

Editorial Office Notes:

RES-19-970.R2

ORIGINAL ARTICLE

Received 10 November 2019

Invited to revise 24 January and 11 March 2020

Revised 2 and 19 March 2020

Accepted 14 April 2020

Associate Editor: Michael Keane

Senior Editor: Lutz Beckert

Publication fee waiver: NO

Volume: 25

Opportunities to diagnose fibrotic lung diseases in routine care: A primary care cohort study

Mark G. Jones*¹, Christopher R.T. Hillyar*², Anjan Nibber³, Alison Chisholm⁴, Andrew Wilson⁵, Toby M. Maher^{6,7}, Alan Kaplan⁸, David Price^{9,10}, Simon Walsh⁷, Luca Richeldi^{1,11} on behalf of the Respiratory Effectiveness Group's Interstitial Lung Disease Working Group

1. National Institute for Health Research Biomedical Research Centre and Clinical and Experimental Sciences, University of Southampton, Southampton, United Kingdom
2. Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom
3. University of Oxford Medical School, Oxford, United Kingdom
4. Respiratory Effectiveness Group, Cambridge, United Kingdom
5. Norwich Medical School, University of East Anglia, Norwich, UK
6. Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK
7. Fibrosis Research Group, National Heart and Lung Institute, Imperial College London, London, UK
8. Department of Family and Community Medicine, University of Toronto, Ontario, Canada
9. Centre of Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom
10. Observational and Pragmatic Research Institute, Singapore
11. Unità Operativa Complessa di Pneumologia, Università Cattolica del Sacro Cuore, Fondazione Policlinico A. Gemelli, Rome, Italy

* M.G.J. and C.R.T.H. contributed equally to this manuscript.

Correspondence:

Name: Professor Luca Richeldi

Address: Unità Operativa Complessa di Pneumologia,

Università Cattolica del Sacro Cuore,

Fondazione Policlinico A. Gemelli,

Rome, Italy

Email: luca.richeldi@policlinicogemelli.it

SUMMARY AT A GLANCE

We analysed a primary care clinical cohort database to investigate respiratory symptoms and healthcare use in the 10 years prior to a diagnosis of pulmonary fibrosis. Utilisation progressively increased in the years prior to diagnosis, suggesting multiple opportunities for diagnosis at an earlier stage.

ABSTRACT

Background and objective: Temporal trends of healthcare use in the period before a diagnosis of pulmonary fibrosis are poorly understood. We investigated trends in respiratory symptoms and lower respiratory healthcare resource utilisation (HRU) in the 10 years prior to diagnosis.

Methods: We analysed a primary care clinical cohort database (UK Optimum Patient Care Research Database) and assessed patients aged ≥ 40 years who had an electronically coded diagnosis of pulmonary fibrosis between 2005–2015 and a minimum two years continuous medical records prior to diagnosis. Exclusion criteria consisted of electronic codes for recognised causes of pulmonary fibrosis such as CTD, sarcoidosis or extrinsic allergic alveolitis.

Results: Data for 2223 patients were assessed. Over the 10 years prior to diagnosis of pulmonary fibrosis there was a progressive increase in HRU across multiple lower respiratory (LR)-related domains. Five years before diagnosis, 18% of patients had multiple healthcare contacts for LR

complaints; this increased to 79% in the year before diagnosis, with 38% of patients having five or more healthcare contacts.

Conclusion: There are opportunities to diagnose pulmonary fibrosis at an earlier stage; research into case-finding algorithms and strategies to educate primary care physicians is required.

Short title: Earlier diagnosis of pulmonary fibrosis

Key Words: Pulmonary Fibrosis, Cough, Respiratory function tests, Clinical Epidemiology, Clinical Respiratory Medicine

INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a form of chronic, progressive fibrotic lung disease. It affects 14–43 people per 100,000, with a mean age at presentation of 66 years [1–4]. Median post-diagnostic survival of IPF is 2–5 years, with a 5-year survival of 20% [3,5]. There is a lack of public awareness, yet IPF kills more people per year than leukaemia or ovarian cancer, with the number affected increasing [6].

Historically, treatment of IPF was limited to symptom management and palliation, potentially providing little motivation for earlier diagnosis. However, even without approved pharmacological treatment, delayed access to specialist interstitial lung disease (ILD) services increases risk of death [7]. Recent approval of two effective anti-fibrotic treatments (nintedanib and pirfenidone), which slow disease progression, heralds a new era for patients with IPF [8,9]. Thus, the importance of an early and accurate diagnosis is clear [10].

While early diagnosis and treatment of IPF is now widely advocated [11,12], how this can be achieved remains less clear. Precision medicine and genomic techniques are being widely researched to identify blood- or lung-specific biomarkers to enhance IPF diagnosis [13]. However, such methodologies are available only in specialist centres, access to which first requires suspicion, or diagnosis, of pulmonary fibrosis. A challenge to the specialist referral pathway is the limited understanding of the natural history of pulmonary fibrosis; it is thought patients may be symptomatic for five years before a formal diagnosis is made [14].

As all patients with pulmonary fibrosis (including IPF) in the United Kingdom (UK) present first in primary care, primary care medical records provide an important resource for understanding temporal patterns of healthcare system utilisation (HRU) prior to diagnosis. This study assessed trends in HRU over 10 years prior to diagnosis of pulmonary fibrosis to inform potential strategies for earlier identification.

METHODS

Data source and study approvals

The Optimum Patient Care Research Database (OPCRD) contains anonymised, longitudinal medical records for patients registered at primary care practices across the UK [17]. It includes demographic, lifestyle, diagnostic and HRU data recorded from primary and secondary care. At the time of the study it comprised records for approximately 2.5 million patients registered across 525 UK-based primary care practices.

The OPCRD is approved by the Health Research Authority of the UK NHS for clinical research use (REC reference: 15/EM/0150). Access was approved by the OPCRD's Anonymised Data Ethics and Protocol Transparency Committee (approval code, ADEPT0616). The study protocol was developed by an independent steering committee of the Respiratory Effectiveness Group (REG) and registered with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP; registration number EUPAS12086) [18].

Study design and population

This historical cohort study included an observation period of 2–10 years (as available) prior to an index date at which patients received a first diagnostic code for pulmonary fibrosis. Data extraction cut-off date was December 31st 2015. Cases were identified from diagnostic (Read) code lists developed by the REG ILD Working Group and aligned with published IPF-related research conducted in the UK [6,14]. Acknowledging variations in coding practice between healthcare professionals, cases were labelled as “pulmonary fibrosis-clinical syndrome” (PFCS) in the presence of a code considered diagnostic of pulmonary fibrosis and the absence of codes associated with ILD, such as a connective tissue disease (CTD), sarcoidosis or extrinsic allergic alveolitis (EAA). A subpopulation of patients with

only IPF-specific diagnostic codes—here termed “IPF clinical syndrome” (IPFCS)—was identified. Diagnostic codes used as inclusion criteria for the PFCS and IPFCS cohorts are provided in Supplementary Table S1 [6,14]. Codes contributing to PFCS or IPFCS were recorded in primary and secondary care.

Eligible patients were aged ≥ 40 years at diagnosis and diagnosed with PFCS between 2005–2015. Patients with less than two years of continuous medical records immediately prior to diagnosis were excluded. Exclusion of patients with CTD, sarcoidosis and EAA (codes in Supplementary Table S2) was agreed by the REG ILD Working Group.

Outcome measures

Characteristics at diagnosis

Characteristics of the study population: demographic (age, sex, anthropomorphic measures); lifestyle (smoking history); clinical features (lung function [FVC, FEV1/FVC] recorded closest to diagnosis); common comorbidities (chronic obstructive pulmonary disease [COPD], asthma and lung cancer, as defined by codes rather than functional tests); and use of obstructive lung disease (OLD) pharmacotherapy in the year preceding PFCS diagnosis.

Healthcare resource utilisation

Temporal trends in respiratory symptoms were evaluated by assessment of population-level per-patient annualised cough events coded through primary care consultations over 10 years prior to PFCS diagnosis. Functional impairment was evaluated using the mini-Medical Research Council (mMRC) score captured through a systematic data capture protocol. Availability of mMRC scores reflected patients’ willingness to participate in routine data collection rather than their clinical situation.

10-year trends in lower respiratory (LR) HRU were evaluated through annualised per-patient rates of coded events, including: primary care consultations, chest radiography (CXR), hospitalisations and emergency room attendances, and antibiotics and acute oral steroid prescriptions for an LR complaint. LR complaint codes included those for LR tract infections (bronchitis, tracheitis, pneumonia), non-

infective LR conditions (chronic respiratory failure), and respiratory symptoms (breathlessness, cough, wheeze).

Opportunities for earlier PFCS diagnosis were explored by evaluating unique patients with 0, 1, 2, 3, 4, ≥ 5 LR-related primary care contacts in each 1-year period in the decade preceding PFCS diagnosis. Prevalence of potential differential pulmonary diagnoses (COPD, asthma and lung cancer) at time of PFCS diagnosis, and their proximity to PFCS diagnosis, were also explored.

Statistical analysis

Summary statistics (n [%]) for patient characteristics at time of PFCS or IPFCS diagnosis were generated using STATA (version 14). Joinpoint Trend Analysis software from the Surveillance Research Programme of the US National Cancer Institute (available at <http://surveillance.cancer.gov/joinpoint>) was used to test whether non-segmented or segmented regression curves best described the evolution of temporal trends for outcomes (LR symptoms and HRU) over the ten-year period prior to diagnosis [19]. Joinpoints (inflection points) indicated timepoints from which a difference in annual percentage change (APC) in outcomes was observed in a segmented curve.

RESULTS

A total of 2223 PFCS patients (Figure 1), including 743 IPFCS patients, were eligible (Supplementary Figure S1).

Cohort characteristics

The PFCS cohort comprised more men (62.9%) than women, with a mean (standard deviation [SD]) age of 72.6 (9.7) years. Two-thirds (67.2%) were current/former smokers. Cardiorespiratory conditions were the most common comorbidities at time of PFCS diagnosis: ischaemic heart disease (32.4%), COPD (22.6%), lung cancer (18.0%), asthma (13.9%). Mean (SD) FVC was 2.9 (5.7) L and FEV₁/FVC ratio was 0.78 (0.1). Almost one-quarter of patients (23.2%) received a prescription for a short-acting beta₂-agonist one year prior to PFCS diagnosis (Table 1). PFCS and IPFCS patient characteristics were broadly similar (Tables 1 and Supplementary Table S3, respectively). In patients with OLD,

COPD diagnosis occurred at a mean (SD) of 5.1 (6.7) years, and asthma at a mean (SD) of 12.1 (14.1) years, prior to PFCS diagnosis. 40% of patients received asthma diagnosis within five years prior to PFCS diagnosis (60% within 10 years). COPD diagnosis occurred at 4.6 (6.2) years, and asthma at 13.0 (17.4) years, prior to IPFCS diagnosis.

Functional impairment (assessed by mMRC) score prior to PFCS diagnosis were available for 47% of patients. When stratified by time between mMRC assessment and PFCS diagnosis, no pronounced temporal trends were observed, although a slight reduction in patients without signs of functional impairment was apparent over the 10-year study period (Supplementary Figure S2).

Healthcare resource utilisation

The mean pre-observation period for patients with non-missing HRU data ranged from 6–8 years. There was marked increase in cough events in lead up to PFCS diagnosis (Figure 2). Mean (SD) annual per-patient cough event rate increased 10-fold (from 0.06 [0.3] per patient per year [pppy] at 10 years and 0.58 [0.6] pppy one year prior to diagnosis). Joinpoint regression analysis identified a significant increase in cough events two years prior to PFCS diagnosis (APCs for years 10–2 vs 2–1: 19.67% vs 115.09%; $P=0.0002$). In the IPFCS cohort cough event rates increased from 0.05 to 0.13 pppy over the 10-year period (data not shown).

Trends in LR HRU demonstrated that mean (SD) primary care LR consultations increased from 0.4 (1.3) pppy at 10 years prior to PFCS diagnosis, to 4.5 (1.6) pppy one year prior to diagnosis (Figure 3 and Table 2). APCs for LR consultations significantly increased two years prior to PFCS diagnosis (APCs for years 10–2 vs 2–1: 19.96% vs 199.72%; $P=0.0002$). Antibiotic and acute oral steroid prescriptions for LR events increased over the 10-year period (0.08 [0.3] to 0.81 [0.7] pppy and 0.02 [0.2] to 1.07 [0.7] pppy, respectively). APCs for antibiotic and acute steroid prescriptions significantly increased two years (APCs for years 10–2 vs 2–1: 20.70% vs 118.72%; $P=0.0002$) and three years (APCs for years 10–3 vs 3–1: 36.92% vs 151.27%; $P=0.0024$) prior to PFCS diagnosis, respectively. These trends were consistent in the IPFCS cohort (Supplementary Figure S3 and Table S4).

APC for secondary care contacts for LR complaints (hospital admissions or ER attendances) was 0.00% until two years prior to PFCS diagnosis. For years 2–1 prior to PFCS diagnosis, an APC of 700.00% was observed (Figure 3 and Table 2). This was consistent in the IPFCS cohort (Supplementary Figure S3 and Table S4). A 3-fold increase in pneumonia in the two years immediately prior to PFCS diagnosis was also observed (data not shown). Chest X-rays (CXRs) increased 10-fold from 10 years to one year prior to PFCS diagnosis (mean [SD] CXRs: 0.03 [0.2] to 0.40 [0.5] ppy, respectively; Figure 3 and Table 2). APC for CXRs significantly increased two years prior to PFCS diagnosis (APCs for years 10–2 vs 2–1: 18.64% vs 216.07%; $P=0.0002$). In the IPFCS cohort, however, APC for CXRs significantly increased one year earlier, i.e. three years prior to IPFCS diagnosis (Supplementary Figure S3 and Table S4).

All-cause primary care consultations were examined to explore whether increases in LR consultations reflected a general escalation in all-cause HRU due to aging. LR consultations accounted for 10% and 35% of all-cause consultations 10 years and one year prior to PFCS diagnosis, respectively. APC for all-cause consultations significantly increased two years prior to PFCS diagnosis (APCs for years 10–2 vs 2–1: 11.90% vs 46.76%; $P=0.0004$). This increase was driven by LR consultations (APCs for years 10–2 vs 2–1: 19.96% vs 199.72%; $P=0.0002$), while APC for non-LR consultations was constant (11.03%) (Figure S4). The contribution of LR consultations to all-cause consultations was evident in the IPFCS cohort (data not shown).

Ten years prior to PFCS diagnosis, 18% of patients visited primary care for an LR reason. Of these patients, only 6.3% consulted ≥ 2 times and 1.5% ≥ 5 times. Five years later, approximately 30% of patients consulted ≥ 1 times for an LR complaint, 14.4% ≥ 2 times and 3.8% ≥ 5 times. One year prior to PFCS diagnosis, almost all patients (99.9%) consulted ≥ 1 times for an LR complaint, 78.8% ≥ 2 times and 38.0% ≥ 5 times (Figure 4 and Supplementary Table S5). Consistent with this, APCs for two, three, four, and five or more LR consultations significantly increased two years prior to PFCS diagnosis (APCs for 10–2 vs 2–1 years: 12.60% vs 85.47% ($P=0.0018$), 20.96% vs 67.79% ($P=0.0104$), 21.95% vs 141.43% ($P=0.0007$), and 25.35% vs 309.71% ($P=0.0002$), respectively).

DISCUSSION

Our findings show a significant increase in HRU across several domains prior to a PFCS diagnosis: cough, LR consultations, antibiotic and acute steroid prescriptions, and CXR. Eighteen percent of patients made multiple primary care visits for LR complaints five years before PFCS diagnosis; increasing to almost 80% one year before diagnosis, with 38% having five or more primary care contacts. Therefore, opportunities exist for earlier referral for pulmonary fibrosis suspected in primary care.

Whilst the natural history of OLD is growing [20], that of fibrotic lung diseases remains limited. Research into fibrotic lung disease has traditionally been restricted to evaluations conducted within specialist centres; large-scale primary care databases offer an opportunity to study more widely representative and generalisable populations.

We analysed a high-quality dataset widely used in primary care studies [17,20]. Anthropomorphic and lung function measures are consistent with previous IPF cohorts [6,8,9,21]. PFCS diagnosis was based on codes diagnostic for pulmonary fibrosis and the absence of codes associated with ILD. While potentially including patients with a form of pulmonary fibrosis other than IPF, a subgroup analysis of codes specific for IPF demonstrated comparability. These patients need specialist referral for investigation irrespective of final diagnosis. Additional studies are required to validate specific diagnostic groups, including the correlation of primary care diagnostic codes with specialist care diagnosis.

Our findings extend those of Hewson *et al.* who identified increasing breathlessness and cough in the five years prior to IPFCS diagnosis [14]. We identified significant increases in cough events and LR-related HRU two years preceding PFCS diagnosis. This may reflect symptom progression from minor to moderate/severe, with symptoms affecting quality resulting in increased HRU. Increases in LR consultations, antibiotic and acute oral steroid prescriptions suggests infective episodes may precipitate HRU and lead to subsequent diagnosis. Alternatively, acute prescriptions may represent empirical trials by primary care physicians following repeat patient attendances for persistent symptoms of uncertain aetiology.

Thirty eight percent of patients had five or more LR consultations one year preceding IPF diagnosis. This suggests repeated primary care attendances are required for specialist referral. Given the potential for misdiagnosis, we investigated the relationship between respiratory comorbidities and PFCS diagnosis. COPD and asthma diagnoses occurred at five and 12 years before PFCS diagnosis, respectively. While these timelines may suggest alternative respiratory diagnoses, this could reflect misdiagnosis of initial functional impairment due to early-onset PFCS.

Limitations in our data and framework conditions potentially restrict our interpretations. For example, spirometry is not routine in primary care, with COPD or asthma patients more likely to undergo spirometry testing. Secondary care data requires manual reporting and entry into primary care records, resulting in under-reporting and under-estimation of HRU from secondary care. A limitation to the interpretation of any relationship between PFCS diagnosis and preceding comorbidities is that comorbidities were captured through codes rather than functional tests. Miscoding represents a source of potential misdiagnosis. Another consideration is framework conditions that determine drug prescribing policies which change over time and may influence HRU in the lead up to diagnosis. Our analysis did not account for this potential covariate.

Nevertheless, patients consistently report dissatisfaction with time taken to diagnosis [15,16]. Studies investigating how to effectively increase awareness of pulmonary fibrosis among primary care physicians are required. This may include re-emphasis of routine lung auscultation to check for crackles in older patients presenting with repeated LR complaints over a short time period. In older populations, a joint spirometric and lung auscultation approach may have value for obstructive and fibrotic lung disease diagnosis. Future research should seek to link primary and secondary care data to focus on the development of a pulmonary fibrosis risk algorithm for integration within clinical decision management systems.

Data availability statement:

Individual participant data is not available.

Disclosure statement:

The study protocol was developed by REG on behalf of its independent ILD Working Group. The dataset was provided by Optimum Patient Care Ltd and the analysis conducted by REG on behalf of the ILD Working Group who reviewed the results and approved the development of this manuscript. The corresponding author had full access to all of the data and accepts responsibility for their submission for publication.

T.M.M. has, via his institution, received industry-academic funding from GlaxoSmithKline R&D and UCB and has received consultancy or speakers fees from Apellis, Astra Zeneca, Bayer, Blade Therapeutics, Boehringer Ingelheim, Bristol-Myers Squibb, Galapagos, GlaxoSmithKline R&D, Indalo, Novartis, Pliant, ProMetic, Respivnat, Roche, Samumed and UCB.

A.K. is on the advisory board or speakers bureau for Astra Zeneca, Behring, Boehringer Ingelheim, Covis, Griffols, GSK, Merck, Novartis, Paladin, Purdue, Teva, Trudell.

S.W. has received speakers fees from Boehringer Ingelheim, Roche, Galapagos, Bracco, Sanofi-Genzyme and is on the advisory boards for Boehringer Ingelheim, Galapagos, The Open Source Imaging Consortium and OncoArendi Therapeutics.

D.P. has board membership with Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, Circassia, Mylan, Mundipharma, Novartis, Regeneron Pharmaceuticals, Sanofi Genzyme, Teva Pharmaceuticals; consultancy agreements with Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Mylan, Mundipharma, Novartis, Pfizer, Teva Pharmaceuticals, Theravance; grants and unrestricted funding for investigator-initiated studies (conducted through Observational and Pragmatic Research Institute Pte Ltd) from AstraZeneca, Boehringer Ingelheim, Chiesi, Circassia, Mylan, Mundipharma, Novartis, Pfizer, Regeneron Pharmaceuticals, Respiratory Effectiveness Group, Sanofi Genzyme, Teva Pharmaceuticals, Theravance, UK National Health Service; payment for lectures/speaking engagements from AstraZeneca, Boehringer Ingelheim, Chiesi, Cipla, GlaxoSmithKline, Kyorin, Mylan, Merck, Mundipharma, Novartis, Pfizer, Regeneron Pharmaceuticals, Sanofi Genzyme, Teva Pharmaceuticals; payment for manuscript preparation from Mundipharma, Teva Pharmaceuticals; payment for the development of educational materials from Mundipharma, Novartis; payment for travel/accommodation/meeting expenses from AstraZeneca, Boehringer Ingelheim, Circassia, Mundipharma, Novartis, Teva Pharmaceuticals; funding for patient enrolment or

completion of research from Chiesi, Novartis, Teva Pharmaceuticals; stock/stock options from AKL Research and Development Ltd which produces phytopharmaceuticals; owns 74% of the social enterprise Optimum Patient Care Ltd (Australia and UK) and 74% of Observational and Pragmatic Research Institute Pte Ltd (Singapore); and is peer reviewer for grant committees of the Efficacy and Mechanism Evaluation programme, and Health Technology Assessment.

L.R. has received personal fees from Biogen, Sanofi-Aventis, Roche, Celgene, Nitto, FibroGen, Promefior, Pliant Therapeutics, Asahi Kasei, Toray, Global Blood Therapeutics, Acceleron and Immune Works, grants and personal fees from Boehringer Ingelheim, outside the submitted work.

Author contributions:

Conceptualization: A.C., M.G.J., D.P., L.R. Data curation: C.R.T.H., A.N. Formal Analysis: C.R.T.H., A.N., M.G.J., A.C., L.R., Methodology: A.C., T.M.M., A.W., S.W., A.K., D.P., L.R., Writing – original draft: A.C., M.G.J., A.N., C.R.T.H. Writing – reviewing and editing: all authors.

References

1. Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2006; 174: 810-6.
2. Pérez ERF, Daniels CE, Schroeder DR, Sauver JS, Hartman TE, Bartholmai BJ, Yi ES, Ryu JH. Incidence, prevalence, and clinical course of idiopathic pulmonary fibrosis: a population-based study. *Chest.* 2010; 137: 129-37.
3. Meltzer EB, Noble PW. Idiopathic pulmonary fibrosis. *Orphanet J Rare Dis.* 2008; 3: 8.
4. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Varvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE, Kondoh Y, Myers J, Muller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schunemann HJ. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med.* 2011; 183: 788-824.
5. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. *Lancet.* 2017; 389: 1941-52.
6. Navaratnam V, Fleming K, West J, Smith C, Jenkins R, Fogarty A, Hubbard RB. The rising incidence of idiopathic pulmonary fibrosis in the UK. *Thorax.* 2011; 66: 464-467.

7. Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. *Am J Respir Crit Care Med*. 2011; 184: 842-7.
8. Richeldi L, Du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, Cottin V, Flaherty KR, Hansell DM, Inoue Y, Kim DS, Kolb M, Nicholson AG, Noble PW, Selman M, Taniguchi H, Brun M, Le Maulf F, Girard M, Stowasser S, Schlenker-Herceg R, Disse B, Collard HR. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med*. 2014; 370: 2071-82.
9. King Jr TE, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster L, Lederer DJ, Nathan SD, Pererira CA, Sahn SA, Sussman R, Swigris JJ, Noble PW. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med*. 2014; 370: 2083-92.
10. The Lancet Respiratory Medicine. The changing landscape of idiopathic pulmonary fibrosis. *Lancet Respir Med*. 2014; 2: 507.
11. Bonella F, Wijsenbeek M, Molina-Molina M, Duck A, Mele R, Geissler K, Wuyts W. European IPF Patient Charter: unmet needs and a call to action for healthcare policymakers. *Eur Respir J*. 2016; 47: 597-606.
12. Antoniou KM, Symvoulakis EK, Margaritopoulos GA, Lionis C, Wells AU. Early diagnosis of IPF: time for a primary-care case-finding initiative? *Lancet Respir Med*. 2014; 2: e1.
13. Martinez FJ, Chisholm A, Collard HR, Flaherty KR, Myers J, Raghu G, Walsh SL, White ES, Richeldi L. The diagnosis of idiopathic pulmonary fibrosis: current and future approaches. *Lancet Respir Med*. 2017; 5: 61-71.
14. Hewson T, McKeever T, Gibson J, Hubbard R, Hutchinson J. Timing of onset of symptoms in people with idiopathic pulmonary fibrosis. *Thorax*. 2018; 73: 683-685
15. Schoenheit G, Becattelli I, Cohen AH. Living with idiopathic pulmonary fibrosis: an in-depth qualitative survey of European patients. *Chron Respir Dis*. 2011; 8: 225-31.
16. Collard HR, Tino G, Noble PW, Shreve MA, Michaels M, Carlson B, et al. Patient experiences with pulmonary fibrosis. *Respir Med*. 2007; 101: 1350-4.
17. Optimum Patient Care Research Database (OPCRD). <http://optimumpatientcare.org/database-overview/> (accessed 7 November 2018).
18. The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). <http://www.encepp.eu/encepp/studiesDatabase.jsp>. (accessed 7 November 2018).

19. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine*. 2000; 19: 3.
20. Jones RC, Price D, Ryan D, Sims EJ, von Ziegenweidt J, Mascarenhas L, Burden A, Halpin DM, Winter R, Hill S, Kearney M, Holton K, Moger A, Freeman D, Chisholm A, Bateman ED. Opportunities to diagnose chronic obstructive pulmonary disease in routine care in the UK: a retrospective study of a clinical cohort. *Lancet Respir Med*. 2014; 2: 267-76.
21. Maher TM, Oballa E, Simpson JK, Porte J, Habgood A, Fahy WA, Flynn A, Molyneaux PL, Braybrooke R, Divyateja H, Parfrey H, Rassi D, Russel AM, Saini G, Renzoni EA, Duggan AM, Hubbard R, Wells AU, Lukey PT, Marshall RP, Jenkins RGI. An epithelial biomarker signature for idiopathic pulmonary fibrosis: an analysis from the multicentre PROFILE cohort study. *Lancet Respir Med*. 2017; 5: 946-55.

FIGURE LEGENDS

Figure 1. Pulmonary fibrosis clinical syndrome (PFCS) eligibility flow diagram.

OPCRD: Optimum Patient Care Research Database; CTD: connective tissue disease; RA: rheumatoid arthritis

Figure 2. Joinpoint regression analysis of cough-related events over 10 years prior to PFCS diagnosis.

APC: annual percent change; Joinpoint: point of change in regression trends; pppy: per patient per year; PFCS: pulmonary fibrosis clinical syndrome; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; Mean pre-observation period: 8 years.

Figure 3. Joinpoint regression analysis of temporal HRU trends over 10 years prior to PFCS diagnosis.

Temporal HRU trends in the 10 years prior to PFCS diagnosis for A) lower respiratory (LR) consultations, B) acute antibiotic prescriptions, C), acute steroid prescriptions, D) LR hospital contacts, and E) chest X-rays. APC: annual percent change; Joinpoint: point of change in regression trends; pppy: per patient per year; PFCS: pulmonary fibrosis clinical syndrome; Number of patients

with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; mean pre-observation period: 8 years.

Figure 4. Joinpoint regression analysis of LR healthcare contacts over 10 years prior to PFCS diagnosis.

Temporal HRU trends showing percentage of patients with A) 0, B) 1, C) 2, D) 3, E) 4, and F) ≥5 LR consultations in the 10 years prior to PFCS diagnosis; APC: annual percent change; Joinpoint: point of change in regression trends; PFCS, pulmonary fibrosis clinical syndrome; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; Mean pre-observation period: 8 years

Table 1. Baseline characteristics at time of PFCS diagnosis

Characteristic	n = 2,223
Male sex, n (%)	1,399 (62.9)
Age at index date (years), mean (SD)	72.6 (9.7)
BMI (kg/m²), mean (SD)	27.3 (7.0)
Smoking status	
Never	688 (32.8)
Current	280 (13.4)
Former	1,127 (53.8)
Comorbidities, n (%)	
COPD	503 (22.6)
Asthma	309 (13.9)
Ischaemic heart disease	720 (32.4)
Heart failure	249 (11.2)
Hypertension	118 (5.3)
Myocardial infarction	258 (11.6)
Lung Cancer	401 (18.0)
Sleep Apnoea	18 (0.8)
GERD*	170 (7.7)
Anxiety and depression	55 (2.5)
Any prescriptions in year prior to PFCS diagnosis, n (%)	
SABA	514 (23.2)
SAMA	133 (6.0)
ICS	196 (8.8)
ICS/LABA	210 (9.4)
Lung Function	
Without comorbid COPD	n = 245
FVC (L), mean (SD)	2.9 (5.7)
FEV ₁ /FVC ratio, mean (SD)	0.78 (0.1)
With comorbid COPD	n = 74
FVC (L), mean (SD)	2.5 (0.9)
FEV ₁ /FVC ratio, mean (SD)	0.72 (0.2)

PFCS: pulmonary fibrosis clinical syndrome; * Active comorbidities defined as recorded within two years prior to diagnosis; SABA: short-acting beta-agonist; SAMA: short-acting anti-muscarinic; ICS: inhaled corticosteroid; GERD: gastroesophageal reflux disease

Table 2. Temporal trends in healthcare resource utilisation in the 10 years prior to PFCS diagnosis

Year prior to PFCS diagnosis

	10	9	8	7	6	5	4	3	2	1
n (non-missing)	1474	1533	1598	1686	1772	1861	1944	2019	2087	2223
Primary care, mean (SD) pppy										
LR consultations	0.36 (1.3)	0.44 (1.3)	0.51 (1.4)	0.61 (1.5)	0.68 (1.7)	0.80 (1.9)	0.98 (2.0)	1.23 (2.3)	1.57 (0.9)	4.45 (1.6)
Prescriptions issued for LR complaints, mean (SD) pppy										
Antibiotics	0.08 (0.3)	0.11 (0.5)	0.12 (0.5)	0.14 (0.5)	0.17 (0.6)	0.19 (0.6)	0.27 (0.7)	0.31 (0.8)	0.37 (0.5)	0.81(0.7)
Acute oral steroid	0.02 (0.2)	0.02 (0.2)	0.05 (0.3)	0.05 (0.3)	0.07 (0.4)	0.07 (0.4)	0.12 (0.5)	0.15 (0.7)	0.52 (0.4)	1.07 (0.6)
Secondary care, mean (SD) pppy										
LR hospital admissions*	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.3 (0.1)	0.08 (0.2)
LR hospital admissions (sen)†	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.05 (0.1)
Diagnostics, mean (SD) pppy										
Chest X-Rays	0.03 (0.2)	0.04 (0.2)	0.05 (0.3)	0.06 (0.3)	0.06 (0.3)	0.07 (0.3)	0.08 (0.3)	0.10 (0.4)	0.14 (0.2)	0.40 (0.5)

PFCS: pulmonary fibrosis clinical syndrome; SD: standard deviation; pppy: per patient per year; * Lower respiratory (LR) code recorded on same day as inpatient admission; † LR code recorded within 14 days of an inpatient admission; Sen: sensitivity; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; Mean pre-observation period: 8 years.

Never received a PFCS diagnosis = 2,486,416

Diagnosis prior to 1990 = 125

Age <40 years at index date = 52

Exclude patients with ever diagnosis of CTD, sarcoidosis, allergic alveolitis, RA = 1,148

OPCRD all patients
n = 2,489,964

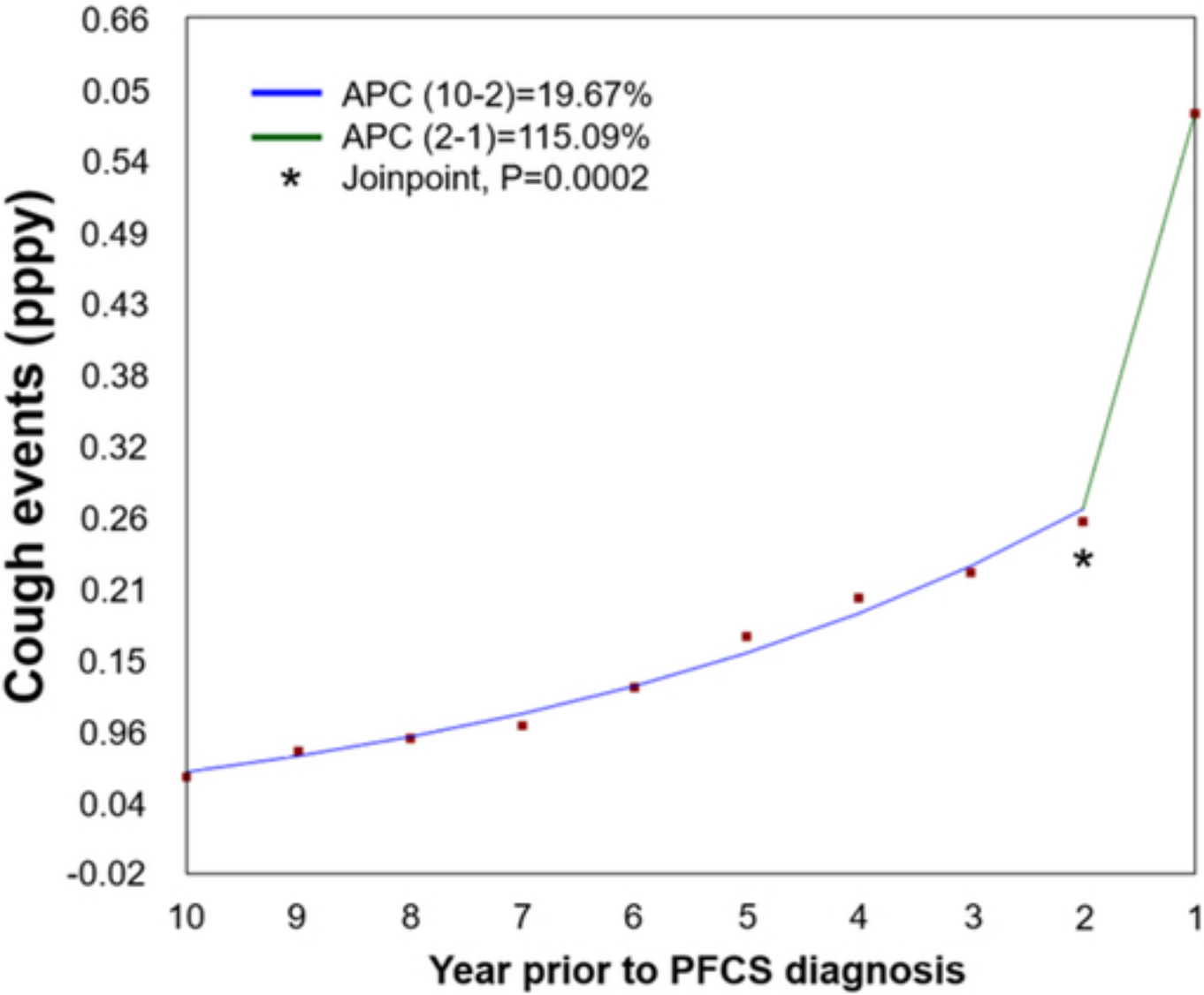
Diagnosis code for PFCS
n = 3,548

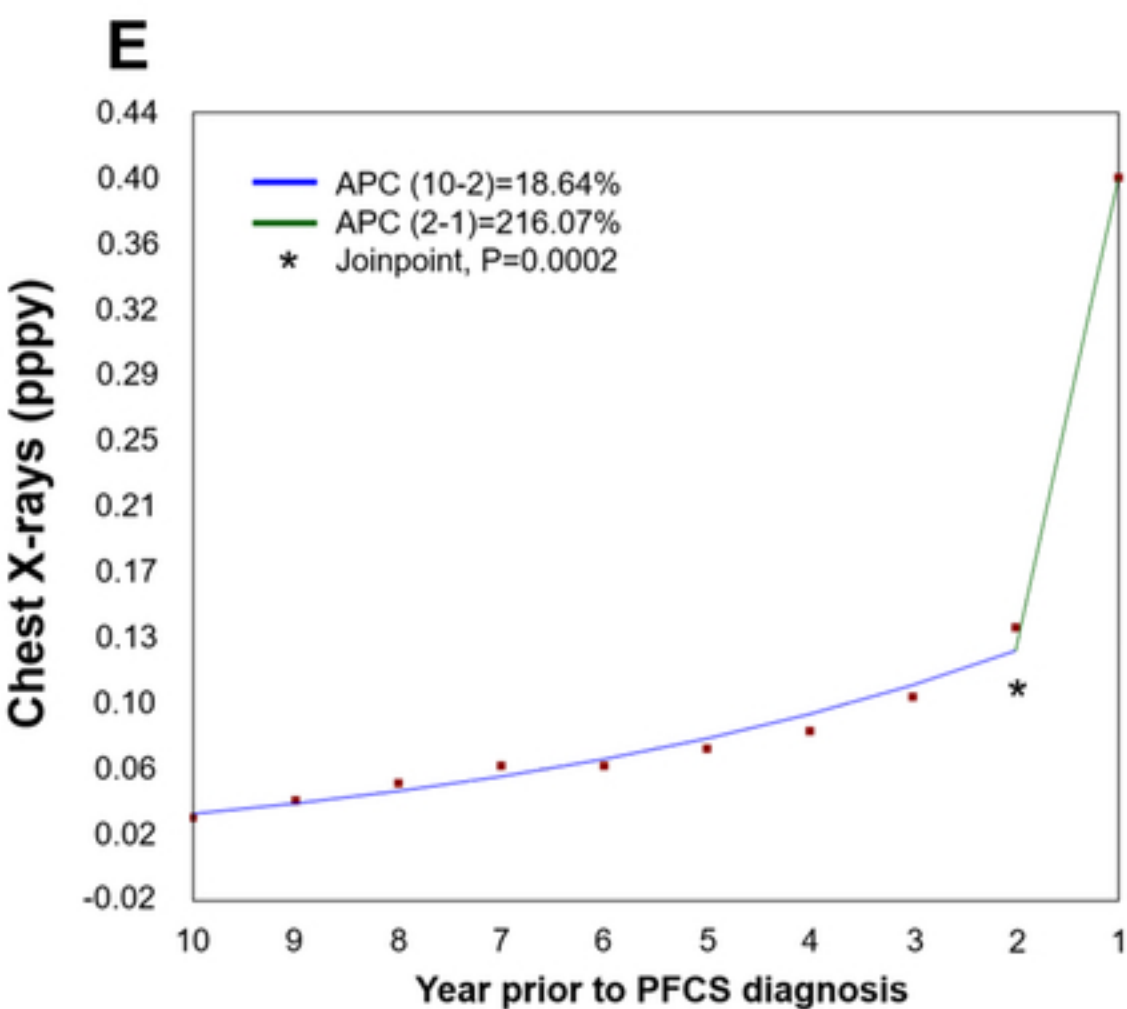
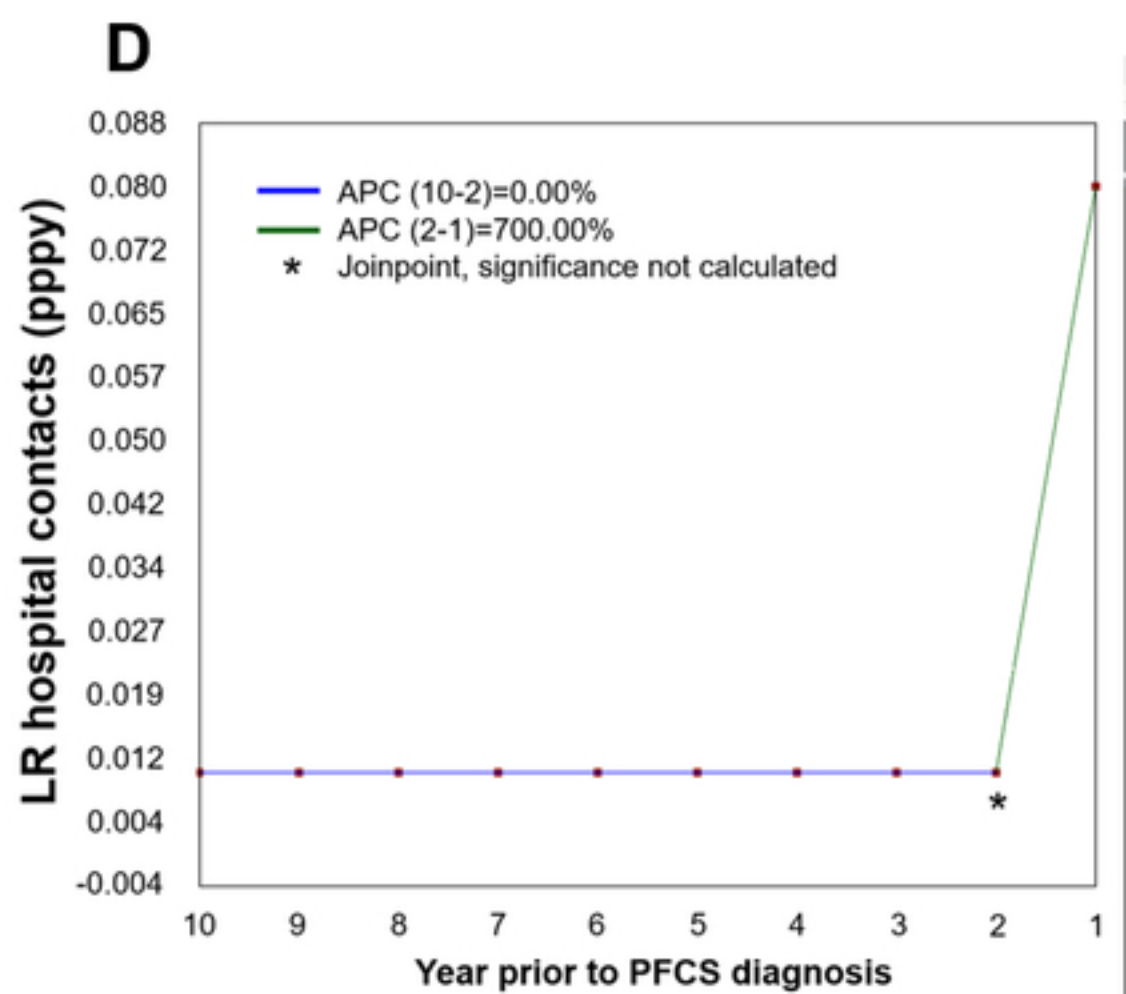
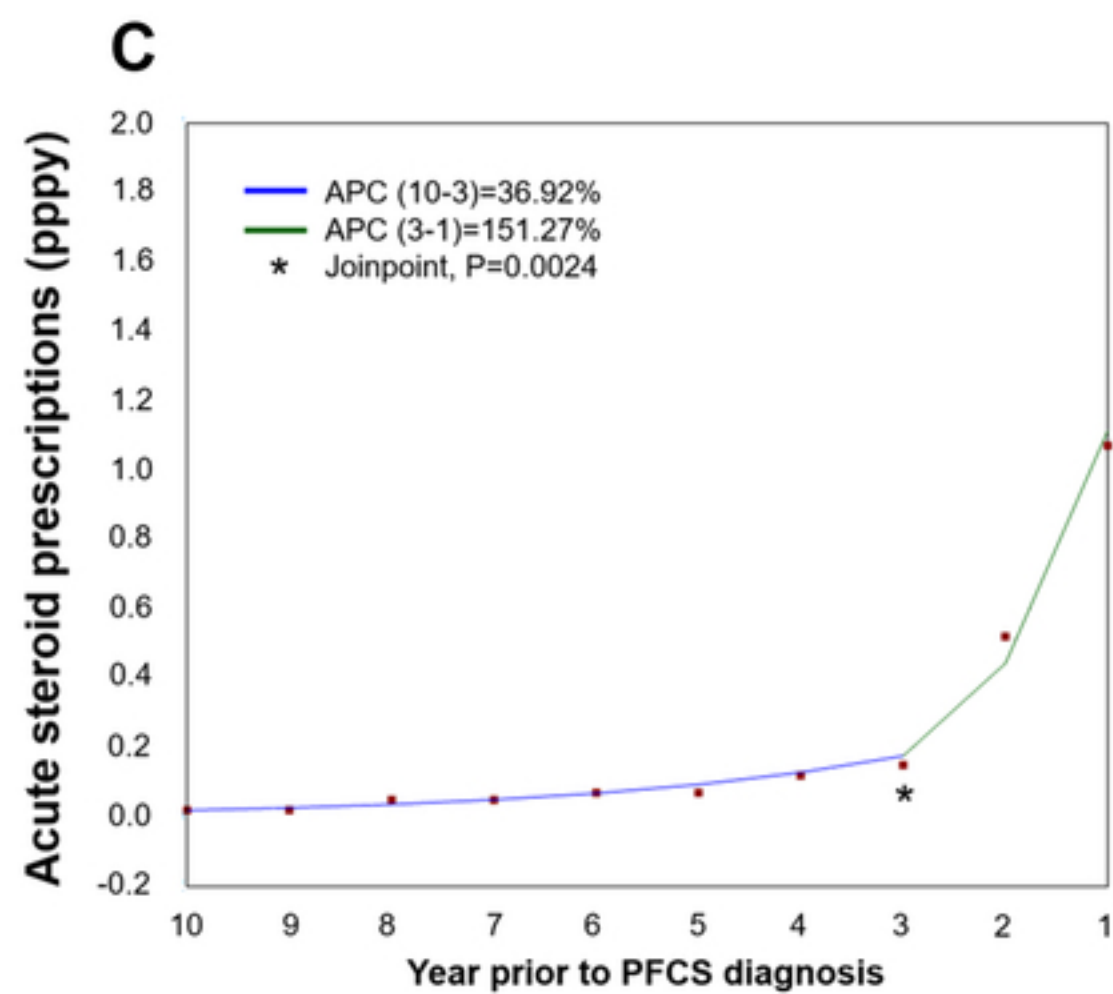
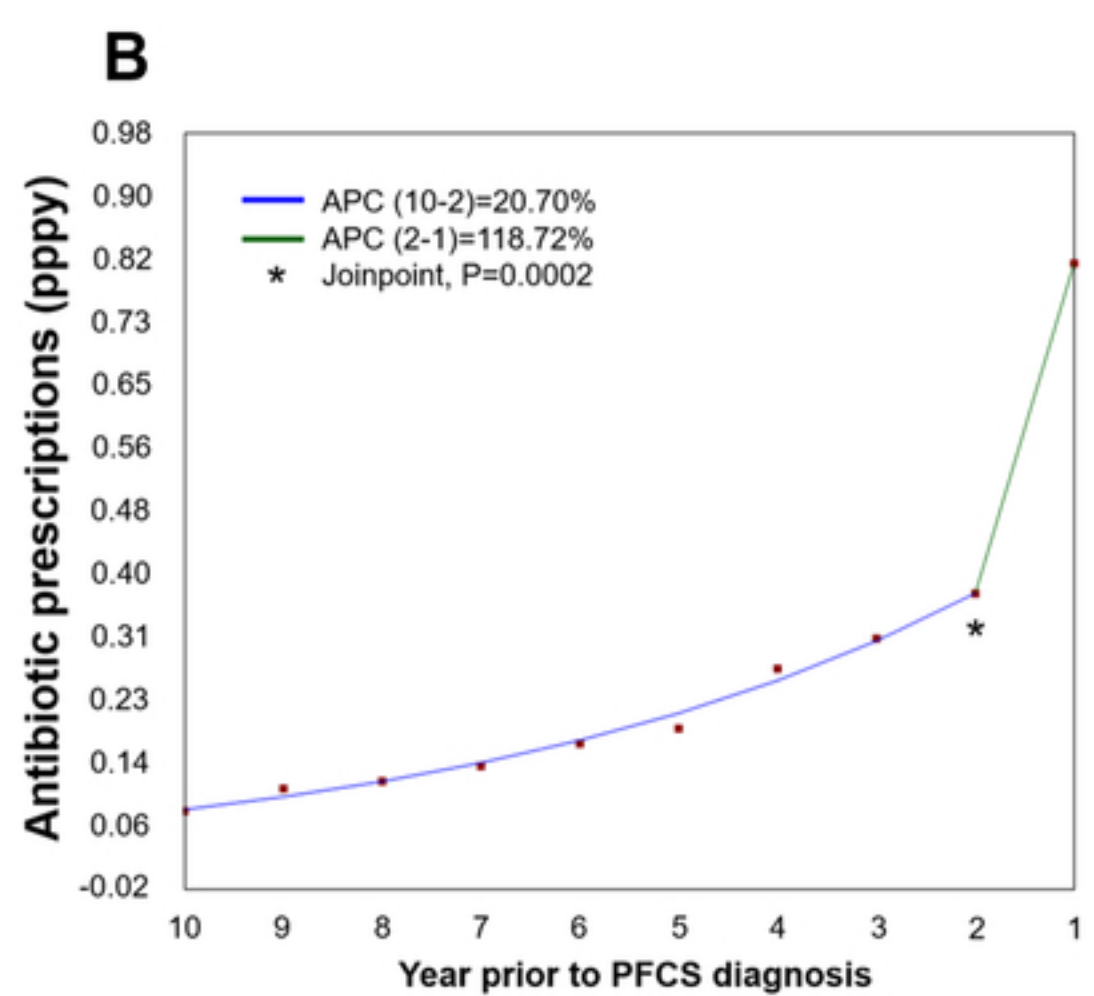
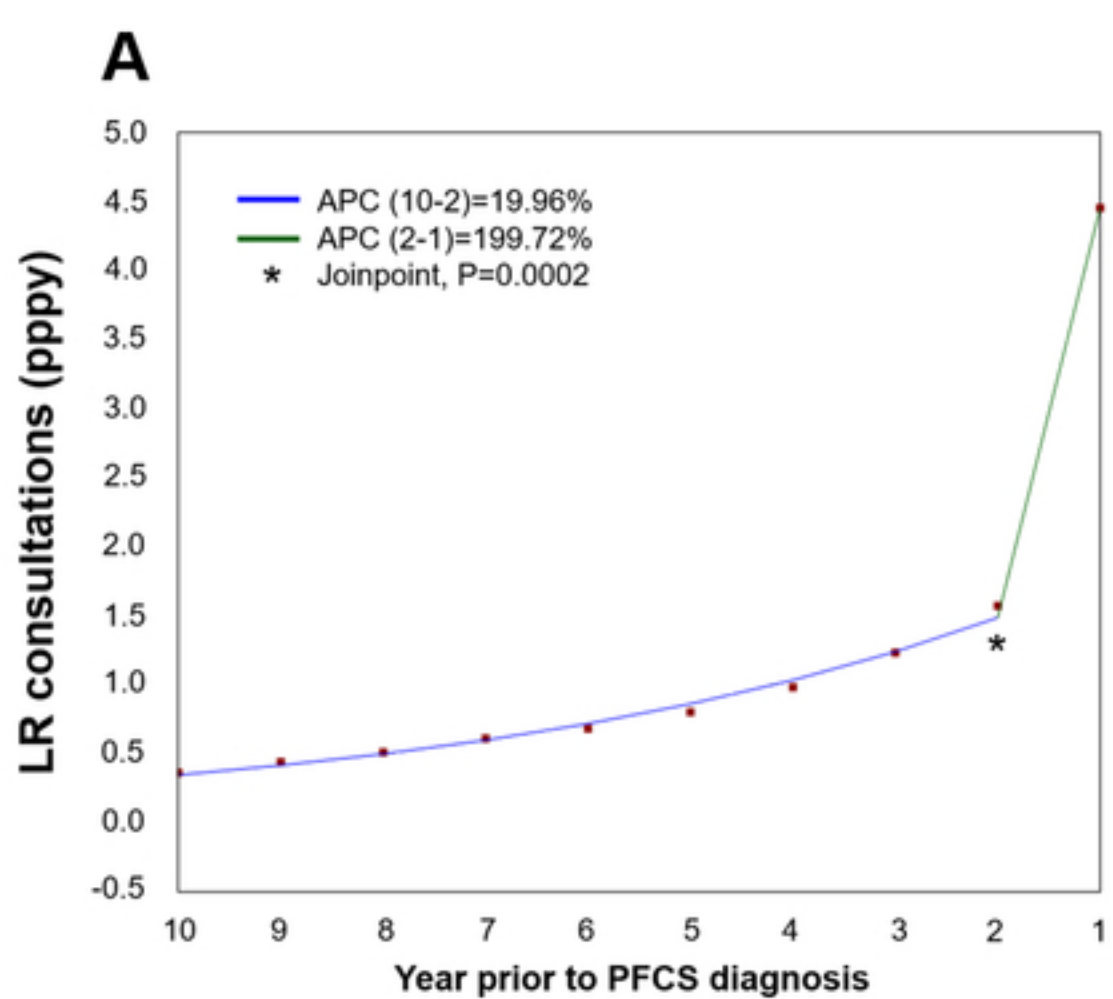
Diagnosis received after 1990
n = 3,423

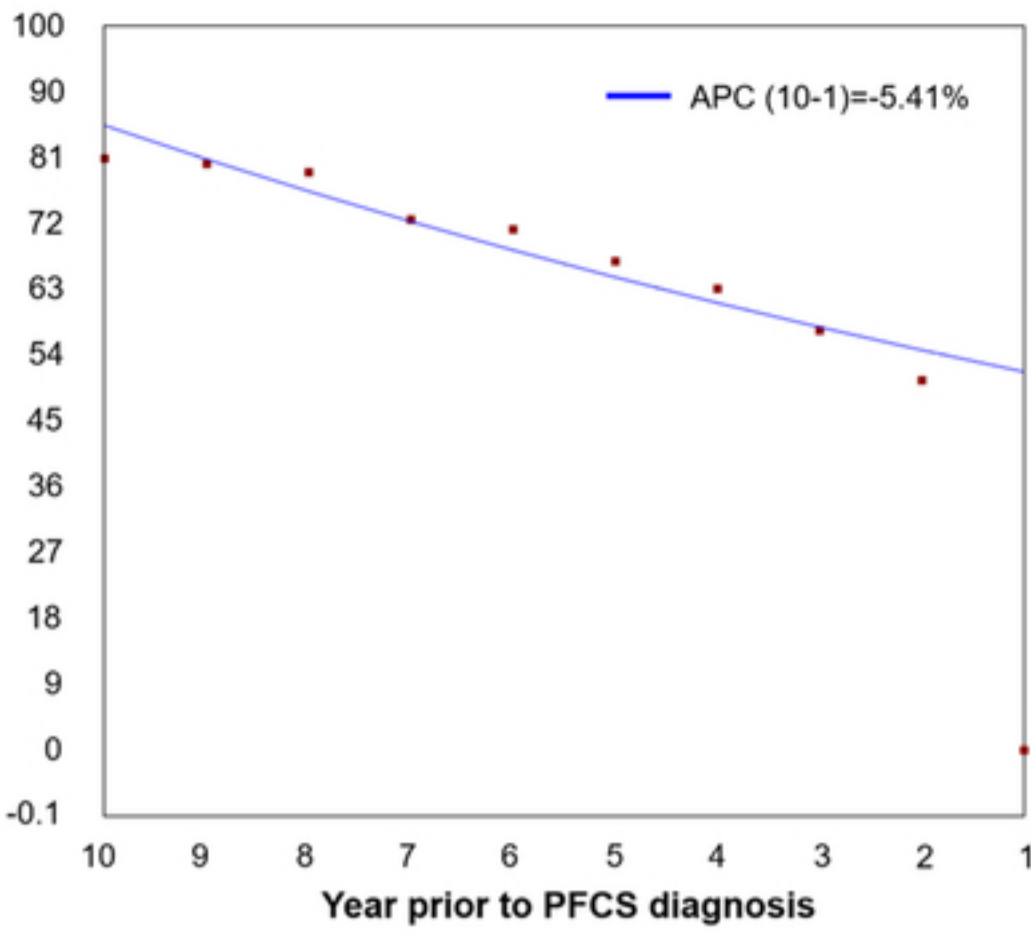
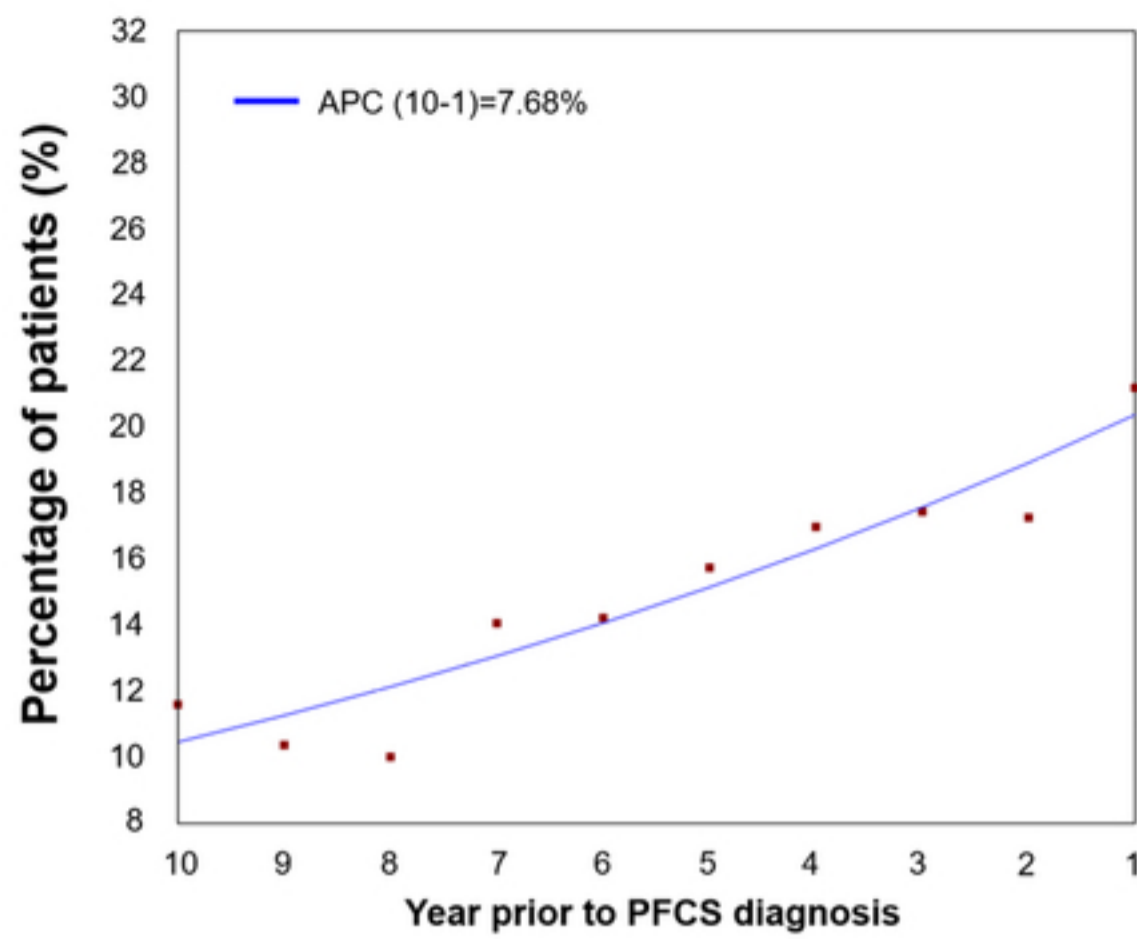
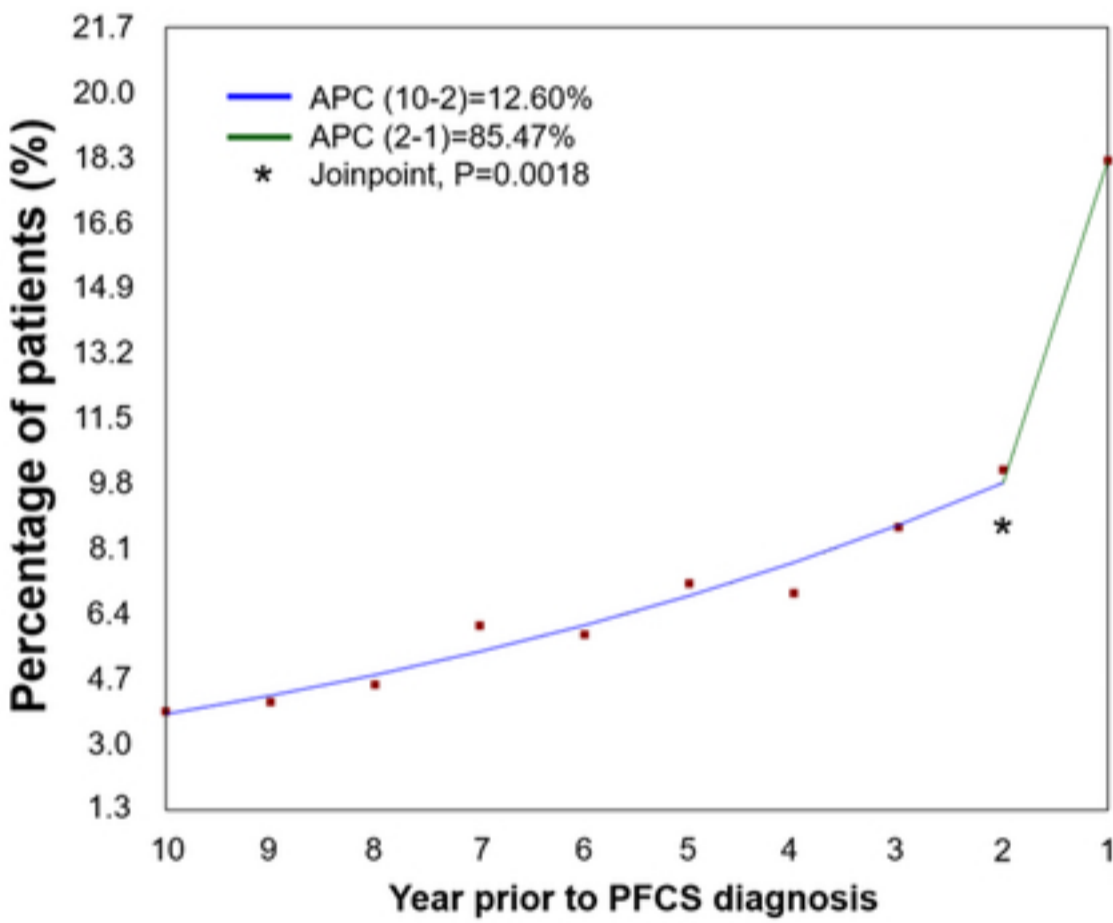
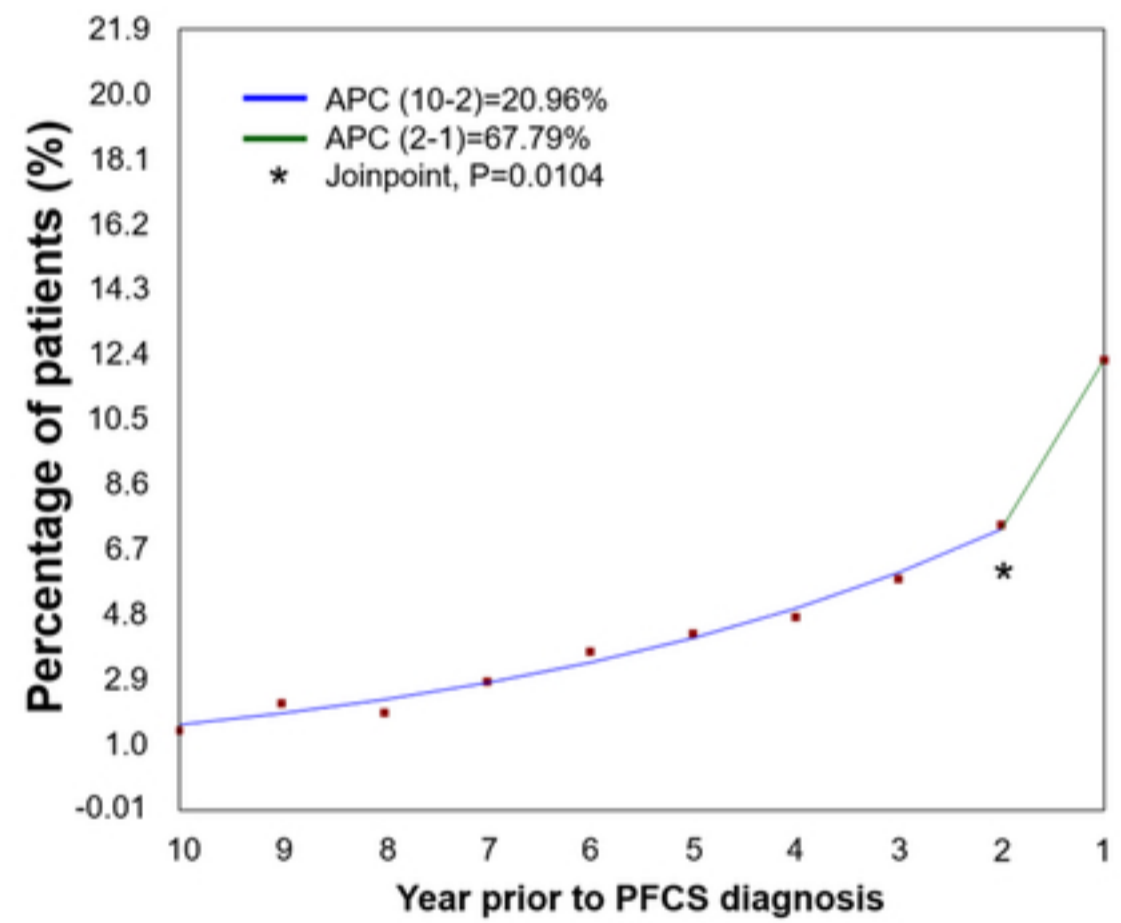
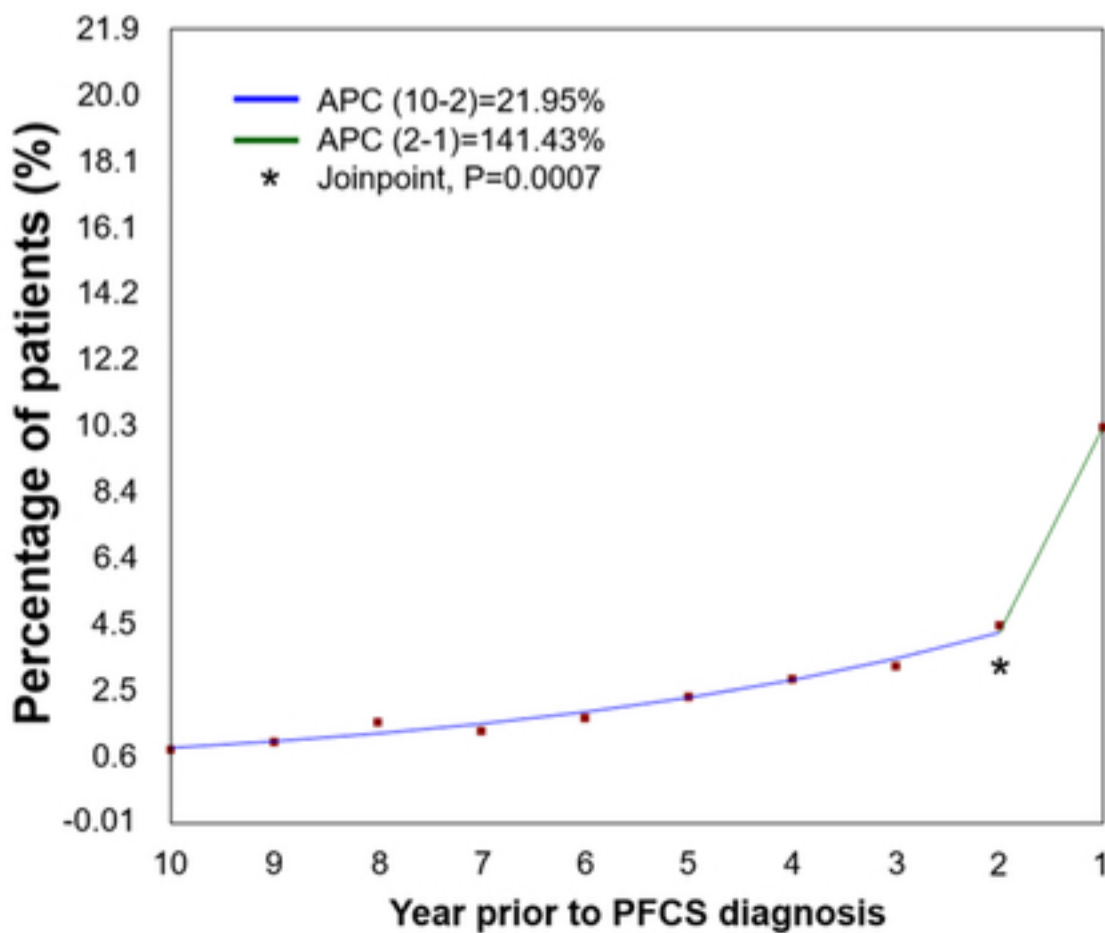
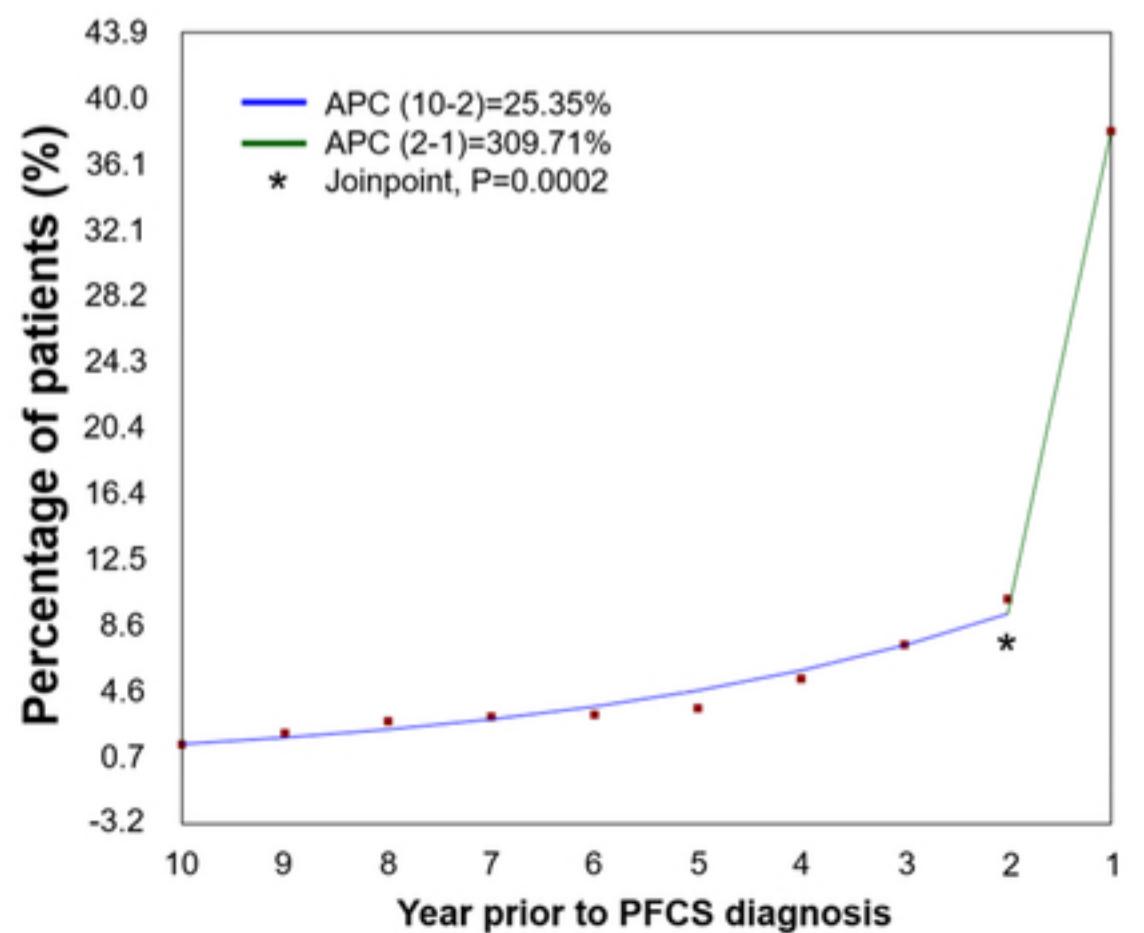
Age ≥40 years at index date
n = 3,371

PFCS cohort
n = 2,223







A 0 LR healthcare contacts**B 1 LR healthcare contact****C 2 LR healthcare contacts****D 3 LR healthcare contacts****E 4 LR healthcare contacts****F ≥5 LR healthcare contacts**

SUPPLEMENTARY INFORMATION

Opportunities to diagnose fibrotic lung diseases in routine care: A primary care cohort study

Mark G. Jones^{*1}, Christopher R.T. Hillyar^{*2}, Anjan Nibber³, Alison Chisholm⁴, Andrew Wilson⁵, Toby M. Maher^{6,7}, Alan Kaplan⁸, David Price^{9,10}, Simon Walsh⁷, Luca Richeldi^{1,11} on behalf of the Respiratory Effectiveness Group's Interstitial Lung Disease Working Group

1. National Institute for Health Research Biomedical Research Centre and Clinical and Experimental Sciences, University of Southampton, Southampton, United Kingdom
2. Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom
3. University of Oxford Medical School, Oxford, United Kingdom
4. Respiratory Effectiveness Group, Cambridge, United Kingdom
5. Norwich Medical School, University of East Anglia, Norwich, UK
6. Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK
7. Fibrosis Research Group, National Heart and Lung Institute, Imperial College London, London, UK
8. Department of Family and Community Medicine, University of Toronto, Ontario, Canada
9. Centre of Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom
10. Observational and Pragmatic Research Institute, Singapore
11. Unità Operativa Complessa di Pneumologia, Università Cattolica del Sacro Cuore, Fondazione Policlinico A. Gemelli, Rome, Italy

* These authors contributed equally to this manuscript and are joint-first authors.

Figure S1-Idiopathic pulmonary fibrosis clinical syndrome (IPFCS) cohort patient eligibility flow diagram

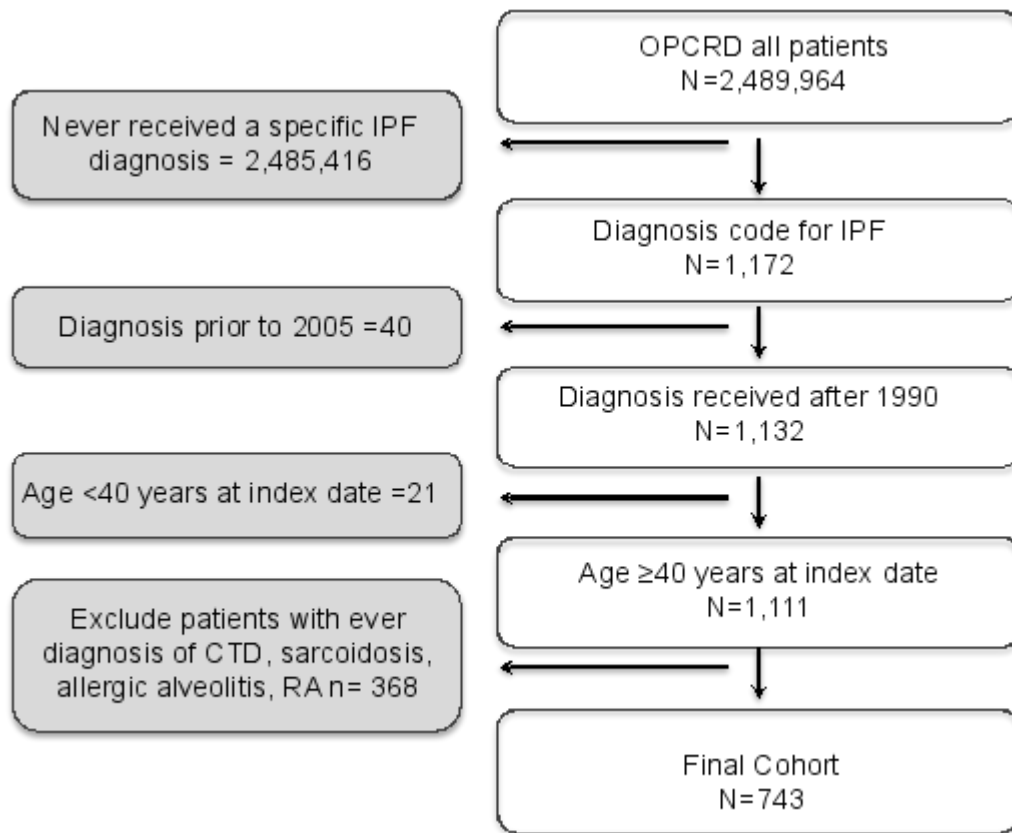
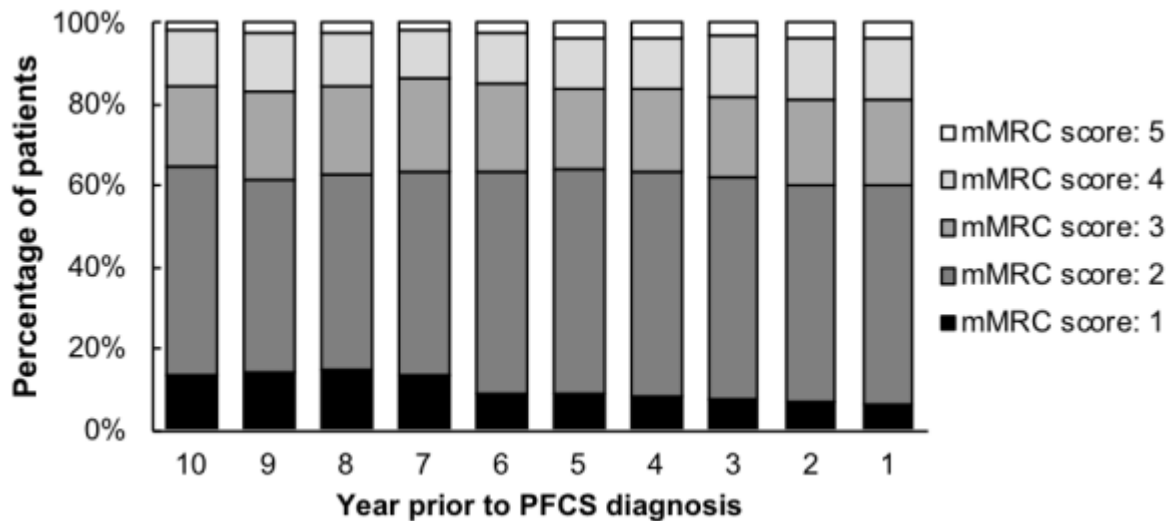


Figure S2. Functional impairment (mMRC score) over 10 years prior to PFCS diagnosis

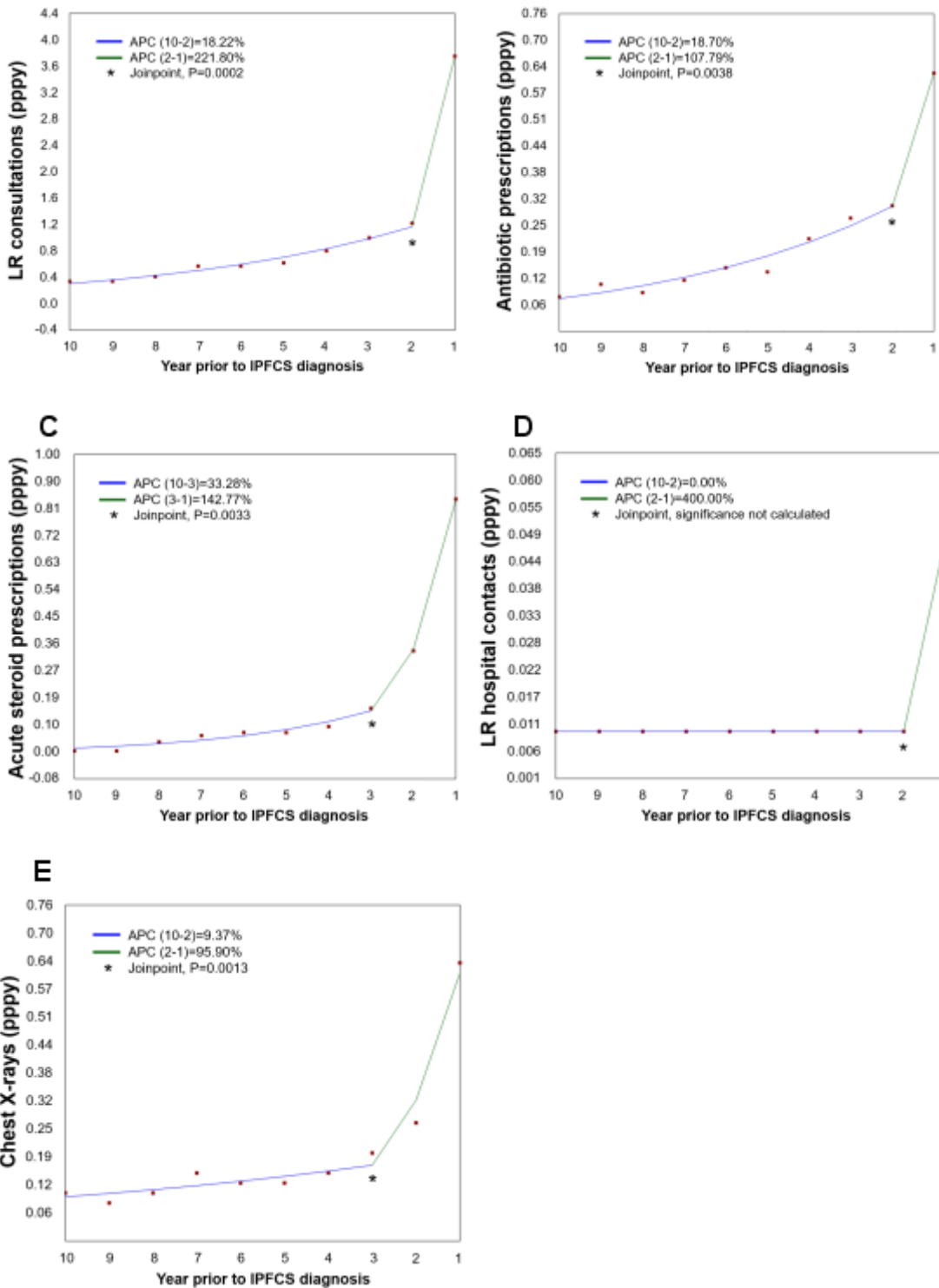


PFCS: pulmonary fibrosis clinical syndrome; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 183, 233, 289, 377, 648, 784, 861, 946, 999, 1034; Mean pre-observation period: 6 years.

Figure S3- Joinpoint regression analysis of temporal HRU trends over 10 years prior to PFCS diagnosis

A

B

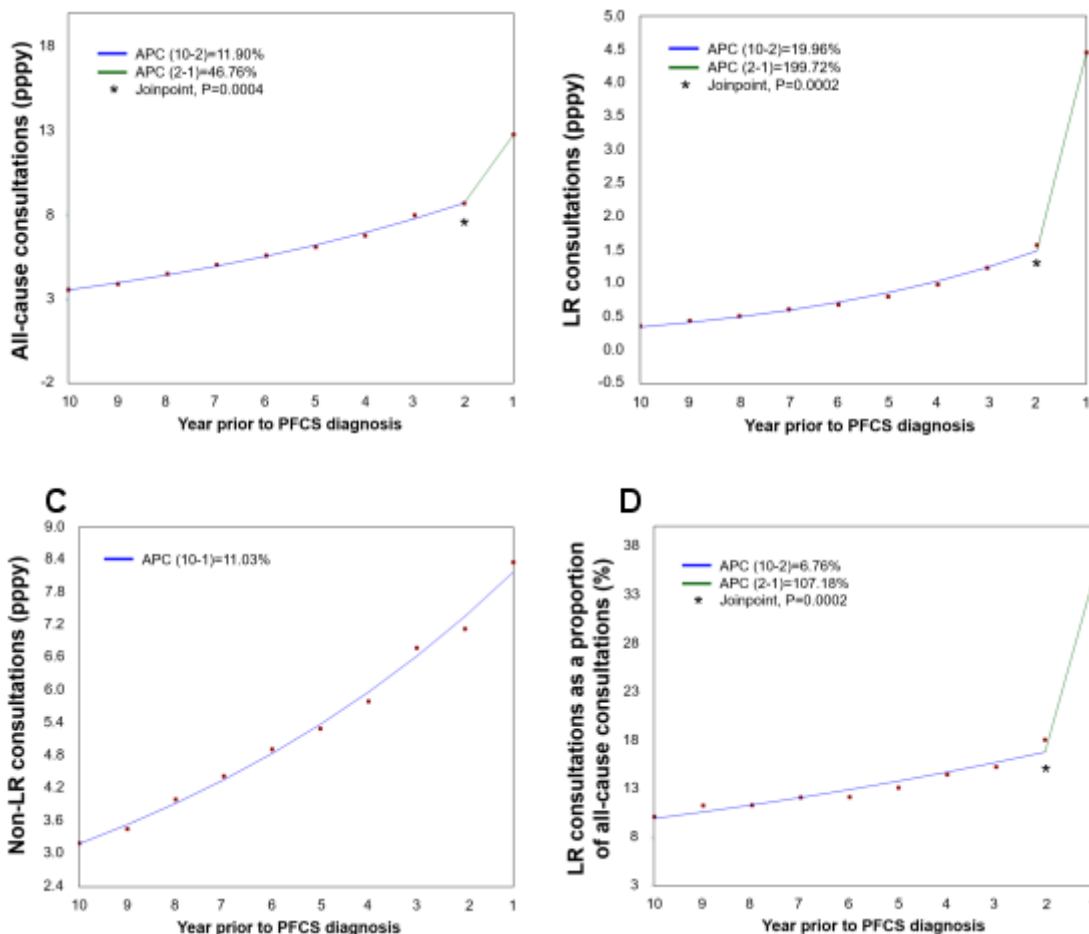


Temporal HRU trends in the 10 years prior to IPFCS diagnosis for A) lower respiratory (LR) consultations, B) acute antibiotic prescriptions, C) acute steroid prescriptions, D) LR hospital contacts, and E) chest X-rays. APC: annual percent change; Joinpoint: point of change in regression trends; IPFCS, idiopathic pulmonary fibrosis clinical syndrome; pppy: per patient per year; Number of patients with non-missing data in year 10 to year 1 prior to IPFCS diagnosis: 473, 488, 513, 540, 570, 601, 629, 658, 708, 743; Mean pre-observation period: 8 years.

Figure S4- Relative contribution of LR consultations to all-cause consultations over 10 years prior to PFCS diagnosis

A

B



Temporal HRU trends in the 10 years prior to PFCS diagnosis for A) all-cause consultations, B), lower respiratory (LR) consultations, C) non-LR consultations, D) LR consultations as a proportion of all-cause consultations; APC: annual percent change; Joinpoint: point of change in regression trends; pppy: per patient per year; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; mean pre-observation period: 8 years.

Table S1- Diagnostic Code list for PFCS and IPFCS cohorts

Read Code	Read Term
H563100*	Diffuse pulmonary fibrosis
H563200*	Pulmonary fibrosis
H563.00	Idiopathic fibrosing alveolitis
H563.12†	Cryptogenic fibrosing alveolitis
H563z00	Idiopathic fibrosing alveolitis NOS
H563300	Usual interstitial pneumonitis
H563.13	Idiopathic pulmonary fibrosis
XE0Yb	Idiopathic pulmonary fibrosis
X102v	Usual interstitial pneumonitis
H563z	Idiopathic fibrosing alveolitis NOS

PFCS cohort included all Read Codes shown above
 IPFCS cohort included all Read Codes except those designated *
 † includes H563.11 (Hamman-Rich syndrome)

Table S2- Exclusion Read code list for pulmonary fibrosis aetiologies

Read code	Description
N040G00	Rheumatoid arthritis of subtalar joint
N040D00	Rheumatoid arthritis of knee
N04y012	Fibrosing alveolitis associated with rheumatoid arthritis
N043100	Acute polyarticular juvenile rheumatoid arthritis
N040R00	Rheumatoid nodule
N040C00	Rheumatoid arthritis of sacro-iliac joint
N043300	Monarticular juvenile rheumatoid arthritis
2G27.00	O/E-hands-rheumatoid spindling
N040000	Rheumatoid arthritis of cervical spine
Nyu1200	[X]Other specified rheumatoid arthritis
N040A00	Rheumatoid arthritis of DIP joint of finger
N040400	Rheumatoid arthritis of acromioclavicular joint
N040700	Rheumatoid arthritis of wrist
N040800	Rheumatoid arthritis of MCP joint
N044.12	Nodular fibrositis of chronic rheumatic disease
N040S00	Rheumatoid arthritis - multiple joint
Nyu1100	[X]Other seropositive rheumatoid arthritis
N040M00	Rheumatoid arthritis of IP joint of toe
N040F00	Rheumatoid arthritis of ankle
N044.00	Chronic post-rheumatic arthropathy
N040K00	Rheumatoid arthritis of 1st MTP joint
N040N00	Rheumatoid vasculitis
N040P00	Seronegative rheumatoid arthritis
N040900	Rheumatoid arthritis of PIP joint of finger
N040H00	Rheumatoid arthritis of talonavicular joint
N040T00	Flare of rheumatoid arthritis
N040J00	Rheumatoid arthritis of other tarsal joint
N042200	Rheumatoid nodule
N040300	Rheumatoid arthritis of sternoclavicular joint
N042.00	Other rheumatoid arthropathy + visceral/systemic involvement
N043.00	Juvenile rheumatoid arthritis - Still's disease
N047.00	Seropositive erosive rheumatoid arthritis
N040L00	Rheumatoid arthritis of lesser MTP joint
N042000	Rheumatic carditis
N045500	Juvenile rheumatoid arthritis
N043200	Pauciarticular juvenile rheumatoid arthritis
14G1.00	H/O: rheumatoid arthritis
N040E00	Rheumatoid arthritis of tibio-fibular joint
N04y000	Rheumatoid lung
F396400	Myopathy due to rheumatoid arthritis
N040100	Other rheumatoid arthritis of spine
N040200	Rheumatoid arthritis of shoulder
N04..00	Rheumatoid arthritis and other inflammatory polyarthropathy
N04X.00	Seropositive rheumatoid arthritis, unspecified
N040500	Rheumatoid arthritis of elbow
N043z00	Juvenile rheumatoid arthritis NOS
N040600	Rheumatoid arthritis of distal radio-ulnar joint
N040B00	Rheumatoid arthritis of hip
H570.00	Rheumatoid lung
N040.00	Rheumatoid arthritis
N040Q00	Rheumatoid bursitis
N042100	Rheumatoid lung disease
Nyu1000	[X]Rheumatoid arthritis+involvement/other organs or systems
Nyu1G00	[X]Seropositive rheumatoid arthritis, unspecified

N043000	Juvenile rheumatoid arthropathy unspecified
F371200	Polyneuropathy in rheumatoid arthritis
N042z00	Rheumatoid arthropathy + visceral/systemic involvement NOS
N001000	Progressive systemic sclerosis
Q471100	Scleroderma in newborn
F396600	Myopathy due to scleroderma
Nyu4500	[X]Other forms of systemic sclerosis
M210z00	Circumscribed scleroderma NOS
N001.12	Systemic sclerosis
H572.00	Lung disease with systemic sclerosis
N001.00	Scleroderma
M210400	Linear scleroderma
M210000	Unspecified circumscribed scleroderma
N001100	CREST syndrome
M210.00	Circumscribed scleroderma
M154200	Lupus erythematosus migrans
N....00	Musculoskeletal and connective tissue diseases
N000000	Disseminated lupus erythematosus
Nyu4400	[X]Other dermatomyositis
N003.11	Poikilodermatomyositis
F396100	Myopathy due to disseminated lupus erythematosus
N003100	Dermatopolymyositis in neoplastic disease
N....11	Connective tissue diseases
K01x400	Nephrotic syndrome in systemic lupus erythematosus
N000400	Systemic lupus erythematosus with pericarditis
H57y100	Lung disease with polymyositis
Nyu4300	[X]Other forms of systemic lupus erythematosus
N000z00	Systemic lupus erythematosus NOS
F371000	Polyneuropathy in disseminated lupus erythematosus
H57y400	Lung disease with systemic lupus erythematosus
Nz...00	Musculoskeletal and connective tissue diseases NOS
Nyu4F00	[X]Mixed connective tissue disease
Myu7800	[X]Other local lupus erythematosus
M154600	Lupus erythematosus unguium mutilans
M154300	Lupus erythematosus nodularis
M154.00	Lupus erythematosus
N000.00	Systemic lupus erythematosus
Nyu4800	[X]Dermato(poly)myositis in neoplastic disease CE
M154000	Lupus erythematosus chronicus
Nyu4E00	[X]Dermatopolymyositis, unspecified
M154700	Subacute cutaneous lupus erythematosus
N00..11	Collagen diseases
N000500	Neonatal lupus erythematosus
M154400	Lupus erythematosus profundus
N00z.00	Collagen disease NOS
M154100	Discoid lupus erythematosus
N004.00	Polymyositis
N000300	Systemic lupus erythematosus with organ or sys involv
M154z00	Lupus erythematosus NOS
M154500	Lupus erythematosus tumidus
N003X00	Dermatopolymyositis, unspecified
N003000	Juvenile dermatomyositis
N00y.00	Other specified diffuse collagen diseases
K01x411	Lupus nephritis
N003.00	Dermatomyositis
F4D3300	Eyelid discoid lupus erythematosus
AD5..00	Sarcoidosis
AD50.00	Sarcoidosis of lung
AD51.00	Sarcoidosis of lymph nodes
AD52.00	Sarcoidosis of lung with sarcoidosis of lymph nodes
AD53.00	Sarcoidosis of skin

AD54.00	Sarcoidosis of inferior turbinates
AD55.00	Sarcoid arthropathy
Cyu0600	[X]Sarcoidosis of other and combined sites
F013.00	Meningitis due to sarcoidosis
F326300	Multiple cranial nerve palsies in sarcoidosis
F374900	Polyneuropathy in sarcoidosis
F396500	Myopathy due to sarcoidosis
G558300	Sarcoid heart disease
G5y7.00	Sarcoid myocarditis
H57y200	Pulmonary sarcoidosis
J63A.00	Hepatic granulomas in sarcoidosis
N233200	Myositis in sarcoidosis
Nyu8900	[X]Myositis in sarcoidosis classified elsewhere
H4...11	Pneumoconioses
H40..00	Coal workers' pneumoconiosis
H42..00	Silica and silicate pneumoconiosis
H420.00	Talc pneumoconiosis
H42z.00	Silica pneumoconiosis NOS
H43..00	Pneumoconiosis due to other inorganic dust
H43z.00	Pneumoconiosis due to inorganic dust NOS
H45..00	Pneumoconiosis NOS
Hyu4000	[X]Pneumoconiosis due to other dust containing silica
Hyu4100	[X]Pneumoconiosis due to other specified inorganic dusts
H41..00	Asbestosis
H41z.00	Asbestosis NOS
H35..00	Extrinsic allergic alveolitis
H35y.00	Other allergic alveolitis
H35yz00	Other allergic alveolitis NOS
H35z.00	Allergic alveolitis and pneumonitis NOS
H35z000	Allergic extrinsic alveolitis NOS
H35zz00	Allergic alveolitis and pneumonitis NOS

Table S3- Baseline characteristics of patients in idiopathic pulmonary fibrosis clinical syndrome (IPFCS) cohort

Baseline characteristic	IPFCS cohort (n = 743)
Male sex, n (%)	484 (65.1)
Age at index date (years), mean (SD)	72.2 (9.6)
BMI (kg/m²), mean (SD)	26.9 (4.9)
Smoking status	
Never	223 (32.8)
Current	94 (13.8)
Former	364 (53.5)
Comorbidities[†], n (%)	
COPD	142 (19.1)
Asthma	109 (14.6)
Ischaemic heart disease	218 (29.3)
Heart failure	81 (10.9)
Hypertension	29 (3.9)
Myocardial infarction	72 (9.7)
Lung cancer	109 (14.7)
Sleep apnoea	6 (0.8)
GERD*	40 (5.4)
Anxiety and depression	18 (2.4)
Time from comorbidity diagnosis to IPFCS diagnosis (years), mean (SD)	
COPD	4.6 (6.2)
Asthma	13.0 (17.4)
Any prescriptions in year prior to IPFCS diagnosis, n (%)	

SABA, n (%)	131 (17.6)
SAMA, n (%)	43 (5.8)
ICS, n (%)	54 (6.3)
ICS/LABA, n (%)	54 (6.3)
Lung Function	IPFCS Cohort Subgroup (n = 76)
FVC (L), mean (SD)	
All patients	2.5 (0.9)
Patients with comorbid COPD	2.6 (0.9)
Patients without comorbid COPD	3.1 (6.8)
FEV₁/FVC, mean (SD)	
All patients	0.81 (0.1)
Patients with comorbid COPD	0.70 (0.2)
Patients without comorbid COPD	0.81 (0.1)

IPFCS: idiopathic pulmonary fibrosis clinical syndrome; † Comorbidities defined as ever-recorded; * Active comorbidities defined as recorded within the last 2 years; SABA: short-acting beta-agonist; SAMA: short-acting antimuscarinic; ICS: inhaled corticosteroid; GERD: gastroesophageal reflux disease; mMRC: mini Medical Research Council

1

Table S4-Temporal trends in healthcare resource utilisation over 10 years prior to IPFCS diagnosis (IPFCS cohort)

	Year prior to IPFCS diagnosis									
	10	9	8	7	6	5	4	3	2	1
n (non-missing)	473	488	513	540	570	601	629	658	708	743
Primary care, mean (SD) per patient per year										
LR consultations	0.34 (1.0)	0.34 (1.0)	0.41 (1.2)	0.57 (1.6)	0.57 (1.4)	0.62 (1.35)	0.80 (1.8)	1.0 (2.0)	1.22 (0.9)	3.75 (1.5)
Prescriptions issued for LR complaints, mean (SD) per patient per year										
Antibiotics	0.08 (0.4)	0.11(0.4)	0.09 (0.4)	0.12 (0.4)	0.15 (0.5)	0.14 (0.4)	0.22 (0.7)	0.27 (0.8)	0.3 (0.3)	0.62 (0.5)
Acute oral steroid	0.01 (0.2)	0.01(0.1)	0.04 (0.3)	0.06 (0.4)	0.07 (0.4)	0.07 (0.4)	0.09 (0.5)	0.15 (0.7)	0.34 (0.4)	0.84 (0.8)
Secondary care, mean (SD) per patient per year										
LR hospital admissions*	0.01 (0.2)	0.01 (0.2)	0.01 (0.2)	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)	0.01 (0.2)	0.01 (0.1)	0.01 (0.1)	0.05 (0.2)
LR hospital admissions (sen)†	0.01 (0.23)	0.01 (0.2)	0.01 (0.2)	0.01 (0.1)	0.01 (0.1)	0.01 (0.2)	0.02 (0.2)	0.01 (0.2)	0.01 (0.01)	0.07 (0.2)
Diagnostics, mean (SD) per patient per year										
Chest X-Rays	0.04 (0.2)	0.03 (0.1)	0.04 (0.2)	0.06 (0.3)	0.05 (0.3)	0.05 (0.2)	0.06 (0.3)	0.08 (0.3)	0.11 (0.2)	0.27 (0.4)

PFCS: pulmonary fibrosis clinical syndrome; SD: standard deviation; * Lower respiratory (LR) code recorded on same day as inpatient admission; sen, sensitivity; † LR code recorded within 14 days of an inpatient admission; number of patients with non-missing data in year 10 to year 1 prior to IPFCS diagnosis: 473, 488, 513, 540, 570, 601, 629, 658, 708, 743; Mean pre-observation period: 8 years.

1

Table S5- Distribution of lower respiratory healthcare consultations over 10 years prior to PFCS diagnosis

	Year prior to PFCS diagnosis									
	10	9	8	7	6	5	4	3	2	1

N, non-missing	1474	1533	1598	1686	1772	1861	1944	2019	2087	2223
Number (%) of LR consultations per year										
0	1191 (82.0)	1228 (81.9)	1262 (81.3)	1222 (74.8)	1261 (73.6)	1243 (69.3)	1226 (66.7)	1158 (62.0)	1055 (50.1)	2 (0.1)
≥1	171 (11.7)	159 (10.6)	160 (10.3)	237 (14.5)	252 (14.7)	293 (16.3)	330 (17.9)	352 (18.8)	360 (17.2)	471 (21.1)
≥2	57 (3.9)	63 (4.2)	73 (4.7)	103 (6.3)	104 (6.1)	134 (7.5)	135 (7.3)	175 (9.3)	212 (10.2)	405 (18.2)
≥3	21 (1.4)	34 (2.3)	31 (2.0)	48 (2.9)	66 (3.9)	79 (4.4)	92 (5.0)	118 (6.3)	155 (7.4)	272 (12.2)
≥4	12 (0.8)	16 (1.1)	26 (1.7)	23 (1.4)	31 (1.8)	44 (2.5)	56 (3.0)	66 (3.5)	93 (4.4)	228 (10.2)
≥5	22 (1.5)	33 (2.2)	46 (3.0)	53 (3.2)	58 (3.4)	68 (3.8)	105 (5.7)	150 (8.0)	212 (10.1)	845 (38.0)

PFCS: pulmonary fibrosis clinical syndrome; Number of patients with non-missing data in year 10 to year 1 prior to PFCS diagnosis: 1474, 1533, 1598, 1686, 1772, 1861, 1944, 2019, 2087, 2223; Mean pre-observation period: 8 years.