BMJ Open Respiratory Research

Study investigating the generalisability of a COPD trial based in primary care (Salford Lung Study) and the presence of a Hawthorne effect

Alexander Pate,[1] Michael Barrowman,[1] David Webb,[2] Jeanne M Pimenta,[2] Kourtney J Davis,[3] Rachael Williams,[4] Tjeerd Van Staa,[1,5] Matthew Sperrin[1]

To cite: Pate A, Barrowman M, Webb D, *et al*. Study investigating the generalisability of a COPD trial based in primary care (Salford Lung Study) and the presence of a Hawthorne effect. *BMJ Open Resp Res* 2018;**5**:e000339. doi:10.1136/bmjresp-2018-000339

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2018-000339).

Received 17 July 2018
Revised 25 September 2018
Accepted 10 October 2018

For numbered affiliations see end of article.

Correspondence to
Dr Alexander Pate;
alexander.pate@manchester.ac.uk

ABSTRACT

Introduction Traditional phase IIIb randomised trials may not reflect routine clinical practice. The Salford Lung Study in chronic obstructive pulmonary disease (SLS COPD) allowed broad inclusion criteria and followed patients in routine practice. We assessed whether SLS COPD approximated the England COPD population and evidence for a Hawthorne effect.

Methods This observational cohort study compared patients with COPD in the usual care arm of SLS COPD (2012–2014) with matched non-trial patients with COPD in England from the Clinical Practice Research Datalink database. Generalisability was explored with baseline demographics, clinical and treatment variables; outcomes included COPD exacerbations in adjusted models and pretrial versus peritrial comparisons.

Results Trial participants were younger (mean, 66.7 vs 71.1 years), more deprived (most deprived quintile, 51.5% vs 21.4%), more current smokers (47.5% vs 32.1%), with more severe Global initiative for chronic Obstructive Lung Disease stages but less comorbidity than non-trial patients. There were no material differences in other characteristics. Acute COPD exacerbation rates were high in the trial population (98.37th percentile).

Conclusion The trial population was similar to the non-trial COPD population. We observed some evidence of a Hawthorne effect, with more exacerbations recorded in trial patients; however, the largest effect was observed through behavioural changes in patients and general practitioner coding practices.

Key messages

- The trial population was similar to the non-trial chronic obstructive pulmonary disease (COPD) population; we observed evidence of a Hawthorne effect, with more exacerbations recorded in trial patients, but this effect was mitigated by supporting evidence from secondary analyses.
- There was further evidence of a Hawthorne effect through behavioural changes in patients and general practitioner coding practices.
- This study develops novel methods to evaluate the presence of a Hawthorne effect operating for a trial, such as the Salford Lung Study in COPD, which was conducted in the setting of everyday clinical practice; to our knowledge, this is the first study of its kind.

INTRODUCTION

Conventional double-blind, randomised controlled trials (RCTs) test the efficacy and safety of an intervention to identify the presence and size of a pharmacological effect. They often have highly restricted entry criteria to reduce heterogeneity, intensive follow-up schedules and monitoring, and proactively encourage adherence to study medication. The generalisability of such trials to everyday practice is therefore questionable.[1 2] This could be overcome by conducting trials with more inclusive entry criteria in an environment that reflects everyday clinical practice. However, the true generalisability of such 'pragmatic' trials is unknown.

The Salford Lung Study in chronic obstructive pulmonary disease (SLS COPD) was a phase IIIb RCT conducted in UK primary care that evaluated the clinical effectiveness and safety of initiating once-daily inhaled fluticasone furoate/vilanterol (FF/VI) 100/25 µg versus continuing usual care (UC) in patients with COPD and exacerbation history.[3] The trial was conducted in Salford and surrounding areas in Greater Manchester, England, which has an established electronic health record (EHR) system, connecting primary-care and secondary-care EHRs so that participants could be closely monitored with minimal intrusion. Novel characteristics of the trial included broad inclusion/minimal exclusion criteria, treatment administration in routine clinical practice, few protocol-mandated clinic visits, patients accessing their medication through their usual general practitioner (GP)/pharmacy, and control patients continuing on UC, which could be

modified at their GP's discretion.[4] While SLS COPD was designed to test effectiveness in the routine care setting, two concerns were raised with regard to extrapolating findings to everyday clinical practice. First, the trial population may not be representative of the wider COPD population; and second, study participant behaviour as assessed by the Hawthorne effect[5 6] may introduce bias, thereby affecting outcomes in both treatment arms.[7]

The Hawthorne effect is a phenomenon whereby participants or practitioners modify their behaviour due to an awareness of being observed.[8 9] Several RCTs have explored this effect,[10 11] with mixed findings on when and how it operates. New concepts and techniques are required to test for this phenomenon in a consistent manner.

The present study had two aims: to evaluate how representative the SLS COPD population was of the wider COPD population in England; and to evaluate the potential Hawthorne effect in the trial setting by comparing COPD outcomes in the UC arm with those in the Clinical Practice Research Datalink (CPRD) primary care database, and by comparing COPD outcomes and other measures (primary care contact, COPD prescription use and treatment switching) over time in the UC arm.

METHODS

Study design and participants

This retrospective, observational cohort study compared SLS COPD patients with selected cohorts of patients with COPD in a non-trial population. The trial cohort comprised patients randomised to the UC arm and data were obtained from patients' primary care EHRs and from the trial database. The comparator cohort was derived from the CPRD, a provider of primary care data from a broadly representative sample of practices in England.[12] The main comparator population was restricted to practices outside Greater Manchester (to avoid duplicating patients enrolled in the trial). We then selected CPRD patients who had a coded COPD diagnosis in primary care (codelist found at https://www.gsk-clinicalstudyregister.com/), ≥1 day of up-to-standard follow-up registration in CPRD, were aged ≥40 years, and who were eligible for linkage to Hospital Episode Statistics (HES), death records from the Office for National Statistics and patient postcode-derived Index of Multiple Deprivation (IMD) 2010 quintiles (a measure of socioeconomic status (SES)).[13] A third cohort of CPRD patients from within Greater Manchester was retained for sensitivity analyses. Although linkage to HES was an eligibility criterion, analyses involving secondary care data are not presented herein. CPRD patients were then assigned index dates by matching to SLS COPD UC patients based on SLS randomisation dates (5 April 2012–24 October 2014). At the assigned index date, CPRD patients were required to meet the main trial inclusion/exclusion criteria (except for excluding patients with life-limiting conditions, which could not be replicated using primary care data) and have 1 year of data before the index date. Further details of the matching strategy are described in online supplementary appendix 1. Final index date-matched comparator cohorts (matching ratios 3:2 and 12:1), hereafter referred to as CPRD-GM and CPRD-xGM, comprised patients who were or were not registered at a practice in Greater Manchester, respectively. Patients were followed for 1 year after the index date (randomisation date), or to death or loss to follow-up, whichever occurred first, to mimic the planned 12-month duration of the trial.

The study protocol (online supplementary appendix 2) was approved by the CPRD Independent Scientific Advisory Committee, protocol 15_059, by GSK's Real World Evidence and Epidemiology Protocol Review Forum, and by the ENCePP (EUPAS10376). SLS COPD was conducted in accordance with the International Conference on Harmonisation Good Clinical Practice guidelines and the provisions of the 2008 Declaration of Helsinki. All patients provided written informed consent. Although this study was based in part on data from the CPRD (obtained under licence from the UK Medicines and Healthcare products Regulatory Agency), the interpretation and conclusions are those of the authors alone.

Outcomes, exposures and confounders

The generalisability of SLS COPD was assessed using an analysis of the following covariates in the UC arm and the CPRD-xGM cohort: sex; age; smoking status; body mass index (BMI); IMD 2010 quintiles; Global initiative for chronic Obstructive Lung Disease (GOLD) stage (2007 classification scheme[14]); current medication group; history of comorbid conditions including cardiovascular disease, cerebrovascular disease, depression, anxiety, asthma, pneumonia, gastro-oesophageal reflux disease and peptic ulcer disease; Charlson Comorbidity Index (a predictive measure of mortality based on comorbidity[15] with COPD removed); number of COPD exacerbations in the previous 12 months; percentage of predicted forced expiratory volume in 1 s (FEV_1% predicted); FEV_1:forced vital capacity (FVC) ratio; Medical Research Council (MRC) dyspnoea score; and history of influenza and pneumococcal vaccinations.

For the purposes of assessing the Hawthorne effect, the rate of acute exacerbations of COPD (AECOPD) episodes over the 12-month period after the randomisation date (trial cohort) or index date (CPRD cohort) was chosen as the primary endpoint. COPD exacerbations were defined using a published algorithm[15] and comprised two parts: identifying events and generating episodes. Events were those that met any criteria from the validated algorithm[16]; events occurring close together were considered likely to be related and were combined into exacerbation episodes (hereafter referred to as 'AECOPD episodes'). Additional outcomes considered were hospitalised pneumonia; a 'strict' (more specific) definition of AECOPD episodes based on acute exacerbation medical codes only; time to first AECOPD; number of days of primary

care contact; number of trial-related prescription items, a binary variable to indicate whether patients switched treatment class during the trial; and mortality. Full definitions of the primary outcome and secondary outcomes are provided in online supplementary appendix 1.

Where possible, variables for the trial cohort were derived using the EHRs to maximise comparability with the CPRD groups. In particular, validated exacerbations reported in the trial database were not used in our analyses; rather, AECOPD episodes were derived algorithmically[16] to maintain comparability by using the EHRs from both sources. History of asthma, rather than current asthma, was used as a covariate in the models due to difficulties in differentiating historical versus current asthma in EHR data and for consistency with other comorbidities, which were all assessed as historical variables. Online supplementary table E1 details the data source used to derive variables for the trial cohort.

Statistical methods

To handle missing data, a single stochastic regression imputation was applied (multiple imputation with m=1).[17] Each cohort was imputed separately, with each model comprising all confounders and the main outcome variable. Imputed variables were IMD, FEV_1% predicted, FEV_1:FVC ratio, MRC dyspnoea score, BMI and smoking status. A complete case analysis was conducted as a sensitivity analysis.

Representativeness of SLS COPD

To evaluate the representativeness of SLS COPD, distributions of baseline demographics, clinical variables, prescribed COPD medications and outcomes were summarised for the trial UC arm and the matched CPRD-xGM cohort. We considered whether the distribution of variables in the trial was unusual in the context of variability between anonymised local authorities (LAs) in the CPRD (LAs are regional areas in the UK of comparable size with Salford). For continuous and binary variables, an empirical 2.5th–97.5th percentile range of LA means (proportions; 95% window) was constructed from CPRD-xGM. If the single mean value from the trial fell outside this range, it was considered unusual. For categorical variables, χ^2 test was performed testing each LA against the reference distribution (all other LAs combined). Test statistics from each test were used to construct a 0th–95th percentile range, and the trial value was deemed unusual if its test statistics lay outside this range.

Comparing outcomes during trial with non-trial population (testing for Hawthorne effect in the UC arm)

Multilevel models[18] were fitted to compare outcomes in the trial and CPRD-xGM cohorts. These models had two levels: patient and LA. Random intercepts were included at the LA level. The point estimates for the random intercepts associated with each LA in CPRD-xGM were combined to generate an empirical 2.5th–97.5th percentile range. Each point estimate is the relative rate of the LA compared with the average. If the estimate for the trial random intercept lay outside this 2.5–97.5 range, it was considered unusual. For the primary outcome (rate of AECOPD episodes), as well as several secondary outcomes (hospitalised pneumonia episodes, strict definition of AECOPD episodes, primary care contact days and number of trial-related prescription items), a Poisson multilevel model was applied to the data. Poisson models were used as the number of AECOPD episodes is a count variable; furthermore it allows adjustment for the time at risk (necessary as episodes and follow-up may vary in length) in this setting.[18 19] A Cox multilevel (frailty) model was fit to the time until the first AECOPD episode and the time to mortality.[20] A logistic multilevel model was fit to the binary variable indicating treatment switching.[18] In these cases the point estimates of the random effects can be interpreted as HRs and ORs compared with the average.

All models included the same set of covariates for consistency. Covariates were included if the likelihood ratio tests in the univariate analyses indicated that they were statistically significant predictors of exacerbations. GOLD stage was not considered for adjustment, as it is derived directly from FEV_1% predicted and FEV_1:FVC ratio, which were included. All continuous variables were modelled with linear and quadratic terms to allow for simple deviations from linearity. For parsimony, interactions between covariates were not considered. Groupings for categorical variables were chosen using pre-existing guidelines (eg, Charlson score).[21] In the Poisson models overdispersion was assessed; if present, generalised Poisson models were applied to the data.[22] Time at risk was incorporated as an offset in Poisson models and was incorporated as censoring in the Cox models. Kaplan-Meier plots of univariate Cox models were produced to assess the proportional hazards assumption.

'Difference in difference' comparison of primary and secondary outcomes before and during SLS COPD (further context for Hawthorne effect)

All primary and secondary outcomes, except hospitalised pneumonia, mortality and time to first exacerbation (excluded due to lack of data before trial or not methodologically possible), were compared using a 'difference in difference' approach in the year before and during the trial.[23] A 'period' variable was introduced, indicating whether the outcome was calculated in the year before or after the index date (ie, during the trial period). Multilevel models[18] (Poisson or logistic, as appropriate; see above) were applied; random intercepts were included at both the patient level (as there were two observations per patient) and the LA level. A random coefficient for the period variable was also included at the LA level and point estimates for the random coefficients were then calculated for each LA. These estimates are relative rates (or ORs for the binary outcome) comparing the two time

periods within each LA. The point estimates from each LA were combined to create a 2.5th–97.5th percentile range, and the position of the trial's random coefficient within this range was of interest.

All cohort derivation and analyses, with the exception of the 'difference in difference' comparison, were independently programmed by the University of Manchester and GSK using SAS/STAT V.9.4 software[24] for Windows. SAS and all other SAS Institute product or service names are registered trademarks or trademarks of SAS Institute (Cary, North Carolina, USA). The 'difference in difference' comparison of primary and secondary outcomes before and during SLS COPD (further context for Hawthorne effect) was conducted using R V.3.4.0[25]; this was programmed at the University of Manchester and the code was reviewed by quality control analysts at GSK.

RESULTS

Study population

The SLS COPD cohort comprised all patients in the UC arm (n=1403). An exclusion flow chart for the CPRD cohorts is provided in figure 1; the main comparison cohort, CPRD-xGM, comprised 16 758 patients. Imputed data were compared with complete data, and no significant differences were observed (online supplementary tables E2, E3 and E4).

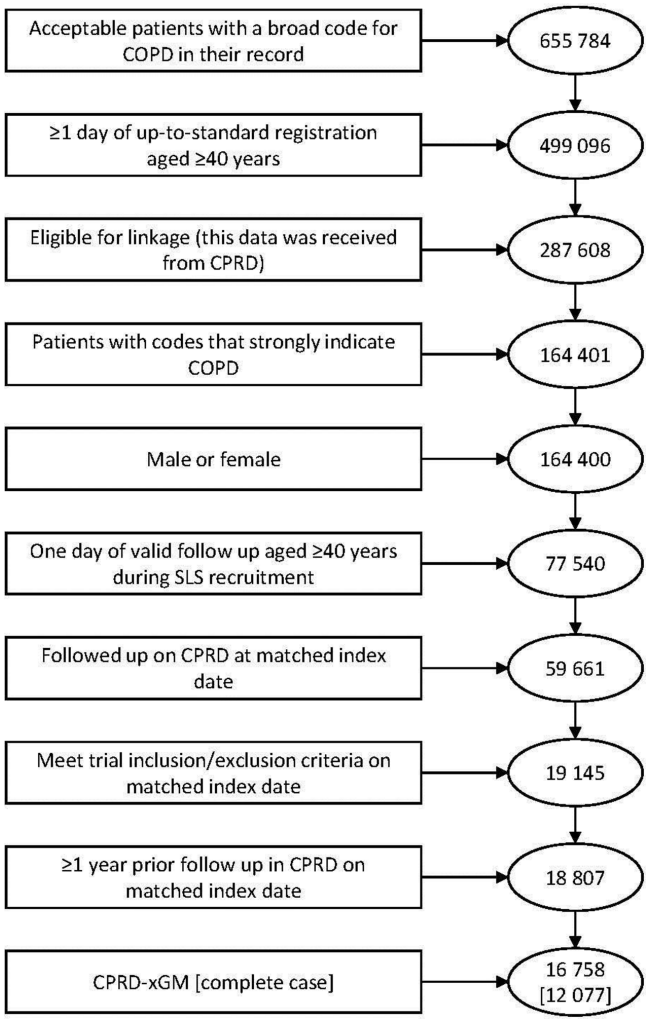

Figure 1 Flow chart for inclusion in CPRD cohort. COPD, chronic obstructive pulmonary disease; CPRD, Clinical Practice Research Datalink; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester; SLS, Salford Lung Study.

Baseline comparisons

The summary statistics for a selected group of covariates are presented in table 1 (other covariates are presented in online supplementary table E5). Figure 2 provides a graphical representation of table 1 for some key covariates. Both the CPRD-GM and trial cohorts were more deprived than the CPRD-xGM cohort, and the trial cohort contained more current smokers than the CPRD-GM and CPRD-xGM cohorts. The severity of airflow limitation in SLS COPD patients assessed according to GOLD stage was more severe than both CPRD cohorts. However, MRC dyspnoea scores of the trial cohort were less severe than the CPRD cohorts.

Figure 3 indicates variables in the trial cohort that were deemed unusual with respect to regional variation in CPRD patients. Age was the only continuous variable considered unusual, with the mean SLS COPD patient age below the 2.5th percentile; comorbidity histories, vaccination history, previous COPD exacerbations, FEV_1% predicted and FEV_1:FVC ratio were all within the usual range (figure 3A). All categorical variables except BMI were deemed unusual (figure 3B).

Outcome comparisons

The rate of AECOPD episodes per person-year was higher in the trial cohort (rate, 1.91; 95% CI 1.83 to 1.99) than the CPRD-GM cohort (rate, 1.63; 95% CI 1.57 to 1.69) and the CPRD-xGM cohort (rate, 1.53; 95% CI 1.51 to 1.56; table 2). This was also seen for the 'strict' AECOPD definition, where rates were higher in trial patients. The mortality rate was lower in trial patients, while primary care usage was higher, with almost twice as many COPD medication prescriptions per patient per year and a lower proportion of treatment switching versus CPRD-GM and CPRD-xGM patients (table 2).

The results from the adjusted multilevel models of outcomes are presented in table 3. These random effects were expressed as relative rates in Poisson models, as HRs in Cox models and as ORs in logistic models. Generalised Poisson models were used for all count outcomes, as the data were overdispersed. For AECOPD episodes, the trial's random effect fell at the 98.37th and 96.70th percentiles for the Poisson and Cox models, respectively. This indicates unusual exacerbation rates and a high, but not unusual, HR for the trial versus CPRD. The primary analysis was also carried out using the cohorts with complete data (online supplementary table E6), with no large differences found (94.06th and 94.01th

Table 1 Clinical and demographic characteristics of CPRD-GM/CPRD-xGM patients and SLS COPD participants

Variable	Category	CPRD-GM	CPRD-xGM	SLS COPD
n		2049	16 758	1403
Total follow-up (mean)	Years	1841.60 (0.90)	14 587 (0.87)	1370.92 (0.98)
Sex	Female	1180 (57.59%)	8595 (51.29%)	671 (47.83%)
Age	Mean (95% CI)	69.77 (69.30 to 70.23)	71.12 (70.96 to 71.29)	66.73 (66.21 to 67.25)
	Median (range, 2.5%–97.5%)	69.96 (47.57–89.22)	71.41 (48.14–90.21)	67.00 (46.00–85.00)
SES IMD 2010 quintiles	Missing	0 (0.00%)	8 (0.05%)	8 (0.57%)
	5 (least deprived)	127 (6.20%)	2499 (14.91%)	72 (5.13%)
	4	219 (10.69%)	3428 (20.46%)	105 (7.48%)
	3	283 (13.81%)	3348 (19.98%)	202 (14.40%)
	2	456 (22.25%)	3897 (23.25%)	294 (20.96%)
	1 (most deprived)	964 (47.05%)	3578 (21.35%)	722 (51.46%)
Current medication (prescriptions in the last 3 months)	None of the below treatments in the last 3 months*	165 (8.05%)	1693 (10.10%)	145 (10.33%)
	LABA only	32 (1.56%)	252 (1.50%)	19 (1.35%)
	LAMA only	174 (8.49%)	1480 (8.83%)	143 (10.19%)
	ICS only	114 (5.56%)	758 (4.52%)	62 (4.42%)
	LABA/LAMA	25 (1.22%)	177 (1.06%)	18 (1.28%)
	LAMA/ICS	24 (1.17%)	235 (1.40%)	56 (3.99%)
	LABA/ICS	594 (28.99%)	4529 (27.03%)	337 (24.02%)
	LABA/LAMA/ICS	921 (44.95%)	7634 (45.55%)	623 (44.40%)
Current asthma		485 (23.67%)	4018 (23.98%)	298 (21.24%)
Comorbidities (history)	Anxiety	572 (27.92%)	3661 (21.85%)	301 (21.45%)
	Asthma	1213 (59.20%)	10 083 (60.17%)	755 (53.81%)
	CVD	387 (18.89%)	3222 (19.23%)	238 (16.96%)
	Depression	767 (37.43%)	5466 (32.62%)	344 (24.52%)
	GORD/peptic ulcer	571 (27.87%)	4120 (24.59%)	355 (25.30%)
	Pneumonia	282 (13.76%)	2635 (15.72%)	147 (10.48%)
Exacerbation history in the previous 12 months†	Events	3262	24 892	2372
	Rate (95% CI)	1.80 (1.74 to 1.86)	1.66 (1.64 to 1.68)	1.94 (1.86 to 2.02)
FEV_1%	Mean (95% CI)	56.72 (55.83 to 57.61)	55.84 (55.53 to 56.16)	60.30 (59.17 to 61.43)
	Median (range, 2.5%–97.5%)	55.80 (22.68–95.23)	55.07 (22.34–95.97)	60.90 (24.30–98.90)
	Missing (%)	12.79	14.25	21.53
FEV_1: FVC (%)	Mean (95% CI)	61.06 (60.24 to 61.88)	60.51 (60.22 to 60.80)	54.39 (53.58 to 55.19)
	Median (range, 2.5%–97.5%)	60.60 (32.20–95.00)	60.00 (31.00–95.70)	54.80 (28.65–79.09)
	Missing (%)	23.13	21.58	21.53
GOLD stage	Missing	479 (23.38%)	3783 (22.57%)	217 (15.47%)
	0 (FEV_1:FVC≥70)	451 (22.01%)	3589 (21.42%)	147 (10.48%)
	1 (FEV_1:FVC<70, FEV_1%≥80)	70 (3.42%)	522 (3.11%)	84 (5.99%)
	2 (FEV_1:FVC<70, 50≤FEV_1%<80)	527 (25.72%)	4347 (25.94%)	522 (37.21%)
	3 (FEV_1:FVC<70, 30≤FEV_1%<50)	422 (20.60%)	3528 (21.05%)	332 (23.66%)
	4 (FEV_1:FVC<70, FEV_1%<30)	100 (4.88%)	989 (5.90%)	101 (7.20%)
Smoking	Never	165 (8.05%)	1349 (8.05%)	59 (4.21%)
	Ex	1177 (57.44%)	10 033 (59.87%)	678 (48.33%)
	Current	707 (34.50%)	5376 (32.08%)	666 (47.47%)

Continued

Table 1 Continued

Variable	Category	CPRD-GM	CPRD-xGM	SLS COPD

*Data from electronic health records.
†Exacerbation history is treated as a rate per person-year.
CPRD-GM, Clinical Practice Research Datalink in Greater Manchester; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester; CVD, cardiovascular and cerebrovascular diseases (specifically heart failure, myocardial infarction and stroke); FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GOLD, Global initiative for chronic Obstructive Lung Disease; GORD, gastro-oesophageal reflux disease; ICS, inhaled corticosteroid; IMD, Index of Multiple Deprivation; LABA, long-acting β2-agonist; LAMA, long-acting muscarinic antagonist; SES, socioeconomic status; SLS COPD, Salford Lung Study in chronic obstructive pulmonary disease.

percentiles, respectively). The percentile for the rate of strict AECOPD episodes was similar (97.36th percentile). The number of primary care contact days was unusually high, while mortality was unusually low for the trial cohort; rates of treatment switching and prescription counts were not unusual. Descriptive modelling illustrated that the crude/unadjusted rate of hospitalised pneumonia was lower in the SLS COPD UC cohort, but the rate of pneumonia in SLS COPD was similar to the CPRD in the fully adjusted multilevel models (online tables E7 and E8).

The results from the comparison of exacerbation rates before and during SLS COPD (difference in difference analysis), using fully adjusted multilevel models, are presented in table 4. The change in exacerbation rate in the trial cohort approximated that in the CPRD-xGM cohort (64.17th percentile), whereas for the strict AECOPD definition the trial cohort was deemed unusual (100th percentile), indicating a large increase in recording of AECOPD codes. There was a large drop in the rate of COPD-related prescriptions in SLS COPD during the trial compared with CPRD (0th percentile) and a decrease in treatment switching (0.27th percentile). Despite the rate of contact with primary care in the year during the trial being unusually high, the change in primary care contact days was similar to the average change observed in CPRD patients (57.06th percentile).

DISCUSSION

This study has contextualised SLS COPD by evaluating the representativeness of the patient population and the potential Hawthorne effect of the trial. We found similarity between the trial UC population and the potentially trial-eligible population across England in terms of sex, comorbidity histories, vaccination history, previous

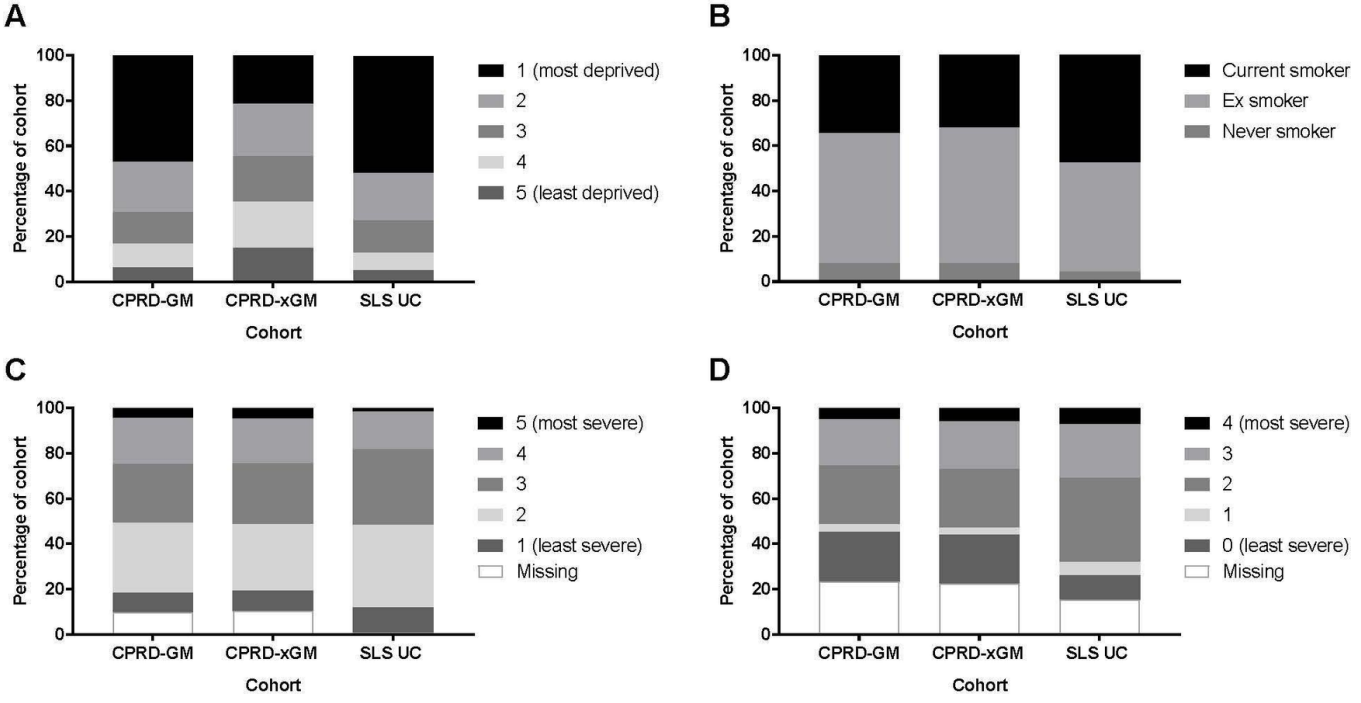

Figure 2 Stacked bar charts for key predictor variables (representativeness of SLS COPD). (A) IMD 2010 quintiles; (B) smoking status; (C) MRC dyspnoea score; (D) GOLD stage. COPD, chronic obstructive pulmonary disease; CPRD-GM, Clinical Practice Research Datalink in Greater Manchester; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester; GOLD, Global initiative for chronic Obstructive Lung Disease; IMD, Index of Multiple Deprivation; MRC, Medical Research Council; SLS, Salford Lung Study; SLS UC, Salford Lung Study usual care arm.

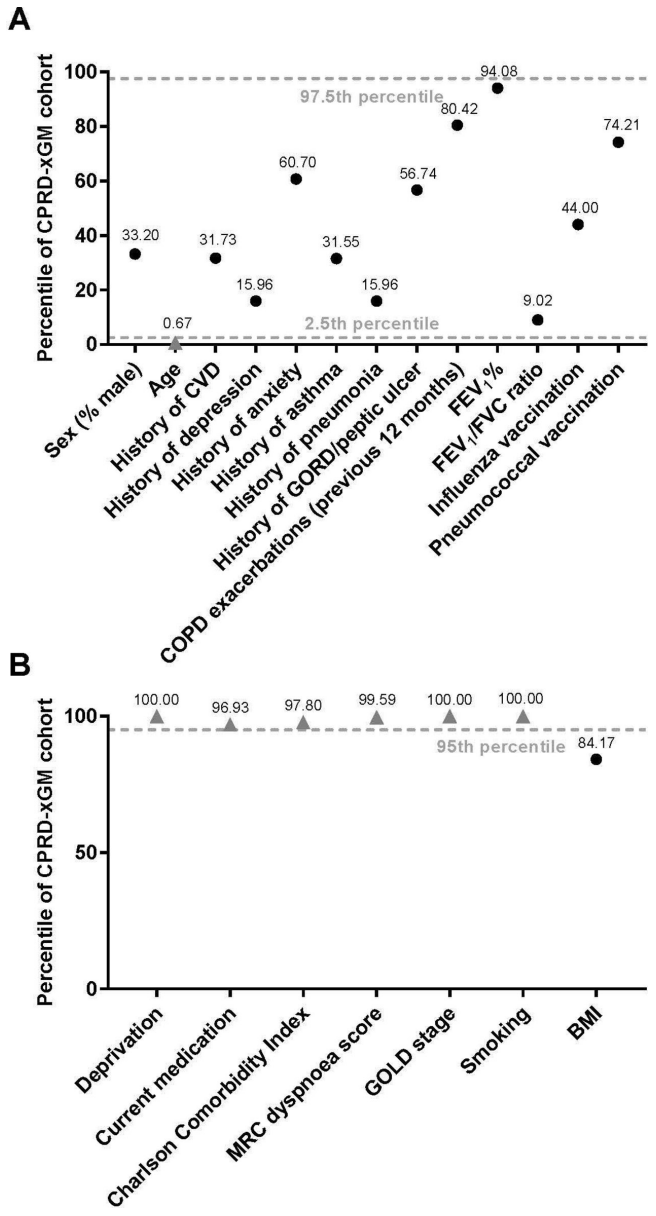

Figure 3 Percentiles of (A) continuous or binary variables and (B) categorical variables in the trial, in the context of regional variation (representativeness of SLS COPD). Variables considered unusual are shown as grey triangles. BMI, body mass index; COPD, chronic obstructive pulmonary disease; CPRD-xGM, Clinical Practice Research Datalink outside of greater Manchester; CVD, cardiovascular and cerebrovascular diseases (specifically heart failure, myocardial infarction and stroke); FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GOLD, Global initiative for chronic Obstructive Lung Disease; GORD, gastro-oesophageal reflux disease; MRC, Medical Research Council; SLS COPD, Salford Lung Study in chronic obstructive pulmonary disease.

COPD exacerbations, FEV_1% predicted, FEV_1:FVC ratio and BMI, but with some differences in age (younger), SES (more deprived), smoking status (higher), current medication (higher), Charlson index (lower), MRC dyspnoea score (lower) and GOLD stage (higher). Most of these differences are reflective of differences between patients in Greater Manchester versus the rest of England (eg, SES scores).[13] SLS COPD showed unusually high rates of current smoking, even compared with CPRD-GM patients. However, overall the trial cohort and the cohort of patients with COPD across England were broadly comparable. History of asthma is high in these cohorts, which may reflect poorly diagnosed COPD or childhood wheeze, while current asthma is much lower.

Concerning the Hawthorne effect evaluation, the rate of AECOPD episodes among SLS UC patients was high (98.37th percentile), but did not appear to be triggered by the trial; exacerbation rate did not change by an unusual amount after the trial began (64.17th percentile). There was, however, a significant increase in the rate of strictly defined AECOPD episodes recorded (100th percentile). This provides some evidence of a Hawthorne effect whereby exacerbations were more likely to be detected in the trial population and/or more likely to be explicitly recorded.

Mortality was comparatively low in the trial (0th percentile); however, we were unable to replicate the preferential selection of patients with fewer life-limiting conditions when selecting the CPRD cohort. There was an unusual drop in treatment switching during the year of the trial (0.27th percentile), which could be explained by rationalisation of multiple treatments into a single form at the study start date. There was also a reduction in prescriptions from the pretrial year to the trial year (0th percentile), possibly also due to treatment rationalisation, but also GPs bringing forward prescriptions to the SLS COPD start date or changes in prescription length. Of note, in the trial, the index date was a GP visit; patients in the CPRD cohort will not have had a comparable exposure on that date that may have precipitated a change in treatment regimen. These results highlight the importance of the 'difference in differences' analysis, as the rate of prescriptions during the trial was comparatively high (87.88th percentile), but had decreased compared with the previous year. This is in contrast to the rate of primary care contact, which appeared unusually high in the primary analysis, but was consistent in the years before and during the trial.

A strength of this study is that we used data, and variable definitions, that were broadly comparable between the trial and routine-care contexts, and a validated algorithm to define the primary outcome of AECOPD episodes. However, the EHR data collected in routine care, used in both the trial and comparator cohorts, present challenges for use in clinical research. In particular, we were unable to reliably ascertain medication adherence variables such as medication possession ratio because of incomplete capture of detailed prescription information.[26] EHR data of trial patients were derived from both EMIS and VISION software, whereas CPRD data comprised only VISION practices. Therefore, some differences could be explained by software—for example, higher coding rates with EMIS (online supplementary tables E9 and E10).

Table 2 Crude counts and rates of outcome variables (testing for the presence of a Hawthorne effect)

Variable	Category	CPRD-GM	CPRD-xGM	SLS COPD
n		2049	16 758	1403
Count of AECOPD episodes*	Number of events	2693	20 184	2288
	Rate per person-year (95% CI)	1.63 (1.57 to 1.69)	1.53 (1.51 to 1.56)	1.91 (1.83 to 1.99)
Count of strict definition of AECOPD† episodes	Number of events	583	3497	766
	Rate per person-year (95% CI)	0.32 (0.30 to 0.35)	0.24 (0.24 to 0.25)	0.58 (0.54 to 0.62)
Mortality	Number of cases	144	1145	24
	Rate per 1 000 person-years (95% CI)	78.19 (65.94 to 92.06)	78.49 (74.01 to 83.17)	17.51 (11.22 to 26.05)
Number of days contact with primary care	Count	48 167	399 291	66 308
	Rate per person-year (95% CI)	26.19 (25.96 to 26.43)	27.37 (27.29 to 27.46)	49.11 (48.74 to 49.49)
Countof COPD-related prescription items‡	Count	25 433	187 396	32 133
	Rate per person-year (95% CI)	13.81 (13.64 to 13.98)	12.85 (12.79 to 12.90)	23.44 (23.18 to 23.70)
Treatment switching§	% switch at some point in follow-up	36.16	40.77	34.67

For full definitions of above outcomes, see online supplementary appendix 1.
*AECOPD events were those that met any criteria from the validated algorithm.[16]
†The strict definition of AECOPD events was strictly medical codes for acute exacerbation of COPD.
‡Prescription of any treatment counted as usual care in SLS COPD.
§Switching from one treatment class to another during follow-up.
AECOPD, acute exacerbations of COPD; COPD, chronic obstructive pulmonary disease; CPRD-GM, Clinical Practice Research Datalink in Greater Manchester; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester; SLS COPD, Salford Lung Study in chronic obstructive pulmonary disease.

Table 3 Random intercept for SLS COPD placed in distribution of random intercepts of local authorities from CPRD (testing for presence of Hawthorne effect)

Variable	CPRD-xGM 2.5th percentile	CPRD-xGM median value	CPRD-xGM 97.5th percentile	SLS COPD value	SLS COPD percentile	Unusual flag
Count of AECOPD episodes*†	0.90	0.99	1.13	1.14	98.37	Yes
Time until first AECOPD episode†‡	0.89	1.00	1.17	1.14	96.71	No
Count of strict definition of AECOPD episodes†§	0.36	1.01	2.62	2.59	97.36	No
Mortality‡	0.91	1.00	1.15	0.61	0	Yes
Number of days contact with primary care*	0.85	1.00	1.30	1.75	100	Yes
Count of trial-related prescription items*¶	0.80	1.01	1.26	1.17	87.88	No
Treatment switching**††	0.66	0.99	1.55	0.69	4.83	No

For full definitions of above outcomes, see online supplementary appendix 1.
For Poisson models, the random intercepts represent relative rates; for Cox models, they represent HRs; for logistic models, they represent ORs.[18 20] SLS COPD is deemed unusual with respect to a particular variable if the random intercept for the SLS lies outside of the 2.5th–97.5th percentile range of the distribution of random intercepts of the local authorities in CPRD. Variables adjusted for, following significant univariate likelihood ratio tests, were sex, age, socioeconomic status, medication group defined from the previous 3 months, history of depression, anxiety, asthma, pneumonia, and gastro-oesophageal and peptic ulcer disease, COPD exacerbation history in the previous 12 months, MRC dyspnoea score, pneumococcal vaccine, FEV_1%, FEV_1:FVC%, and smoking status.
*Poisson model.
†AECOPD events were those that met any criteria from the validated algorithm.[16]
‡Cox model.
§The strict definition of AECOPD events was strictly medical codes for acute exacerbation of COPD.
¶Prescription of any treatment counted as usual care in SLS COPD.
**Logistic model.
††Switching from one treatment class to another during follow-up.
AECOPD, acute exacerbations of COPD; COPD, chronic obstructive pulmonary disease; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; MRC, Medical Research Council; SLS COPD, Salford Lung Study in chronic obstructive pulmonary disease.

Table 4 Random coefficient for SLS COPD placed in distribution of random coefficients of local authorities from CPRD (self-controlled comparison within the trial before and during the trial period)

Variable	CPRD-xGM 2.5th percentile	CPRD-xGM median value	CPRD-xGM 97.5th percentile	SLS COPD value	SLS COPD percentile	Unusual flag
Count of AECOPD*†	0.88	1.00	1.15	1.02	64.17	No
Count of strict definition[n] of AECOPD*‡	0.81	0.96	1.21	1.34	100	Yes
Number of days contact with primary care*	0.97	1.07	1.19	1.08	57.06	No
Count of trial-related prescriptions§	0.95	1.03	1.08	0.63	0	Yes
Treatment switching¶**	0.60	0.63	0.65	0.59	0.27	Yes

For full definitions of above outcomes, see online supplementary appendix 1.
For Poisson models, the random intercepts represent relative rates; for logistic models, they represent ORs. SLS COPD is deemed unusual with respect to a particular variable if the random intercept for the SLS lies outside of the 2.5th–97.5th percentile range of the distribution of random intercepts of the local authorities in CPRD.
*Poisson model.
†AECOPD events were those that met any criteria from the validated algorithm.[16]
‡The strict definition of AECOPD events was strictly codes for acute exacerbation of COPD.
§Prescription of any treatment counted as usual care in SLS COPD.
¶Logistic model.
**Switching from one treatment class to another during follow-up.
AECOPD, acute exacerbations of COPD; COPD, chronic obstructive pulmonary disease; COPD, Salford Lung Study in chronic obstructive pulmonary disease; CPRD-xGM, Clinical Practice Research Datalink outside of Greater Manchester.

However, the comparison of outcomes in the year pretrial and peritrial would not have been affected by variation in the EHR systems.

Our decision to determine whether a variable was unusual based on a 95% interval is arbitrary; therefore, the percentile in which the observation for membership of the trial cohort lies is important. A strength of this method is the transparent approach of displaying the percentile in addition to the binary classification, allowing readers to draw their own inferences.

In conclusion, we found broad similarity between the enrolled SLS COPD UC cohort and a wider trial-eligible COPD cohort across England on most measures. We observed that AECOPD episode rates were relatively high in the trial, indicative of a potential Hawthorne effect, although this was mitigated by the pretrial and peritrial analysis. The main evidence of a Hawthorne effect was observed through behavioural changes—for example, coding practices or number of COPD medications prescribed by GPs. In future studies similar to this, it may be preferable to focus on accurately measuring behavioural factors of physicians. We also recommend performing a pretrial and peritrial comparison of outcomes of interest within patients to reduce some potential biases.

Overall, this study supports the generalisability of SLS COPD results and comparative effectiveness of FF/VI when use becomes routine. There is a small body of literature exploring the generalisability or transportability of trial results, using sampling weights to adjust estimates of interest.[27 28] However, to our knowledge, ours is the first study of its kind, comparing both patient characteristics and outcomes with regional variation across England. Given that EHR-enabled real-world trials are becoming more feasible[29 30] and relevant to inform decision making by regulators, health technology assessment bodies, providers and payers, these companion cohort studies are becoming increasingly important[31 32] and should be conducted wherever possible to assess the generalisability of open-label trials and to inform the design, operations and analytic methods development for future studies.

Author affiliations
[1]Farr Institute, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[2]Real World Evidence and Epidemiology, GlaxoSmithKline, Uxbridge, UK
[3]Real World Evidence and Epidemiology, GlaxoSmithKline, Collegeville, Pennsylvania, USA
[4]Clinical Practice Research Datalink, Medicines and Healthcare products Regulatory Agency, London, UK
[5]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, Utrecht, The Netherlands

Acknowledgements The authors thank the members of the GSK clinical team including Loretta Jacques, Susan Collier and David Leather, and the CHESS Scientific Steering Committee. Medical writing support in the form of editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copyediting, referencing, and graphic services was provided by Emma Landers, PhD, of Gardiner-Caldwell Communications (Macclesfield, UK) and was funded by GSK.

Contributors AP: involved with data handling and planning/leading the statistical analyses, codrafted the manuscript with author MS, and incorporated comments from other authors. MB: data processing, data analysis and editing of the manuscript. DW: contributed to study design, data analysis and interpretation, and manuscript writing/review. JMP: contributed to study design, protocol writing, interpretation of results and manuscript development. KJD: contributed to study design, data analysis plan, data interpretation and editing of the manuscript. RW: contributed to study design, data extraction, interpretation of study results and manuscript review/revision. TVS: contributed to study design, data analysis and interpretation, and manuscript development. MS: contributed to study design and analysis plan, supervised the data analysis, contributed to data interpretation, codrafted the manuscript with author AP, and revised the manuscript critically for intellectual content. All authors approved the final version of the manuscript for submission. The corresponding author had full access to the study data. Analyses were led by the academic partners (AP, MB, TVS and MS), who made the final decision to submit for publication. All authors vouch for the accuracy and completeness of the data/analyses. Manuscript

drafting was led by AP and MS, and all authors collaborated to prepare the final content for publication.

Funding This analysis was funded by GSK (study PRJ2282/201491; EUPAS registration number EUPAS10376). SLS COPD was also funded by GSK (HZC115151; NCT01551758). Medical and treatment codelists for this study will be available at https://www.gsk-clinicalstudyregister.com. The sponsor, GlaxoSmithKline, contributed to the study design, analysis plan, analyses and data interpretation.

Competing interests AP: grants, personal fees and non-financial support from GSK during the conduct of the study. MB: grants, personal fees and non-financial support from GSK during the conduct of the study. DW: employed by and holds stocks in GSK. JMP: employed by and holds stocks in GSK. KJD: employed by and holds stocks in GSK. RW: grants from GSK during the conduct of the study and grants from various organisations, outside the submitted work; employed by CPRD. CPRD received funding from GSK for access to the CPRD data and research services used in this study. CPRD also received payments from the University of Manchester for access to data and research services for studies outside the submitted work. CPRD is a research organisation offering interventional and observational research services. TVS: grants from GSK during the conduct of the study, and grants from National Osteoporosis Society, outside the submitted work. MS: grants from GSK during the conduct of the study.

Patient consent Not required.

Ethics approval This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data are provided by the patients and collected by the NHS as part of their care and support. The Office for National Statistics (ONS) is the provider of the ONS data contained within the CPRD data. Hospital Episode data and the ONS data (2014) are reused with the permission of the Health and Social Care Information Centre. All rights reserved.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data in this study cannot be published as we do not have permission.

Author note Quality control: This study is registered in the European Post Authorisation Safety studies registry of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP/SDPP/10376). The methodology in this study has been awarded with the ENCePP study seal approval.

REFERENCES

1. Herland K, Akselsen JP, Skjønsberg OH, *et al*. How representative are clinical study patients with asthma or COPD for a larger "real life" population of patients with obstructive lung disease? *Respir Med* 2005;99:11–19.
2. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365:82–93.
3. Vestbo J, Leather D, Diar Bakerly N, *et al*. Effectiveness of fluticasone furoate-vilanterol for COPD in clinical practice. *N Engl J Med* 2016;375:1253–60.
4. New JP, Bakerly ND, Leather D, *et al*. Obtaining real-world evidence: the Salford Lung Study. *Thorax* 2014;69:1152–4.
5. Parsons HM. What Happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science* 1974;183:922–32.
6. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. *BMJ* 2015;351:h4672.
7. Dal-Ré R. Could phase 3 medicine trials be tagged as pragmatic? A case study: The Salford COPD trial. *J Eval Clin Pract* 2018;24:258–61.
8. Konstantinou GN. Pragmatic trials: how to adjust for the 'Hawthorne effect'? *Thorax* 2012;67:562.
9. Berthelot JM, Le Goff B, Maugars Y. The Hawthorne effect: stronger than the placebo effect? *Joint Bone Spine* 2011;78:335–6.
10. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267–77.
11. McCarney R, Warner J, Iliffe S, *et al*. The Hawthorne Effect: a randomised, controlled trial. *BMC Med Res Methodol* 2007;7:30.
12. Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
13. Department for Communities and Local Government. The english indices of deprivation, 2010. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf [Accessed 4 Jan 2017].
14. Soriano JB, Lamprecht B, Ramírez AS, *et al*. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. *Lancet Respir Med* 2015;3:443–50.
15. Quan H, Li B, Couris CM, *et al*. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol* 2011;173:676–82.
16. Rothnie KJ, Müllerová H, Hurst JR, *et al*. Validation of the recording of acute exacerbations of COPD in UK primary care electronic healthcare records. *PLoS One* 2016;11:e0151357.
17. SAS. *SAS/STAT 9.2 user's guide - the MI procedure*. Cary, NC: SAS Institute Inc, 2008.
18. Goldstein H. *Multilevel statistical models*. 4th edn. West Sussex: John Wiley & Sons Ltd, 2011.
19. Cameron AC, Trivedi PK. *Regression analysis of count data*. Cambrdige, UK: Econometric Society Monographs - Cambrdige University Press, 2013.
20. Austin PC. A tutorial on multilevel survival analysis: methods, models and applications. *Int Stat Rev* 2017;85:185–203.
21. Khan NF, Perera R, Harper S, *et al*. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Fam Pract* 2010;11:1.
22. Joe H, Zhu R. Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution. *Biom J* 2005;47:219–29.
23. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA* 2014;312:2401–2.
24. SAS Institute Inc. *SAS, Version 9.4*. Cary, NC: SAS Institute Inc, 2014.
25. R Development Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008.
26. Crowe M. Do you know the difference between these adherence measures? Available from: http://www.pharmacytimes.com/contributor/michael-crowe-pharmd-mba-csp-fmpa/2015/07/do-you-know-the-difference-between-these-adherence-measures [Accessed 17 Aug 2017].
27. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. *Am J Epidemiol* 2010;172:107–15.
28. Westreich D, Edwards JK, Lesko CR, *et al*. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol* 2017;186:1010–4.
29. van Staa TP, Dyson L, McCann G, *et al*. The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. *Health Technol Assess* 2014;18:1–146.
30. Zannad F, Pfeffer MA, Bhatt DL, *et al*. Streamlining cardiovascular clinical trials to improve efficiency and generalisability. *Heart* 2017;103:1156–62.
31. Sherman RE, Anderson SA, Dal Pan GJ, *et al*. Real-world evidence - what is it and what can it tell us? *N Engl J Med* 2016;375:2293–7.
32. Jarow JP, LaVange L, Woodcock J. Multidimensional evidence generation and FDA regulatory decision making: defining and using "Real-World" data. *JAMA* 2017;318:703–4.