PATIENT-LED RESEARCH COLLABORATIVE

UCSF Department of Epidemiology & Biostatistics

Algorithm dependence of patient phenotypes in Long COVID

A patient-led, multi-method clustering of 6031 patients using 162 self-reported symptoms

Alison Cohen, PhD, MPH

Assistant Professor of Epidemiology & Biostatistics

University of California San Francisco

Patient-Led Research Collaborative

November 5, 2026

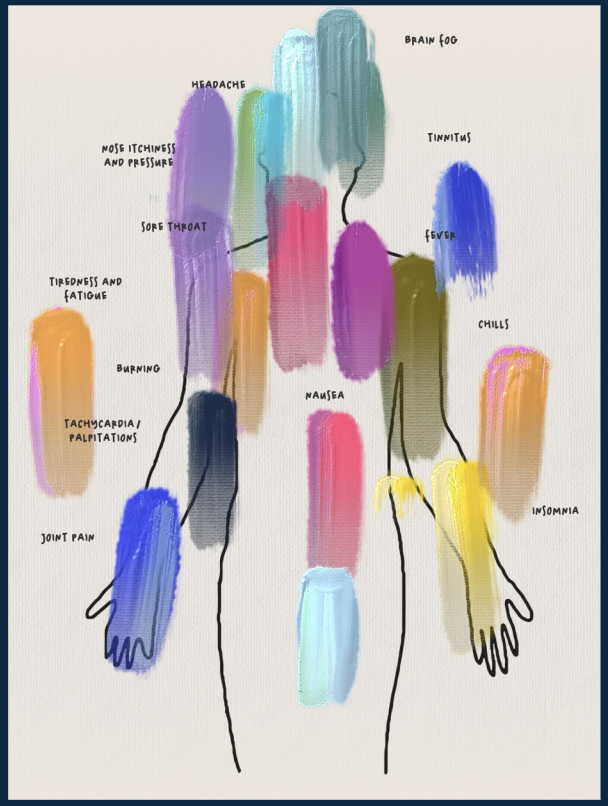
Oxford Open Immunology, 2026, 7(1), iqag010
https://doi.org/10.1093/oxfimm/iqag010
Published: 13 June 2026
Original Research

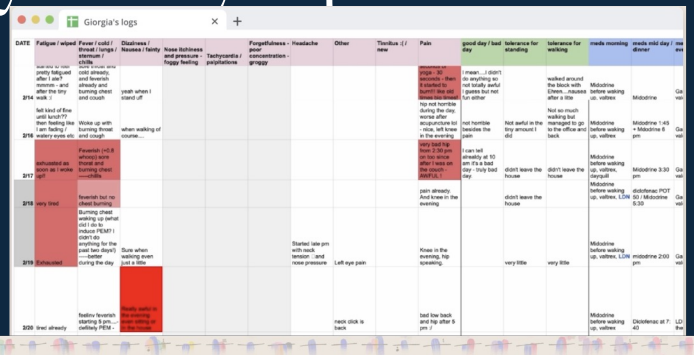
OXFORD

Algorithm dependence of patient phenotypes in Long COVID: a patient-led, multi-method clustering of 6031 patients using 162 self-reported symptoms

Tessa D. Green[1,2,*], Chris McWilliams[1,3,†], Leonardo de Figueiredo[1,†], Letícia Soares[1], Beth Pollack[1,4], Alison K. Cohen[1,5], Tan Zhi-Xuan[6,7], Tess Falor[1], Hannah E. Davis[1]

Long COVID's panoply of symptoms

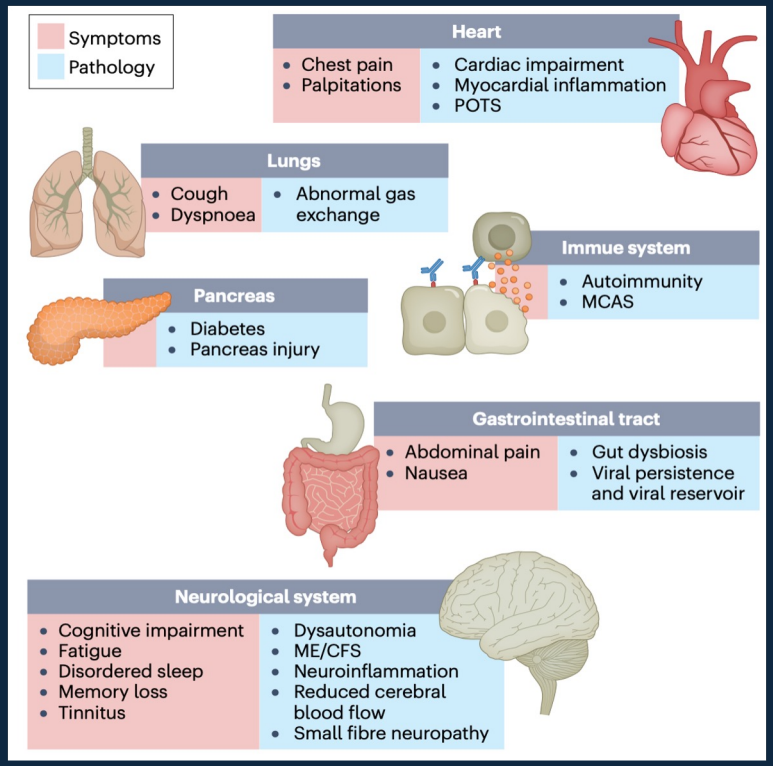


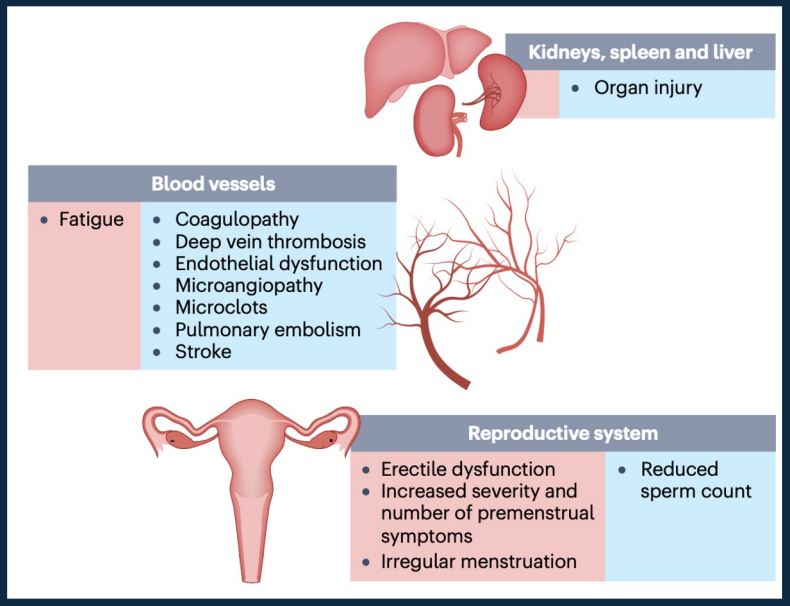




Excerpts from Giorgia Lupi's (2023) NYT op-ed: https://www.nytimes.com/interactive/2023/12/14/opinion/my-life-with-long-covid.html

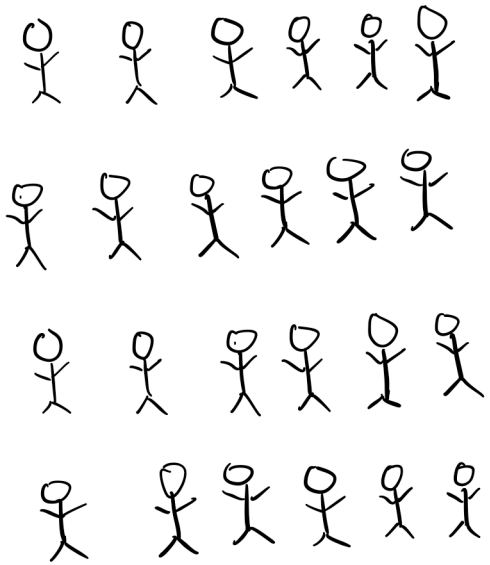
Diversity of symptoms & endpoints





Davis, McCorkell, Vogel, & Topol, 2023

Patient phenotypes guide treatment response, study design, and subtype discovery

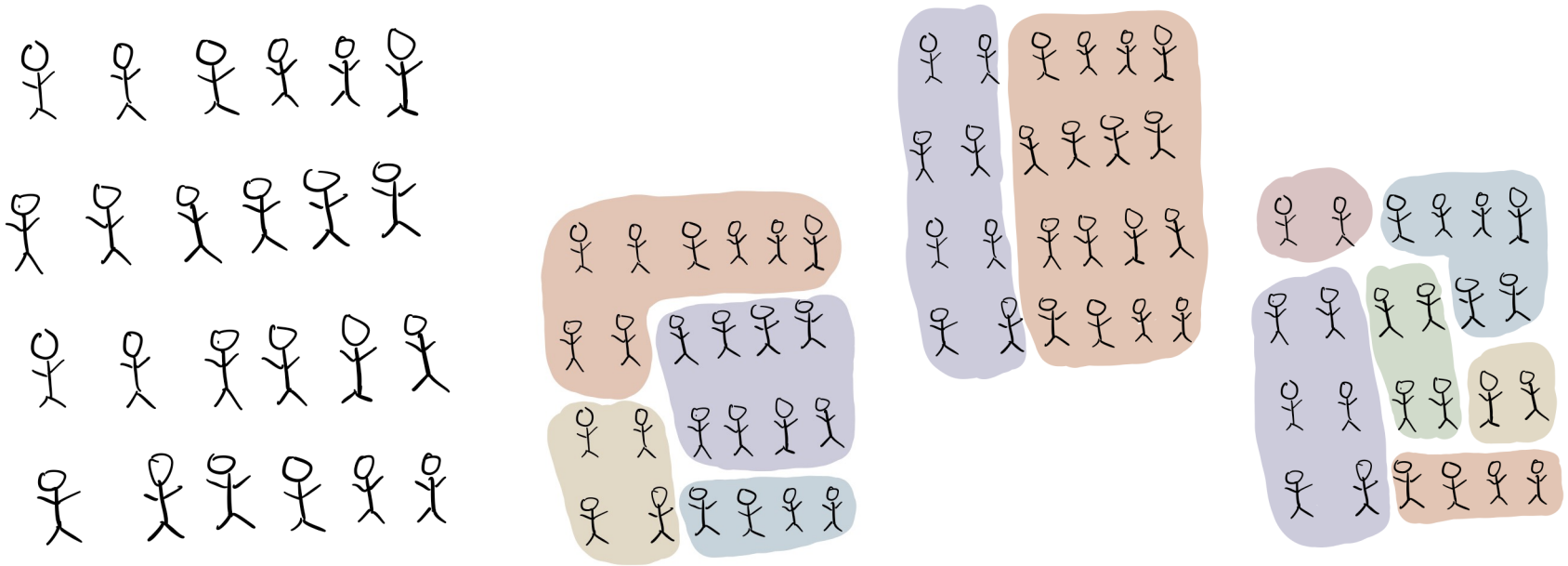


- What subgroups exist?
- Are they reproducible?
- Do they have clinical meaning?

How can we ensure that discovered phenotypes are robust and reproducible?

PATIENT-LED RESEARCH COLLABORATIVE

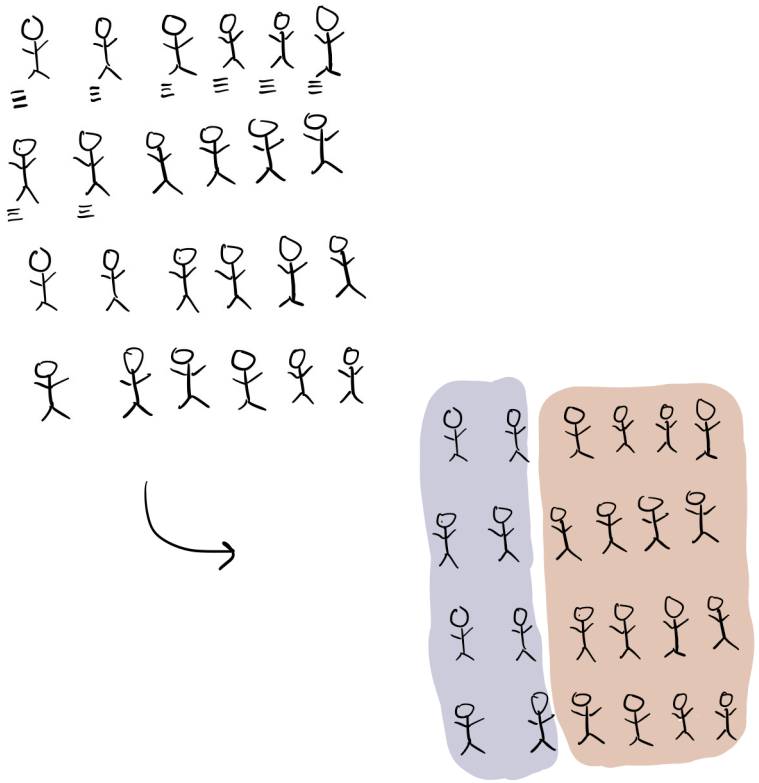
Patient phenotypes guide treatment response, study design, and subtype discovery.

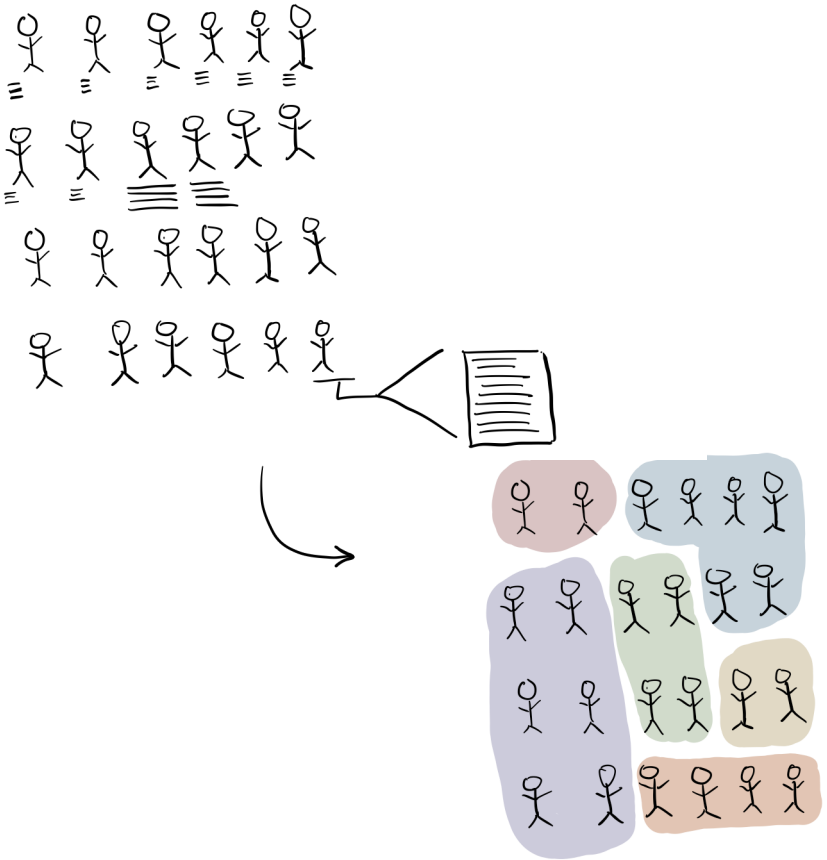


How can we ensure that discovered phenotypes are robust and reproducible?

PATIENT-LED RESEARCH COLLABORATIVE

Fewer or different symptoms can change the observed groups





Our broad patient-led symptom dataset

- **6031** adults with ≥ 90 days of ongoing illness
- **162** self-reported symptoms across 10 domains
- Input: symptom present at any surveyed timepoint
- Additional data: cognitive/physical PEM severity, age, gender



RESEARCH PAPER · Volume 38, 101019, August 2021 · *Open Access*

eClinicalMedicine
Part of THE LANCET *Discovery Science*

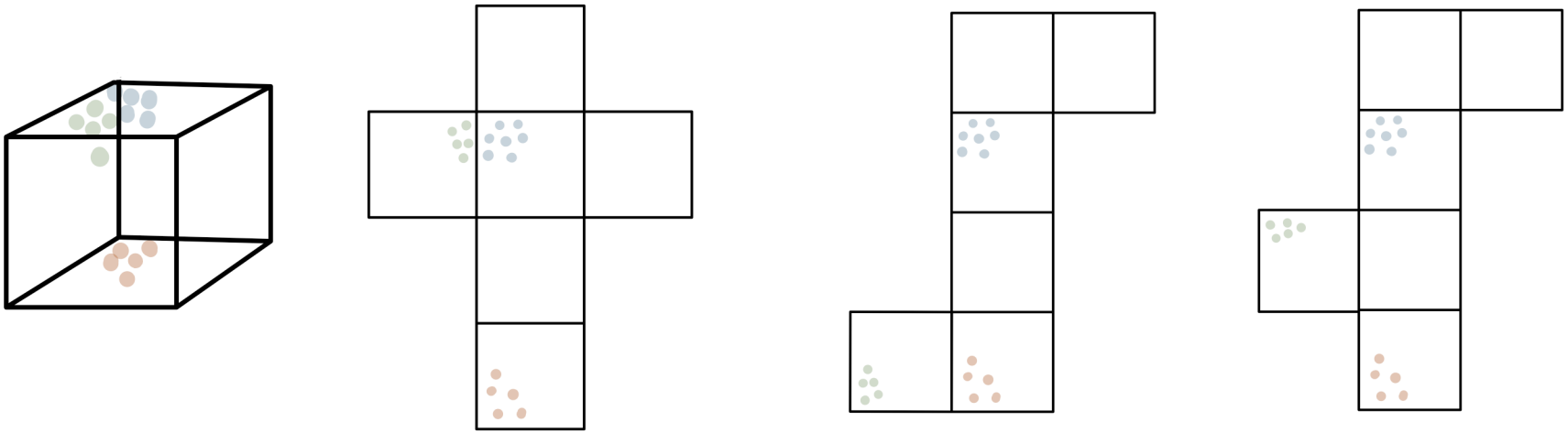
Characterizing long COVID in an international cohort: 7 months of symptoms and their impact

Hannah E. Davis [a,1] · Gina S. Assaf [a,1] · Lisa McCorkell [a,1] · Hannah Wei [a,1] · Ryan J. Low [a,b,1] · Yochai Re'em [a,c,1] · Signe Redfield [a] · Jared P. Austin [a,d] · Athena Akrami [a,b,1] Show less

The study was approved by the UCL Research Ethics Board.

PATIENT-LED RESEARCH COLLABORATIVE

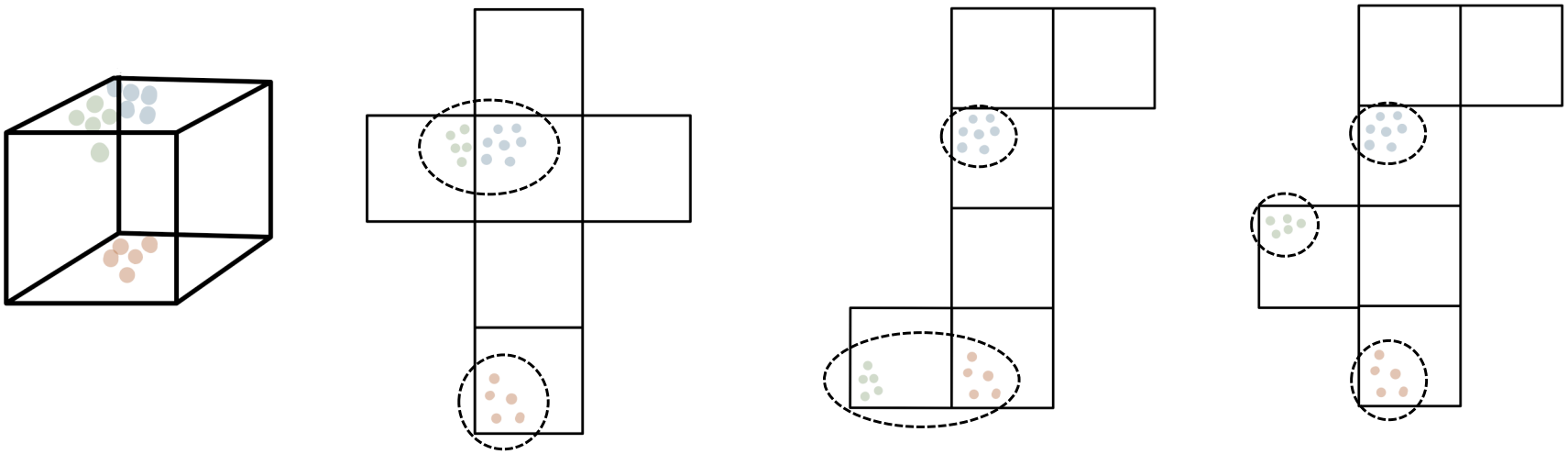
Clustering high-dimensional data is hard



The choice of how we put it into lower dimensions has a huge impact

PATIENT-LED RESEARCH COLLABORATIVE

Clustering high-dimensional data is hard



The choice of how we put it into lower dimensions has a huge impact

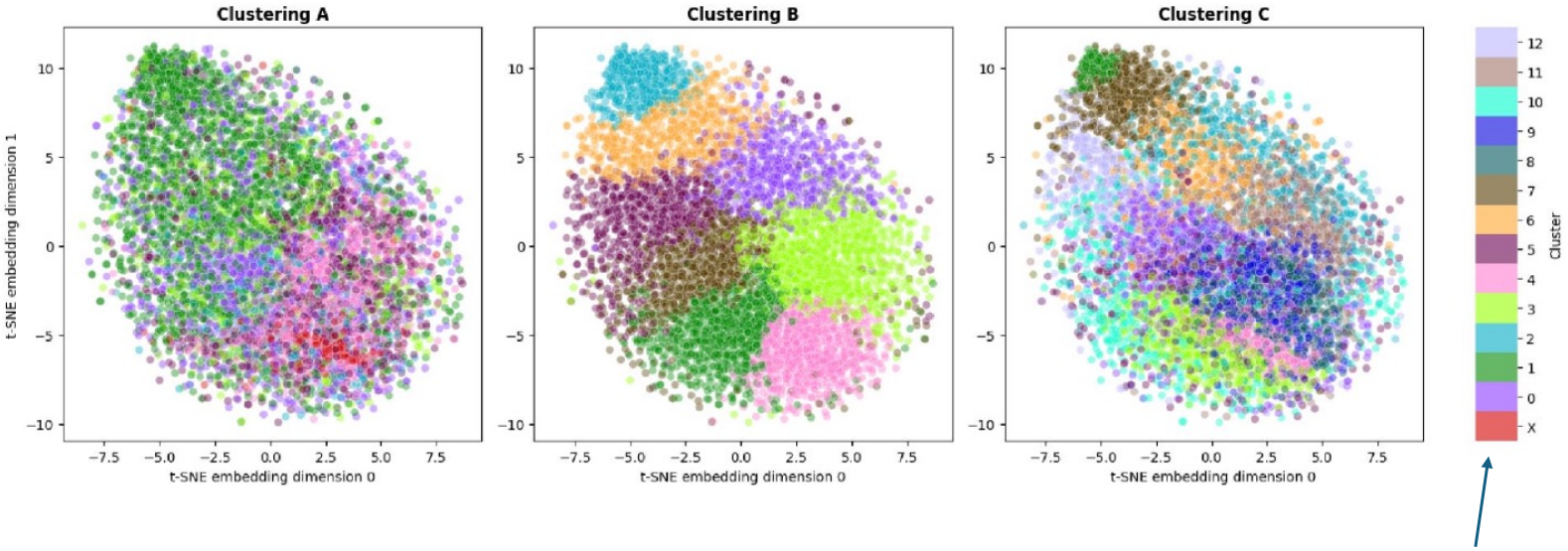
PATIENT-LED RESEARCH COLLABORATIVE

We used three approaches with the same dataset

Method	How it groups patients	Result
A Autoencoder + HDBSCAN	Learn a compressed representation, then find dense groups	6 clusters + noise
B Ensemble	Combine multiple UMAP + k-means solutions into a consensus	8 clusters
C Latent class analysis	Model symptom probabilities within latent classes	13 classes

PATIENT-LED RESEARCH COLLABORATIVE

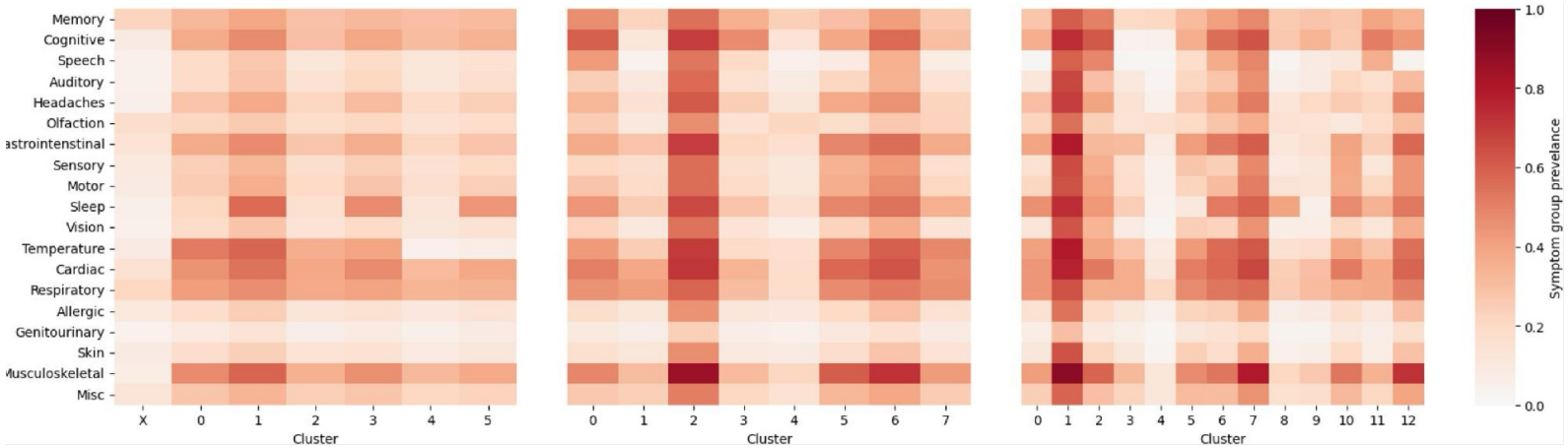
Each clustering visually distinct in 2D tSNE



0.13–0.40 Adjusted mutual information across method pairs

Cluster numbers and colors are not consistent across the clusters- this just illustrates a lot of variation

Different methods lead to different numbers of clusters with different symptom characteristics



Autoencoder and HDBSCAN:
6 clusters + noise

Ensemble:
8 clusters

LCA:
3 clusters

PATIENT-LED RESEARCH COLLABORATIVE

Description of clusters from approach A

Table 2 Summary of clustering method A.

Cluster	Name	Size	Symptom count	Women	Age	Physical PEM Severity	Cognitive PEM Severity
AX	Outliers	222	16	0.667	47.6	0.619	0.271
A0	Temperature dysregulation (chills and flushing sweats). Cardio-respiratory (shortness of breath, heart palpitations and dizziness). Cognition (short term memory and attention/ concentration).	1125	44	0.795	47.3	0.773	0.549
A1	Largest cluster. Highest proportion of women. Highest symptom burden. Severe insomnia and sleep interruption. Severe physical fatigue and PEM.	**2126**	**60**	**0.830**	46.7	**0.811**	**0.636**
A2	Elevated temperature. Generally elevated cardio-respiratory symptoms.	321	35	0.791	45.9	0.747	0.485
A3	Very generalized cluster. Symptom severity scattered over: cardio, cognition, short term memory, sleep disturbances, temperature dysregulation and headaches	498	46	0.791	46.3	0.773	0.594
A4	Lowest symptom burden of non outlier clusters. Lowest severity of cognitive PEM. Notably low sleep disturbance.	850	28	0.714	**47.7**	0.719	0.480
A5	Elevated insomnia, short term memory	889	37	0.721	47.1	0.735	0.530
Cohort		6031	42	0.78	47.0	0.778	0.637

For each numeric value the cluster with the highest value is shown in bold. Descriptive names have been assigned to each cluster based on characteristic symptoms, however more details are given in the text. Age corresponds to the midpoint average of the age class of patients within a cluster. Women lists the fraction of patients whose reported gender was woman. Cognitive and physical PEM are averaged reported cognitive and physical symptom severity during PEM respectively, calculated as the average of patients who reported experiencing the reported form of PEM. The largest value in each metadata column is highlighted in **bold**.

PATIENT-LED RESEARCH COLLABORATIVE

Description of clusters from approach B

Table 3 Similar to Table 2, but for clustering B.

Cluster	Name	Size	Symptom count	Women	Age	Physical PEM Severity	Cognitive PEM Severity
B0	Speech, memory, cognition, sleep and eye/ vision (neurological).	615	53	0.833	45.1	0.806	0.739
B1	Cognitive PEM much less severe than physical PEM. Shortness of breath, chest burning pain, short term memory, heart palpitations.	872	30	0.748	47.6	0.740	0.478
B2	Highest symptom burden. Severe physical and cognitive PEM. Highest proportion of women. Sensory sensitivity, difficulty communicating verbally and processing information, neuropathy, itchy eyes and blurred vision.	506	**89**	**0.866**	45.5	**0.876**	**0.786**
B3	Largest cluster. Cognitive dysfunction symptoms most enriched. Cognition, long term memory, speech, PEM, and fatigue.	**1105**	34	0.693	47.3	0.749	0.654
B4	Lowest symptom burden. Highest average age. Lowest PEM severity. Altered smell and taste including loss, respiratory.	887	21	0.692	**50.3**	0.676	0.479
B5	Nausea, temperature dysregulation, sleep disturbance, paresthesia and vibrations, palpitations, dizziness and vertigo.	727	54	0.834	45.0	0.794	0.611
B6	Paresthesia, speech, vision problems, neuropathy, tremors, temperature dysregulation.	611	69	0.856	45.7	0.848	0.736
B7	Temperature dysregulation, dry cough, loss of smell, taste and appetite, shortness of breath.	708	40	0.836	47.2	0.763	0.554
Cohort		6031	42	0.78	47.0	0.778	0.637

PATIENT-LED RESEARCH COLLABORATIVE

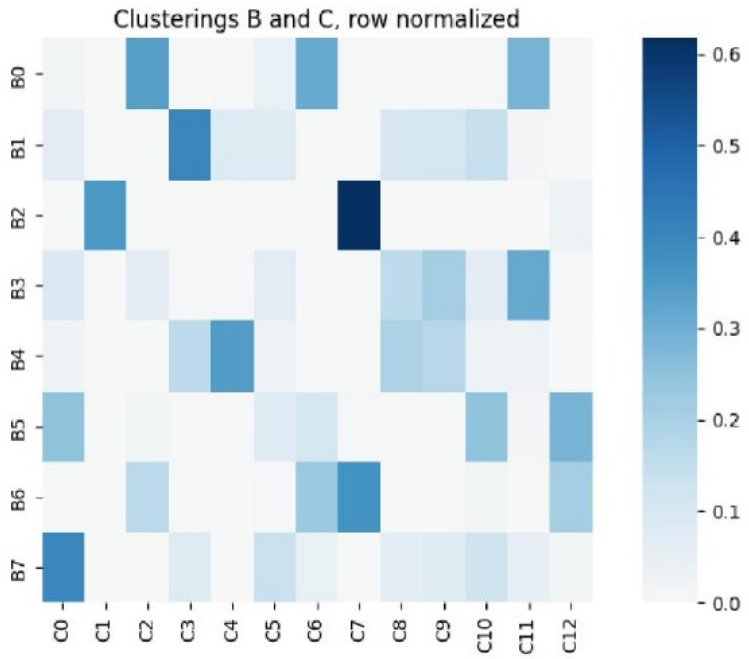
Description of clusters from approach C

Table 4 Similar to Table 2, but for clustering C.

Cluster	Name	Size	Symptom Count	Women	Age	Physical PEM Severity	Cognitive PEM Severity
C0	Temperature dysregulation, sleep disturbance, cognition (attention/concentration) and shortness of breath	**656**	40	0.794	45.2	0.752	0.595
C1	High symptom burden with cognitive dysfunction, difficulty reading and communicating, and sensory and allergic symptoms; Elevated musculoskeletal symptom count; high physical and cognitive PEM severity	178	**101**	0.848	44.0	**0.897**	**0.818**
C2	Difficulty reading and communicating, high prevalence of cognitive symptoms	399	56	0.769	45.7	0.804	0.772
C3	Low symptom burden, cognitive symptoms reduced with short term memory difficulty still reported	549	27	0.745	47.9	0.701	0.416
C4	Lowest symptom burden, Lowest cognitive PEM severity, reduced frequency of PEM, cognitive symptoms reduced	389	11	0.640	47.6	0.637	0.394
C5	Chest tightness, dizziness/vertigo, shortness of breath, gasping for air while oxygen normal	352	44	0.835	46.8	0.783	0.593
C6	Impaired reading and communication, elevated temperature and chills/sweats, sleep difficulty and elevated average musculoskeletal symptom count	449	58	**0.902**	44.2	0.821	0.703
C7	High symptom burden with impaired speech, auditory and visual comprehension, blurred vision, pain: muscle aches, muscle spasms, neuralgia and skin burning without associated rash.	544	78	0.881	45.8	0.854	0.757
C8	Insomnia, difficulty falling asleep and waking up in night, lower symptom burden	490	23	0.673	**49.4**	0.732	0.539
C9	Consistent sleepers with PEM and cognitive impairment, lower symptom burden	534	24	0.738	47.3	0.696	0.574
C10	Insomnia, tremors, sensory features: tingling/prickling, skin burning without rash, vibrations.	509	45	0.800	46.9	0.751	0.555
C11	Speech and cognition with moderate symptom burden	608	38	0.745	46.3	0.767	0.682
C12	High symptom burden with reduced cognitive/speech, elevated musculoskeletal symptom count	374	63	0.813	45.3	0.820	0.675
Cohort		6031	42	0.78	47.0	0.778	0.637

PATIENT-LED RESEARCH COLLABORATIVE

Cross-method agreement is limited but can still be informative



0.13–0.40

Adjusted mutual information across method pairs

Similar labels contain different sets of patients

Figure 3, B–C comparison (cropped), Green et al., 2026. AMI: 0 ≈ chance agreement, 1 = identical partitions.

PATIENT-LED RESEARCH COLLABORATIVE

Some clinical patterns recur across methods

High-burden group	Patients	Mean symptom count
A1	2,126	60
B2	506	89
C1	178	101

Greater symptom burden

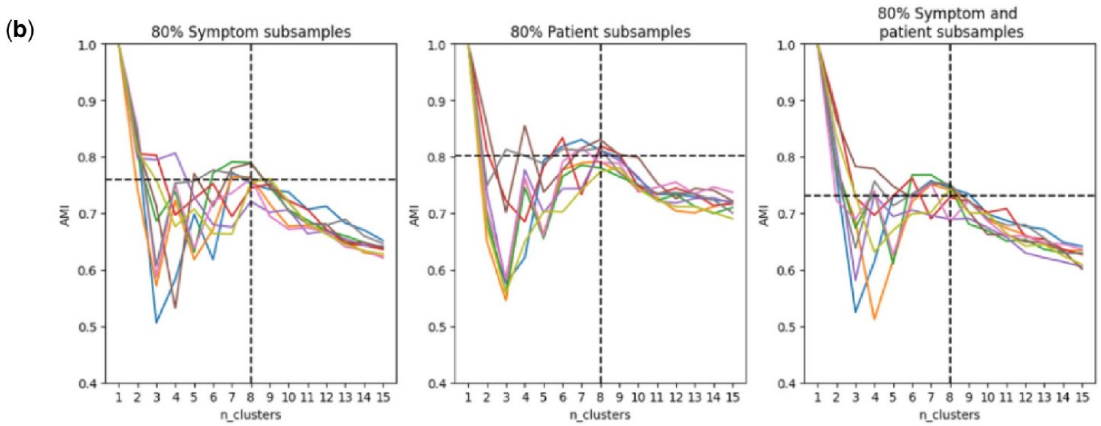
More severe reported PEM
Higher proportion of women

Speech / cognitive difficulties recur, as does PEM with little reported sleep disturbance

Tables 2–4 and Comparative analysis, Green et al., 2026. Group-level associations in this cohort.

PATIENT-LED RESEARCH COLLABORATIVE

Robustness depends on what we measure



Within Method B

Subsampling agreement: mean AMI 0.74–0.82

Within Method C

Fewer symptoms favored fewer classes

A shorter inventory can hide distinctions

Figure 2b (cropped) and reported LCA sensitivity analysis, Green et al., 2026.

PATIENT-LED RESEARCH COLLABORATIVE

Other notable findings

- People experiencing severe post-exertional malaise were a key phenotype- and sometimes there were multiple phenotypes that separated PEM experiences further (e.g., patients reporting PEM with minimal sleep disturbance, cognitive PEM, physical PEM)
- Subgroups with prominent speech and cognitive-linguistic difficulties emerged across multiple methods.
- Shorter symptom lists may miss important heterogeneity and produce smaller, more stable but potentially less-representative clusters.

PATIENT-LED RESEARCH COLLABORATIVE

Conclusion: exercise caution

- Long COVID phenotypes can be sensitive to the methodological approach and the dataset (including number of variables)
- Consider using multiple methods within the same dataset when trying to identifying Long COVID phenotypes, and focus on what phenotypes consistently emerge across all
  - For our study, this was a high-symptom-burden group that experienced post-exertional malaise

PATIENT-LED RESEARCH COLLABORATIVE

For more info, our publication

JOURNAL ARTICLE

Algorithm dependence of patient phenotypes in Long COVID: a patient-led, multi-method clustering of 6031 patients using 162 self-reported symptoms

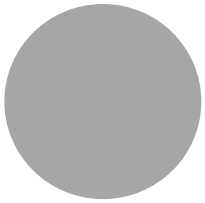
Tessa D Green ✉, Chris McWilliams, Leonardo de Figueiredo, Letícia Soares, Beth Pollack, Alison K Cohen, Tan Zhi-Xuan, Tess Falor, Hannah E Davis Author Notes

Oxford Open Immunology, Volume 7, Issue 1, 2026, iqag010, https://doi.org/10.1093/oxfimm/iqag010

Published: 13 June 2026 **Article history** ▾


Tessa D Green


Chris McWilliams


Leonardo de Figueiredo


Letícia Soares


Beth Pollack


Alison K Cohen


Tan Zhi-Xuan


Tess Falor


Hannah E Davis

Contact: alison.cohen@ucsf.edu or tessa.d.green@proton.me

This research was supported in part by Kanro (grant to PLRC) and individual supporters of PLRC.
Thanks to Tessa Green and Haya Bakdounes for helping craft the slides!