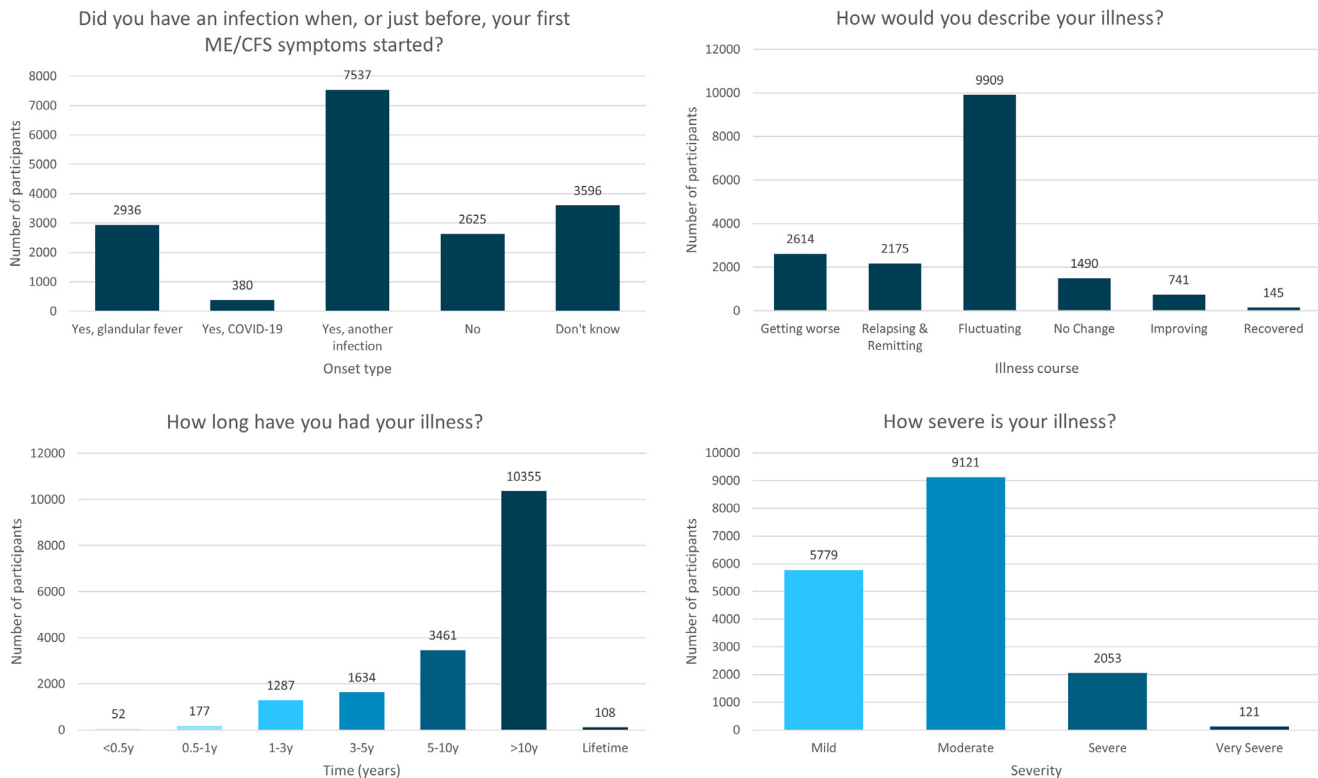

Figure 1. Onset type, illness course, duration of illness and severity of the DecodeME cohort. Numbers of DecodeME participants reporting whether they had an infection prior to ME/CFS onset (top left) as well as their illness' course (top right), duration (bottom left) and severity (bottom right; *n* = 17,074 participants).

26% (n=1,953) respectively of respondents with these potential triggers reported a positive laboratory test confirming the infectious agent.

Over half (58.0%; n=9,909) indicated that their ME/CFS is "Fluctuating (my symptoms vary day to day but don't go away)", 12.7% (n=2,175) describe their symptoms as "Relapsing and remitting (good periods with no symptoms alternating with symptomatically bad periods)" and 15.3% (n=2,614) indicate their symptoms are "Getting worse" (Figure 1).

Most (61.3%; n=10,463) participants have had ME/CFS for over 10y, and 81.6% (n=13,924) over 5y (Figure 1). Together, study participants have experienced over 1.3x10^5 years of ME/CFS symptoms.

50.6% (n= 8,637) of participants (all with self-reported ME/CFS diagnosed by a health professional) reported two or more comorbid conditions, most commonly irritable bowel syndrome (IBS; 41.3%; n=7,052), clinical depression (32.4%; n=5,537) and fibromyalgia (29.5%; n=5,043), anaemia (14.1%; n=2,402) and hypothyroidism (12.8%; n=2,178) (Figure 2). Fibromyalgia and IBS occur together with ME/CFS for 18.0% (n=3,073) of participants (Figure 2B). 22.6% (n=3,865) report no comorbidities.

DecodeME participants' most frequent symptom is post-exertional malaise, a cardinal symptom of ME/CFS[3], followed by unrefreshing sleep, confusion or brain fog, fatigue, muscle pain and gut symptoms (Figure 3). Almost

all answered that once they had exceeded their energy limit their change in symptoms lasts "a long time, which can be more than 24 hours" (97.5%; n=16,649) and agreed that their fatigue affected them both physically and mentally (96.2%; n=16,433). For 88.7% (n=15,142), their fatigue occurs more than half of the time and 87.3% (n=14,921) report their fatigue as disabling.

ME/CFS after glandular fever mostly affects adults a decade after peak incidence

We first analysed participants' ages at ME/CFS onset. Onsets occurring >5y ago do not allow fine resolution of their dates, especially for those responding "Between 5 and 10 years" or "Over 10 years" to "How long have you had your illness?" Consequently, we only considered participants reporting onsets within the last 5y (n=3,150). The median age of this group was 40y, IQR=31y-51y, implying that most participants' onsets occurred between 25y and 50y of age. This is older than the peak incidence of glandular fever in the UK (15y-19y old)[15]. Rather than most participants reporting ME/CFS onset in the last 5y after glandular fever being in their early twenties, as expected, they were a decade older (median ages 30.5y [IQR=23y-41y]). This difference is consistent with adolescents being less likely, than older people, to develop ME/CFS after glandular fever.

ME/CFS comorbidities and symptoms are sex- and age-biased

The substantial number of males participating in DecodeME (n=2,827) allowed the study to reveal previously unreported
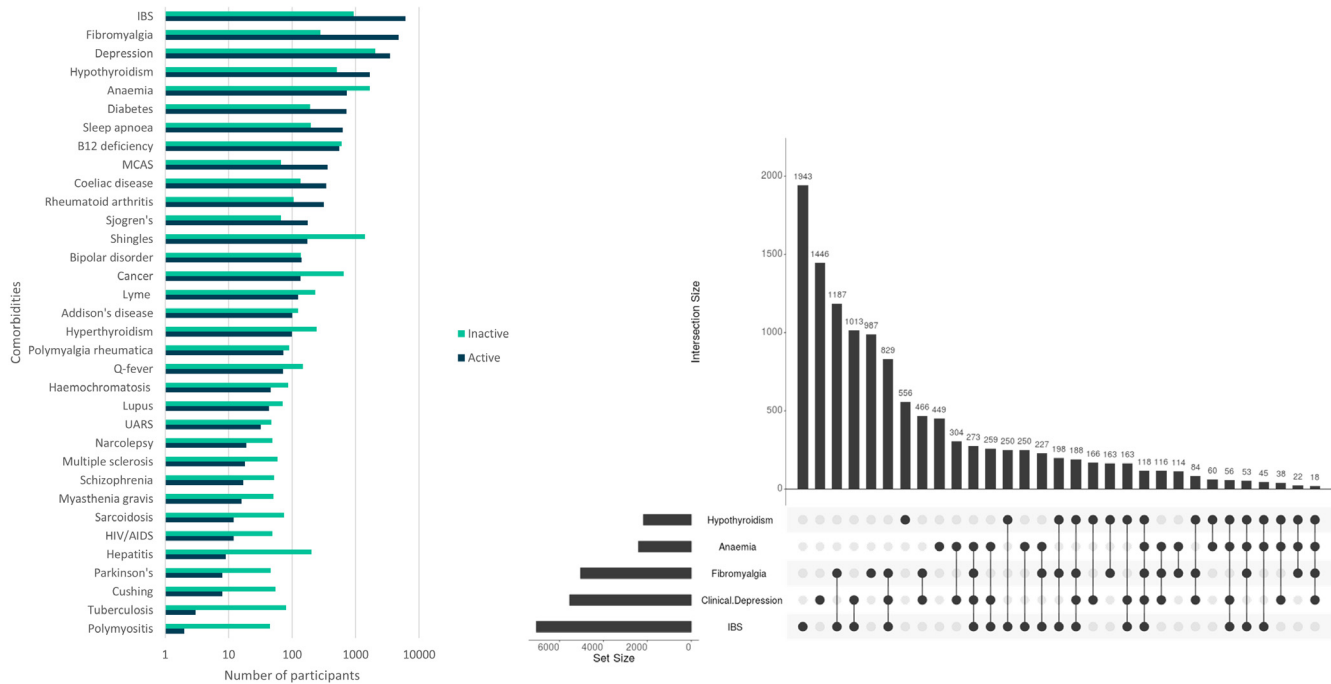
Figure 2. Numbers of DecodeME participants reporting conditions co-occurring with ME/CFS (comorbidities); total, 17,074 participants. In (**A**) numbers are shown in log_10-scale and those with active or inactive comorbidities are indicated in blue or green, respectively. The UpSet plot[16] (**B**) shows numbers of participants with five conditions that most frequently co-occur with ME/CFS. These either co-occur together with others (indicated by filled circles linked by lines) or else separately (filled circles not linked by lines).
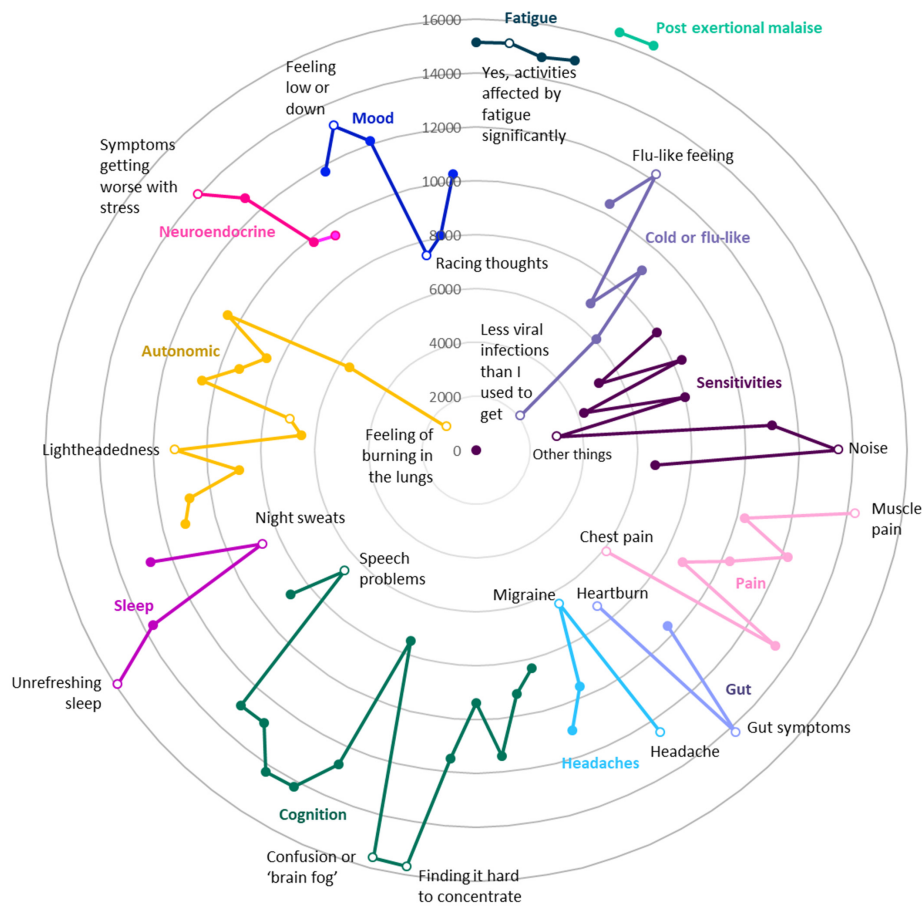
Figure 3. Numbers of DecodeME participants reporting symptoms (Radar chart); total, 17074 participants. Most frequently reported symptoms are furthest from these circles' centre. Twelve different groups of questions are indicated in separate colours; for each symptom group, the most and least frequently reported symptoms are listed and indicated as unfilled circles. With reference to the DecodeME questionnaire (www.decodeme.org.uk/app/uploads/2022/08/DecodeME-Questionnaire.pdf) the questions (Q) are, clockwise: Fatigue (Q8-answer 3 [Q8-3], Q9-3, Q3-1, Q10-1), Post-exertional malaise (PEM, Q12-1 AND Q13-1), Cold or flu-like (Q14-4, -2, -1, -5, -6, -3), Sensitivities (Q15-1, -2, -3, -5, -7, -9, -4, -6, -8), Pain (Q16-4, -6, -5, -3, -2, -7, -1), Gut (Q17-1, -2, -3), Headaches (Q18-1, -4, -2, -3), Cognition (Q19-15, -7, -8, -9, -12, -3, -1, -2, -6, -10, -5, -13, -4, -14, -11), Sleep (Q20-4, -3, -2, -1), Autonomic (Q21-3, -6, -11, -10, -9, -4, -2, -1, -5, -7, -12, -8), Neuroendocrine (Q22-3, -1, -2, -4), and Mood (Q23-2, -3, -1, -5, -6, -4).

sex-biases in comorbidities or symptoms. Females with ME/CFS reported more comorbidities and symptoms than males in the DecodeME questionnaire. Two-thirds (66.7%; $n=$ 9,507) of females, but a half (52.7%; $n=1,489$) of males, reported at least one active comorbidity; similarly 39.2% ($n=5,588$) of females and 28.6% ($n=809$) of males reported at least one inactive comorbidity. Female participants reported, on average, more symptoms than males (42 versus 36).

To test more formally for an association between age and sex and each symptom we used logistic regression and a Bonferroni correction to adjust for multiple testing (Methods). This identified 62 of 80 symptoms as significantly female-biased, and 61 as biased towards younger age (Figure 4). Female-bias is evident across all symptom types (Figure 4). Females were significantly more likely to report fatigue 'often, repeatedly, or all the time' ($p=5.9\times10^{-4}$; age $p=1.1\times10^{-13}$), and

more likely to report post-exertional malaise after physical or mental activity ($p=2.8\times10^{-4}$; age $p=4.2\times10^{-7}$).

Factors, comorbidities and symptoms associated with ME/CFS severity

Being female, increasing age and being over 10y from ME/CFS onset are each separately associated with severity in the DecodeME cohort (sex: $p=4.5\times10^{-4}$; age: $p<2.2\times10^{-16}$; years since ME/CFS onset: $p=1.6\times10^{-6}$). These results are from a comparison of those with mild ME/CFS (34%; $n=5,779$) against the remaining 66% ($n=11,295$) with moderate, severe or very severe illness. Testing for all 68 co-occurring (active and inactive) comorbidities, and including both age and sex as covariates in the model, 6 active comorbidities were significantly associated with severity. In order of decreasing significance these were: fibromyalgia ($p<2\times10^{-16}$), clinical depression ($p<2\times10^{-16}$), irritable bowel syndrome ($p=5.7\times10^{-12}$), mast cell
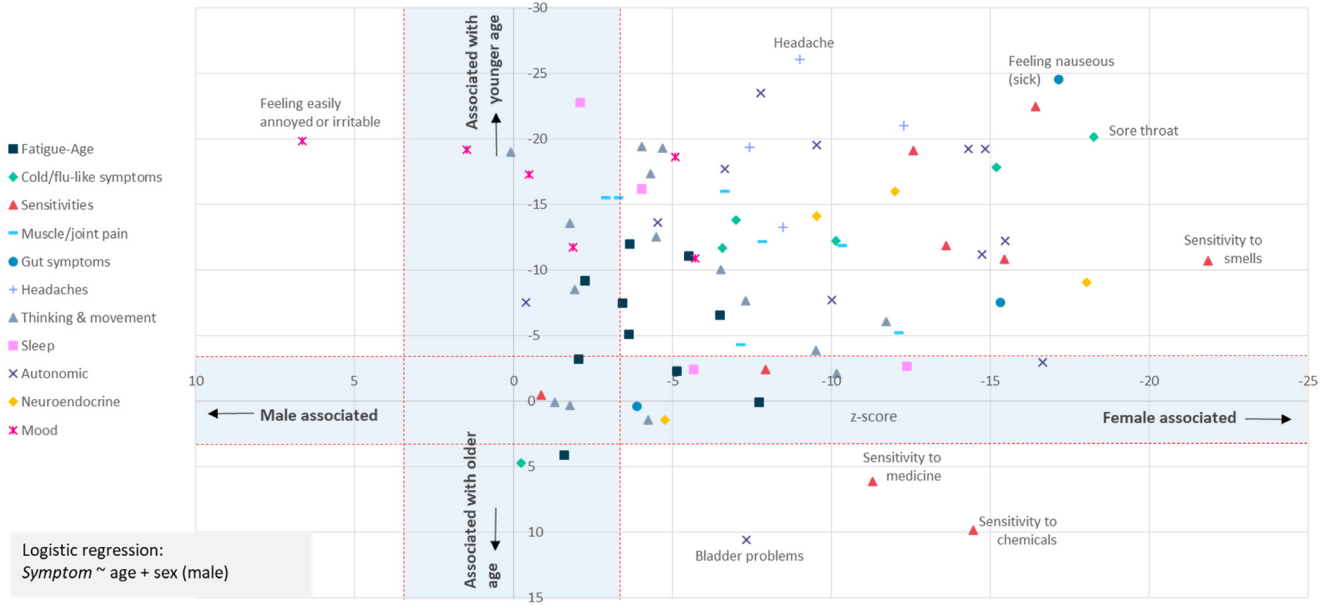
Figure 4. Most symptoms are strongly associated with female sex at birth and younger age. The question asked was: "In the last 6 months, have you had any of the symptoms below often, repeatedly, or constantly? Please mark any that apply." Sex-biased (X-axis) and/or Age-biased (Y-axis) associations in a logistic regression analysis (*Symptom* ~ age + sex + intercept) are shown as data points. Data points within the blue-shaded areas are not significant after accounting for 82 tests (*p*<0.05/82, or |Z|<3.427. Only one symptom ("Feeling easily annoyed or irritable") was male-biased; 3 symptoms (sensitivities to chemicals or medicine, or bladder problems) were associated with older age. Results for 80 symptoms are shown.
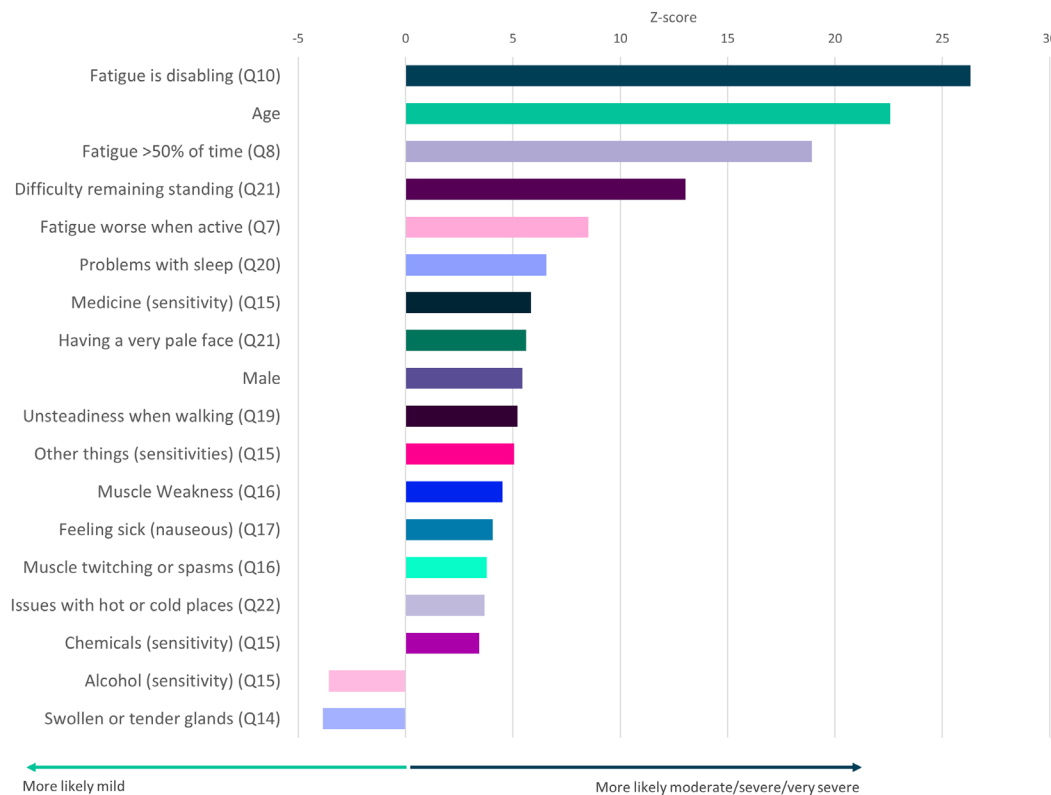
Figure 5. Questionnaire responses that significantly associate with ME/CFS symptom severity. Z-scores are shown for symptoms that significantly associate with severity (*p*<0.05 after Bonferroni correction for 82 tests, including age and sex). Here severity is defined by self-report of moderate or severe or very severe symptoms versus self-report of mild symptoms (see Figure 1). Responses to questions 14 and 15 (Q14, Q15) are significantly associated with mild symptoms. Responses relate to DecodeME Questionnaire questions (e.g. question 10, Q10).
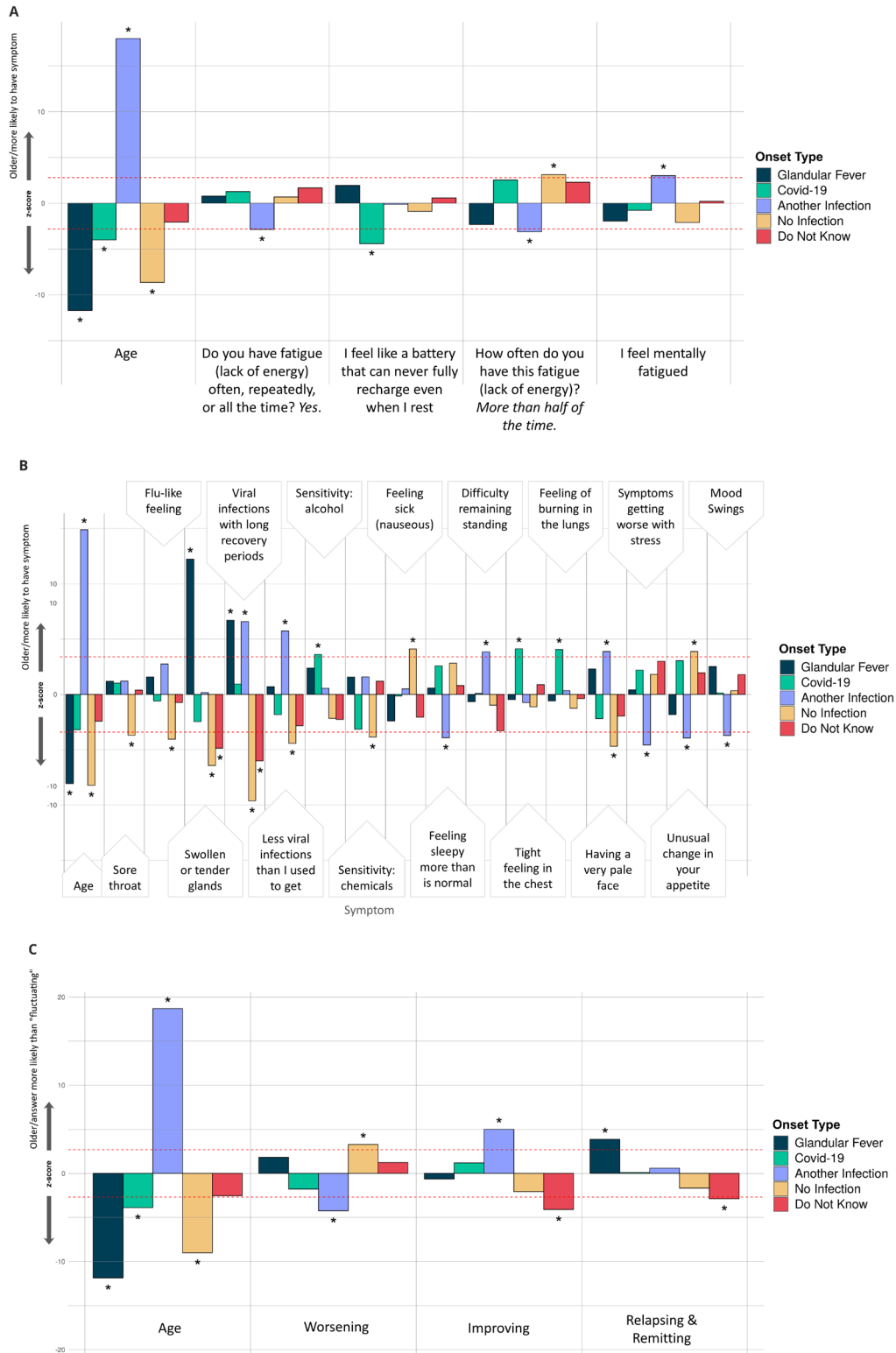
Figure 6. Associations of symptoms or age to 5 ME/CFS onset types: (**A**) Fatigue symptoms (10 tests), (**B**) Non-fatigue symptoms (74 tests), and (**C**) Illness course descriptions (7 tests). These were considered in a logistic regression model of the form OnsetType ~ age + sex + symptoms/descriptions and an intercept. A covariate is only shown if it survived Bonferroni multiple testing correction ($p<0.05$) per regression for one or more symptom/description. Significant associations are indicated with an asterisk (*); their Z-scores lie outside of non-significant values, bounded by the red dashed lines, after Bonferroni multiple testing correction. The z-score (Y-axis) is the effect-size estimate in standard deviation units.
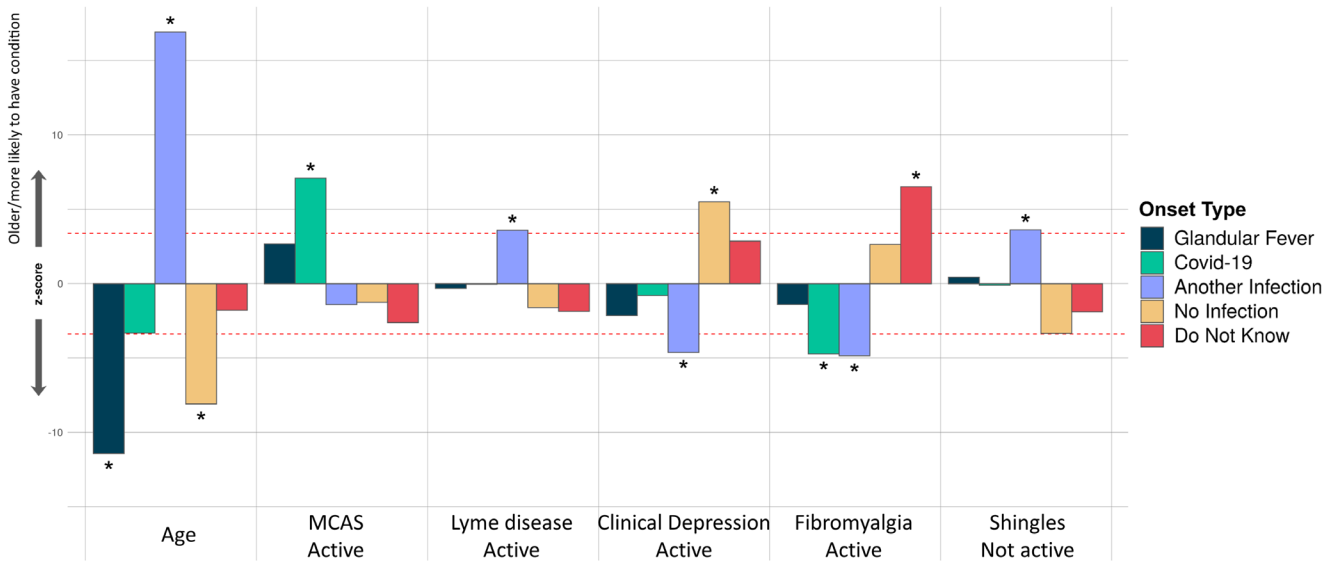
Figure 7. Associations of comorbidities or age to 5 ME/CFS onset types. Thirty-four comorbidities were considered in a logistic regression model of the form OnsetType ~ age + sex + comorbidities and an intercept. A covariate is only shown if it survived Bonferroni multiple testing correction (p<0.05) for one or more onset type. Active and inactive comorbidities were considered independently: Active, if the condition has given symptoms in the past 6 months, or Inactive, if the condition has not given symptoms in the past 6 months, either because it has died down or treatment has controlled it. The z-score (Y-axis) is the effect-size estimate in standard deviation units.

activation syndrome ($p=1.8\times10^{-11}$), diabetes ($p=9.5\times10^{-10}$) and sleep apnoea ($p=5.2\times10^{-8}$). Severity was also associated with a single inactive comorbidity, hypothyroidism ($p=1.6\times10^{-5}$).

Testing all symptoms simultaneously with sex and age, showed strong association between ME/CFS severity and 18 factors including fatigue, age, difficulty remaining standing, and sleep problems (Figure 5). Finally, participants describing their illness as relapsing and remitting were significantly less likely to report their illness as moderate, severe or very severe than those reporting fluctuating symptoms ($p<2.2\times10^{-16}$).

The type of infectious or non-infectious disease onset does not explain these strong and pervasive sex-biases because proportions of females were not significantly different across the five onset types (83.1%-84.5%; $\chi^2 = 1.707$, df = 4, $p = 0.79$).

Length of illness, symptoms and comorbidities differ by onset type

A feature that strongly distinguished among the five onset types was longevity of participants' ME/CFS symptoms. Participants reporting an infection at onset were more likely to have had ME/CFS symptoms for over 10y than those reporting no infection at onset (66.8% [n=7,246] vs. 45.1% [n=1,183]). This is despite their similar ages (medians 54y [IQR=43y-64y] and 52y [IQR=41y-62y], respectively).

The statistical significance of this difference is strong. When testing for association between those with an infection around the time of ME/CFS onset and duration (<10y vs. >10 years

since time of onset), age and sex, only association with duration was significant ($p = 4\times10^{-67}$). This relative paucity of participants not reporting an infection around the time of onset of their ME/CFS over 10y ago is unexpected, and not easily explained by historic variation in ME/CFS triggers because association with age was not significant in this analysis ($p > 0.05$). When analysed separately, each onset type was not associated with participants' sex at birth, when including age and ME/CFS duration over 10y in the analysis.

Significant differences occurred between the 5 ME/CFS onset types and 4 fatigue symptoms (Figure 6A), 16 other symptoms (Figure 6B) and 3 different types of illness course (Figure 6C). Those with glandular fever onset were significantly more likely than others to report swollen or tender glands and viral infections with long recovery periods within the last 6 months, and to experience relapsing and remitting symptoms (relative to 'Fluctuating', the majority response). Others with COVID-19 infection at ME/CFS onset preferentially reported a tight feeling in the chest, sensitivity to alcohol and a feeling of burning in the lungs. Participants with other types of infection onset more frequently reported feeling mentally fatigued, feeling fatigued less than half the time, and difficulties remaining standing, and less frequently reported feeling more sleepy than is normal, having worsening symptoms (relative to 'Fluctuating'), unusual changes in appetite and mood swings.

Participants reporting an infectious onset (when compared to those who did not) were also significantly more likely to

report: improving symptoms, relapsing/remitting, or recovered (relative to 'Fluctuating') symptoms, and less likely to report worsening symptoms (again, relative to 'Fluctuating'). They were more likely, among other things, to report viral infections with long recovery periods, fewer viral infections than they used to get, and having a pale face. Other symptoms that were significantly more likely to be reported by participants without an identified infection at onset were fatigue more than half the time, reduced libido, and unusual changes in appetite. They were also less likely to report symptoms common during infection: flu-like feelings, and swollen or tender glands.

Those with an infection around the time of onset of ME/CFS more frequently reported symptoms typical of infection in the last 6 months, whereas those reporting no infection at onset less frequently indicated these symptoms. This was unexpected because of the long time-lag between onset (mostly >10y ago) and participants' recent questionnaire responses. Even though most participants report a long interval between their onset of ME/CFS (mostly >10y ago) and their recent symptoms characteristic of infection, our results cannot distinguish between whether these recent symptoms are a natural consequence of their ME/CFS onset, for example because of viral persistence in some individuals[17], or else they are independent of onset.

In our final analysis, we tested for association between participants' onset type and their comorbidities, age and sex. Only younger age, rather than any comorbidity, was significantly associated with glandular fever onset (Figure 7). Among all onset types, only coronavirus disease 2019 (COVID-19 caused by SARS-CoV-2) infection was significantly associated active Mast Cell Activation Syndrome (MCAS), i.e. MCAS symptoms within the previous 6 months. COVID-19 related onset was also negatively associated with active fibromyalgia. Onset with another infection was positively associated with inactive Shingles or active Lyme disease, and negatively associated with fibromyalgia or clinical depression. Onset without reported infection at onset was significantly associated with recent clinical depression symptoms; and, onset with unknown infection status was significantly associated with active fibromyalgia as a comorbidity (Figure 7).

In summary, we report significant associations to five onset types derived from participants' responses to the question 'Did you have an infection when, or just before, your first ME/CFS symptoms started?':

1. '**Yes, glandular fever**' (17%; *n*=2,936): These participants were more likely to report swollen or tender glands and viral infections with long recovery periods, and to experience relapsing and remitting symptoms.

2. '**Yes, COVID-19**' (2%; *n*=380): These participants were more likely to report having Mast Cell Activation Syndrome, a tight feeling in the chest or a burning feeling in the lungs. Mast cell activation symptoms are prevalent in Long-COVID[18] but this condition is rarely diagnosed in people with ME/CFS[19] although perhaps

because only recently have MCAS diagnostic criteria been defined[20].

3. '**Yes, another infection**' (44%; *n*=7,537): These participants were more likely to be mentally fatigued, to report viral infections needing long recovery periods, and to have had Shingles in the past or symptomatic Lyme disease in the last 6 months. They were also less likely than others to report active clinical depression or fibromyalgia. Over 100 types of infections have been reported to occur at ME/CFS onset[11].

4. '**No**' (i.e. no infection at onset; 16%; *n*=2,625): These participants were more likely to report fatigue more than half of the time, to feel nauseous, and to have recent clinical depression symptoms.

5. '**Don't know**' (21%; *n*=3,596): These were more likely to report fibromyalgia as a comorbidity, and less likely to report cold or flu-like symptoms, improving or relapsing and remitting symptoms.

Discussion

DecodeME questionnaire responses from *n*=17,074 participants reveal how people, who report being diagnosed with ME/CFS, do not form a single homogeneous group. Although this was long suspected[21], it had not previously been substantiated using a large country-wide cohort ascertained using a single protocol. More specifically, the cohort's heterogeneity was most evident in four respects: (1) large and statistically significant differences among five ME/CFS onset types, relating to their different associations to symptoms, comorbidities and illness severity; (2) the greater likelihood of participants to have longstanding (>10y) ME/CFS symptoms if they report an infection at onset; (3) substantial differences between females and males in their symptoms and comorbidities; and (4) greater disease severity for those who are female, older and/or have had ME/CFS for >10y.

This initial DecodeME cohort has a comparable age-distribution to previous USA-based studies[22–24], the reported comorbidities (Figure 2) are similar to those of a previous study[25], and proportions of participants reporting glandular fever or another infectious disease around the time of onset are similar to those previously reported[11,25,26]. The DecodeME cohort's females outnumber males by over five-to-one, which is one of the highest female-bias among those with ME/CFS yet reported internationally[3,9,27–32]. Cohorts of these previous questionnaire studies numbered in the hundreds. DecodeME's larger cohort thus provides robust statistical support to these previous findings from less well-powered studies.

Studies involving hundreds of participants previously concluded that ME/CFS exhibits few sex differences in illness patterns[33,34]. Smaller studies indicated older age as associated with greater ME/CFS symptom severity, but other studies found no such association (reviewed in 12,35). These previously limited cohort sizes did not permit comprehensive analysis. In a previous study, three symptoms were reported significantly more often by females than males: fever, swollen glands,

and sore throat[34]. In our study, we replicated these findings, and found a further 59 of 80 ME/CFS symptoms that are also female-biased. Our analyses additionally found 61 symptoms biased towards younger age, with only 5 biased towards older age.

The raw number of symptoms may not be meaningful, however, as symptoms can be overlapping, and people with ME/CFS may, over time, pace sufficiently to avoid triggering some symptoms or may begin to describe their symptoms with fewer labels, particularly when interventions are not available to treat each symptom effectively. Indeed, rather than younger participants reporting increased severity, we found that being female, older and over 10y from onset are all risk factors for ME/CFS severity.

The median time to receive a clinical diagnosis in the UK is 2 years which is reflected in DecodeME participants' responses. Specifically, participants whose illness started within the last 1-3y or 0.5-1y (n = 1,287 and 177) were respectively 21% and 57% fewer per year than the study's participants from the 3-5y recruitment interval (n=1,634).

Despite its large cohort size (N=17,074), extensive community reach and use of paper, as well as online, questionnaires, the analysis presented here – of the December 2022 DecodeME data freeze – has four main limitations. First, recruitment is restricted to participants over the age of 16y, which limited investigation of paediatric ME/CFS. Second, when asking participants if they were diagnosed by a health professional we did not require clinical confirmation of reported answers. Nevertheless, our extensive engagement with participants and the internal consistency of their responses encourage us to believe that questionnaire answers have been given in good faith, noting that inconsistent responses may result from respondents' ME/CFS symptoms including their cognitive dysfunction. Third, most ME/CFS symptoms are not independent of one another. Consequently, multicollinearity should be borne in mind for those analyses considering multiple symptoms in the same analysis. Fourth, regrettably DecodeME has not yet been successful in recruiting proportionately from minoritised groups. There is little consensus on whether ME/CFS prevalence differs among these and other groups[36]. Other recruitment and representativeness biases are also possible, as with all research cohorts.

A previous study indicated that ME/CFS onset type associates with severity[37] although this was not replicated by our larger study. Instead, we identified large numbers of comorbidities and symptoms that are each more likely to be reported by participants with a specific onset type.

These onset types reveal differences amongst those with ME/CFS regarding their symptoms and comorbidities (Figure 4). However, these distinctions are not absolute. For example, those reporting no infection at onset (Type 4, above) are not cleanly distinguished from all others by active clinical depression. Rather, they were the only onset type that was more likely

to report this diagnosis (25.4%; n=667) than all other participants were (19.6%; n=2829). Similarly, Type 3 ("other infection") contains a higher proportion (9.4%; n=710) of those who report inactive shingles, than all other participants (7.3%; n=692). Shingles is caused by reactivation of latent varicella-zoster virus (a herpesvirus). People with herpes zoster infection are known to have a significantly higher risk of ME/CFS up to at least 6 years[38] fuelling speculation that varicella-zoster virus infection is a cause of ME/CFS that may be prevented by vaccination. 2.5% of ME/CFS cases have been attributed to varicella-zoster virus infection[11]. We note that among those reporting no infection around the time of onset (Type 4) some may have developed ME/CFS secondary to an infection without an obvious acute phase, such as can occur with Epstein-Barr virus[39]. However, we are unable to test this hypothesis here.

ME/CFS' poor long-term prognosis, its severe symptoms – especially for older females, its profound impact on the quality of life of people with ME/CFS and their family members[9,32], and its high population prevalence (>0.2%)[1] present formidable healthcare and research challenges. Considering that 64% (n=10,853) of DecodeME participants reported an infection around the time of onset, any vaccination against the major infectious agents triggering ME/CFS, including Epstein-Barr virus[40], SARS-CoV-2[41] and influenza viruses[42] may help reduce ME/CFS incidence in the future, especially for individuals more susceptible to severe disease, or those more likely to be exposed to the infectious agents.

Formal investigation of each onset type's significance is now warranted. In order to address effectively the devastating impact that ME/CFS has on the millions of people worldwide affected, the research community and policy-makers will need to give sustained focus on disease classification and aetiology. It is for this reason that DecodeME is seeking to identify genetic factors causal of altered ME/CFS risk[14] and will do so for infectious versus non-infectious onset participants separately and combined, if final recruitment numbers allow. Recruitment to the DecodeME study is ongoing. Results will be updated after the project's recruitment phase has ended.

Data availability

Extended Data is available from the Open Science Framework (OSF) website: https://osf.io/rgqs3/. This Project site contains CC-BY license DecodeME Study Documents: (i) the DecodeME Questionnaire (version 6) annotated by question identifier (Qid) which were used in the logistic regression analyses detailed in (ii) the regressionResults.txt file. DecodeME anonymised data allowing investigation of this study's consented data are available to researchers by managed access via a Data Access Committee, https://www.decodeme.org.uk/faqs/who-will-be-able-to-use-my-data-and-sample/. This committee consists of a scientist, a patient and a charity representative who strictly control access to the data. DecodeME's anonymised and consented data are only shared with studies that meet high standards and whose academic or industrial researchers agree to treat its data with respect and to keep it secure.

Acknowledgements

References

1. Nacul LC, Lacerda EM, Pheby D, *et al.*: **Prevalence of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) in three regions of England: a repeated cross-sectional study in primary care.** *BMC Med.* 2011; **9**: 91.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Baker R, Shaw EJ: **Diagnosis and management of chronic fatigue syndrome or myalgic encephalomyelitis (or encephalopathy): summary of NICE guidance.** *BMJ.* 2007; **335**(7617): 446–8.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Institute of Medicine: **Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness.** Washington, DC: The National Academies Press, 2015.
   PubMed Abstract | **Publisher Full Text**

4. Hickie I, Davenport T, Wakefield D, *et al.*: **Post-infective and chronic fatigue syndromes precipitated by viral and non-viral pathogens: prospective cohort study.** *BMJ.* 2006; **333**(7568): 575.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Katz BZ, Shiraishi Y, Mears CJ, *et al.*: **Chronic fatigue syndrome after infectious mononucleosis in adolescents.** *Pediatrics.* 2009; **124**(1): 189–93.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Sukocheva OA, Maksoud R, Beeraka NM, *et al.*: **Analysis of post COVID-19 condition and its overlap with myalgic encephalomyelitis/chronic fatigue syndrome.** *J Adv Res.* 2022; **40**: 179–96.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Ayoubkhani D, Bermingham C, Pouwels KB, *et al.*: **Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study.** *BMJ.* 2022; **377**: e069676.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Reyes M, Dobbins JG, Nisenbaum R, *et al.*: **Chronic Fatigue Syndrome Progression and Self-Defined Recovery.** *J Chronic Fatigue Syndr.* 1999; **5**(1): 17–27.
   Publisher Full Text

9. Hvidberg FM, Brinth LS, Olesen AV, *et al.*: **The Health-Related Quality of Life for Patients with Myalgic Encephalomyelitis / Chronic Fatigue Syndrome (ME/CFS).** *PLoS One.* 2015; **10**(7): e0132421.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. Tyson S, Stanley K, Gronlund TA, *et al.*: **Research priorities for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): the results of a James Lind alliance priority setting exercise.** *Fatigue.* 2022; **10**(4): 200–211.
    Publisher Full Text

11. Jason LA, Yoo S, Bhatia S: **Patient perceptions of infectious illnesses preceding Myalgic Encephalomyelitis/Chronic Fatigue Syndrome.** *Chronic Illn.* 2022; **18**(4): 901–10.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Cairns R, Hotopf M: **A systematic review describing the prognosis of chronic fatigue syndrome.** *Occup Med (Lond).* 2005; **55**(1): 20–31.
    PubMed Abstract | **Publisher Full Text**

13. Carruthers BM, Jain AK, De Meirleir KL, *et al.*: **Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome.** *J Chronic Fatigue Syndr.* 2003; **11**(1): 7–115.
    Publisher Full Text

14. Devereux-Cooke A, Leary S, McGrath SJ, *et al.*: **DecodeME: community recruitment for a large genetics study of myalgic encephalomyelitis / chronic fatigue syndrome.** *BMC Neurol.* 2022; **22**(1): 269.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Kuri A, Jacobs BM, Vickaryous N, *et al.*: **Epidemiology of Epstein-Barr virus infection and infectious mononucleosis in the United Kingdom.** *BMC Public Health.* 2020; **20**(1): 912.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Lex A, Gehlenborg N, Strobelt H, *et al.*: **UpSet: Visualization of Intersecting Sets.** *IEEE Trans Vis Comput Graph.* 2014; **20**(12): 1983–92.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Rasa S, Nora-Krukle Z, Henning N, *et al.*: **Chronic viral infections in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS).** *J Transl Med.* 2018; **16**(1): 268.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Weinstock LB, Brook JB, Walters AS, *et al.*: **Mast cell activation symptoms are prevalent in Long-COVID.** *Int J Infect Dis.* 2021; **112**: 217–26.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Afrin LB, Self S, Menk J, *et al.*: **Characterization of Mast Cell Activation Syndrome.** *Am J Med Sci.* 2017; **353**(3): 207–15.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Valent P, Akin C, Arock M, *et al.*: **Definitions, criteria and global classification of mast cell disorders with special reference to mast cell activation syndromes: a consensus proposal.** *Int Arch Allergy Immunol.* 2012; **157**(3): 215–25.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Huber KA, Sunnquist M, Jason LA: **Latent class analysis of a heterogeneous international sample of patients with myalgic encephalomyelitis/chronic fatigue syndrome.** *Fatigue.* 2018; **6**(3): 163–78.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Jason LA, Richman JA, Rademaker AW, *et al.*: **A community-based study of chronic fatigue syndrome.** *Arch Intern Med.* 1999; **159**(18): 2129–37.
    PubMed Abstract | **Publisher Full Text**

23. Reyes M, Nisenbaum R, Hoaglin DC, *et al.*: **Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas.** *Arch Intern Med.* 2003; **163**(13): 1530–6.
    PubMed Abstract | **Publisher Full Text**

24. Reeves WC, Jones JF, Maloney E, *et al.*: **Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia.** *Popul Health Metr.* 2007; **5**: 5.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

25. Chu L, Valencia IJ, Garvert DW, *et al.*: **Onset Patterns and Course of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome.** *Front Pediatr.* 2019; **7**: 12.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Naess H, Sundal E, Myhr KM, *et al.*: **Postinfectious and chronic fatigue syndromes: clinical experience from a tertiary-referral centre in Norway.** *In Vivo.* 2010; **24**(2): 185–8.
    PubMed Abstract

27. Collin SM, Bakken IJ, Nazareth I, *et al.*: **Trends in the incidence of chronic fatigue syndrome and fibromyalgia in the UK, 2001-2013: a Clinical Practice Research Datalink study.** *J R Soc Med.* 2017; **110**(6): 231–44.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Bakken IJ, Tveito K, Gunnes N, *et al.*: **Two age peaks in the incidence of chronic fatigue syndrome/myalgic encephalomyelitis: a population-based registry study from Norway 2008-2012.** *BMC Med.* 2014; **12**: 167.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

29. Orji N, Campbell JA, Wills K, *et al.*: **Prevalence of myalgic encephalomyelitis/ chronic fatigue syndrome (ME/CFS) in Australian primary care patients: only part of the story?** *BMC Public Health.* 2022; **22**(1): 1516.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

30. Pheby D, Lacerda E, Nacul L, *et al.*: **A Disease Register for ME/CFS: Report of a Pilot Study.** *BMC Res Notes.* 2011; **4**: 139.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

31. Pendergrast T, Brown A, Sunnquist M, *et al.*: **Housebound versus nonhousebound patients with myalgic encephalomyelitis and chronic fatigue syndrome.** *Chronic Illn.* 2016; **12**(4): 292–307.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

32. Vyas J, Muirhead N, Singh R, *et al.*: **Impact of myalgic encephalomyelitis/ chronic fatigue syndrome (ME/CFS) on the quality of life of people with ME/CFS and their partners and family members: an online cross-sectional survey.** *BMJ Open.* 2022; **12**(5): e058128.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

33. Buchwald D, Pearlman T, Kith P, *et al.*: **Gender differences in patients with chronic fatigue syndrome.** *J Gen Intern Med.* 1994; **9**(7): 397–401.
    PubMed Abstract | **Publisher Full Text**

34. Tseng CL, Natelson BH: **Few Gender Differences Exist Between Women and Men with Chronic Fatigue Syndrome.** *Journal of Clinical Psychology in Medical Settings.* 2004; **11**(1): 55–62.
    Publisher Full Text

35. Joyce J, Hotopf M, Wessely S: **The prognosis of chronic fatigue and chronic fatigue syndrome: a systematic review.** *QJM.* 1997; **90**(3): 223–33.
    PubMed Abstract | **Publisher Full Text**

36. Dinos S, Khoshaba B, Ashby D, *et al.*: **A systematic review of chronic fatigue, its syndromes and ethnicity: prevalence, severity, co-morbidity and coping.** *Int J Epidemiol.* 2009; **38**(6): 1554–70.
    PubMed Abstract | **Publisher Full Text**

37. Sharpe M, Hawton K, Seagroatt V, *et al.*: **Follow up of patients presenting with fatigue to an infectious diseases clinic.** *BMJ.* 1992; **305**(6846): 147–52.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

38. Tsai SY, Yang TY, Chen HJ, *et al.*: **Increased risk of chronic fatigue syndrome following herpes zoster: a population-based study.** *Eur J Clin Microbiol Infect Dis.* 2014; **33**(9): 1653–9.
    PubMed Abstract | **Publisher Full Text**

39. Womack J, Jimenez M: **Common questions about infectious mononucleosis.** *Am Fam Physician.* 2015; **91**(6): 372–6.
    PubMed Abstract

40. Zhong L, Krummenacher C, Zhang W, *et al.*: **Urgency and necessity of Epstein-Barr virus prophylactic vaccines.** *NPJ Vaccines.* 2022; **7**(1): 159.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

41. Shaheen N, Shaheen A: **Long-term sequelae of COVID-19 (myalgic encephalomyelitis): An international cross-sectional study.** *Medicine (Baltimore).* 2022; **101**(45): e31819.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

42. Magnus P, Gunnes N, Tveito K, *et al.*: **Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is associated with pandemic influenza infection, but not with an adjuvanted pandemic influenza vaccine.** *Vaccine.* 2015; **33**(46): 6173–7.
    PubMed Abstract | **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ❓ ✔️

Version 2

Reviewer Report 13 July 2023

https://doi.org/10.3310/nihropenres.14591.r29746

❓
Sarah Tyson

Division of Nursing, Midwifery & Social Work, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, England, UK

Thank you for asking me to review this resubmission. The authors have obviously undertaken a lot of work and the paper is much improved. Congratulations. It is now much easier to follow and the scientific merit is clear. I am happy to accept the changes made, except I would like the authors to reconsider one point, which is now starker given the helpful revisions in the paper.

I am not really clear whether the authors have answered the original research question – are there different types of ME/CFS? Are they saying –yes, illness onset defines different types of ME? There seems to be some leaning towards this interpretation but it isn't stated explicitly. If so, I am not convinced. Not because of anything in the method or analysis (as far as I can tell, that is fine), but because they started with the assumption that onset type (and they predefined the types of onset) was the definitive defining variable. I'm no statistician, but I think in a cohort and dataset of this size, one is bound to find some associations and significant differences. The issue is the importance of those found.

The essential section of the discussion regarding the importance of these findings and their potential application in future research and clinical practice is missing, possibly because the authors aren't sure themselves.

I think this comes down to the founding basis of the paper; that it started with illness type as the defining characteristic, and had pre-defined the types of onset. What was the rationale for this choice? There is a logic to it, but would other characteristics provide clearer differentiation, or be more relevant to future research efforts or clinical applications? Viral/non-viral onset; severity of symptoms or disability; outcome/course of illness (in terms of recovery, fluctuations deterioration etc); or symptomology according to the body system(s) affected are also obviously clinically applicable and meaningful. I'm not saying that illness onset isn't a (or maybe THE) way to categorise ME, but that to be adopted in practice or research it needs to show a clearer differentiation between types than other possible dependent/ defining variables. I.e. it needs to be the best way of typing PwME, not just a possible way of doing so.

From what I can gather, the upshot of a section on the importance and application of the findings might be something along the lines of …. *"The findings show that sex and age, although associated with symptomology, co-morbidities and disease severity, are not factors that define different types of ME. Type of onset does show some significant differences in terms of symptomology, co-morbidity and illness severity, however these differences are not absolute and using 'type of onset' to define different types of ME would be premature. Further research is needed to investigate the accuracy with which 'type of onset' (probably with additional variables) can categorise pwME, and also to explore other potential ways defining variables, before comparing these with onset type to establish the most effective way to categorise type of ME."*

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: pwME, clinical researcher academic

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 13 July 2023

https://doi.org/10.3310/nihropenres.14591.r29745

✔ **Suzanne Vernon** ⓘD

The Bateman Horne Center of Excellence, Salt Lake City, UT, USA

Thank you for making the revisions to this interesting manuscript.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical research of ME/CFS and PASC

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 25 May 2023

https://doi.org/10.3310/nihropenres.14555.r29219

Suzanne Vernon (iD)
The Bateman Horne Center of Excellence, Salt Lake City, UT, USA

This is a study uses responses to the DecodeME questionnaire to identify ME/CFS subtypes based on self reported type of infection onset that coincided with ME/CFS symptom onset. The choices included "Yes, glandular fever", "Yes, COVID-19", "Yes, another infection", "No", or "Don't know". Responses to this question forms the basis for the symptom analysis and subtyping.

The authors claim that more comorbidity's in females have not been previously reported. However, there are studies by Jason et al [1] and Jones et al [2] that also reported this. Including these important citations would further strengthen the findings of this paper. A couple of citations are provided for your consideration. There are many more that would support these results.

Involvement of patients and the public in the study design is important and adds essential lived experience(s) to the overall DecodeME study design. (This section of the methods should be written more succinctly). However, it is not clear why a standardized and well-validated questionnaire(s), for example the DePaul Symptom Questionnaire (DSQ) was not used rather than or in addition to the DecodeME Questionnaire. The DSQ is validated for ME/CFS, has been used worldwide, and is available in a number of languages. It is also used in the collection of core data elements for ME/CFS (as developed by NIH/NINDS). The use of a new questionnaire that has not been validated for ME/CFS limits the generalizability of these results. Finally, the inclusion of people with comorbid diseases that are exclusionary could explain symptoms limits the ability to replicate these results.

How does a paper questionnaire help decrease or eliminate bias?

The results describe 80+ symptoms yet the DecodeME Questionnaire asks about 12 major symptoms. What accounts for this discrepancy. If each item listed under each major symptom was considered as a separate system, how was collinearity accounted for?

What was the positive test for those that responded "Yes, other infection"?

Duplicated text in the results, 2nd to last paragraph in Results.

The authors point out in the discussion that previous work associated onset type with disease severity and that this study did not find this. Could this be due to the lack of clear severity and frequency assessment of symptoms in the DecodeME Questionnaire?

The major strength of this paper is the >17,000 respondents.

The major weaknesses of this paper are the use of a new ME/CFS questionnaire and the inclusion of exclusionary diseases that could explain the reported symptoms.

References

1. Jason LA, Helgerson J, Torres-Harding SR, Carrico AW, et al.: Variability in diagnostic criteria for chronic fatigue syndrome may result in substantial differences in patterns of symptoms and disability.*Eval Health Prof*. 2003; **26** (1): 3-22 PubMed Abstract | Publisher Full Text

2. Jones JF, Lin JM, Maloney EM, Boneva RS, et al.: An evaluation of exclusionary medical/psychiatric conditions in the definition of chronic fatigue syndrome.*BMC Med*. 2009; **7**: 57 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical research of ME/CFS and PASC

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 18 May 2023

https://doi.org/10.3310/nihropenres.14555.r29220

❌ **Sarah Tyson**

Division of Nursing, Midwifery & Social Work, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, England, UK

The results of the DeCode have been eagerly awaited, even the interim results so there is bound to be a lot of interest in this paper. However I feel it needs some more work to make it as easy to understand as possible. The lack of clarity stems primarily (I think) from a uncertainty about the aim/objective/research questions addressed in this paper. It has a feel of something of a fishing exercise, and having gotten some results, the authors are keen to tell people about them (which is admirable) but aren't really clear about what and why. The methods section, in particular needs more details so the reader can understand what was done, why and how.

To look at each detail in turn:

<u>Abstract</u>
- The abstract needs a clear aim/objective/res question and further methodological detail. At present there is no info about the questionnaire, except that there was one. An indication of the selection criteria, recruitment methods, and details of the questionnaire (what does it measure, how?) And the analysis methods (which need to be linked research questions) are needed. In the results section there needs to be some data and the unique findings made clearer. It is hardly news that ME/CFS is a heterogeneous condition.

- The plain English summary includes rather lot of jargon. For example – "For many people, ME/CFS means a substantial impairment of their activity levels, a high level of disability and a poor quality of life" rather than *"for many people, ME/CFS means disability and poor quality of life .."* NB. The PSP included a range of HCPs not just Doctors,

<u>Methods</u>
- I appreciate that PPI has been central to the DeCode project but there needs to be further details of the actual methods used. The 3rd paragraph of the methods isn't really relevant to what was done in this project and can be removed.

- The description of the questionnaire needs to include its aim; an outline of the content (what does it ask about?), its size/length, types of questions and response format(s). The details of how it was developed and validated and any other psychosomatics would be useful. The reader needs to know what data were collected and how, and be reassured that the information obtained was relevant and accurate ie psychometrically robust.

- The paragraph on the 'cohort' describes the process of completing part of the questionnaire. It would be easier to understand who was recruited to the study and how by listing the selection criteria, and stating the recruitment strategies- where and how were people recruited? I think it was largely via social media and support groups which is fine. Just say so.

- The paragraph on significance testing needs to be replaced by an explanation of the analysis used, linked to the objectives/ research questions. It might be helpful to get a statistician to help with this. I see that logistic regression model(s) was/were used. This needs further detail, particularly the dependent and independent variables. There is reference to figures which are not explained and I cannot really follow. Also Onset Type, symptoms (in groups with different numbers), severity, illness courses, 'Fluctuating', the majority response, and comorbidities are unexplained, undefined and unconnected to any objective or research question. This section is key to understanding the whole paper so needs to clearly describe what was analysed, why and how.

Results

- The 1st paragraph (after the 1st sentence) describes what participants did rather than what was found, and so belongs in the method section. The results need to start with a summary description of the sample – the number, mean age; sex ratio; time since diagnosis +/or duration of symptoms. Then the main co-morbidities; frequency of symptoms etc. The text needs to be understandable without the tables/figures and vice versa. The figures /tables add extra detail but there needs to summary in the text.

- The results section needs to only include what was found in this study. Reflections on how they relate to other studies (or not) are for the discussion (see paragraph 2).

- In the summary of results, try to avoid copy and pasting the results as they were churned out. Eg rather than "Participants were asked: "Did you have an infection when, or just before, your first ME/CFS symptoms started?" with five possible responses: (i) Yes, glandular fever (*n*=2,936), (ii) Yes, COVID-19 (*n*=380), (iii) Yes, another infection (*n*=7,537), (iv) No (*n*=2,625), or (v) Don't know (*n*=3,596; Figure 3). Proportions of people in the first 3 categories reporting a positive laboratory test of their infection prior to ME/CFS were 68.4%, 50.5% and 25.9%, respectively" – which is lengthy and unfocussed Summarise this for the reader by saying (for example) " *Most (n, %) reported an infectious onset to their symptoms, most commonly glandular fever (17%, n=xx) this was glandular fever, followed by covid-19 (n,%). However only 68% and 51% respectively of respondents with these triggers reported a positive laboratory test confirming the diagnosis"*.

- Do not include reference/comparison to other studies in the results section; this is for the discussion section.

- Several of the sentences are rather convoluted and a bit more plain English would not go amiss. For example, I eventually managed to work out that "Participants who started their illness within the last 1–3y or 0.5–1y numbered 1,287 or 354, respectively. These numbers are 57% and 21% fewer, per year, than the study's 1,634 participants from the 3–5y recruitment interval." Actually meant *"Most* 1,634 *(xx%) were diagnosed 3-5 years previously, followed by 1287 (x%) with 1-3 year duration and 354 (X%) were more recently diagnosed (0.5-1 year)"*.

- Always include both the number and % of participants being referred to.

- Avoid any interpretation in the results.

- Whenever presenting an average value an indication of the variability (whether standard deviations for means, or inter-quartile ranges for medians) is needed.

- Why was a sub-group used to assess incidence? I don't think 'incidence' is the proper term. The reference to the frequency with which ME/CFS occurs within a defined general population and this dataset doe sot include that data.

- My comments above also apply to the rest of the results. At present they are difficult to read and to work out the main message from the rests. The reader is left to work out the main gist of the results themselves and it is far from easy.

Discussion

- The 1st paragraph of the discussion needs to summarise the main findings; the headline news in sufficient detail to the reader to have an insight into the significance of the results. Merely saying ME/CFS is heterogeneous with substantial differences in onset; symptoms, comorbidities and severity and sex differences isn't new, we already know this. The headline news needs some (summary) details about what the differences were.

- Then you compare your findings with previous studies – which is great. Make sure all the comparisons and interpretation are here and not in the results. Also try to include any possible explanation for differing/contrasting results (usually due to sample size, method of data collection or selection criteria)

- The limitations section is good.

- The list of symptoms and other characteristics associated with different onset types are new data and belong in the results section – and would help a great deal to understand the results. Although I am a bit confused - why/how was onset type used to sub-group pwME? why not look at cluster/co-occurrence of symptoms and/or co-morbidities +/- onset? It may be that this was done and onset was the strongest factor, if so it needs to be explained more clearly.

- What is the clinical or significance/ implications of the findings? Are you saying onset is such an important/explanatory factor that it should be used to group different types of ME/CFS? Or that they seem the best bet (if so, I didn't pick that up) or what?

- I'm presuming there isn't a conclusion section where the authors relate the findings to the original objectives /research questions and whether they have been fulfilled, because the journal specifically excludes one. If not, please add a conclusion.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
No

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: pwME, clinical researcher academic

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.