UNPACKING THE BURDEN: A MULTIVARIATE EPIDEMIOLOGICAL MODEL EXPLAINING THE SPREAD OF HIGH-IMPACT CHRONIC PAIN.

By

Christophe Tanguay-Sabourin

MSc Thesis Submission

Integrated Program in Neuroscience

McGill University

In Partial Fulfilment of the Requirements

for the Degree of Master of Science

July 2021

Acknowledgement

First, I would like to express my sincerest appreciation to my supervisor, Dr. Etienne Vachon-Presseau who was the best supervisor I could have hoped for. As a mentor both in and out of academia, he supported my goals to pursue a career as a physician-scientist from the very beginning of my research. He thought me to always strive for the highest standards of science while striving to maintain a healthy life balance along the way.

I would also like to thank Dr. Mathieu Roy who supervised my undergraduate research and recommended Dr. Vachon-Presseau to me. Dr. Mathieu Roy always generated a lot of excitement for pain research, and I'm grateful that I continued having him in my research circle during my master's.

In addition, I would like to thank all the lab members of the lab, both previous (ST & HS) and current (GG, MF, AZ, EW, JN) who's been part of creating a fantastic research environment.

I would like to acknowledge to all the McGill's and Montreal's institutions beyond the Integrated Program in Neuroscience that supported my technical training and learning the knowledge were required both in pain and neuroscience for my current thesis: Department of Psychology, Faculty of Dentistry, McConnell Brain Imaging Centre Brain, Alan Edwards Centre for Research on Pain, Healthy Brain for Healthy Lives, and Montreal Brainhack. This acknowledgment also includes the financial support from both Healthy Brain for Healthy Lives and McGill's Faculty of Medicine and Health Sciences.

Last, I would like to thank my partner, whom I had the chance to meet early on during my degree, and who've supported me all along the way.

Abstract

TITLE: Unpacking the burden: a multivariate epidemiological model explaining the spread of high-impact chronic pain

Pain is the primary reason individuals seek healthcare and the leading utilization of health care resources. Yet, the biological, psychological, and social characteristics that determines the experience of pain have rarely, or never, been studied together in a large population. Here, we derive a multivariate epidemiological model of explaining the spread of pain from 10 categories of clinical features (99 total items) in 500,000 participants from the UK Biobank. This model was characterized by poor mental health, sociodemographics, and physical health aligned with the biopsychosocial framework of pain. The model also predicted the occurrence of distal anatomical pain sites and high-impact secondary outcomes, including health ratings, opioid medication usage, and inability to work. Moreover, the model was associated with a series of biological markers, such as a polygenetic risk score, elevated immune-inflammatory markers, and pre-determined brain functions associated with persisting pain. Together, our results establish a framework that characterizes widespread pain and predicts the likelihood of chronic pain to spread across body sites.

Statement of Contribution: We found that individuals with higher risk of severe pain were at risk for greater spread of pain across the body.

Table of Contents

Introduction

Chronic Pain. In the past twenty years, various national studies have estimated that chronic pain affects 20-25% of the global population, with approximately one out of every ten people developing chronic pain every year. (Breivik, Collett, Ventafridda, Cohen, & Gallacher, 2006; Johannes, Le, Zhou, Johnston, & Dworkin, 2010; Schopflocher, Taenzer, & Jovey, 2011). The prevalence of chronic pain is even higher in middle-aged and elderly populations (Rustøen, Wahl, Hanestad, Lerdal, Paul & Miaskowski, 2005). In Canada, this high prevalence translates to an estimated $6 billion/year in medical treatments and $37 billion/year in lost productivity (Lynch, 2011). In the US, the cost of chronic pain has been estimated at $634 billion/year, which is more than heart disease, cancer, and diabetes combined (Gaskin, & Richard, 2012). The estimated price of pain has been especially apparent in patients described as having high-impact chronic pain, defined as persistent chronic pain with substantial restriction of life activities. This has been denoted, for instance, in musculoskeletal conditions, where chronic pain is rarely confined to a single body site (Carnes et al., 2007). A recent US analysis of the Medical Expenditures Panel Survey regarding chronic spinal-related pain conditions suggests that high-impact chronic pain presents a two-fold increase in the personal health-cost and a four-fold increase in the use of pain-related opioids with an average dosage five times the morphine equivalent in mg compared to low-impact chronic pain (Herman et al., 2019).

Given its high prevalence and important societal cost, chronic pain has evolved from a common comorbidity to a prominent health concern of its own and has led public health professionals to suggest that pain should be considered a global public health priority (Goldberg, & McGee, 2011). The overall disease burden associated with pain, however, remains difficult to assess, with recent studies highlighting the impact of chronic pain on an increasing number of

dimensions of life, from one's cognitive functions (McCracken & Iverson, 2001) to one's social entourage (Bannon et al., 2021; Rice, Smith & Blyth, 2016).

Chronic Pain Development. How does chronic pain differ from occasional aches and pains? Chronic pain refers to a subjective report of pain persisting past the expected healing time, generally lasting more than three months (Bonica & Hoffman, 1954, Merskey & Bogduk, 1994). Past the three-month mark, the likelihood of the pain persisting for months to years increases drastically.

Unfortunately, the underlying etiology of most chronic pain conditions remains poorly understood, as do the mechanisms behind the acute to chronic transition (Treede et al. 2019). The traditional *injury model* is not adequate to explain the occurrence of pain and its chronicity. Studies reveal a discrepancy between the state of the injury and the subjective pain intensity. For example, little to no relationship appears between radiographic osteoarthritis of the knee and reports of knee joint pain (Bedson, & Croft, 2008) or between abdominal radiograph and nontraumatic abdominal pain (Ahn et al., 2002). In one of the most prevalent forms of chronic pain, low back pain, only a small portion of patients (<5%) report a traumatic event or an injury that could have contributed to their pain development (Vlaeyen, et al. 2018). This unclear etiology is not unique to low back pain as most types of chronic pain do not present a single specific etiology, but rather a mix of them.

Chronic pain can develop and have been studied through various settings including acute traumas at the emergency department, chronic postsurgical pain following surgeries, and in its frequent occurrence in various chronic illnesses (diabetes, arthritis, heart diseases, etc.). This has motivated additional research into the risk and resilience factors for the development of chronic pain (Katz, & Seltzer, 2009). Unfortunately, few empirical studies have assessed the risk of chronic

pain development due to the high cost and low compliance in primary care settings. Systematic reviews reveal psychosocial variables provide the most prospective in explaining the transition from acute-to-chronic pain (Hasenbring, Hallner & Klasen, 2001). However, our ability to predict the development of chronic pain for an individual in a large, diversified cohort has been restricted to acceptable discrimination (i.e. 0.6-0.66 AUC) in chronic back pain, one of the most prevalent form of chronic pain (Traeger, et al. 2016; Stevans, et al. 2021).

Pain management. Pain is the leading utilization of health care resources and is considered central to our healthcare system. However, shifting focus from acute diseases toward chronic diseases has put pressure on the health care system to deliver quality and cost-effective care in face of the increasing prevalence of chronic diseases (Goldberg, & McGee, 2011). This change in medical care has been accompanied by a need for patients to shift from passive recipients of care to active partners in care (Holamn, H. & Lorig, K., 2000). A substantial part of this care has been done using opioid analgesics, postoperative pain, and acute traumas. However, do not appear to provide benefits for the management of chronic pain in the long-term (Volkow, & McLellan, 2016). Additionally, opioids are associated with psychological distress and potential risk for abuse (Carr & Goudas, 1999; Volkow, & McLellan, 2016).

Considering the poorly understood etiology and the lack of targeted treatment, multidisciplinary treatment approaches have been found to work best for chronic pain. In addition to opioid-derived medications, this includes an array of off-label medications (anticonvulsants, antidepressants, and others; Guzmán et al. 2001; Juniper, Le & Mladsi, 2009) paired with interventions (physical exercise therapy, relaxation techniques, and cognitive behavior therapy among many others; Hoffman, Papas, Chatkoff & Kerns, 2007). This approach to pain management, now considered standard in chronic pain care, aims to directly target the reported

pain symptoms. However, this approach must be done early because individuals past the acute-to-chronic transition phase (3-6 months) are at greater risk of developing psychiatric comorbidities due to their pain.

Comorbidities have also been recognized as having a central role in the accumulating risk of chronic pain. A cross-sectional study conducted in the general population of New Zealand demonstrated that a quarter of the chronic pain population reported 2 or more comorbid physical conditions, including conditions such as arthritis, anxiety/depression, diabetes, and others (Dominick, Blyth, & Nicholas, 2012). Such idea is well aligned with the framework of comorbid load where these medical conditions appeared to increase the risk of chronic pain and predispose the development of other physical or mental health conditions. Comorbid load, in turn, also predisposes or reinforces the maintenance of the original chronic pain condition and potentially, at risk for the development of further chronic pain sites.

The high prevalence of comorbidities and risk factors has also motivated further research toward a biopsychosocial model of pain, acknowledging the synergic contributions of physiological, psychological, and sociocultural factors to the development and maintenance of chronic pain (Lumley et al., 2011). Large interindividual differences exist between common determinants of health (negative affect, sleep, mood, physical activity, substance use, etc.) and the report of pain, which have translated to the development of various pain-specific questionnaires (e.g. pain-related sleep, pain-related mood, etc.) that highlight these differences. Unfortunately, these questionnaires present certain shortcomings as they cannot be applied outside the pain clinic, nor can they be compared to ratings of pain-free individuals. In addition, the high prevalence of patients with co-occurring chronic pain sites makes it difficult for these questionnaires to monitor multiple pain sites.

Moreover, work on the synthesis of this biopsychosocial model in research and clinical practice has been suboptimal (Pincus et al., 2013). This synthesis requires integration of the high-dimensional patient's biopsychosocial variables, which can be achieved using a multivariable analysis framework. This method, when done in a general population, can provide a more accurate picture of a patient's pain-associated burden.

Statement of Problem. A large body of research has highlighted the role of a biopsychosocial model for chronic pain playing a synergic role in the experience of pain. This has been done through various meta-analyses and reviews highlighting independently the various contributions of psychological, social, and biological factors contributing to pain. Yet, the extent to which these factors integrate and the relative contribution of each factor in a model determining chronic pain remains mostly unknown. Providing a biopsychosocial framework characterizing chronic pain in the general population would provide a useful basis to better understand chronic pain and quantifying individual risks of developing chronic pain.

Here, I will discuss the pain sensitization occurring in widespread pain conditions and the associated spreading of pain across multiple sites. Then, I will highlight the role of stress-induced psychopathology and the common biological pathway with pain. Finally, I will end with the social determinants of pain and their roles in predisposing chronic pain patients to present worse symptoms.

Background

Pain sensitization occurring in widespread pain conditions. Pain is defined by the *International Association for the Study of Pain* (IASP) as "an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" (Treede et al., 2019). This is in contrast with *nociception*, which refers to the neural encoding of impending or actual tissue damage from nociceptors responding to *pain* signals (i.e. noxious stimulation). The sensitization (i.e increased excitability or long-term potentiation) of these neurons from repeated nociception is typically accompanied by an increased response to painful stimuli, referred to as *hyperalgesia* (i.e in comparison to analgesia, being a decrease response), and the experience of pain from stimulus normally not painful, referred to as *allodynia*.

Both processes have been recognized as part of this sensitization, a major theme and theory of chronic persisting pain presenting the most reconciliation between animal and human research. One course that could explain this sensitization is a windup state (Ji, Nackley, Huh, Terrando & Maixner, 2018) occurring at the periphery (i.e. peripheral sensitization) that would elicit a response through the central nervous system (i.e. central sensitization). Note however, the central sensitization does not require a windup response to occurs (Wolf, 1996). This process is also known to occur among other pain classifications, including acute pain types such as nociceptive pain (e.g. visceral or somatic) and chronic pain types such as neuropathic (e.g. nerve lesion) or inflammatory pain conditions (e.g. arthritis). However, newer classifications include neuroplastic (i.e. idiopathic pain) conditions. While these conditions have proven meaningful in preclinical research to derive animal models of various chronic pain conditions, growing evidence demonstrates an overlap in pain classifications, for example neuropathic pain conditions often presenting inflammatory responses (Marchand, Perretti & McMahon, 2005).

As a result, sensitization appears to remain a major regrouping framework for all pain conditions. Peripheral and central sensitization also suggest that the wind-up state can predispose the excitation of nociceptors in proximal pain sites. A persistent (i.e. chronic) wind-up state could potentially explain how pain can spread across body sites as well as the occurrence of pain at certain sites without a clear trigger or source. The sensitization response is also accompanied by low-grade systematic inflammation and immune response typically seen in chronic widespread pain conditions (arthritis conditions, inflammatory bowel syndromes, etc.) that are known to generate fatigue, a distinctive feature of widespread pain (Louati, & Berenbaum, 2015). Immune cells and the release of cytokines and chemokines have been identified as contributing (directly or indirectly) to the pain (Marchand, Perretti & McMahon, 2005).

In human experimental research, the hypersensitivity that accompanies the sensitization process has been studied using quantitative sensory testing (QST). When repeated noxious stimuli were presented at shorter intervals, each of these stimuli elicited a greater pain response (referred as *temporal* summation of pain). Alternatively, noxious stimulation over a greater body region elicited greater pain responses (referred as *spatial* summation of pain). Both summation responses appear to be consistent with the sensitization process. Finally, QST has also led to discoveries regarding clinical pain. The conditioned pain modulation paradigm (also known as diffuse noxious inhibitory control in rodents) where noxious stimulation can be suppressed under a simultaneous other noxious stimulation (i.e. “pain inhibits pain”). This pain inhibition typically appeared reduced in various idiopathic (i.e. unknown cause) chronic pain conditions including irritable bowel syndrome, temporomandibular disorders, fibromyalgia, and tension headaches, and chronic widespread back pain (in comparison to local back pain). Small effects of similar directions were

also observed in other conditions, including chronic fatigue syndrome and atypical facial pain (Yarnitsky, 2010; Gerhardt, et al. 2017).

Central sensitization syndrome is routinely integrated in the clinical practice to explain various chronic pain conditions characterized by pain across multiple sites. These conditions include but are not limited to fibromyalgia (FM; the most stereotypical syndrome), temporomandibular disorder (TMD), irritable bowel syndrome (IBS), vulvodynia, myalgic encephalomyelitis/chronic fatigue syndrome (CFS), interstitial cystitis or painful bladder syndrome, endometriosis, chronic tension-type headaches, migraine headache, and chronic lower back pain (Maixner et al. 2016). This has also led to clinical tools including the Central Sensitization Inventory (CSI) with CNS-mediated symptoms such as fatigue, sleep, cognitive or mood impairments, traumas (i.e. stress) and other psychological problems as well as the Widespread Pain Index (WPI) for the number of pain sites reported. Both these tools are routinely used to characterized central sensitization in various chronic pain conditions (e.g. irritable bowel syndrome and temporomandibular disorders, among others; Neblett et al., 2013) as well as fibromyalgia specifically (Wolfe, Egloff & Häuser, 2016).

Central sensitization syndrome has also extended to other psychiatric or non-pain-specific conditions such as PTSD, chronic fatigue syndrome, restless leg syndrome (all of which presented heighted experimental pain responses). The similar symptoms along with the co-occurrence of these conditions potentially indicates shared processes at-play. This idea is also supported using antidepressant and antiepileptic (relaxants) drugs for pain relief, suggesting shared biological mechanisms.

The manifestation of “centrally-mediated” symptoms frequently occurring in chronic multisite or widespread pain conditions (e.g. as described by the CSI) has been increasingly

recognized in animal research. Experimental lesions inflicted on animals to model chronic pain has commonly involved testing withdrawal reflex as a primary outcome to assess mechanical sensitivity. Done over a few days, this would represent the equivalent of 3 months in a mouse's lifetime. However, it has become increasingly common now to test these animals on affective and cognitive dimensions (e.g. cognitive, anxiety, depressive behaviors in mice using various tasks, see Gregory et al., 2013 for a review), to reflect dimensions on which humans are also suffering. The lack of consideration for these dimensions during experimental lesions has been hypothesized to be an explanation for poor translation into pharmacological analgesics treatments from preclinical models.

Spreading of pain across multiple sites. An increasingly common and recent idea in the field of pain has been the presence of chronic overlapping pain conditions (COPCs). The theoretical framework of COPCs highlights the mechanistic definition of most chronic pain conditions at the expense of an anatomical perspective from these sites (Maixner et al., 2016). By displaying the high co-prevalence of various chronic pain conditions, Maixner et al. proposed that some common etiology and therefore, genetic, and biopsychosocial determinants may overlap between these chronic coinciding pain conditions. Furthermore, greater count of COPCs was associated with stronger negative moods, perceived stress, and pain catastrophizing (Fillingim, 2020). Moreover, few studies have examined number of pain sites to measure spreading. A study by Kamaleri et al., found that number of pain sites was higher in smokers, individuals with less physical activity and higher body mass index as well as reduced overall health, sleep quality and psychological health (Kamaleri et al., 2008). This approach allowed researchers to look at the widespreadness of the pain not as a binary state bounded by cut-off of number of pain sites, but as a continuum regarding its impact on health determinants. Another study of the same cohort in

Norway revealed that the number of pain sites predicted work disability, with sick leave and general health being strong intermediate variables (Kamaleri, Natvig, Ihlebaek, & Bruusgaard, D., 2009).

More importantly, even fewer studies have examined the process of spreading occurring from one original pain site spreading to various pain sites across the body. A study looking at musculoskeletal conditions has investigated the prognosis of chronic single-site pain and the risk of spreading to other musculoskeletal sites (low-back, knee or neck/shoulder pain; Andersen, Clausen, Carneiro, & Holtermann, 2012). The spread of pain across high-resolution of anatomical sites (up to 36 pre-defined anatomical areas) also reveals an association with pain intensity, pain interference and pain duration. As the authors suggests, the work highlights the need for early treatment in widespread pain conditions, as pain sites may worsen with time when left untreated (Gerdle, Fischer, Cervin, & Ringqvist, 2021). Finally, the idea of spread across sites and across different pain conditions becomes clear from the detailed anatomical pathways behind them (see Affaitati et al., 2020 for details). The work of Affaitati et al. (2020) provides a review of the co-occurrence behind various chronic pain conditions: visceral pain (ischemic heart disease, irritable bowel syndrome, dysmenorrhea/endometriosis, and urinary pain), fibromyalgia, musculoskeletal pain, and headache. While recognizing that the pathophysiological association of these pain conditions is complex and multifactorial, the authors highlight that effective treatment in one of these pain sites or conditions will most likely improve the symptoms in the others pain conditions. Alternatively, one may expect that ineffective treatment targeting one site may be partially due to the persistence or maintenance of pain at these other sites.

Overall, this process could explain the worsening and development from one site of pain to multiple sites of pain as well as the occurrence of central sensitization symptoms in chronic pain conditions typically considered to be single site (e.g. low back pain, migraine, etc.).

The role of stress-induced psychopathology in chronic pain. Overall psychopathology has been studied depression and other mood disorders but have increasingly recognized to play a role in pain, in part due to the role of chronic stress presenting an important etiological factor for many chronic pain syndromes and symptoms exacerbations (Jennings, Okine, Roche, & Finn, 2014). Pain is typically recognized as having sensory-discriminative (i.e. detection mechanisms, reactive nociception referred as pain intensity), affective-motivational (i.e. reinforcement of avoidant behaviour toward any pain or injury referred as pain unpleasantness) and evaluative-cognitive (thoughts, beliefs, appraisals and distraction) dimensions (Melzack & Casey, 1968). Chronic stress-induced hyperalgesia has been shown to interact with each of psychological dimensions of pain (i.e. sensory-discrimination, affective-motivational and evaluative-cognitive) in part due to high comorbidity with stress-related psychiatric conditions. For example, substantial evidence has linked depression to chronic pain. One study has linked predisposition to depression as a risk factor of an increased likelihood of pain chronicity in patients following pain onset (Fishbain, Cutler, Rosomoff & Rosomoff, 1997). A national survey in 2007 (n = 91 347) with a 2-year follow-up (n = 55 690) observing low back pain patients found that the severity of depression interrelated (40%) with the intensity of pain, suggesting an intimate relationship between the severity of comorbid symptoms (Meyer, Cooper & Raspe, 2007).

Some researchers have even suggested chronic stress and chronic pain to be two sides of the same coin given the substantial conceptual and physiological overlaps (Abdallah & Geha,

2017). Just like stress, pain is adaptive in its acute form and maladaptive in its chronic form. Stress and pain are typically managed and treated very similarly (rest, physical exercise, mindfulness, etc.). Stress has therefore also been theorized to have a role in the acute-to-chronic transition and potentially, the extension and generalization of pain to other body parts as presented by the sensitization process (Hannibal & Bishop, 2014). Another accompanying process is sickness response which emerges from pro-inflammatory cytokines and can lead on its own to depressive behaviours (Dantzer et al., 2008). Overall, both physiological stressors (e.g. pain) and psychological stressors (e.g. life traumas or adversity) may appear to induce a similar stress-induced psychopathology, which suggest they may share common biological pathways (Bair, Robinson, Katon, & Kroenke, 2003; Vachon-Presseau, 2018).

Central and peripheral biological processes involved in pain. Pain is known integrate information across various biological systems, from the periphery (molecular and physiological) to the central nervous system (spinal cord and brain). The final conscious experience of pain emerges at the level of the brain integrating information from ascending pathways while simultaneously acting on descending pathways to the spinal cord. Being the last step of the ascending tract, the brain appears to be an intuitive biological system to study for pain as it may be the most closely linked to self-report. Brain-based biomarkers have therefore been hypothesized to provide better clinical validity on the complex experience of pain than peripheral measures (Davis et al., 2017). One major finding from neuroimaging studies has been the dissociation of the sensory dimension (intensity; more closely linked to the somatosensory cortex) and the affective dimension (unpleasantness; more closely linked in the anterior cingulate cortex) of pain using Positron Emission Tomography (PET; Rainville et al., 1997). This affective dimension explains how pain and mood can coexist and interact. However, more recent evidence demonstrates that

this representation of pain is much more complex. Psychological (i.e. mood) and physiological pain both appear to involve the same brain regions, notably the insular cortex, prefrontal cortex, anterior cingulate, thalamus, hippocampus, and amygdala (Meerwijk, Ford, & Weiss, 2013). This is in addition to larger scale shared brain-networks, like the default mode network, which is known to be involved in chronic pain as well as other psychiatric conditions (depression, anxiety, bipolar, schizophrenia, PTSD).

On way to deal with the complex elicited response of pain in the brain has been to rely on multivariate patterns to study pain, as one brain region may not be enough to quantify the experience of pain. This approach has been used with Functional Magnetic Resonance Imaging (fMRI). When exposed to experimental pain, fMRI can reliably discriminate painful heat stimulation from non-painful heat stimulation using multivariate patterns (Wager et al., 2013). This Neurological Pain Signature (NPS) relied on wide distributed pattern of brain regions including the brain stem, insula, and somatosensory cortices, and more. However, patients presenting with chronic pain do not report different processing of pain compared to healthy controls. When controlled for pain sensitivity, meta-analyses do not report differences in the expression of pain in the brain (Xu et al., 2021). Rather, individuals suffering from chronic pain experience a greater baseline of pain activity in the brain compared to healthy individuals, as observed with sensitization response. In parallel, pain perception has also been found to be reliably discriminable independently from nociception. This discrimination was done using a signature (SIIPS; Stimulus Intensity Independent Pain Signature; Woo et al., 2017) trained on the residuals or unexplained variance of the Neurological Pain Signature. They found that this new signature (independent on the noxious intensity yet still associated with the perception of pain) appears to

follow the anatomy typically associated more closely to the affective component of pain, with an anatomical distribution resembling the one observed other mood disorders.

Finally, a third signature integrating interindividual differences in brain activity (i.e. at rest, resembling what is done in patients) combined with the introduction of an experimental model of tonic pain using capsaicin appears to provide a model of pain that is generalizable to both experimental pain in healthy participants and clinical pain in chronic pain patients in independent cohorts (Lee et al., 2021). More importantly, the brain representation obtained from the signature also appeared to resemble those observed in chronic clinical pain much more closely than the previous Neurological Pain Signature.

One area that may allow the dissociation of pain from its affective dimension is the periphery. While inflammatory response is not typically salient in most chronic pain conditions, the persistence of chronic pain is typically associated with systemic low-level inflammation. However, recent works have now emerged suggesting that mood disorders and other psychiatric conditions also appear to present an increased inflammatory response (Yuan et al., 2019). While these effects are somewhat smaller, and less consistent, they provide evidence of low-level inflammation in these conditions. The inflammatory process is believed to lead to decreased permeability of the blood-brain-barrier (BBB), leading to an increase in the accessibility of the CNS for immune cells (Lee & Giuliani, 2019). As expected, an elevated cortisol response has been shown in both chronic pain and affective disorders (Vachon-Presseau et al., 2013, Burk et al., 2003). It remains unclear how pain and intense traumas may have evolved to induce an immune-inflammatory response. Finally, interindividual differences in predisposing genetic risk also converge to some degree in chronic pain and mood disorders. As mentioned by Diatchenko and colleagues, "[…] Pain perception is one of the most complicated measurable traits because it is an

aggregate of several phenotypes associated with peripheral and central nervous system dynamics, stress responsiveness and inflammatory state. As a complex trait, it is expected to have a polygenic nature shaped by environmental pressures" (Diatchenko et al., 2007). One known genetic factor is catechol-o-methyltransferase (COMT) gene polymorphism involved in the breakdown stress-related neurotransmitters and hormones, which have been associated with the more severe experience of pain (i.e. pain catastrophizing) but also with schizophrenia, bipolar, and depressive disorders (Diatchenko et al., 2006; Craddock, Owen, & O'Donavan, 2006). Another genetic risk is mitochondrial genetic polymorphisms (i.e. involved in energy metabolism) that have been found to be associated with increased risk for fibromyalgia as well as depressive disorders (van Tilburg et al., 2020; Allen et al., 2018). It becomes clear that across various biological levels both affective mood disorders and chronic pain present convergent representation in their biological systems.

Social determinants of pain. Recent work has also extended the idea of allostatic load past traumas and symptoms to the environmental conditions of the individuals facing these events. These societally induced stressors may result in environmental challenges which surpass the individual's coping ability, resulting in allostatic overload. While research has increasingly started to combine biological and psychological measures to better understand chronic pain, social factors remain commonly overlooked. Researchers may control for social factors, but do not typically examine them, despite the various health outcomes associated with socioeconomic status (SES) and demographics/occupational status.

The World Health Organization refers to SES as the social gradient of health, running from the top to the bottom of the socioeconomic spectrum. Denoted by a graded association between health and all levels of SES, this measure includes factors such as income, education, and profession. Additionally, demographic factors such as one's gender, age, and ethnicity also play a

role in the socioeconomic gradient of health as they contribute to one's overall status and position in society.

Interestingly, this gradient is found across many chronic diseases including heart failure, arthritis, type 2 diabetes, ulcers, and certain cancers, all of which can present with debilitating pain diagnoses such as chest pain, osteoarthritis, diabetic neuropathy, burning stomach, and cancer pain. In parallel, extensive literature describes low SES as a strong potential risk factor for chronic pain. Population-wide studies reveal large socioeconomic disparities in pain across gender, education, and income. Pain levels are lower for men than women and decrease across increasing education and income quartiles (Grol-Prokopczyk, 2017). A study examining the experience of pain showed that lower SES was associated with a higher prevalence of severe pain, including a greater number of painful body sites, greater pain intensity, and/or greater feelings of being disabled by pain. Even at the same intensity of pain and the same number of painful body sites, the study reported that participants at the lowest as compared to the highest SES were two to three times more likely to feel disabled by their pain (Dormer et al., 2011). This is in addition to the greater risk of developing multiple psychiatric comorbidities: generalized stress disorder, panic disorder, depression, and post-traumatic stress disorder along with their pain commonly observed in lower SES populations (McWilliams, Cox, & Enns, 2003; Miech, et al., 1999).

Demographic variables also contribute to the effects mentioned above within socioeconomic gradient, with pain typically being more pronounced in women (Grol-Prokopczyk, 2017). Pain-related disparities between men and women are both biological (sex) and socio-cultural (gender) in nature. Strong biological evidence indicates that males experience stronger opioid analgesia (Fillingim & Ness, 2000). Yet, these differences are also the recipient of social influences. For instance, questionnaires assessing gender roles in pain expectation reveal that even

after controlling for sex, self-identified gender still predicts pain tolerance (Wise et al, 2002). It remains unclear, however, the extent to which we can dissociate the contribution of SES from gender, as women typically present lower SES compared to men. A very similar case can be made with ethnicity, where pain is found more prevalent and disabling among non-Caucasian ethnicities (Grol-Prokopczyk, 2017). However, a study suggests that this race effect appears to be strongly mediated by socioeconomic disparities. While non-Caucasian populations are typically the first targets of this socioeconomic gradient, presenting lower income, education, and occupational role, non-Caucasian individuals living in neighborhoods of greater SES do not seem to present the same pain severity as those of lower SES (Fuentes, Hart-Johnson & Green, 2007).

Toward a theory of chronic pain. The Lancet article *Rethinking Chronic Pain* highlights the need to understand chronic pain as a product of abnormal neural signalling, with biopsychological dimensions requiring a multimodal treatment approach (Kaulitzki, 2017). Here, we hope to integrate two dominant theories of the field of pain to provide a framework to the understanding of pain: i) the integration and synthesis of the biopsychosocial model to explain one's chronic pain and environmental context, and ii) the framework of pain-induced sensitization across various chronic pain conditions along with the assessment of both peripheral and biological determinants. Together, these provide a powerful framework to understand pain and its chronicity across most conditions and diseases.

Rationale

Here, we aim to empirically derive and validate an epidemiological model predicting how high impact chronic pain is spreading over body sites in a large population. This provides the opportunity to explore the high-dimensional representation typically referred in its biopsychosocial model. The overarching goals of this project are to better understand the

determinants of widespread pain and better identify the most important features that may targeted to improve pain management.

Aim #1. **Examining the widespread pain phenotype in the UK Biobank.** We hypothesize that the spreading of chronic pain is a core variable that can be used to phenotype pain patients. We also hypothesize that various chronic pain conditions obtained from self-reported diagnoses will present some degree of spreading of the pain with high prevalence of chronic multi-site pain compared to chronic single-site pain.

Aim #2. Deriving a multivariate epidemiological model explaining the spreading of pain in a large population of nearly 450 000 individuals. We hypothesize that the epidemiological model will allows to capture the high-dimensional burden associated with pain, presenting both vulnerabilities and consequences of living with pain.

Aim #3. Evaluating the performance of the epidemiological model to discriminate the spreading of pain and high-impact secondary outcomes. We hypothesize that our multivariate model of pain-associated burden will explain medium-to-strong variance of an individual's number of self-reported pain sites. We secondly hypothesize that our multivariate model will generalize to a left-out testing set of nearly 50 000 individuals.

Aim #4. Assessing the longitudinal risk for the spreading of pain and high-impact secondary outcomes at 10-year follow-up. We hypothesize that our multivariate model will predict within subject chronification or persistence of chronification. Similarly, we hypothesize this approach will allow to model both the risk for development and persistence of high-impact secondary outcomes (i.e. opioid medication usage and inability to work).

Aim #5. Peripheral and central biological determinants associated with the epidemiological model. We aim to assess the relevant biology and physiology of such model. We

hypothesize that our multivariate model will relate with a computed polygenetic risk score, C-reactive protein, and the a-priori selected tonic pain brain signature.

Method

Aim #1. Examining the widespread pain phenotype in the UK Biobank.

UK Biobank population. The UK Biobank is a prospective epidemiological cohort study comprised of approximately 500,000 middle-aged and older individuals across Great Britain (40-69 years old at recruitment; https://ukbiobank.ac.uk/). Participants provided electronically signed consent, answered questions on socio-demographic, lifestyle, and health-related factors, and completed a range of physical measures (Bycroft et al., 2018). A subset of 50,000 participants underwent an additional follow-up visit at 6-10 years later which included whole-body MRI imaging. While the UK Biobank is known to present a low response rate, the dataset has been shown to present a similar risk factor association compared to other pooled studies across the UK, confirming its generalizability (18 studies, approximately 90 000 individuals with high response rate; Batty, et al. 2020). The present analyses were conducted under the UK Biobank application 20802. All participants provided written, informed consent, and the study was approved by the Research Ethics Committee (REC number 11/NW/0382). Further information on the consent procedure can be found elsewhere (http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=200).

Pain status at baseline visit. Participants were asked if they experienced pain interfering with their usual activity at any of the following major anatomical sites in the past 3 months (head, face, neck or shoulder, back, stomach or abdominal, hip, knee, or pain all over the body; data field 6159). Note that when *pain all over the body* specifically was selected, no other pain sites could be selected. When answering positively to a given site, participants were asked if this pain at this site had lasted longer than 3 months.

This item was used to assess the presence of chronic pain in line with the 3-month classification from the International Association for the Study of Pain (IASP; Treede et al. 2019)

or acute pain, from a negative answer on chronic pain status. Participants with missing data on any of the chronic pain sites were excluded from the data (<2.5% of the population). Finally, the sum of the pain sites experienced (excluding pain all over the body) was used to assess the spreading of pain (i.e. widespread pain) as a continuum. While fine grained representation of specific sites of pain may not have been reported at the initial baseline visit, the proposed body sites in the UK Biobank remained the most prevalent and stereotypical sites in chronic pain.

Online follow-up assessment. Online questionnaires were administered about 8-13 years after the baseline visit to better phenotype chronic pain patients. A total of 167,000 participants filled the *experience of pain* (10 sections; https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/pain_questionnaire.pdf) administered online by the UK Biobank. About 80,000 of these individuals answered affirmatively to a first question regarding the presence of pain or discomfort that has lasted for more than 3 months (i.e. chronic). From this online questionnaire, various pain-related questionnaires were used to validate the importance of increased widespread pain using the same major anatomical sites as those assessed in person. Duration of the pain or discomfort was assessed across three categories: 3-12 months, 1-5 years, 5+ years. Pain ratings were measured as the worst pain experienced in the past 24 hours, to allow comparable measurements between single-site and multi-site pain groups. Pain interference was measured using items adapted from the Brief Pain Inventory (BPI; Tan, et al., 2004) including scales of how much the most bothersome pain has interfered with life across seven dimensions (general activity, enjoyment of life, normal work, sleep, walking ability, mood, and relationships) from 0, beings “does not interfere”, to 10, “completely interferes”. The severity of depressive symptoms was measured using the Patient Health Questionnaire (PHQ-9; Löwe et al.,

2004). Widespread pain symptoms were examined across 3 scales (0-3): waking unrefreshed, fatigue, and cognitive symptoms.

Spreading of pain over multiple sites. Spreading is estimated by looking at the risk of occurrence across pain sites (excluding patients entering *pain all over the body* as their response; these patients will be used as an additional test set, see below). This was measured using odd ratios from the exponential function of a logistic regression coefficient estimated for each combination of sites. The number of sites setting apart each combination was used a measure of their distance (1-7).

Using online follow-up (subsample n = 80 000), the same analysis was performed entering ratings of pain intensity. For each site presenting pain in the past three months, ratings in the past 24 hours (0 being "no pain" and 10 being "pain is as bad as it could be") were assessed. A correlation between pain ratings across sites was measured. To ensure the robustness of the association between co-occurrence and distance, our results were compared to a null model generated from 10,000 permutations tests.

Non-cancer illnesses and medication. Information regarding self-reported non-cancer illness was assessed (cancer or non-cancer illness; data field 20002/9) at each visit. If the participant was uncertain of the type of illness they had, they would describe it to the interviewer (a trained nurse) who would attempt to place it within the coding tree. Self-reported medication and health supplements were recorded at each time visit (data field 20003). The duration of use and dosage of the medications were not collected in the biobank. Medication and health supplements were coded according to Wu et al. (2019) classification. From the 6,745 categories of medications recorded, only 1,809 medications were taken by at least 10 participants in the initial visit (> 0.002 % of reported use) and were manually mapped to their corresponding active

ingredient using online information according to the Anatomical Therapeutical Chemical (ATC) Classification System from the World Health Organization (WHO; for more information, see Wu et al. 2019).

Aim #2. Deriving a multivariate epidemiological model explaining the spreading of pain in a large population of nearly 450,000 individuals.

Clinical features. A total of 99 variables collected at baseline visit were carefully selected based on their relevance to chronic pain. The selection was based on the Prognosis Research Strategy (PROGRESS) group who recently provided a framework for the development of a prognostic model to determine risk profile (Hemingway et al. 2013). Variables were organized along a hierarchical framework made of three distinct dimensions (mental health, physical health, and sociodemographic) grouping ten categories. The mental health dimension was constructed from neuroticism (all individuals items and their sum-score, based on neurotic behaviors closely linked to negative affect), traumas (illness, injury, bereavement, or stress in the last 2 years), and mood (reported frequency from past 2 weeks of certain moods and visits to a general practitioner or psychiatrist for nerves, anxiety, tension, or depression). Physical health was constructed based on physical activity (reported from Metabolic Equivalent Task scores based on the International Physical Activity Questionnaire guidelines; Cassidy et al., 2016), based on sleep (questions regarding duration, napping, snoring, sleeplessness), on substance use (smoking and alcohol use), and on physiological measures (including anthropometric measures, fractures that occurred over the last 5 years, blood pressures, etc.). The sociodemographic dimension was constructed from socioeconomic status (education completed, income, employment, etc.), occupational measures

(individuals present within household, social entourage, and manual or physical job), and demographic information (age, sex, and ethnicity).

Training and testing participants. Among the 99 features selected, participants with more than 20% missing features were excluded, as were individuals who were missing data regarding their pain status. No other exclusion criteria were applied to ensure the findings of the study to be as generalizable as possible to the greater population. This led to a training cohort of participants – those who did not attend a follow-up visit – of 445,132 participants from whom the multivariate model was derived. The testing dataset of those who did attend a follow-up visit included 48,079 participants.

Imputation and normalization of the features. To minimize potential bias from incomplete questionnaires, a data-driven Bayesian ridge regression model was applied for imputation of missing data as a function of all other features in the model using the median as prior. Features were then standardized across the participants by centering mean to zero and scaling the variance to one unit. The same process (exclusion followed by imputation for missing data and standardization with the same mean and variance) was applied separately for the testing dataset.

Deriving a multivariate epidemiological model. Nonlinear Iterative Partial Least Square (NIPALS) regression algorithm (implemented using scikit-learn.org/) was used to derive an epidemiological model that explained the number of pain sites reported at the baseline visit. This was done excluding individuals presenting *pain all over the body* to avoid making assumptions regarding the equivalence of *pain all over the body* and the number of pain sites experienced. This specific algorithm was chosen to reduce the 99 features into a few sets of distinct components associated with self-reported number of pain sites. A common rule-of-thumb in multiple regression suggests that the minimum ratio of sample size per variable is 10:1 with greater ratios equivalent

to greater stability. Here, we observe a 450-fold increase from the recommended ratio, providing us confidence in our stability. Ten-fold cross-validation was used to assess the number of components to use in the model. A total of three components were selected due to the largest increase in the variance explained and the largest decrease in negative mean squared error. Finally, the model was applied in the testing dataset.

Assessment of model categories using a network analysis. Dimensions and categories were estimated using the weighted sum of the clinical features. Categories were integrated using a network analysis approach relying on a thresholded partial correlation network. Partial correlation was used to measure the association between categories while controlling for all other potential edges. A threshold was applied to conserve connections equivalent to a small effect size (partial correlation > 0.1). Nodes were placed using a force-displacement layout (i.e. spring layout, using Fruchterman-Reingold algorithm). Starting in a circular layout and trough iterations, nodes more connected will be placed closer together while those less connected or negatively linked are placed further apart.

Aim #3. Evaluating the performance of the epidemiological model to discriminate the spreading of pain and high-impact secondary outcomes.

Assessment of the model performance at baseline. In both the training and testing datasets, the performance was assessed using number of pain sites and the secondary pain-related outcomes described below. Pearson correlation coefficients were used to measure the association between the epidemiological score obtained from the model and the number of pain sites across chronic and acute sites. To assess the impact of both acute and chronic pain across sites, the score was compared to pain-free participants for each site. Differences in the model score were measured

using both Cohen's d for effect size (using pooled standard deviation due to unequal sample size) and the Area Under the Receiver Operating Characteristic curve (AUC-ROC) for discrimination. The AUC-ROCs were used for estimating the model accuracy because they are i) threshold-unspecific and ii) resilient to class imbalance (inherent to less frequent pain conditions or clinical outcomes). The secondary outcomes examined included: i) overall-health rating (data field 2178), opioid medication use (NO2A from the ATC classification), and inability to work due to sickness or disability (data field 6142).

Cross-sectional analysis of the spreading at baseline visit. The associated risk of spreading was estimated using the individual output of our multivariate model. For each pain site, the risk of occurrence with all other pain sites (i.e. spreading) was measured as a function of this epidemiological risk score. This was done using odds ratios associated with an increase of one unit in the risk score. As mentioned in the co-occurrence of pain sites, statistical significance (p-value) was estimated using 10,000 permutation tests. The dependence on the spatial relationships of the site was assessed as mentioned in *Spreading of pain over multiple sites*, using the distance across the anatomical sites.

Aim #4. Assessing the longitudinal risk for the spreading of pain and high-impact secondary outcomes at 10-year follow-up.

Prognosis of chronic pain and the risk for spreading. While the model was derived cross-sectionally, the model provides a prognostic value for the risk of development, persistence, or worsening of chronic pain in the testing cohort at their second visit at a 10-year follow-up. Here, measured the risk induced by chronic pain for each anatomical site at the baseline visit for the persistence of that same chronic pain site at follow-up using odds ratios. We also measured the

risk induced by a chronic pain site at baseline for the development or the persistence of chronic pain for each of the seven anatomical pain sites at follow-up using odds ratios. Similarly, we used the risk estimated from the multivariate epidemiological model at each chronic pain site to estimate the odds ratio for that same anatomical pain site to persist at follow-up (i.e. odds associated with one unit higher in the model). The same estimated risk at each chronic pain site was used to estimate the odds ratio for the chronic pain to develop or persist for each of the seven anatomical pain sites. Finally, the pattern of spreading was evaluated the same way as in the *Spreading of pain over multiple sites*: using the spatial distance between pain sites and their odds ratio measured.

Risk for Chronic Pain and High-impact pain. Pain status and high-impact secondary outcomes (i.e. opioid medication usage and the inability to work due to sickness or disability) were assessed longitudinally. We examined the risk estimated from our models for participants developing chronic pain at different levels of spreading (i.e. number of sites). This procedure was repeated in 3 groups: chronic pain-free participants, chronic single-site pain participants, and chronic multi-site pain participants (2 sites or more, including *pain all over the body*). Moreover, Cohen's d effect sizes were used to measure the difference in the estimated risk at baseline for opioid medication and inability to work at follow-up. The difference in risk associated with the continuation of opioid medication or the continuation of being unable to work was also measured. Finally, the corresponding discrimination measures were also evaluated using Area Under the Receiver Operating Characteristic curve (AUC-ROC).

Aim #5. Peripheral and central biological determinants associated with the epidemiological model.

Genetic risk. We used a pre-defined polygenic risk score of chronic multisite pain (Khoury et al. 2021) done in the UK Biobank. The score was derived excluding participants used in this current study. We next tested the association between the polygenetic risk score and the burden score using left-out 19,000 individuals from the test set. This method avoid circularity as the polygenic risk score and the burden score were derived from the training set and their associations were only tested in an independent group of individuals.

Immune-inflammatory profile. The UK Biobank haematological data included a complete blood count (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/haematology.pdf). The sample handling and storage has been published (Eliott & Peakman, 2008). Inflammation was estimated using C-Reactive Protein (CRP; data field 30710) through saliva sample and measured by immunoturbidimetric using a high sensitivity analysis on a Beckman Coulter Analyzer. Given the positive skewness of CRP (see https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=30710), logarithmic transformation was applied. Immune cell count included neutrophils, platelets, reticulocytes, basophils, lymphocytes, eosinophiles, monocytes; most of which have been shown to be independently linked to chronic pain and the sickness response and associated depressive profile (Marchand, Perretti, & McMahon, B., 2005; Dantzer et al., 2008).

Brain phenotypes: grey-matter density. Our study built on the minimally preprocessed pipeline designed and carried out by FMRIB, Oxford University, UK (Millet et al., 2019). From the follow-up visits, our study includes approximately 19 000 participants who underwent resting-state functional MRI (fMRI) and structural gray-matter volume (T1w). Preprocessing of T1-weighed brain images included gradient distortion correction, brain extract, followed by

superimposed linear and non-linear registration to the MNI152 standard space (tools part of the FMRIB Software Library). All transformation were combined and applied in a single step to avoid unnecessary interpolation (Miller et al. 2016). We introduce a Morphological Grey-Matter Pattern derived on grey-matter density derived by the UK Biobank Initiative using FMRIB's Automated Segmentation Tool (FAST). This segmentation was done from a Grey-Matter atlas of 139 regions (https://git.fmrib.ox.ac.uk/falmagro/UK_biobank_pipeline_v_1/-/tree/master/templates/GMatlas). Partial least square was used to extract the dominant mode of covariance between grey-matter density and the number of pain sites in a left-out group of participants (approx. n = 21,000), referred as our Morphological Grey-Matter Pattern. In an independent group of 19,000 in which all biological measures were measured, we validated the pattern and its association with number of pain sites as well as with our epidemiological model.

Brain phenotypes: dynamic functional connectivity. Then, the minimally preprocessed resting-state fMRI data from the UK Biobank were analyzed using the following preprocessing steps: motion correction with MCFLIRT (Jenkinson et al., 2002), grand-mean intensity normalization, high pass temporal filter, fieldmap unwarping, and gradient distortion correction. Noise terms were identified and removed using FSL ICA-FIX. Full information on the UK Biobank preprocessing is published (Miller et al. 2016). Additional preprocessing was done using Nipype including 3DDespike (AFNI), registration to the 2mm MNI template (FSL), 6 mm kernel smoothing (Nilearn), resampling to 3-mm (i.e. for storage purposes). A modified Brainnetome atlas (Fan et al. 2016) was used to parcel the brain in 279 distinct regions to apply the weights from the Tonic Pain Signature (ToPS; Lee et al. 2021), a capsaicin-induced tonic pain signature of pain derived from the brain that was associated with both experimental and clinical pain. The modified version includes additional midbrain brainstem and cerebellar regions. Dynamic

connectivity was estimated like the ToPS, using Dynamic Conditional Correlation (DCC). DCC is based on generalized autoregressive condition heteroscedastic (GARCH) and exponential weighted moving average (EWMA) models (https://cocoanlab.github.io/tops/). From this, the ToPS signature response for each participant was estimated. The preprocessing aimed to be as similar as possible to the original study without diverging from the minimally preprocessing data from the UK Biobank. The weights of the signature were thresholded to the top 5% to avoid overfitting and to minimize relation with head motion. Multiple thresholds (1, 5, 10, 15, 20%) were also tested to ensure generalizability. In our study, we assess both the ability of the ToPS to explain the number of pain sites but also our epidemiological model associated with the spreading of pain.

Brain Phenotypes: confound-removal procedure. Covariates included for both resting-state and grey-matter included age, sex, imaging site, position in the scanner, and coil position (X, Y, Z respectively) from our brain features. Head motion was also regressed out from the resting-state brain features. Both covariates and brain features were normalized to zero mean and one unit of variance across participants. A confound-removal procedure was applied by deriving a multivariate regression model to predict each normalized brain feature as a function of the normalized confounds. This procedure was done for each of the brain features, making them orthogonal to confounds.

Statistical tests for significance. To ensure the robustness of small effect sizes between our epidemiological and biological variables, each biological variable assessed in the testing cohort was compared to 10,000 randomly generated null model from which statistical significance test was derived.

Results

Aim #1. Examining the widespread pain phenotype in the UK Biobank.

Pain phenotype was first assessed using anatomical sites reported by the participant in the UK Biobank (**Fig. 1A**). Co-occurrence between pain sites in the training dataset (n = 445,132) was estimated using odds ratios between each combination of anatomical pain sites (**Fig. 1B**) for both acute (≤ 3 months) and chronic (> 3 months) sites. Individuals reporting chronic pain at a given anatomical site were twice to 11 times more likely to report co-occurring chronic pain at another pain site while individuals reporting acute pain were 1.6 to 5.4 times more likely to report co-occurring pain at another pain site. Further examination of this co-occurrence reveals that both acute ($R^2 = 0.31$, $p_{perm} = 0.002$) and chronic ($R^2 = 0.40$, $p_{perm} < 0.0001$) pain sites were associated with co-occurring pain at proximal sites (**Fig. 1C**). Together, these findings suggest that co-occurrence of pain over multiple sites is not random. Instead, chronic pain tends to spread to neighbouring sites and in rare cases, reaches distal sites. As a follow-up analysis, we interrogated the online follow-up data (n = 84,030) by computing the correlation between chronic pain intensity ratings (past 24 hours) across the impact sites (**Fig. 1D**). Pain intensity ratings were correlated between r = 0.24 to r = 0.64, indicating medium to large effect sizes. Importantly, the ratings were also higher and more interrelated for proximal sites ($R^2 = 0.41$, $p_{perm} = 0.0003$; **Fig. 1C**), revealing a concordant pattern of pain spreading with the one observed using co-occurrence of pain site at baseline (**Fig. 1A-C**). Further analyses using odds ratio were used to explore the risk of co-occurrence of acute sites along with chronic pain sites and vice-versa. Neither revealed a relationship with the distance between sites.

Assessment of self-reported doctor-diagnosed non-cancer illnesses was then used to demonstrate that pain co-occurring on multiple sites (multisite pain) presents higher prevalence

many common chronic pain conditions, including fibromyalgia, various types of arthritis, spinal cord injury, rheumatoid conditions, neuropathic condition or nerve injury, and others (**Fig. 1E**). In all these conditions, chronic multisite pain prevalence was higher than chronic single-site pain prevalence. These findings suggest that the pain spread over co-occurring sites is a valid phenotype for chronic pain, which reflects both patterns in pain intensity ratings (**Fig. 1D**) and the clinical diagnoses of pain related pathologies (**Fig. 1F**).

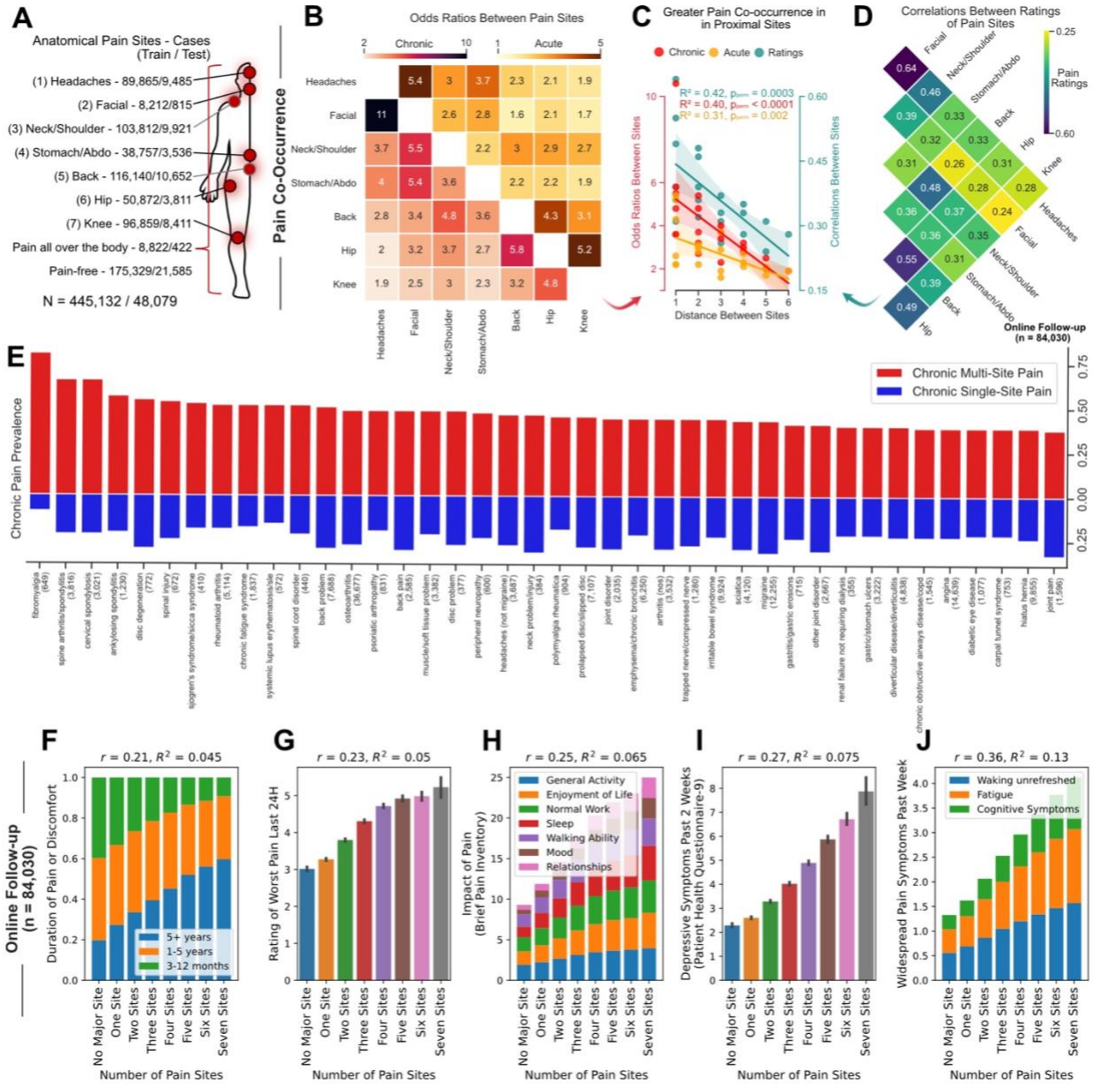


Figure 1. Examining the widespread pain phenotype in the UK Biobank. A. Anatomical body map of pain sites assessed and prevalence of pain for the train and test sets at baseline. **B-D. Cross-sectional analysis of pain-associated spreading. B.** Odd ratios of co-occurrence between pain sites (chronic on lower-diagonal and acute on upper-diagonal) in training dataset at baseline. **C.** The odds ratios of co-occurring pain between two sites were negatively associated with their distances. Significance was determined using 10,000 permutation tests. D. The negative association was validated using correlations of pain intensity ratings from 80,000 participants experiencing chronic pain during an online follow-up. **E**. Top 50 non-cancer illnesses with the highest prevalence of chronic multisite pain (2 sites or more, including pain all over the body; in red) and the prevalence of chronic single-site pain (in blue). **F-J. Validating the spreading of pain phenotype from the online assessment. F.** Duration of pain or discomfort across 3 groups. **G**. Rating of pain on 10 in the last 24 hours. **H.** Interference of pain across 7 dimensions – each on 10 (10 = Pain as bad as you can imagine). **I.** Clinical major depressive questionnaire (PHQ-9; 5-9 ratings associated with mild depression) severity. **J.** Widespread pain major symptoms, 0-3 for each (3 = Pervasive, continuous, life disturbing problems).

Additional assessments through the online questionnaire further confirmed that the spreading of pain across sites is a meaningful phenotype to quantify the impact of pain (**Fig. 1F-J**). Online follow-up among 80 000 patients reporting chronic pain in the 3 months prior to this follow-up suggests that an increase in the number of co-occurring pain sites is associated with greater duration of pain or discomfort of the pain (**Fig. 1F**, $r = 0.21$, $R^2 = 0.045$; overall prevalence for 3-12 months, 1-5 years and 5+ years: across 25.9/38.6/34.9%) and higher ratings of the worst pain over the past 24 hours (**Fig. 1G**, $r = 0.23$, $R^2 = 0.05$; overall mean(std) = 3.9(2.7)). Increasing number of self-reported pain sites were also associated with greater pain interference all-around (**Fig. 1H**, $r = 0.25$, $R^2 = 0.065$; overall mean(std) = 16.3(15.7)) measured across the seven dimensions of the Brief Pain Inventory: general activity, enjoyment of life, normal work, sleep, walking ability, mood, and relationships. Number of pain sites was also associated with stronger depressive symptoms measured using the Patient Health Questionnaire-9 (**Fig. 1I**, $r = 0.27$, $R^2 = 0.075$; overall mean(std) = 3.6(4.1)) as well as poor sleep, fatigue, and cognitive impairment (**Fig. 1J**, $r = 0.36$, $R^2 = 0.13$; overall mean(std) = 2.2(1.7)). Overall, the spreading of pain alone was

associated with a variety of major clinical outcomes regarding chronic pain and its severity. We concluded that the spreading of pain over co-occurring pain sites is an important phenotype of chronic pain patients that has been too often ignored in the previous literature. In this proposal, the number of co-occurring pain sites (spreading of pain) will therefore be used as our primary pain outcome in the following aims.

Aim #2. Deriving a multivariate epidemiological model explaining the spreading of pain in a large population of nearly 450 000 individuals.

Our second aim was to derive a multivariate model that explains and predict chronic pain and it's spread over multiple co-occurring anatomical sites. This was done using Partial Least Square Regression to derive a pain-associated burden index that explains maximal variance about the number of self-reported pain sites in a training population (n = 445,132). The model will be then assessed in an independent group of individuals that were not used to derive the model (testing set; n = 48,079). A total of three components were selected explaining the highest variance while remaining sparse (R^2 and Mean Squared Error across 10-fold cross validation are reported in **Supplementary Fig. 1A**). A classification regarding participants' in-person information was therefore performed to assess the respective contribution of mental health, physical health and sociodemographics. This classification included a series of 99 features forming ten categories regrouped as followed: i) mental health (traumas, neuroticism, and mood), ii) sociodemographics (socioeconomic, occupational, and demographic), and iii) physical health (substance use, sleep, physical activity, and physiology measures; **Fig. 2A**). This model was derived to best explain the number of self-reported pain as a function of the normalized features presented in this classification.

Here, we examine the representation of the pain-associated burden derived from our model by assessing the relative contribution of the three dimensions and the ten categories, as well as their interactions. The weighted sum of the normalized features was use model each dimension and categorie. Overall, the model explained about 18% of the total variance in self-reported number of pain sites. We observed differential but overlapping contribution across the three dimensions (**Fig. 2B,** see validation in **Supplementary Fig. 1C**). Mental health alone explained the most variance (14% of the total 18%), followed by physical health (11% out of 18%), and sociodemographic (2% out of 18%). Furthermore, their combinations (i.e. unions) reveal that substantial overlap exist between vulnerabilities regarding mental health, physical health and sociodemographics to pain from the union of these dimensions being far from additive.

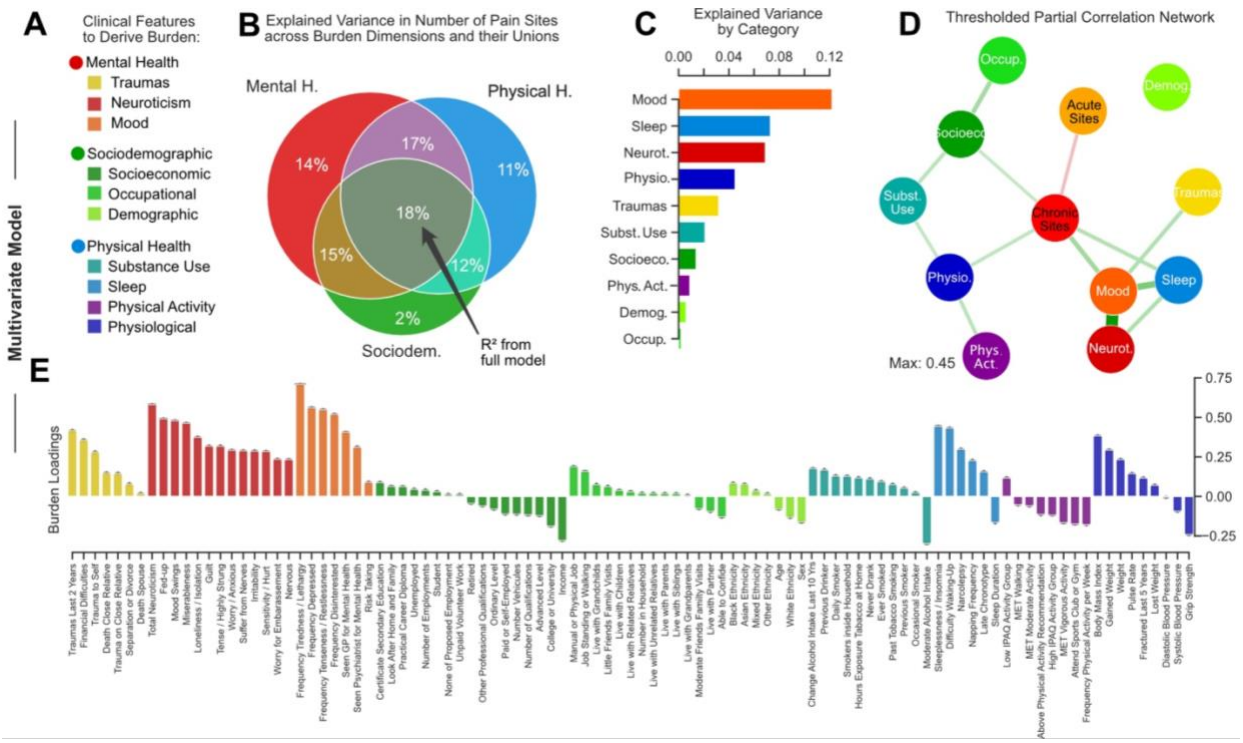


Figure 2. Deriving a multivariate epidemiological model explaining the spreading of pain in a large population of nearly 450,000 individuals (training set). A. Classification of clinical features in 10 categories regrouped across 3 dimensions. **B.** Venn diagram of the explained variance using weighted sum across dimensions of the epidemiological model. **D.** Explained variance of the number of pain sites across categories. **E.** Network analysis using partial correlation with a 0.1 threshold across categories including the number of acute and chronic pain

sites. **E.** Loadings of clinical features from the epidemiological model colored by categories, with confidence interval coefficients displayed with violin plot (see the edge of the bars) using 1,000 bootstrap resampling.

One arising question is how these categories interact together and relate to pain. To better understand the relationship between these pain-related variables, we implemented a network analysis approach. First, the weighted sum was applied across each of the ten categories. Mood alone could explain about 12% of the explained variance, followed by sleep, neuroticism, and physiological measures (**Fig. 2C,** see validation in **Supplementary Fig. 1D**). All categories were significantly associated with number of pain sites ($p < 1e\text{-}180$). Then, these categories were integrated using a partial correlation network analysis approach to examine the dominant pathways between our categories and pain sites (**Fig. 2D**). These connections were evaluated using partial correlation controlling for all potential other edges and a threshold to only conserve major connections (partial-r = 0.1; equivalent to a small effect size). For instance, **Fig. 2D** reveals how neuroticism and traumas were intertangled with mood and sleep who directed impacted chronic pain through major connections. Overall, we observed that mood, sleep, physiological measures, and socioeconomic status were the dominant contributors to the number of chronic pain sites. Importantly, no major contributors were observed for the number of acute pain sites, suggesting that the chronic state of pain specifically was link to the pain-associated burden. Other categories like physical activity, occupational, and substance use presented indirect contribution reinforcing the dominant contributors mentioned above. An important feature of the network displayed in **Fig. 2D** is the relative distance and location of the nodes determined from their connectedness in the network, meaning that chronic pain sites but not acute appeared central with the 10 categories.

Finally, **Fig. 2E** unpacks the high-dimensional representation of the burden across the 99 clinical features included to create the 10 categories displayed in **Fig. 2A-D** by showing their

respective loadings to our multivariate epidemiological model of pain spreading. The weights used to obtain the weighted sum and the overall model can be seen in the supplementary (**Supplementary Fig. 1B**). As a result of these loadings, the epidemiological model now referred as quantifying the burden accompanying the spreading of pain.

Aim #3. Evaluating the performance of the epidemiological model to discriminate the spreading of pain and high-impact secondary outcomes.

The epidemiological model of the burden associated with the spreading of pain was evaluated in both the training (through cross-validation; n = 445,132) and testing (unbiased sample of n = 48,079) cohort data. Of note is that these patients reporting *pain all over the body* were initially excluded from our co-occurring pain sites analyses to avoid making assumptions regarding their equivalence in number of pain sites (i.e. as equivalent to more or less than seven pain sites). These individuals were instead used as an additional validation group, where the participants reported pain on multiple unspecified sites (i.e. widespread). The burden index obtained from the model was evaluated in the training (**Fig. 3A-C**, *upper row show cross validation results*) and generalized in the testing (**Fig. 3A-C**, *lower row*) cohort with medium effects (**Fig. 3C**, *train:* $r = 0.40$, $R^2 = 0.16$, *test*: $r = 0.34$, $R^2 = 0.12$) for the spread of chronic pain sites, and small-to-medium effects (*train*: $r = 0.15$, $R^2 = 0.02$, *test*: $r = 0.15$, $R^2 = 0.02$) association for spread of acute pain sites. Interestingly, while the model was derived to predict the spread of pain unspecific to the state, our model presents an association almost 3-folds larger for chronic pain sites compared to acute. This specificity for chronic states suggests that the clinical features can better capture the spread of chronic pain than acute pain.

We validated this idea by evaluating how discriminable each pain sites were in compared to pain-free individuals using both Cohen's d effect sizes and discrimination performance using Area Under the Receiver Operating Characteristic curve (AUC-ROC). Both in the training and testing group, medium-to-very-large effect sizes in the burden index were found across participants reporting pain and those pain-free which corresponded to acceptable-to-excellent discrimination between participants presenting chronic pain and those pain-free (**Fig. 3B**, *train in red*: Cohen's d = 0.67-1.08 & AUC-ROC = 0.73-0.82; *test in red*: Cohen's d = 0.53-0.97 & AUC-ROC = 0.70-0.81). Similarly, poor-to-large effects sizes and poor-to-near-excellent discrimination were found between participants with acute pain and those pain-free (**Fig. 3B**, *train in orange*: Cohen's d = 0.42-0.83 & AUC-ROC = 0.65-0.77; *test in orange*: Cohen's d = 0.33-0.69 & AUC-ROC = 0.63-0.74). Anatomical body sites presenting greater proportion of chronic multi-site pain compared to chronic single-site pain were most discriminable (e.g. facial and stomach/abdominal pain). Most importantly, we also tested the model to discriminate individuals reporting *pain all over the body* (a category that was not used to derive the model for the reasons mentioned above) compared to pain-free participants. We found very large effect sizes and outstanding or near-to-outstanding discrimination for chronic pain all over the body (**Fig. 3B**, *train in red*: Cohen's d = 1.61 & AUC-ROC = 0.91; *test in red*: Cohen's d = 1.42 & AUC-ROC = 0.88) as well as near-large effect sizes and near-to-excellent discrimination for acute pain all over the body (**Fig. 3B**, *train in orange*: Cohen's d = 0.86 & AUC-ROC = 0.78; *test in orange*: Cohen's d = 0.74 & AUC-ROC = 0.76).

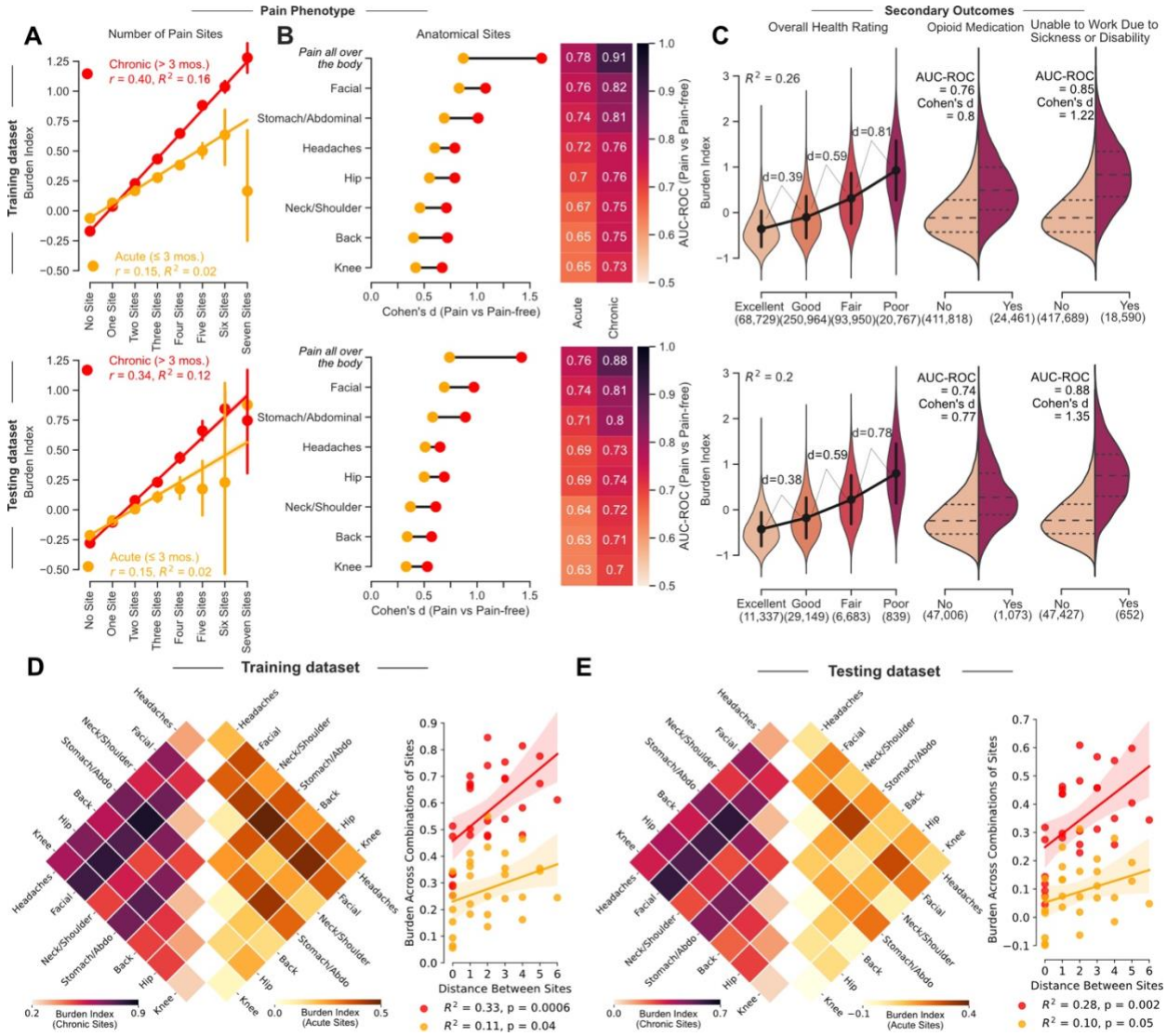


Figure 3. Evaluating the performance of the epidemiological model to discriminate the spreading of pain and high-impact secondary outcomes. A. Predicting the spread of pain from the epidemiological model of the pain-associated burden (i.e. burden index) in the training population (above) and testing cohort (below) excluding participants with *pain all over the body* (used as validation). Mean estimated across pain sites and standard errors plotted. **B.** Cohen's d effect sizes in the burden index for each pain site (acute in orange and chronic in red) compared to pain-free individuals. Heatmap represents the discrimination using AUC-ROC for both acute and chronic pain sites compared to pain-free individuals. **C.** Three pain-related secondary outcomes (health ratings, opioid medication, unable to work due to sickness or disability) were used to assess the clinical impact of pain. Cohen's d effect sizes in the burden (near-medium to large) estimated across self-reported level of overall health ratings. Large Cohen's d effect size and acceptable-to-excellent discrimination in predicting opioid medication use. Very large Cohen's d effect size and excellent-to-outstanding discrimination in predicting inability to work due to sickness or disability. **D-E. Mapping the spread of pain across sites. D.** Projected burden in the training population

across chronic (> 3 months; red in left) and acute (≤ 3 months; yellow in right) pain sites. Association between the distance of the co-occurring pain sites and the burden score. **E.** Association between the distance of the co-occurring pain sites and the burden score in the testing cohort.

Here, we demonstrated that the model systematically discriminated better chronic than acute pain sites compared to pain-free individuals but was most sensitive to discriminate patients whom the pain has spread to across the whole body, in participants presenting chronic pain *all over the body*.

Furthermore, the epidemiological model of the pain-associated burden was also evaluated regarding pain-related high-impact secondary outcomes including: i) overall health rating, ii) opioid medication use (N02A group, according to WHO's ATC taxonomy), and iii) inability to work due to sickness or disability (**Fig. 3E**). These secondary outcomes were used to extent the epidemiological to predict clinically useful outcomes associated with stronger spreading of pain.

Firstly, the ability from the burden index estimated epidemiological model to predict overall health rating was assessed. The model could explain a quarter to a fifth of the overall health rating variance equivalent to large or near large effect sizes with small-to-large effect sizes across all four levels of health ratings (**Fig. 3C**, *train*: $R^2 = 0.26$, Cohen's $d_{Excel\text{-}Good} = 0.39$, $d_{Good\text{-}Fair} = 0.59$, $d_{Fair\text{-}Poor} = 0.81$; *test*: $R^2 = 0.20$, Cohen's $d_{Excel\text{-}Good} = 0.38$, $d_{Good\text{-}Fair} = 0.59$, $d_{Fair\text{-}Poor} = 0.78$). The strong association with overall health ratings confirms that epidemiological model of spreading extent to the contributing role of overall poor health, a feature closely linked to diseases severity in many chronic conditions which often requires some extent of pain management.

Secondly, opioids medication usage (**Fig. 3C**, *middle*) and inability to work (**Fig. 3C**, *right*) were evaluated. These represent major clinical outcomes associated with high-impact chronic pain in need for high accommodations. Individuals using opioid medications ($n_{train/test} = 24{,}461/1{,}073$) presented near-large effect sizes in the burden compared to the rest of the population not using

opioids ($n_{train/test}$ = 411,818/47,006) while the model provided acceptable-to-excellent discrimination (**Fig. 3C**, *train*: Cohen's d = 0.8 & AUC-ROC = 0.76; *test*: Cohen's d = 0.77 & AUC-ROC = 0.74). Moreover, participants unable to work ($n_{train/test}$ = 18,590/652) presented large effect sizes in the burden compared to the rest of the population who were able to work ($n_{train/test}$ = 417,689/47,427) while the model provided excellent-to-outstanding discrimination (**Fig. 3C**, *train*: Cohen's d = 1.22 & AUC-ROC = 0.85; *test*: Cohen's d = 1.35 & AUC-ROC = 0.88). Our model predicted opioid medication and inability to work with acceptable-to-excellent and excellent-to-outstanding discrimination. These results suggests that the individuals at risk for presenting a pain that spread across multiple body sites are also at risk to be in greater need for clinical accommodations to manage their pain, potentially because individuals with high burden may be at risk to experience the pain more severely.

Lastly, we hypothesize that individuals with greater burden would present pain more spread over multiple body sites, including the distal ones. Here, we assess the risk of pain co-occurrence as a function of the burden (**Fig. 3D**, *train in left*; **Fig. 3E**, *test in right*) by evaluating the burden score across each anatomical and co-occurring pain sites. We found that a greater burden was associated with greater risk for the pain to spread distally from the occurring site. While this effect was robust for chronic pain sites in both training and testing datasets (**Fig. 3D,** *train in red*: R^2 = 0.33, p_{perm} = 0.0006 with 10,000 permutations; **Fig. 3E**, *test in red*: R^2 = 0.28, p_{perm} = 0.002), it was only marginally significant for acute pain sites (**Fig. 3D,** *train in orange*: R^2 = 0.11, p_{perm} = 0.04; **Fig. 3E**, *test in orange*: R^2 = 0.10, p_{perm} = 0.05). These results suggest that proximal co-occurring pain will occur most independently from the burden score (as shown in Fig. **3D-E**), while the rarely co-occurring pain between distal sites is most strongly determined by the burden score (as shown in **Fig. 1B**).

Aim #4. Assessing the longitudinal risk for the spreading of pain and high-impact secondary outcomes at 10-year follow-up

We extent our epidemiological model to examine its prognosis value in predicting the spreading of chronic pain in the testing population (n = 48,000) at 10-year follow-up visit using a longitudinal approach. Firstly, we examine the prognosis of chronic pain at follow-up given the chronic pain status and the estimated burden from our epidemiological model at the initial visit. Individuals presenting chronic pain at the initial visit presented higher odds (4.9 to 63 times more likely) to have chronic pain persist at the 10-years follow-up (**Fig. 4A**, see *matrix diagonal*). Importantly, individuals presenting chronic pain at a given anatomical sites were also 0.5 to 7.5 times more likely to develop chronic pain or for the chronic pain to persist in neighbouring (i.e. proximal) sites. Similar as first shown across pain sites, greater odds (logs) for the development or persistence of chronic pain observed in proximal sites (**Fig. 4A**, $R^2 = 0.41$, $p_{perm} < 0.0001$ with 10,000 permutations). This finding suggests that the occurrence of chronic pain induces a substantial to very large risk for the chronic pain to persist in the long-term, but to potentially spread to surrounding sites. Furthermore, a relationship of opposite direction was found in the predicted burden and the spreading of pain across each anatomical site. One unit higher in the burden index was associated with low odds (0.9 to 1.8 times more likely) for individuals presenting chronic pain at a given anatomical site to have their chronic pain persist at 10-year follow-up (**Fig. 4B**, see matrix diagonal). Most importantly, individuals with chronic pain at a given anatomical site were at 1.4 to 3.5 times more likely to develop chronic pain or for the chronic pain to persist in sites further apart (i.e. distal) from the estimated risk (one unit higher in the burden). This is shown with greater odds (log) for the development or persistence of chronic pain observed in distal sites (**Fig. 4B**, $R^2 = 0.26$, $p_{perm} < 0.0001$). Together, these relationships demonstrate a dual process

at play for chronic pain and its spreading. While chronic pain typically induces some risk to spread in close neighbouring sites, the burden accompanying the pain can determine the extent of this spread to other sites.

Secondly, we assess the risk estimated from the epidemiological burden as a function of the severity of the spreading (0 sites to 4+ sites) in three groups determined from their pain status in the initial visit: i) chronic pain-free individuals (**Fig. 4C**, *left*) ii) chronic single-site pain (**Fig. 4C**, *middle*), and iii) chronic multi-site pain (two or more sites, including individuals with *pain all over the body*; **Fig. 4C**, *right*). Greater burden was associated with stronger spreading at the 10-years follow-up across the three respective groups (**Fig. 4C**, *left*, r = 0.18, p = 2.2e-213; **Fig. 4C,** *middle*, r = 0.20, p = 1.7e-96; **Fig. 4C,** *right*, r = 0.26, p = 1.9e-118). Likewise, individuals who would present less spreading of their pain presented lower burden than the rest of their groups.

The model also presented similar prognosis value for high-impact secondary outcomes previously introduces: i) opioid medication use and ii) inability to work due to sickness or disability. Participants using persisting their opioid medication at 10-years follow-up (n = 260) present near-medium effect sizes in the burden compared to those who stopped their medication (n = 798) with the model provided poor discrimination (**Fig. 4D**, Cohen's d = 0.43, p = 1.8e-11 & AUC-ROC = 0.63). In contract, participants who started using opioid medications at follow-up (n = 445) presented medium effect sizes in the burden compared to those who never started (n = 46,205) with the model providing acceptable discrimination (**Fig. 4E**, Cohen's d = 0.58, p = 3.7e-39 & AUC-ROC = 0.7). This finding confirms the idea that the burden estimated from the epidemiological model is a mixture of both predisposition for pain and consequences reinforcing the pain.

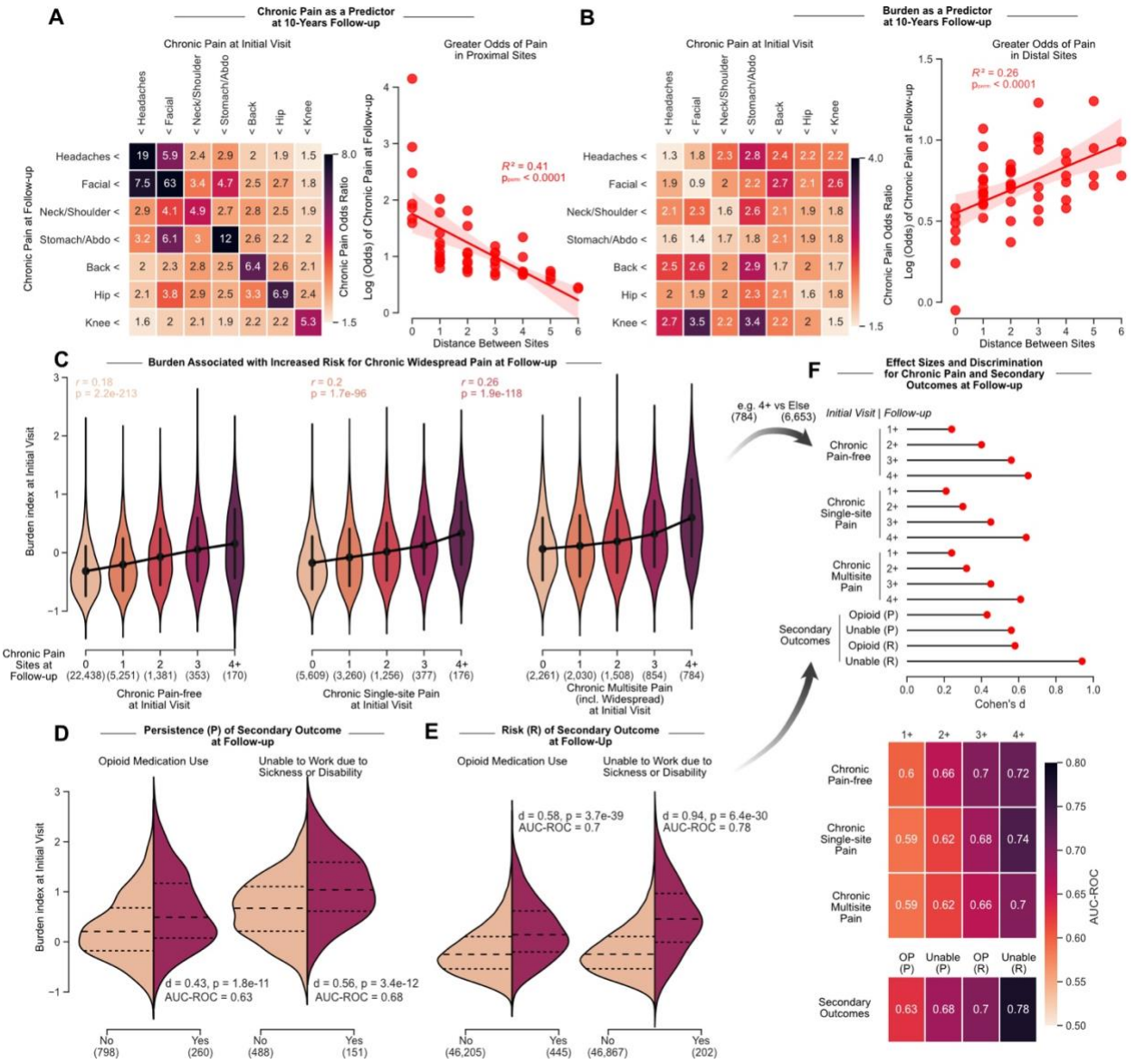


Figure 4. Longitudinal risk for spreading of pain and high-impact secondary outcomes at 10-year follow-up. A-B. Chronic pain induces a risk for pain at proximal sites while the burden induces a risk for pain at distal sites. A. Odds of chronic pain persisting at the follow-up visit from the occurrence across pain sites. **B.** Odds of chronic pain persisting at follow-up predicted by the burden for a given anatomical pain site. **C.** Estimated burden at the initial visit across the number of pain sites for individuals with chronic pain-free, chronic single-site pain, and chronic multisite pain (including *pain all over the body*). **D-E. Secondary outcomes. D.** Risk of opioid medication and inability to work persisting at the follow-up visit predicted from the burden at the initial visit using Cohen's d effect size and AUC-ROC discrimination. **E.** Risk of being prescribed opioid medication and being no longer able to work due to sickness or disability at the follow-up visit estimated from the burden. **F. (Recapitulative) A higher burden is associated with a predisposition to developing more widespread pain.** Cohen's d effect sizes (0.3-0.94)

across comparisons of groups (i.e., 4+ sites vs else). Area Under the Receiving Operating Characteristic Curve for all variables assessed (AUC-ROC from 0.59-0.78).

Additionally, participants who persisted to be unable to work at 10-years follow-up (n = 151) present medium effect sizes in the burden compared to those who came back to work (n = 488) with the model provided poor-to-acceptable discrimination (**Fig. 4D**, Cohen's d = 0.56, p = 3.4e-12 & AUC-ROC = 0.68). In contract, participants who became unable to work at follow-up (n = 202) presented large effect sizes in the burden compared to those didn't became unable (n = 46,867) with the model providing acceptable-to-excellent discrimination (**Fig. 4E**, Cohen's d = 0.94, p = 6.4e-30 & AUC-ROC = 0.78). Overall, our model developed to predict the spreading of pain predicted with poor or poor-to-acceptable discrimination for the persistence of opioid medication usage and inability to work in comparison to acceptable-to-excellent discrimination for their associated risk of occurrence. This finding suggests that the model predict the of persistence of secondary outcomes and even better the risk of occurrence, potentially because the model can discriminate high-risk individuals who will experience the pain more severely.

Overall, we assess the ability for the epidemiological model to discriminate across levels of spreading while including secondary outcomes for comparison (**Fig. 4F**). We found that the burden index presented increasing discrimination across the three groups (chronic pain-free, chronic single-site pain and chronic multisite pain at the initial visit) to discriminate those that will develop multiple sites (**Fig. 4F**, *above* for Cohen's d effect sizes, and *below* for AUC-ROC discriminations). Discriminations of the development of a single pain site or its recovery were poor (AUC-ROC = 0.59-0.60) to acceptable for widespread pain (4 sites or more, including *pain all over the body*; AUC-ROC = 0.70-0.72). These results confirmed what is seen cross-sectionally (**Fig 3D-E)** where chronic pain will spread proximally independently from the burden score, while chronic pain that spreading distally is strongly determined by the burden score (**Fig. 4A-B**).

Aim #5. Peripheral and central biological determinants associated with the epidemiological model.

Three major biological axes were integrated to observe the impact of biological determinants on the expression of the burden derived from our epidemiological model (**Fig. 5A**).

The first axis is genetic risk using an a-priori derived polygenetic risk score from a genome-wide association study on the UK Biobank (Khoury et al., 2021). This score was derived on chronic multisite pain participants while leaving a group of subjects for a test set that were used in the current study. We found that greater scores in the polygenic risk score was associated with greater expression of this burden (**Fig. 5B**, *left*, r = 0.11, p = 6.8e-54). 10,000 Permutation tests were done to assess the robustness of small effect sizes and revealed no overlap between this association and the distribution of randomly generated null models (**Fig. 5B**, *right*, r = 0.11, $p_{perm} < 0.0001$). The genetic risk associated with the burden, however, did not differ substantially from the spreading of pain that present very similar association the polygenic risk score (**Supplementary Fig. 3B**, *left*: r = 0.10, p = 1.1e-40, right: r = 0.10, $p_{perm} < 0.0001$).

The second axis is the immune-inflammatory profile that included C-Reactive protein from saliva samples and immune cell counts from complete blood count. We found that the burden presented a small-to-medium association with C-Reactive Protein (log-transformed; **Fig. 5C**, left: r = 0.20, p = 1.6e-154), and small associations with immune cell counts (**Fig. 5C**, *right*: r = 0.03-0.11). The relationships from C-Reactive protein with the burden was substantially larger (over 2-fold) compared than with the spreading of pain sites (**Supplementary Fig. 3A**, *left*: r = 0.09, p = 2.4e-32; *right*: r = 0.09, $p_{perm} < 0.0001$) suggesting that the burden associated with pain may have more predictive value regarding the associated inflammation of the spreading than pain itself.

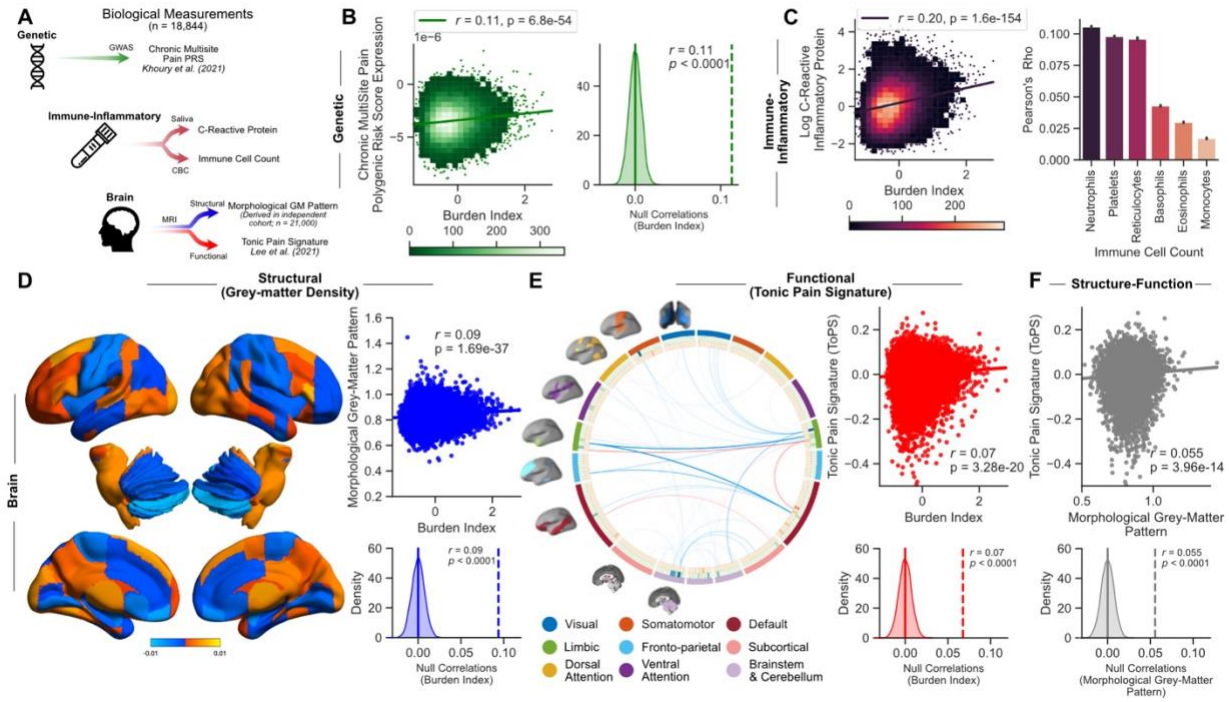


Figure 5. Peripheral and central biological determinants associated with the epidemiological model. A. Schematic of the three major biological axes examined: genetic (polygenic risk score; Khoury et al., 2021), immune-inflammatory markers and brain phenotype (predefined functional signature; Lee et al., 2021 and Morphological Grey-Matter Pattern; see *Supplementary Fig. 2*) in a testing cohort (n = 18,844). **B-C. Peripheral biological determinants. B.** Higher scores in polygenic risk score were associated with greater expression of the burden (left), significance validated with 10,000 permutation tests (right). **C.** Higher c-reactive inflammatory protein and immune cells counts were associated with higher burden scores. **D-F. Central biological determinants. D.** Morphological grey-matter density pattern representation from the weights from a multivariate model derived on a left-out dataset (n = 21,000; see *Supplementary 2*). Higher scores in this morphological grey-matter pattern were associated with higher burden scores. **E.** Validation of the functional tonic pain (ToPS) brain signature associated with experimental and clinical pain. Circle diagrams display the top 5% connectivity weights (positive links in red and negative links in blue) across nine large-scale brain networks. Major edges included limbic-default-frontal parietal networks and were associated with the burden index. **F.** Coupling of between the derived Morphological Grey-Matter Pattern and the pre-determined functional Tonic Pain Signature.

The third axis is the brain phenotype using both structural and functional magnetic resonance imaging (MRI) from grey-matter density and functional connectivity respectively. The structural phenotype was derived on a left-out group of participants, while the functional phenotype was from a capsaicin-induced Tonic Pain Signature (ToPS; Lee et al., 2021), using the

top 5% connectivity links. The grey-matter morphological pattern was derived in a left-out group of participants (n = 21,000) to model the spreading of pain (**Supplementary Fig. 2E-G**) across 139 anatomical regions from the grey-matter density (**Supplementary Fig. 2A**), commonly used to estimate the density of grey-matter cells in brain volume. The pattern presented a bilateral loss of grey-matter density in motor areas including in the sensorimotor, supplementary motor cortex, and cerebellum and a gain of grey-matter density in regions involved into arousal including involved brainstem and subcortical regions as well as the superior frontal gyrus (i.e. overlap with dorsolateral prefrontal cortex; **Fig. 5D,** *left*). Some asymmetry was observed in the anterior cingulate cortex (loss in the anterior and a gain in the posterior cingulate cortex) and in the frontal pole (loss in the left and gain in the right frontal pole). The pattern presented very small associations in both left-out and studied test groups (**Supplementary Fig. 2E**, *train*: r = 0.07, p = 8.65e-22, $p_{perm} < 0.0001$, **Supplementary Fig. 2G**, *test*: r = 0.05, p = 1.29e-11, $p_{perm} < 0.0001$), robust to 10,000 permutation test. However, the epidemiological model captured better the same morphological grey-matter pattern with small or near small associations (**Supplementary Fig. 2F**, *train*: r = 0.11, p = 2.81e-55, $p_{perm} < 0.0001$, **Fig. 5D**, *test*: r = 0.09, p = 1.69e-37, $p_{perm} < 0.0001$). This potentially suggest that the coupling of a loss-of-function in the sensorimotor and accompanied motor areas combined with a gain-of-function in the arousal areas could be weakly associated of the risk in spreading of pain.

The same relationship was found from the pre-determined and published ToPS signature derived specially to predict experimental pain but found to also predict clinical pain. The top 5% connectivity links were characterized by both increased and decreased connectivity within the limbic network as well as increase connectivity between the limbic or subcortical networks with the default-mode network (**Fig. 5E**, *left*). Smaller decreases in connectivity were observed in the

prefrontal-parietal networks with limbic and default-mode networks, as well as between the subcortical network with the brainstem and cerebellum networks. While the ToPS signature presented weak associations with the spreading of pain sites (**Supplementary Fig. 2G**, r = 0.025, p = 6.99e-4, p_{perm} = 0.0003), it presented very-small-to-small association with the epidemiological model (**Fig. 3E**, r = 0.07, p = 3.3e-20, p_{perm} < 0.0001) and quite robust to 10,000 permutations with no overlap. The results with the ToPS were robust across multiple thresholds (1, 5, 10, 15, 20%). Finally, we show assess convergence between measures of the structural and functional brain. We found that individuals presenting a brain morphology more like our grey-matter density pattern were also more at risk to express higher signature responses in the ToPS (**Fig. 3F**, r = 0.055, p = 4.0e-14 p_{perm} < 0.0001). Note that functional MRI results controlled linearly for movement in the scanner (regressed out from each feature), due to high association between mean head motion and the burden index (r = 0.31; more regarding future validations in Discussion).

Discussion

A common but poorly understood phenomenon in conditions with multisite pain or widespread pain is increased pain sensitivity or the occurrence of pain at sites far from the source of pain. This phenomenon is also seen in patients for whom the source of the pain or the etiology is unknown (i.e. considered idiopathic; Vlaeyen, et al. 2018). Even in individuals living with chronic diseases putting them at risk for chronic pain (e.g. osteoarthritis), large discrepancies exist between those without pain to those developing one or multiple sites of pain. Here, we assess the role of a series of clinical features in line with the biopsychosocial model to predict the risk of pain spreading across multiple sites.

First, we showed that various chronic pain conditions present a higher prevalence of chronic multi-site pain than chronic single-site pain, with a tendency for the pain to spread in

proximal sites with more severe symptoms. Second, we derived in a large population a multivariate epidemiological model explaining this spreading of the pain at baseline visit using 99 clinical features regarding major dimensions of life, including one's physical health, mental health, and sociodemographic status. Given the high-dimensional representation of the model presenting both a mixture of predispositions and consequences of the pain, we referred to our model as a burden index. Third, we found that the epidemiological model of the burden could explain cross-sectionally the spread, presented outstanding discrimination for individuals with chronic pain *all over the body*, and predicted high-impact secondary outcomes related to pain such as health ratings, opioids medication usage, and inability to work. The burden could predict the spreading of pain in sites further apart (i.e. distal) and was the highest in anatomical pain sites commonly present higher occurrence of multisite pain (facial, stomach/abdominal and headaches). Fourth, we validated the spreading longitudinally at a 10-years follow-up visit. We found that chronic pain induced risk for the development of chronic pain or its persistence at follow-up in proximal anatomical sites while the risk induced by the epidemiological burden was in distal anatomical sites. The model predicted the development of four pain sites or more as well as the development of high-impact secondary outcomes such as the risk to start opioid medication usage and becoming unable to work due to sickness or disability. Fifth and lastly, elevated biological markers were associated with higher expression of this epidemiological burden of pain. This was found in both peripheral markers (polygenic risk score and immune-inflammatory markers) and central markers (pre-determined brain signatures and patterns associated with pain).

In this research, we introduce a new framework to understand widespread pain and the risk of spreading across pain sites. The online assessment from the experience of pain questionnaires revealed that the severity of pain increases across many dimensions as the number of pain sites

accumulate with greater spread seen in patients who've experienced pain the longest. While the number of pain sites has been used in some studies shown to be associated with poor health and work disability (Kamaleri et al., 2008; Kamaleri, Natvig, Ihlebaek, & Bruusgaard, D., 2009), other studies have shown that pain sites tend to frequently co-occurs (Andersen, Clausen, Carneiro & Holtermann, 2012), no studies have exploited the spatial distance and the risk of co-occurrence associated. We found that while pain co-occurred in neighbouring sites, elevated risks factors was associated with high co-occurrence in pain sites further apart from each other. This was demonstrated in both cross-sectionally and longitudinally and provide some important insight into widespread pain. First, widespread pain should be understood as a continuum by integrating the number of self-reported pain sites associated. Second, the occurrence of any pain site is typically associated with an elevated risk for proximal pain. Third, biopsychosocial determinants such as the burden exploited in this paper can further predispose individuals to increased risk of pain sites further apart from the initial pain site, potentially leading to widespread. One potential limitation is the idea of body overcompensation, which attributes the development of pain to over-reliance overtime to other body parts. However, additional analyses revealed that the spread occurred at similar duration, where chronic pain sites did not induce a spread in acute sites and acute sites did not induce a spread in chronic sites. This suggests that this process is unlikely to emerge from the development of compensating pain sites but rather, the simultaneous development of pain across multiple body sites.

Our study relies on a wholistic and patient-centered approach to provide a first empirical demonstration of the clinical benefits from the integration and synthesis of the biopsychosocial model of pain (Lumley et al., 2011; Pincus et al., 2013). The normalization we used for the clinical features allowed for their comparisons. As aligned with previous findings, psychological

contributors (i.e. mental health regrouping mood, neuroticism, and traumas) were found to be the most predictive of pain and its spreading (Hasenbring, Hallner & Klasen, 2001). However, other categories also provided substantial prediction such as sleep and physiological measures in physical health as well as socioeconomic status were also predictive. Regardless of the occurrence of accompanying medical conditions (e.g. osteoarthritis or other accompanying chronic diseases), a considerable proportion of the total variance in number of self-reported pain sites could be captured from our multivariate biopsychosocial approach. We found substantial overlap between the different dimensions suggesting that integrating more clinical features may provide only marginal gain in predication. The integration of commonly used and clinically readily accessible pain-agnostic measurements used in the derived epidemiological model provided good discrimination for pain outcomes, high-impact secondary outcomes. Most importantly, these discriminations surpass (AUC of 0.70-0.74 for 4 sites or more) past studies where longitudinal models (AUC equivalent of 0.60-0.66) were derived specifically to track risk for chronic pain development in rich diversified cohort (Traeger, et al. 2016; Stevans, et al. 2021).

We hope the epidemiological model provides insight toward a better understanding of the high-dimensional health representation associated with pain. Note that the aim of the presented work is not to assess the directionality (i.e. predisposition or consequence) or to present a case for the best-questionnaires to model pain. Rather, it is to unpack the burden extending beyond just the pain and to demonstrate that conditions with the most severe pain states (e.g. widespread) are more vulnerable and impacted from such burden. The numerous failures in finding new analgesics in clinicals trials to find opioid-alternatives may be due in part to the need to treat and target the accompanying burden of these pain conditions rather than solely the pain itself. This is in addition to the idea ineffective treatment targeting one site may be partially due to the persistence or

maintenance of pain at other sites. As a result, the spreading of pain would represent a meaningful clinical outcome to measure during analgesics clinical trials. Furthermore, the examination of the burden or barriers of a patient associated with their pain provides clear personalized targets in pain management from which clinical efforts can be applied to achieve realized patient outcomes. Moreover, the use of readily available clinical features with little required expertise would potentially be suitable within primary care or another clinical framework while providing acceptable prognosis value.

Extensive literature has suggested that repeated nociceptive activity could lead to abnormal processing of pain in the central nervous system and centrally mediated symptoms typically seen in widespread pain patients. This idea is referred to as a central sensitization process and represents a major theoretical framework of pain (Neblett et al., 2013, Louati, & Berenbaum, 2015, Maixner et al. 2016, Wolfe, Egloff & Häuser, 2016, Ji, Nackley, Huh, Terrando & Maixner, 2018). The epidemiological model explaining the spreading of pain loaded on features commonly associated with central sensitization with the largest weight of the model being the frequency of tiredness feelings (i.e. fatigue) accompanied by large weights in mood, sleep, and traumas (see **Supplementary Fig. 1B**). Additional analysis also revealed that individuals presenting on the higher spectrum of this burden also presents cognitive impairments (McCracken & Iverson, 2001), a symptom typically seen in fibromyalgia and widespread pain conditions. These centrally mediated symptoms explain why an increasing number of conditions (depression, PTSD, low back pain, etc.) have been recognized as conditions with central sensitization syndrome (Maixner et al. 2016).

One alternative view to interpret the model could be that these symptoms linked to central sensitization represent high levels of psychopathology, with those most impaired presented a high

level of fatigue, impaired sleep, and poorer cognitive functions. These symptoms are common and shared across chronic pain and mood disorders (Serafini, Pryce, & Zachariou, 2020). Another alternative view to interpret the model could be that these symptoms are linked to the *sickness response* presenting similar symptomology as central sensitization (includes fatigue, numbness, muscle, and joint aches, and a form of malaise, among other symptoms). The only major distinction is that central sensitization suggests a process being more dominant in the central nervous system while the sickness response acknowledges it as being peripherally produced. However, our model is associated with both peripheral and central processes while slightly better captured by the peripheral inflammation from C-Reactive Protein. However, some work has highlighted the convergence between inflammatory processes seen in infections associated with the sickness response and depression seen in mood disorders (Dantzer et al., 2008). Some overlap appears to exist between the sickness response, overall psychopathology, and their contributions to central sensitization.

We also found the burden to be associated with higher levels of the various biological markers associated with pain. These markers included peripheral measures such as C-Reactive Inflammatory Protein and Polygenic Risk Scores (Khoury et al., 2021) or central measures such as our derived Grey-Matter Morphological Pattern or the Tonic Pain Signature (ToPS; Lee et al., 2021). Interestingly, these markers were better captured by the high-dimensional risk of spreading than the actual self-reported spread of pain sites (see **Supplementary Fig. 2-3**) with certain pain-free individuals at high risk scoring higher in these biological markers. This would suggest that these individuals are biologically (or *nociceptively*, due to lack of pain report) predisposed to experience pain more severely and more widespread if it develops. This predisposition was

confirmed longitudinally from the high-risk pain-free individuals developing more widespread pain (four or more sites out of the seven examined) at 10-years follow-up.

An interesting and important point is that when quantifying interindividual differences in brain structure and function relating to pain, pain-free individuals at high risk can present a brain phenotype like those with pain. However, individuals at high risk were found to be at higher risk to develop pain more widespread at the follow-up visit. This result suggests that the development of any biomarker of pain may also present some form of prognosis utility in pain-free individuals. Note that the effect sizes of brain phenotypes were nowhere near what is expected from a clinically useful biomarker of pain. The studied multivariate brain pattern and signature were done while controlling from various confounds and would require validation in other MRI scanning sequences. It is also important to recognize the presence of some selection bias occurring in the testing population attending the MRI visit at follow-up. We noticed that participants who attended the second visit presented less impaired sleep, lower BMI, and lower scores of this burden overall. This difference was due to the exclusion criteria (i.e. avoiding falling asleep in the scanner, being able to fit in a 3T MRI scanner, or not being claustrophobic) regarding eligibility to the follow-up session. These discrepancies are not specific to our study given that some selection bias can occur in MRI studies. One recent study on participants with anxiety found that those undergoing a brain MRI scan were typically healthier and less anxious (Charpentier et al., 2021) than those participating in the behavioral study. Future studies may also want to further characterize the role of inflammation in the spreading of pain by relying on a wide range of inflammatory markers given that the current study was limited to those offered by the UK Biobank. Other inflammatory markers include Erythrocyte Sedimentation Rate (ESR) and Plasma Viscosity (PV) used in primary care

for diagnosis and monitoring of inflammatory conditions, although C-Reactive Protein presented marginally superior diagnostic accuracy (Watson et al., 2019)

Overall, we develop an epidemiological model explaining the spreading of pain across anatomical sites highlighting the role of various risks factors to predispose and worsen the pain. This model was done on a large population of approximately 450,000 individuals and generalize on a smaller population of 48,000 individuals using solely readily accessible clinical features. We demonstrate the clinical benefits obtained from the synthesis and integration of the biopsychosocial model of pain. Our works establish a framework to understand chronic pain and the risk of spreading over multiple sites. Finally, we put forward the idea that the spreading of pain or its likelihood provides a meaningful target in clinical trials to complement pain ratings.

References

- Abdallah, C. G., & Geha, P. (2017). Chronic pain and chronic stress: two sides of the same coin?. *Chronic Stress*, *1*, 2470547017704763.
- Affaitati, G., Costantini, R., Tana, C., Cipollone, F., & Giamberardino, M. A. (2020). Co-occurrence of pain syndromes. Journal of Neural Transmission, 127(4), 625-646.
- Ahn, S. H., Mayo-Smith, W. W., Murphy, B. L., Reinert, S. E., & Cronan, J. J. (2002). Acute nontraumatic abdominal pain in adult patients: abdominal radiography compared with CT evaluation. *Radiology*, *225*(1), 159-164.
- Allen, J., Romay-Tallon, R., Brymer, K. J., Caruncho, H. J., & Kalynchuk, L. E. (2018). Mitochondria and mood: mitochondrial dysfunction as a key player in the manifestation of depression. *Frontiers in neuroscience*, *12*, 386.
- Andersson, G. B. (1999). Epidemiological features of chronic low-back pain. *The lancet*, *354*(9178), 581-585.
- Andersen, L. L., Clausen, T., Carneiro, I. G., & Holtermann, A. (2012). Spreading of chronic pain between body regions: prospective cohort study among health care workers. European journal of pain, 16(10), 1437-1443.
- Bair, M. J., Robinson, R. L., Katon, W., & Kroenke, K. (2003). Depression and pain comorbidity: a literature review. Archives of internal medicine, 163(20), 2433-2445.
- Bannon, S., Greenberg, J., Mace, R. A., Locascio, J. J., & Vranceanu, A. M. (2021). The role of social isolation in physical and emotional outcomes among patients with chronic pain. *General Hospital Psychiatry*, *69*, 50-54.
- Batty, G. D., Gale, C. R., Kivimäki, M., Deary, I. J., & Bell, S. (2020). Comparison of risk factor associations in UK Biobank against representative, general population based studies

with conventional response rates: prospective cohort study and individual participant meta-analysis. *Bmj*, *368*.

- Bedson, J., & Croft, P. R. (2008). The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC musculoskeletal disorders*, *9*(1), 1-11.
- Bonica, J. J., & Hoffman, J. F. (1954). The Management of Pain with Special Emphasis on the Use of Analgesic Blocks in Diagnosis, Prognosis, and Therapy. *Anesthesia & Analgesia*, *34*(5), 57-58.
- Breivik, H., Collett, B., Ventafridda, V., Cohen, R., & Gallacher, D. (2006). Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European journal of pain*, *10*(4), 287-333.
- Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology*, *30*(9), 846-856.
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., ... & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203-209.
- Carnes, D., Parsons, S., Ashby, D., Breen, A., Foster, N. E., Pincus, T., ... & Underwood, M. (2007). Chronic musculoskeletal pain rarely presents in a single body site: results from a UK population study.
- Carr, D. B., & Goudas, L. C. (1999). Acute pain. *The Lancet*, *353*(9169), 2051-2058.
- Cassidy, S., Chau, J. Y., Catt, M., Bauman, A., & Trenell, M. I. (2016). Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233 110 adults from the UK

Biobank; the behavioural phenotype of cardiovascular disease and type 2 diabetes. *BMJ open*, *6*(3).

- Charpentier, C. J., Faulkner, P., Pool, E. R., Ly, V., Tollenaar, M. S., Kluen, L. M., ... & O'Doherty, J. P. (2021). How representative are neuroimaging samples? Large-scale evidence for trait anxiety differences between fMRI and behaviour-only research participants. Social Cognitive and Affective Neuroscience.
- Cormier, S., Lavigne, G. L., Choinière, M., & Rainville, P. (2016). Expectations predict chronic pain treatment outcomes. *Pain*, *157*(2), 329-338.
- Craddock, N., Owen, M. J., & O'Donovan, M. C. (2006). The catechol-O-methyl transferase (COMT) gene as a candidate for psychiatric phenotypes: evidence and lessons. *Molecular psychiatry*, *11*(5), 446-458.
- Crofford, L. J. (2015). Chronic pain: where the body meets the brain. *Transactions of the American Clinical and Climatological Association*, *126*, 167.
- Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W., & Kelley, K. W. (2008). From inflammation to sickness and depression: when the immune system subjugates the brain. Nature reviews neuroscience, 9(1), 46-56.
- Davis, K. D., Flor, H., Greely, H. T., Iannetti, G. D., Mackey, S., Ploner, M., ... & Wager, T. D. (2017). Brain imaging tests for chronic pain: medical, legal and ethical issues and recommendations. *Nature Reviews Neurology*, *13*(10), 624-638.
- Dominick, C. H., Blyth, F. M., & Nicholas, M. K. (2012). Unpacking the burden: understanding the relationships between chronic pain and comorbidity in the general population. *Pain*, *153*(2), 293-304.

- Dorner, T. E., Muckenhuber, J., Stronegger, W. J., Ràsky, É., Gustorff, B., & Freidl, W. (2011). The impact of socio-economic status on pain and the perception of disability due to pain. *European journal of pain*, *15*(1), 103-109.
- Edwards, R. R. (2005). Individual differences in endogenous pain modulation as a risk factor for chronic pain. *Neurology*, *65*(3), 437-443.
- Elliott, P., & Peakman, T. C. (2008). The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. International journal of epidemiology, 37(2), 234-244.
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., ... & Jiang, T. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral cortex*, *26*(8), 3508-3526.
- Fayaz, A., Croft, P., Langford, R. M., Donaldson, L. J., & Jones, G. T. (2016). Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. *BMJ open*, *6*(6), e010364
- Fillingim, R. B., & Ness, T. J. (2000). Sex-related hormonal influences on pain and analgesic responses. Neuroscience & Biobehavioral Reviews, 24(4), 485-501.
- Fillingim, R. B., Ohrbach, R., Greenspan, J. D., Sanders, A. E., Rathnayaka, N., Maixner, W., & Slade, G. D. (2020). Associations of psychologic factors with multiple chronic overlapping pain conditions. *J Oral Facial pain headache*, *34*, s85-s100.
- Fishbain, D. A., Cutler, R., Rosomoff, H. L., & Rosomoff, R. S. (1997). Chronic pain-associated depression: antecedent or consequence of chronic pain? A review. *The Clinical journal of pain*, *13*(2), 116-137.

- Freburger, J. K., Holmes, G. M., Agans, R. P., Jackman, A. M., Darter, J. D., Wallace, A. S., ... & Carey, T. S. (2009). The rising prevalence of chronic low back pain. *Archives of internal medicine*, *169*(3), 251-258.
- Fuentes, M., Hart-Johnson, T., & Green, C. R. (2007). The association among neighborhood socioeconomic status, race and chronic pain in black and white older adults. *Journal of the National Medical Association*, *99*(10), 1160.
- Gaskin, D. J., & Richard, P. (2012). The economic costs of pain in the United States. *The Journal of Pain*, *13*(8), 715-724.
- Gatchel, R. J. (2004). Comorbidity of chronic pain and mental health disorders: the biopsychosocial perspective. *American Psychologist*, *59*(8), 795.
- Gerdle, B., Fischer, M. R., Cervin, M., & Ringqvist, Å. (2021). Spreading of pain in patients with chronic pain is related to pain duration and clinical presentation and weakly associated with outcomes of interdisciplinary pain rehabilitation: a cohort study from the Swedish Quality Registry for Pain Rehabilitation (SQRP). Journal of Pain Research, 14, 173.
- Gerhardt, A., Eich, W., Treede, R. D., & Tesarz, J. (2017). Conditioned pain modulation in patients with nonspecific chronic back pain with chronic local pain, chronic widespread pain, and fibromyalgia. Pain, 158(3), 430-439.
- Goldberg, D. S., & McGee, S. J. (2011). Pain as a global public health priority. *BMC public health*, *11*(1), 1-5.
- Gregory, N. S., Harris, A. L., Robinson, C. R., Dougherty, P. M., Fuchs, P. N., & Sluka, K. A. (2013). An overview of animal models of pain: disease models and outcome measures. *The Journal of Pain*, *14*(11), 1255-1269.

- Grol-Prokopczyk, H. (2017). Sociodemographic disparities in chronic pain, based on 12-year longitudinal data. *Pain*, *158*(2), 313.
- Guzmán, J., Esmail, R., Karjalainen, K., Malmivaara, A., Irvin, E., & Bombardier, C. (2001). Multidisciplinary rehabilitation for chronic low back pain: systematic review. *Bmj*, *322*(7301), 1511-1516
- Hannibal, K. E., & Bishop, M. D. (2014). Chronic stress, cortisol dysfunction, and pain: a psychoneuroendocrine rationale for stress management in pain rehabilitation. *Physical therapy*, 94(12), 1816-1825.
- Hasenbring, M., Hallner, D., & Klasen, B. (2001). Psychological mechanisms in the transition from acute to chronic pain: over-or underrated?. *Schmerz (Berlin, Germany)*, *15*(6), 442-447.
- Herman, P. M., Broten, N., Lavelle, T. A., Sorbero, M. E., & Coulter, I. D. (2019). Health care costs and opioid use associated with high-impact chronic spinal pain in the United States. *Spine*, *44*(16), 1154-1161.
- Hemingway, H., Croft, P., Perel, P., Hayden, J. A., Abrams, K., Timmis, A., ... & Riley, R. D. (2013). Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. Bmj, 346.
- Holman, H., & Lorig, K. (2000). Patients as partners in managing chronic disease: partnership is a prerequisite for effective and efficient health care.
- Hoffman, B. M., Papas, R. K., Chatkoff, D. K., & Kerns, R. D. (2007). Meta-analysis of psychological interventions for chronic low back pain. *Health psychology*, *26*(1), 1.
- International Narcotics Control Board. (2002). Report of the International Narcotics Control Board for 2001.

- Johannes, C. B., Le, T. K., Zhou, X., Johnston, J. A., & Dworkin, R. H. (2010). The prevalence of chronic pain in United States adults: results of an Internet-based survey. *The Journal of Pain, 11*(11), 1230-1239.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841.
- Jennings, E. M., Okine, B. N., Roche, M., & Finn, D. P. (2014). Stress-induced hyperalgesia. *Progress in neurobiology, 121*, 1-18.
- Ji, R. R., Nackley, A., Huh, Y., Terrando, N., & Maixner, W. (2018). Neuroinflammation and central sensitization in chronic and widespread pain. *Anesthesiology, 129*(2), 343-366.
- Juniper, M., Le, T. K., & Mladsi, D. (2009). The epidemiology, economic burden, and pharmacological treatment of chronic low back pain in France, Germany, Italy, Spain and the UK: a literature-based review. *Expert opinion on pharmacotherapy, 10*(16), 2581-2592.
- Kamaleri, Y., Natvig, B., Ihlebaek, C. M., Benth, J. S., & Bruusgaard, D. (2008). Number of pain sites is associated with demographic, lifestyle, and health-related factors in the general population. European Journal of Pain, 12(6), 742-748.
- Kamaleri, Y., Natvig, B., Ihlebaek, C. M., & Bruusgaard, D. (2009). Does the number of musculoskeletal pain sites predict work disability? A 14-year prospective study. European journal of pain, 13(4), 426-430.
- Katz, J., & Seltzer, Z. E. (2009). Transition from acute to chronic postsurgical pain: risk factors and protective factors. *Expert review of neurotherapeutics, 9*(5), 723-744.
- Kaulitzki, S. (2017). Rethinking chronic pain. *Lancet, 391*, 1391-454.

- Lee, C. H., & Giuliani, F. (2019). The role of inflammation in depression and fatigue. *Frontiers in immunology*, *10*, 1696.
- Lee, J. J., Kim, H. J., Čeko, M., Park, B. Y., Lee, S. A., Park, H., ... & Woo, C. W. (2021). A neuroimaging biomarker for sustained experimental and clinical pain. *Nature medicine*, *27*(1), 174-182.
- Louati, K., & Berenbaum, F. (2015). Fatigue in chronic inflammation-a link to pain pathways. Arthritis research & therapy, 17(1), 1-10.
- Lumley, M. A., Cohen, J. L., Borszcz, G. S., Cano, A., Radcliffe, A. M., Porter, L. S., ... & Keefe, F. J. (2011). Pain and emotion: a biopsychosocial review of recent research. *Journal of clinical psychology*, *67*(9), 942-968.
- Lynch, M. E. (2011). The need for a Canadian pain strategy.
- Marchand, F., Perretti, M., & McMahon, S. B. (2005). Role of the immune system in chronic pain. *Nature Reviews Neuroscience*, *6*(7), 521-532.
- Maixner, W., Fillingim, R., Sigurdsson, A., Kincaid, S., & Silva, S. (1998). Sensitivity of patients with painful temporomandibular disorders to experimentally evoked pain: evidence for altered temporal summation of pain. *Pain*, *76*(1-2), 71-81.
- Maixner, W., Fillingim, R. B., Williams, D. A., Smith, S. B., & Slade, G. D. (2016). Overlapping chronic pain conditions: implications for diagnosis and classification. The Journal of Pain, 17(9), T93-T107.
- McCracken, L. M., & Iverson, G. L. (2001). Predicting complaints of impaired cognitive functioning in patients with chronic pain. *Journal of pain and symptom management*, *21*(5), 392-396.

- McWilliams, L. A., Cox, B. J., & Enns, M. W. (2003). Mood and anxiety disorders associated with chronic pain: an examination in a nationally representative sample. *Pain*, *106*(1-2), 127-133.
- Meerwijk, E. L., Ford, J. M., & Weiss, S. J. (2013). Brain regions associated with psychological pain: implications for a neural network and its relationship to physical pain. Brain imaging and behavior, 7(1), 1-14.
- Melzack, R., & Casey, K. L. (1968). Sensory, motivational, and central control determinants of pain: a new conceptual model. *The skin senses*, *1*, 423-43.
- Merskey, H. and Bogduk, N. (1994). Classification of Chronic Pain. 2nd Edition, IASP Task Force on Taxonomy. IASP Press, Seattle. http://www.iasp-pain.org/Education/content.aspx?ItemNumber=1698
- Meyer, T., Cooper, J., & Raspe, H. (2007). Disabling low back pain and depressive symptoms in the community-dwelling elderly: a prospective study. *Spine*, *32*(21), 2380-2386.
- Miech, R. A., Caspi, A., Moffitt, T. E., Wright, B. R. E., & Silva, P. A. (1999). Low socioeconomic status and mental disorders: A longitudinal study of selection and causation during young adulthood. *American journal of Sociology*, *104*(4), 1096-1131.
- Miech, R. A., Caspi, A., Moffitt, T. E., Wright, B. R. E., & Silva, P. A. (1999). Low socioeconomic status and mental disorders: A longitudinal study of selection and causation during young adulthood. *American journal of Sociology*, *104*(4), 1096-1131.
- Neblett, R., Cohen, H., Choi, Y., Hartzell, M. M., Williams, M., Mayer, T. G., & Gatchel, R. J. (2013). The Central Sensitization Inventory (CSI): establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample. The Journal of Pain, 14(5), 438-445.

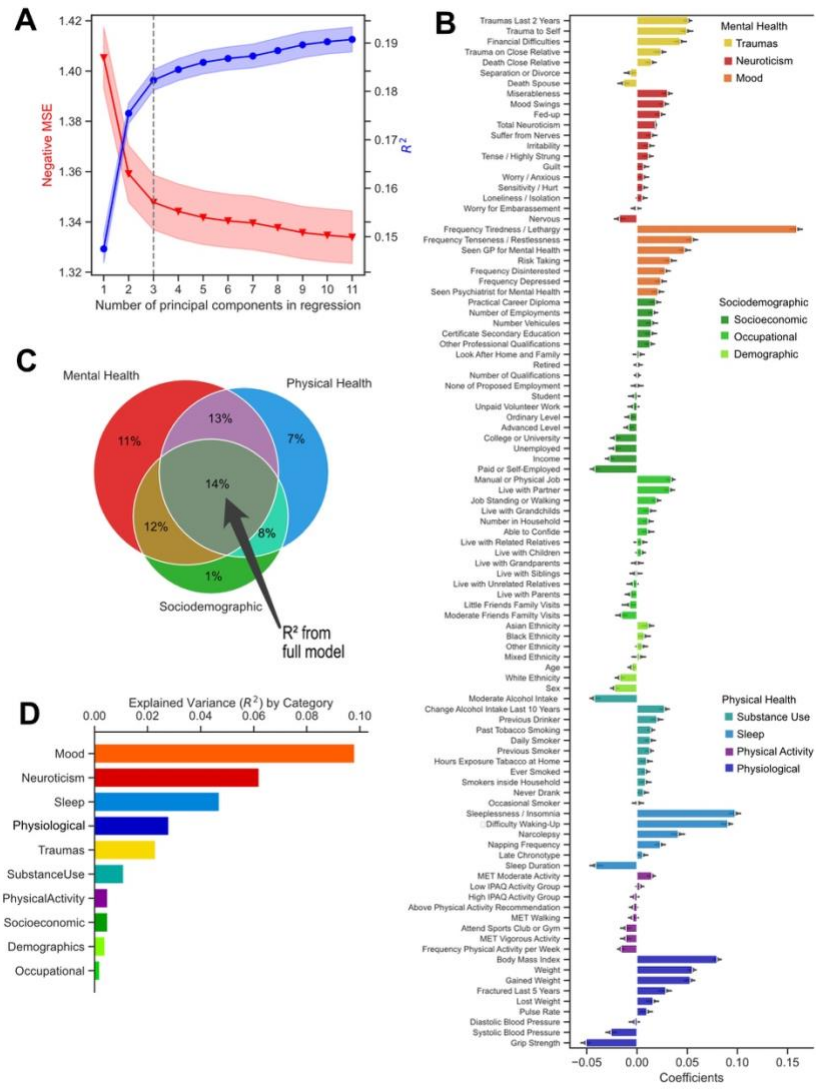
- Ohayon, M. M., & Schatzberg, A. F. (2003). Using chronic pain to predict depressive morbidity in the general population. *Archives of general psychiatry*, *60*(1), 39-47.
- Pincus, T., Kent, P., Bronfort, G., Loisel, P., Pransky, G., & Hartvigsen, J. (2013). Twenty-five years with the biopsychosocial model of low back pain—is it time to celebrate? A report from the twelfth international forum for primary care research on low back pain. *Spine*, *38*(24), 2118-2123.
- Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science, 277(5328), 968-971.
- Rice, A. S., Smith, B. H., & Blyth, F. M. (2016). Pain and the global burden of disease. *Pain*, *157*(4), 791-796.
- Rustøen, T., Wahl, A. K., Hanestad, B. R., Lerdal, A., Paul, S., & Miaskowski, C. (2005). Age and the experience of chronic pain: differences in health and quality of life among younger, middle-aged, and older adults. *The Clinical journal of pain*, *21*(6), 513-523.
- Schopflocher, D., Taenzer, P., & Jovey, R. (2011). The prevalence of chronic pain in Canada. *Pain research and management*, *16*(6), 445-450.
- Stevans, J. M., Delitto, A., Khoja, S. S., Patterson, C. G., Smith, C. N., Schneider, M. J., ... & Saper, R. B. (2021). Risk factors associated with transition from acute to chronic low back pain in US patients seeking primary care. JAMA network open, 4(2), e2037371-e2037371.
- Tan, G., Jensen, M. P., Thornby, J. I., & Shanti, B. F. (2004). Validation of the Brief Pain Inventory for chronic nonmalignant pain. The Journal of Pain, 5(2), 133-137.

- Treede, R. D., Rief, W., Barke, A., Aziz, Q., Bennett, M. I., Benoliel, R., ... & Wang, S. J. (2019). Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain*, *160*(1), 19-27.
- Vachon-Presseau, E., Roy, M., Martel, M. O., Caron, E., Marin, M. F., Chen, J., ... & Rainville, P. (2013). The stress model of chronic pain: evidence from basal cortisol and hippocampal structure and function in humans. *Brain*, *136*(3), 815-827.
- Vachon-Presseau, E., Berger, S. E., Abdullah, T. B., Huang, L., Cecchi, G. A., Griffith, J. W., ... & Apkarian, A. V. (2018). Brain and psychological determinants of placebo pill response in chronic pain patients. *Nature communications*, *9*(1), 1-15.
- Vachon-Presseau, E. (2018). Effects of stress on the corticolimbic system: implications for chronic pain. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *87*, 216-223.
- van Tilburg, M. A., Parisien, M., Boles, R. G., Drury, G. L., Smith-Voudouris, J., Verma, V., ... & Diatchenko, L. (2020). A genetic polymorphism that is associated with mitochondrial energy metabolism increases risk of fibromyalgia. Pain, 161(12), 2860-2871.
- Vlaeyen, J. W., Maher, C. G., Wiech, K., Zundert, J. V., Meloto, C. B., Diatchenko, L., . . . Linton, S. J. (2018). Low back pain. *Nature Reviews Disease Primers*, 4(55).
- Volkow, N. D., & McLellan, A. T. (2016). Opioid abuse in chronic pain—misconceptions and mitigation strategies. *New England Journal of Medicine*, *374*(13), 1253-1263.
- Yach, D., Hawkes, C., Gould, C. L., & Hofman, K. J. (2004). The global burden of chronic diseases: overcoming impediments to prevention and control. *Jama*, *291*(21), 2616-2622.

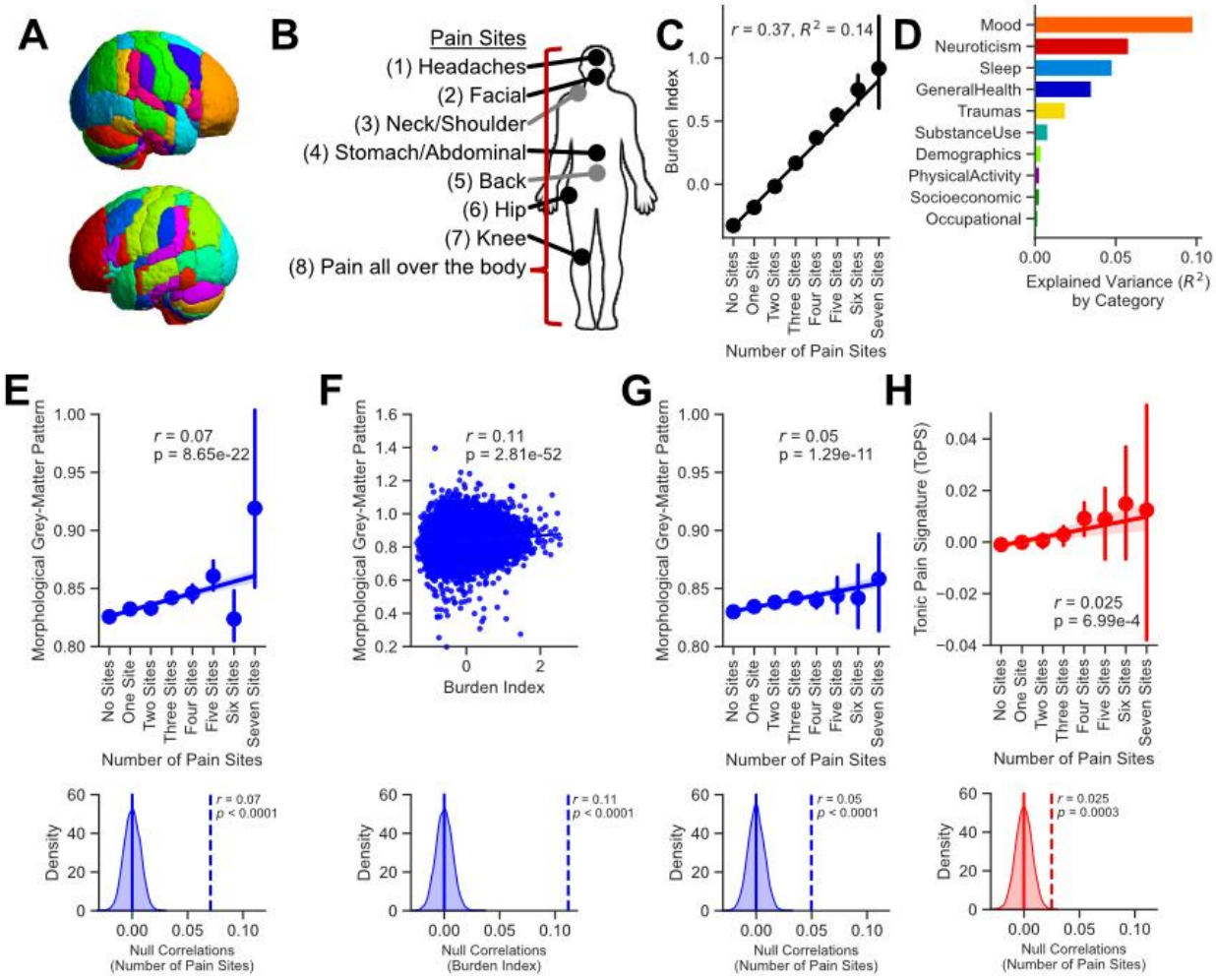
- Yarnitsky, D. (2010). Conditioned pain modulation (the diffuse noxious inhibitory control-like effect): its relevance for acute and chronic pain states. *Current Opinion in Anesthesiology*, *23*(5), 611-615.
- Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C. W., & Kross, E. (2013). An fMRI-based neurologic signature of physical pain. *New England Journal of Medicine*, *368*(15), 1388-1397.
- Watson, J., Jones, H. E., Banks, J., Whiting, P., Salisbury, C., & Hamilton, W. (2019). Use of multiple inflammatory marker tests in primary care: using Clinical Practice Research Datalink to evaluate accuracy. *British journal of general practice*, *69*(684), e462-e469.
- Wise, E. A., Price, D. D., Myers, C. D., Heft, M. W., & Robinson, M. E. (2002). Gender role expectations of pain: relationship to experimental pain perception. *Pain*, *96*(3), 335-342.
- Woolf, C. J. (1996). Windup and central sensitization are not equivalent. *Pain*, *66*(2), 105-108.
- Wolfe, F., Egloff, N., & Häuser, W. (2016). Widespread pain and low widespread pain index scores among fibromyalgia-positive cases assessed with the 2010/2011 fibromyalgia criteria. *The journal of rheumatology*, *43*(9), 1743-1748.
- Woo, C. W., Schmidt, L., Krishnan, A., Jepma, M., Roy, M., Lindquist, M. A., ... & Wager, T. D. (2017). Quantifying cerebral contributions to pain beyond nociception. *Nature communications*, *8*(1), 1-14.
- Wu, Y., Byrne, E. M., Zheng, Z., Kemper, K. E., Yengo, L., Mallett, A. J., ... & Wray, N. R. (2019). Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nature communications*, *10*(1), 1-10.

- Xu, A., Larsen, B., Henn, A., Baller, E. B., Scott, J. C., Sharma, V., ... & Satterthwaite, T. D. (2021). Brain Responses to Noxious Stimuli in Patients With Chronic Pain: A Systematic Review and Meta-analysis. JAMA network open, 4(1), e2032236-e2032236.
- Yuan, N., Chen, Y., Xia, Y., Dai, J., & Liu, C. (2019). Inflammation-related biomarkers in major psychiatric disorders: a cross-disorder assessment of reproducibility and specificity in 43 meta-analyses. Translational psychiatry, 9(1), 1-13.

List of Supplementary Figures

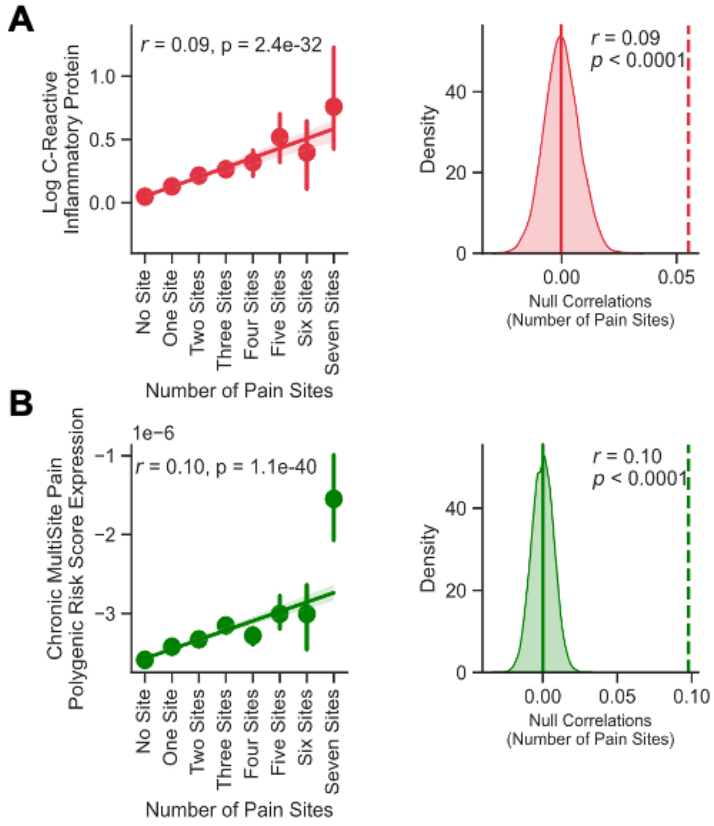


Supplementary Figure 1. Deriving the burden and its validation in the testing cohort. A. Explained variance (R2) and negative mean squared error (Negative MSE) in 10-fold cross-validation across components in the PLS model. **B.** Coefficients of the PLS model applied to obtain the burden score. **C.** Validation of the Venn diagram of the explained variance using weighted sum across dimensions of the burden in the testing cohort. **D.** Explained variance of the number of pain sites across categories in the testing cohort.



Supplementary Figure 2. Validation of the burden at the 10-year follow-up visit and its association with brain determinants. A. Grey-Matter atlas (139 regions) defined in the UK Biobank using FMRIB's Automated Segmentation Tool (FAST). **B.** Anatomical body map used to sum the number of pain sites. **C.** Validation of the association between the number of pain sites and the burden index of the testing cohort at the 10-year follow-up. **D.** Explained variance of the number of pain sites across categories in the testing cohort at the 10-year follow-up. **E-F. Training a morphological grey-matter pattern in a left-out group of participants (n = 21,000). E.** Association of the morphological grey-matter pattern with the number of pain sites in the training group (above) with 10,000 permutation tests (below) to assess the robustness of these

small effect sizes. **F.** Association with the burden index in the training group. Test of the difference between two dependent correlations revealed the Morphological Grey-Matter Pattern to be more associated with the burden index than the number of pain sites ($z = 5.3$, $p < 1 \times 10^{-5}$). **G.** Association of the number of pain sites in the testing group (association with burden in testing is shown in **Fig. 3D**.). **H.** Association of the functional Tonic Pain Signature (ToPS) with the number of pain sites in the testing group (association with burden in testing is shown in **Fig. 3E.**). The ToPS also appeared to be more associated with the burden index than the number of pain sites ($z = 5.1$, $p < 1 \times 10^{-5}$).



Supplementary Figure 3. Association of inflammation and genetic risk with widespread pain. **A.** Association of Log C-Reactive Protein with the number of pain sites (left) with 10,000 permutation tests (right). **B.** Association of Chronic Multisite Pain Polygenic Risk Score with the number of pain sites (left) with permutation testing.