Universiteit
Leiden
The Netherlands

Promoting early recognition of persistent somatic symptoms in primary care

Kitselaar, W.M.

Citation
Kitselaar, W. M. (2023, June 27). *Promoting early recognition of persistent somatic symptoms in primary care*. Retrieved from https://hdl.handle.net/1887/3628068

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3628068

Note: To cite this publication please use the final published version (if applicable).

Chapter 7

General discussion

General discussion

The studies presented in this thesis intend to provide comprehensive insight into possibilities to predict persistent somatic symptoms (PSS) and how to reuse routine primary care data extracted from electronic medical records (EMR) for this purpose. The included studies are aimed at supporting and promoting early identification of patients with PSS and facilitate early intervention. Based on the findings presented in this thesis it is possible to identify patients with PSS with moderate to high accuracy at least two years prior to diagnosis. Moreover, the compiled studies provide direction for more precise data-based prediction. In the current chapter, the study findings are outlined and discussed in relation to each other. First, the main findings of the studies in this thesis will be listed, to form an overview of key messages. Thereafter, the main issues regarding early identification of PSS based on routine primary care data are discussed.

Main findings

1. Predictors of PSS onset are multifold and cannot be reduced to a single domain of the dynamic biopsychosocial model (**chapter 2**). Exploring complaints from different domains is therefore paramount to improve care for patients with PSS.
2. The majority of general practitioners (GPs) reports to require additional tools or other support for classification of PSS and consultation with patients at risk of PSS (**chapter 3**).
3. GPs use a wide range of methods for registration of PSS in their electronic medical records (EMRs; self-reported results from **chapter 3** are supported by observation in **chapter 4**). These registration methods can differ greatly between GPs.
4. Identifying patients within the broad spectrum of PSS in routine primary care data (**chapter 4**) is possible by using a combination of methods (including clinical codes of PSS-syndromes, episode descriptions, and recorded outcomes of screening tools).
5. Mental health-related registrations in routine primary care data, such as psychological ICPC-codes, referrals to mental health care, and psychopharmacological prescriptions, can be adequately re-used to predict diagnosis of common PSS syndromes (IBS, FM, and CFS) (**chapter 5**).

6. There are both overarching and distinct mental health registrations for PSS syndromes (IBS, CFS, and FM) in routine primary care data (**chapter 5**).
7. Patients with medium to high risk of PSS can be identified in routine primary care data two to seven years prior to diagnosis (**chapter 6**). Highly accurate recognition of risk of PSS appears to be impeded by general registration irregularities and missing data.
8. Theory- and data-driven methods show similar performance in the ability to predict PSS diagnosis based on extracted and anonymized routine primary care data (**chapter 6**). Benefits of using more complex, time-consuming methods may therefore be limited.

Selecting patients with PSS in routine primary care data

The primary challenge towards predictive modeling of PSS based on routine care data is the lack of a gold standard method of classifying PSS. As demonstrated in **chapter 3**, GPs reportedly use a variety of methods to register PSS and an unambiguous clinical code for PSS is not available. This does not only limit the reusability of routine primary care data for PSS research purposes but is also problematic for many GPs in daily clinical practice. Moreover, besides the problematic lack of a clinical code, GPs report to have difficulties in classifying and identifying PSS, which is further corroborated by other studies [1-3] and studies showing a significant diagnostic delay in PSS-syndromes.[4-6] Consequently, it would be undesirable to solely rely on GPs or the clinical codes they use for selecting patients with PSS in the routine care data. Previous EMR studies used a variety of approaches to select patients with PSS, including inquiring GPs about specific patients and a combination of clinical (symptom) codes with exclusion of comorbid mental or medical conditions.[7-11] Since research shows that patients with PSS often have comorbid conditions [12,13] (see also, **chapter 4**) the latter is undesirable. Based on this knowledge the studies in **chapter 3** and **chapter 4** were directed at developing a data-based classifier for PSS.

In **chapter 3** registration behavior was gathered via a survey that included ecologically valid methods of inquiry. Based on the findings, the most viable methods were selected and effectiveness was explored in **chapter 4.** This resulted in a combination of methods

that identifies a group of patients that approaches the prevalence rate of PSS in the general population. The identification consists of clinical coding of PSS-syndromes, unstructured episode descriptions of PSS, and recorded outcomes of screening questionnaires. Results show that the use of a combination of these methods is crucial, since GPs vary in their approach towards registration of complaints and disorders. In all, because the use of additional screening questionnaires or the exclusion of comorbid chronic medical and mental conditions are not required, a data-based classifier of PSS can facilitate the use of large routine primary care databases for research on the broad spectrum of PSS. Still, this approach towards identification of PSS in routine care data is cumbersome and reusability of EMR data can be facilitated by incorporating a single general code for PSS in classification systems such as the ICPC or ICD. Such a code could also benefit GPs, since they report that the lack of a single unambiguous code is problematic for them (**chapter 3**). However, such a code should be accompanied with adequate universally accepted guidelines and definitions, which has been proven to be a large challenge in the international field of PSS.[14]

The challenge of creating a multidomain risk profile using routine primary care data

In **chapter 2**, an extensive review of cohort studies is presented mapping multidomain predictors of PSS. This was executed in order to evaluate which data in primary care EMRs would be relevant to include in our predictive model. This review shows that although risk factors from the biomedical domain are currently dominant in the literature, factors from all domains of the dynamic biopsychosocial model (i.e., biological, psychological, interpersonal, contextual, and health behavior) have shown to be significant predictors of PSS-onset. Still, investigations of routine care data showed a predominance of biomedical registrations and predictors, which impedes the creation of a fully multidomain risk profile using currently available routine primary care data. This is especially evident in the prediction model presented in **chapter 6** in which all available primary care data was utilized (i.e., the full model) and PSS was predicted at least two years prior to diagnosis. Based on LASSO regression (a machine learning technique that is able to handle large amounts of data and incorporates predictor reduction), the EMR-based predictors mainly consist of biomedical variables, with limited representation of

psychosocial factors. However, in contrast, the prediction models presented in **chapter 5** based on the same data and similar analysis shows that PSS syndrome (i.e., IBS, CFS, and FM) diagnosis can be predicted with high accuracy based on mental health-related registrations. In contrast to **chapter 6**, this study only investigated mental health-related registrations and candidate predictors were constructed based on registration directly prior to PSS diagnosis. Based on these findings it is hypothesized that psychosocial problems are under registered and/or recognized by GPs, unless the patient presents with persistent problems. This may also be related to findings from other studies that indicate that patients may not readily visit their GP with mental health problems (for instance because of cultural reasons).[15-19] Thus, arguably, in routine primary care data registrations of psychosocial problems that are present prior to (severe or recurring) somatic symptom onset could be missing in the data. Consequently, the somewhat limited performance of the models in **chapter 6** could be related to high levels of missing data in the psychological and social domains or limited consultations prior to PSS diagnosis.

Further considerations for a data-based early classification of PSS

Although lacking registrations of mental health and psychosocial problems seems a plausible explanation for the differences in performance of the models presented in **chapter 5** and **chapter 6**, there are some other explanations that should be considered. First, the use of a 2-year prediction gap in **chapter 6**, compared to no prediction gap in **chapter 5** (i.e., gathering candidate predictors directly before PSS index date) is likely to affect the results. EMRs may contain limited relevant data, because data is dependent of patients visiting the GP and GP's registration behavior. On the other hand, since PSS is often accompanied by an accumulation of recurring complaints, the repeated out-patient visits or number of symptoms should be indicative of emerging PSS. In **chapter 6**, where LASSO regression indicated that consultation frequency may be explained by other variables, valuable evidence is given to explain the relationship between consultation frequency and PSS, as indicated by a large body of research. That is, factors such as repeated imaging referrals, and multiple pain sites and symptoms, are predictors that may explain increased consultation frequency. This sheds light on the importance of

correctly interpreting some of these behavior- or context-related predictors (such as, HCU and unemployment). Thus, while these are also found in the large body of literature on predictors of PSS (**chapter 2)**, it should be noted that these predictors are rather a consequence of the accumulated disease burden than an actual predictor leading to PSS. Specification of predictors that may lead to undesirable contextual or behavioral aberration may be especially useful to distinguish PSS from other disorders that are accompanied by high consultation frequency. Thus, focusing primarily on consultation frequency may limit specificity of PSS prediction and identifying related latent factors may be necessary to increase predictive accuracy. Second, the broader definition of PSS in **chapter 6** (investigating the broad spectrum of PSS) compared to **chapter 5** (having unique PSS-syndrome classifications as outcomes) may affect predictive accuracy, because of increasing heterogeneity between patients. For instance, patients with IBS and FM display distinct somatic symptom presentations (bowel problems and widespread musculoskeletal pain, respectively). Furthermore, research indicates that the duration of diagnostic delays may be different between PSS subtypes.[4-6] This is further corroborated by findings from **chapter 3**, where GPs report more competency and willingness to diagnose IBS compared to CFS and FM (for which they are more likely to refer to specialist care). Different durations in diagnostic delay may also affect heterogeneity between patients with PSS. **Chapter 5** corroborates this finding by identifying a difference in predictive accuracy between the three prediction models for IBS, FM, and CFS, and some discrepancies in predictors (although the latter is mainly the case for CFS). Since diagnostic delays can cause a large number of problems, including inducing psychological problems [20] and complicating the physician-patient relationship,[16,21,22] it is likely this also affects EMR registrations.

Promoting integrative care for patients with PSS

The limited registration of factors beyond the biomedical domain appears to have a twofold effect on PSS classification, both affecting physician-based classification and data driven classification. While research clearly indicates that the origin of PSS is multidomain in nature (**chapter 2**), the predictors derived from the early prediction model in **chapter 6** indicate that GPs do not make use of an integrative understanding of

the patients' problems (i.e., relating combined biological, behavioral, psychological, interpersonal, and contextual factors to the health of the patient). On the other hand, the lack of information also limits the performance of the early prediction model. Although, based on these findings, it could be argued that routine primary care data is not suitable for predictive modeling of PSS, the moderate accuracy of the model may have potential to break the cycle of under-recognition by physician and algorithm. Especially because overloaded work hours and prioritization of potential life-threatening diseases, limits the options of GPs to inquire about psychosocial factors – irrespective of having the appropriate training or not. Therefore, a simple data-based clinical decision rule may have the potential to sift or identify patients that require and may benefit from an integrative approach. This is in line, with the current trend towards proactive population health management: identify patients at risk for adverse health events like ineffective and counterproductive specialist referrals and expose them to less invasive interventions.[23,24] The implementation of a clinical decision rule into EMR software, flagging patients with increased risk of developing PSS, would enable (earlier) referral to interdisciplinary health care resources for further assessment. For GPs, earlier referral could reduce time investment and refocusing on the exclusion of possible life-threatening pathophysiology.

The contributions of machine learning techniques

In this thesis, a variety of statistical methods were used, including machine learning techniques (**chapter 6**). Previous research has shown that temporal pattern mining and relative grounding of lab results of structured data could be effectively employed to improve model performance and reconfirm and identify new predictors.[25,26] Therefore, to increase the likelihood of finding predictors that were thus far unidentified, both techniques were employed. Although the machine learning techniques did not improve the performance of the model, it did identify some known (i.e., referral to radiology) and novel plausible predictors (i.e., stable lab results for lymphocytes, thyroid, and systolic blood pressure), which validates the effectivity of the methods. Finally, the predictions were modeled using logistic LASSO regression as a form of supervised machine learning. While studies show that LASSO logistic regression generally performs well for predictive

modeling,[26,27] recent studies show that more advanced machine learning techniques may result in better performing models.[28] However, compared to more advanced machine learning algorithms in which logical explanations of models are often lost due to the black box phenomenon (i.e., lack of interpretability), regression is generally seen as more comprehensible. Although research has been directed at improving the interpretability of more advance machine learning techniques,(see for example [29]) regression is generally deemed more suitable for the use in clinical populations.[30-34]

Methodological considerations

The results and the implications of this thesis should be viewed in the light of several strengths and limitations. With the use of routine primary care data come both great opportunities and challenges. In recent years the increased quality of routine primary care data (i.e., both registration quality and technical advances) has provided many opportunities for scientific research. The data is generally low-cost and provides relatively easy access to rich, ecologically valid, longitudinal data from large populations.[35] On the other hand, registration of (especially) non-biomedical health information may be inconsistent (**chapter 3-6**) and depends highly on the patient's decision to visit the GP with a particular problem and the GP's or practice personnel's registration behavior. One of the major challenges of reusing routine care data is the methodological handling of missing data.[36-38] While data collected in a standardized way is generally missing at random and imputation techniques may be safely used, imputation for routine care data is less straightforward. In routine care data, it is common practice to assume that "missing data" means a factor is not present.[37,39] However, this is disputed by findings from **chapter 3-6**, that imply that especially data beyond the biomedical domain is likely to be sparsely recorded. Due to these considerations, imputation was not used for the reported studies.

A major strength of this thesis is the extensive research towards the aim of early identification of patients with PSS based on routine primary care data. The steps taken in this process highlights several factors that should be considered for future studies. First, **chapter 5** shows that there may be differences in predictors for PSS-subtypes and that the performance of all models in **chapter 6** (which includes a broad spectrum of PSS) is

markedly lower than the models in **chapter 5**. This could indicate that subtype-specific differences in predictors or registrations may impact the performance of the models. Second, although the survey in GPs (**chapter 3**) provided elaborate insight into registration behavior of Dutch GPs regarding PSS as an outcome, the survey lacked information on other outcomes (e.g., the registration of possible psychosocial predictors) that may be relevant to PSS. It should however be noted that, the results of the survey may be quite generalizable, even to countries using other classification systems, such as the ICD-11. This is indicated by studies that show that German GPs, who operate an ICD coded system, also have difficulties in registration and classification of PSS.[2,3,40] Finally, in hindsight, preliminary investigations should also have included investigations of differences in diagnostic delays between PSS-subtypes. The current literature does not contain information on syndrome specific diagnostic delays in primary care, which would be needed to evaluate how this may affect looking at the broad spectrum of PSS.

In all, the investigations in this thesis increased the validity of the defined candidate predictors and the outcome employed for the final modeling (**chapter 6**). Even so, due to the nature of the data, misclassification is inevitable and a major limitation to this research. Firstly, the outcome was not externally validated, and the prevalence rate was somewhat lower than prevalence in the general population. Second, candidate predictors in all models were compiled based on data with high levels of (nonrandom) missings. Although the applied design aimed to control for registration irregularities by compiling candidate predictors based on a variety of sources (such as ICPC, ATC, referrals, and lab results), success was limited for the desired early prediction model. Nonetheless, these results do reflect best the current clinical practice, since the data available is the data available to the GP, which increases the generalizability of the results.

Clinical and societal implications

As described above, the implication of the clinical decision rule that can be derived from **chapter 6** has the potential to promote proactive population health management. Patients at risk for adverse health events (e.g., ineffective and counterproductive specialist referrals) are identified and exposed to less invasive investigations and

interventions). As such, the clinical decision rule indicates which patients require and would potentially benefit from an integrative care approach. Due to the multidomain origin of PSS, the implementation of an integrative approach is expected to result in early identification of PSS. Furthermore, the implementation of an integrative approach for the at-risk population could potentially impact the way physicians perceive PSS, improve consultations, and improve understanding between physician and patient. Studies have shown that an integrated approach to health has many benefits for patients, physicians, and society, for instance by increasing perceived quality of care and increasing survival rates in cancer.[41-43]

Chapter 3 shows that many GPs report a need for more support in the diagnosis and classification of PSS. This indicates that GP training should be improved with more attention to consultations and classifications related to PSS. Improvements in GP training should for instance include training in communication skills that facilitate a broader integrative inquiry of problems.[21,22,44,45] Additionally, since PSS have a problematic history of being burdensome to clinical care, reframing of PSS is desirable. Johansen et al., 2017 makes a strong case for reframing using experience-based knowledge from senior GPs and integrate models from different disciplines.[1]

Chapter 3-6 show a reported and observed lack of an unambiguous coding scheme for PSS. The simplest example to this ambiguity is the lack of a singular accepted clinical code for PSS. To optimize the utility of EMR data for clinical practice and research, PSS requires more globally accepted uniform coding schemes (that will increase interrater reliability). While a simple way towards the development of such a scheme has proven difficult in the past, the collaboration between groups of experts such as EURONET-SOMA, ICPC and ICD workgroups may be necessary.

In **chapter 5** (i.e., prediction of the three common PSS syndromes IBS, FM, and CFS, shortly prior to diagnosis) the algorithm's performance is sufficient for clinical implementation. Implementation of the algorithm in the GP's EMR software could support the GP in more prompt classification and treatment. Especially for FM, which has marked long diagnostic delays,[4] implementation could impact patients greatly, possibly

leading to more proactive intervention and consequential lessening of the disease burden.

In sum, the implementation of integrated care, an unambiguous coding scheme, and support for GPs (including but not limited to a clinical decision rule) is needed to improve care for patients with PSS and decrease the burden of PSS on the health care system.

Future directions for data-based early recognition of PSS in primary care

The findings of this thesis provide a road map towards early identification of patients with PSS in primary care using data from EMRs. While the findings are promising, at present concise data-based identification of PSS diagnoses is limited. To improve predictive modeling for PSS with the current state of data, some promising approaches remain. Firstly, optimal utilization of unstructured (i.e., free text) data could possibly improve existing models. For the present study there was only limited accessibility and since GPs may be more prone to unstructured registration of factors beyond the medical domain (**chapter 3**). Natural language processing may assist in changing such data in quantifiable factors.[46] Second, although previous efforts showed limited success,[47] advances in the field of semantic enrichment (i.e., targeting irregularities in registrations), may improve future models. Finally, since different PSS syndromes have unique lengths of diagnostic delay, using different timelines for candidate predictor selection may enhance heterogeneity of predictor data. Future research could also employ simple data-based methods to identify patients at medium to high risk of PSS and test whether this, in combination with widely available screening questionnaires (i.e., 4DSQ, SSD-12, PHQ-15) can support the GP in early recognition of high PSS risk. Such a wide classification could be beneficial for the current trend towards more preventative health care and proactive population health management,[48,49] since measurable and controllable problems may be especially prevalent in patients with an elevated risk of PSS. Ultimately, the goal would be to improve the GPs understanding of the patient from different perspectives – even beyond the better known two-track policy (i.e., exploring both physical and mental health problems), rather towards a multi-track policy of integrated care (i.e., exploring problems from a biopsychosocial perspective and beyond). Thus, based on a simple algorithm implemented in the GP's EMR, the GP would

have a clearer direction for what patients the integrated approach may be most important. As a results, this may alleviate both the burden on the patient as well as increasing long-term time efficiency for primary care.

In conclusion

Patients with somatic symptoms generally visit their GP to find a cause and treatment for their symptoms. Most GPs consider identifying a biomedical cause, an appropriate treatment, and if needed adequate referral of patients to secondary health care as their primary job. However, some somatic symptoms may be caused by a complex interplay between multidomain factors through which it is not possible to find a single biomedical cause for symptoms. Problems arise when symptoms persist or are aggravated without a well understood biomedical cause in line with the presentation (i.e., in the case of PSS). Besides obvious burden of disease on patient and health care costs related to repeated consultations and testing, this also puts a strain on GPs who reportedly do not have adequate training and tools to specifically identify patients with PSS early. Identifying early on whose somatic symptoms may not be explainable by a biomedical pathology but by problems from multiple biopsychosocial domains is key to improve care for patients with PSS. These patients may benefit from an integrated treatment approach (i.e., targeting a combination of biological, psychological, social, interpersonal, and contextual factors that influence the patients' health), also if they have identifiable comorbidity. This thesis shows that relatively simple data-based algorithms may help to identify patients at risk of PSS at an earlier stage. This suggests that a data-based clinical decision algorithm can provide support for GPs in early identification of PSS. With early identification, GPs can possibly direct patients at risk of PSS on track for an integrated treatment approach that may reduce the disease burden of both patient and the health care system.

References

1 Johansen ML, Risor MB. What is the problem with medically unexplained symptoms for GPs? A meta-synthesis of qualitative studies. Patient Educ Couns 2017; 100: 647–54.

2 Lehmann M, Pohontsch NJ, Zimmermann T, Scherer M, Löwe B. Diagnostic and treatment barriers to persistent somatic symptoms in primary care - representative survey with physicians. BMC Fam Pract 2021; 22. DOI:10.1186/S12875-021-01397-W.

3 Murray AM, Toussaint A, Althaus A, Löwe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: A systematic review of barriers to diagnosis in primary care. J Psychosom Res 2016; 80: 1–10.

4 Gendelman O, Amital H, Bar-On Y, et al. Time to diagnosis of fibromyalgia and factors associated with delayed diagnosis in primary care. Best Pract Res Clin Rheumatol 2018; 32: 489–99.

5 Comiskey C, Larkan F. A national cross-sectional survey of diagnosed sufferers of myalgic encephalomyelitis/chronic fatigue syndrome: pathways to diagnosis, changes in quality of life and service priorities. Ir J Med Sci 2010; 179: 501–5.

6 Varenna M, Crotti C, Ughi N, Zucchi F, Caporali R. Determinants of Diagnostic Delay in Complex Regional Pain Syndrome Type 1: An Observational Study of 180 Consecutive New Cases. J Clin Rheumatol 2021; 27: E491–5.

7 Smith RC, Gardiner JC, Armatti S, et al. Screening for high utilizing somatizing patients using a prediction rule derived from the management information system of an HMO: A preliminary study. Med Care 2001; 39: 968–78.

8 Morriss R, Lindson N, Coupland C, Dex G, Avery A. Estimating the prevalence of medically unexplained symptoms from primary care records. 2012. DOI:10.1016/j.puhe.2012.05.008.

9 den Boeft M, van der Wouden JC, Rydell-Lexmond TR, de Wit NJ, van der Horst HE, Numans ME. Identifying patients with medically unexplained physical symptoms in electronic medical records in primary care: A validation study. BMC Fam Pract 2014; 15: 109.

10 van Westrienen PE, Pisters MF, Veenhof C, de Wit NJ. Identification of patients with moderate medically unexplained physical symptoms in primary care with a five years follow-up. BMC Fam Pract 2019; 20. DOI:10.1186/S12875-019-0950-7.

11 Sitnikova K, Pret-Oskam R, Dijkstra-Kersten SMA, et al. Management of patients with persistent medically unexplained symptoms: A descriptive study. BMC Fam Pract 2018; 19. DOI:10.1186/s12875-018-0791-9.

12 Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Arch Gen Psychiatry 2005; 62: 903–10.

13 Waal MWM de, Arnold IA, Eekhof JAH, Hemert AM van. Somatoform disorders in general practice: Prevalence, functional impairment and comorbidity with anxiety and depressive disorders. The British Journal of Psychiatry 2004; 184: 470–6.

14 Weigel A, Hüsing P, Kohlmann S, et al. A European research network to improve diagnosis, treatment and care for patients with persistent somatic symptoms: Work report of the EURONET-SOMA conference series. In: Journal of Psychosomatic Research. Elsevier Inc., 2017: 136–8.

15 Salmon P, Dowrick CF, Ring A, Humphris GM. Voiced but unheard agendas: qualitative analysis of the psychosocial cues that patients with unexplained symptoms present to general practitioners. The British Journal of General Practice 2004; 54: 171.

16 Peters S, Rogers A, Salmon P, et al. What do patients choose to tell their doctors? Qualitative analysis of potential barriers to reattributing medically unexplained symptoms. J Gen Intern Med 2009; 24: 443–9.

17 Vasiliadis HM, Tempier R, Lesage A, Kates N. General practice and mental health care: Determinants of outpatient service use. Canadian Journal of Psychiatry 2009; 54: 468–76.

18 Fleury MJ, Grenier G, Bamvita JM, Caron J. Determinants and patterns of service utilization and recourse to professionals for mental health reasons. BMC Health Serv Res 2014; 14. DOI:10.1186/1472-6963-14-161.

19 Löwe B, Gerloff C. Functional Somatic Symptoms Across Cultures: Perceptual and Health Care Issues. Psychosom Med 2018; 80: 412–5.

20 Zhang Y, Liang D, Jiang R, et al. Clinical, psychological features and quality of life of fibromyalgia patients: a cross-sectional study of Chinese sample. Clin Rheumatol 2018; 37: 527–37.

21 Salmon P. Conflict, collusion or collaboration in consultations about medically unexplained symptoms: The need for a curriculum of medical explanation. Patient Educ Couns 2007; 67: 246–54.

22 Weiland A, Blankenstein AH, van Saase JLCM, et al. Training Medical Specialists to Communicate Better with Patients with Medically Unexplained Physical Symptoms (MUPS). A Randomized, Controlled Trial. PLoS One 2015; 10. DOI:10.1371/JOURNAL.PONE.0138342.

23 Bradley PS. Implications of Big Data Analytics on Population Health Management. Big Data 2013; 1: 152–9.

24 Lobach DF, Kawamoto K, Anstrom KJ, et al. Proactive Population Health Management in the Context of a Regional Health Information Exchange Using Standards-Based Decision Support. AMIA Annual Symposium Proceedings 2007; 2007: 473.

25 Richter AN, Khoshgoftaar TM. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. Artif Intell Med 2018; 90: 1–14.

26 Kop R, Hoogendoorn M, Teije A ten, et al. Predictive modeling of colorectal cancer using a dedicated pre-processing pipeline on routine electronic medical records. Comput Biol Med 2016; 76: 30–8.

27 Sarraju A, Ward A, Chung S, Li J, Scheinker D, Rodríguez F. Machine learning approaches improve risk stratification for secondary cardiovascular disease prevention in multiethnic patients. Open Heart 2021; 8. DOI:10.1136/OPENHRT-2021-001802.

28 Półchłopek O, Koning NR, Büchner FL, Crone MR, Numans ME, Hoogendoorn M. Quantitative and temporal approach to utilising electronic medical records from general practices in mental health prediction. Comput Biol Med 2020; 125: 103973.

29 Wongvibulsin S, Wu KC, Zeger SL. Improving Clinical Translation of Machine Learning Approaches Through Clinician-Tailored Visual Displays of Black Box Algorithms: Development and Validation. JMIR Med Inform 2020; 8. DOI:10.2196/15791.

30 Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA - Journal of the American Medical Association 2017; 318: 517–8.

31 Maddox TM, Rumsfeld JS, Payne PRO. Questions for Artificial Intelligence in Health Care. JAMA - Journal of the American Medical Association 2019; 321: 31–2.

32 Shortliffe EH, Sepúlveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. JAMA - Journal of the American Medical Association 2018; 320: 2199–200.

33 Lipton ZC. The Doctor Just Won't Accept That! 2017; published online Nov 19. http://arxiv.org/abs/1711.08037 (accessed Sept 7, 2022).

34 Beam AL, Kohane IS. Big data and machine learning in health care. JAMA - Journal of the American Medical Association 2018; 319: 1317–8.

35 Casey JA, Schwartz BS, Stewart WF, Adler NE. Electronic Health Records and Population Health Research. Front Public Health Serv Syst Res 2016; 5: 15–22.

36 Callahan A, Shah NH, Chen JH. Research and Reporting Considerations for Observational Studies Using Electronic Health Record Data. Ann Intern Med 2020; 172: S79–84.

37 Wells BJ, Nowacki AS, Chagin K, Kattan MW. Strategies for handling missing data in electronic health record derived data. EGEMS (Wash DC) 2013; 1: 7.

38 Mack C, Su Z, Westreich D. Types of Missing Data. 2018. https://www.ncbi.nlm.nih.gov/books/NBK493614/ (accessed Sept 7, 2022).

39 Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006; 59: 1087–91.

40 Pohontsch NJ, Zimmermann T, Jonas C, Lehmann M, Löwe B, Scherer M. Coding of medically unexplained symptoms and somatoform disorders by general practitioners – an exploratory focus group study. BMC Fam Pract 2018; 19: 129.

41 Baxter S, Johnson M, Chambers D, Sutton A, Goyder E, Booth A. The effects of integrated care: a systematic review of UK and international evidence. BMC Health Serv Res 2018; 18. DOI:10.1186/S12913-018-3161-3.

42 Fulton JJ, LeBlanc TW, Cutson TM, et al. Integrated outpatient palliative care for patients with advanced cancer: A systematic review and meta-analysis. Palliat Med 2019; 33: 123–34.

43 Coates D, Coppleson D, Schmied V. Integrated physical and mental healthcare: an overview of models and their evaluation findings. Int J Evid Based Healthc 2020; 18: 38–57.

44 Houwen J, Lucassen PLBJ, Stappers HW, et al. How to learn skilled communication in primary care MUS consultations: a focus group study. Scand J Prim Health Care 2021; 39: 101–10.

45 Houwen J, Lucassen PLBJ, Stappers HW, Assendelft WJJ, Olde Hartman TC, van Dulmen S. Improving GP communication in consultations on medically unexplained symptoms: a qualitative interview study with patients in primary care. Br J Gen Pract 2017; 67: e716–23.

46 Edgcomb JB, Zima B. Machine Learning, Natural Language Processing, and the Electronic Health Record: Innovations in Mental Health Services Research. Psychiatr Serv 2019; 70: 346–9.

47 Hoogendoorn M, Szolovits P, Moons LMG, Numans ME. Utilizing uncoded consultation notes from electronic medical records for predictive modeling of colorectal cancer. Artif Intell Med 2016; 69: 53–61.

48 Williams AA. The Next Step in Integrated Care: Universal Primary Mental Health Providers. J Clin Psychol Med Settings 2020; 27: 115–26.

49 Prince M, Patel V, Saxena S, et al. No health without mental health. Lancet 2007; 370: 859–77.