Adv Ther (2021) 38:4847–4858
https://doi.org/10.1007/s12325-021-01827-2

ORIGINAL RESEARCH

Key Learnings from Running an Extension Study to a Real-World Effectiveness Trial: The Extended Salford Lung Study

Wilhelmine Meeraus · Mark Fry · Richard Yeatman · Jeanne M. Pimenta · Jamila Astrom · Alan Barth · Sheila McCorkindale · Rupert Jones · David Leather

Received: April 20, 2021 / Accepted: June 10, 2021 / Published online: August 6, 2021

ABSTRACT

Introduction: The Salford Lung Studies (SLS) were real-world randomised controlled trials set within UK primary care that assessed the effectiveness and safety of initiating once-daily fluticasone furoate/vilanterol versus continuing usual care in patients with chronic obstructive pulmonary disease or asthma. Data were collected for a relatively short period, limiting the study of long-term outcomes. To broaden the capture of SLS patients' data, we undertook the Extended SLS (Ext-SLS), aiming to better understand the patient disease journey and the effects of treatment in a real-world setting, through collection of patient-level data. Here, we present study design information and the challenges and learnings gathered in creating the Ext-SLS.

Methods: The Ext-SLS was intended to augment the SLS by collecting retrospective and prospective (up to 10 years from consent) primary and secondary care electronic health record (EHR) data and patient questionnaires. After ethics approval, general practitioners (GPs) obtained consent from SLS patients remotely (mean 3.2 years post-SLS completion). To facilitate GPs identifying eligible patients, a novel EHR-based approach flagged SLS patients who were alive and registered with their original GP. An automated system sent consent forms/questionnaires to patients. Medical data were collected via EHRs; primary care data were extracted from GPs' systems whilst secondary care data were sourced from the UK NHS.

Results: Of the 75 GP sites from the SLS, 35 (47%) declined Ext-SLS participation leaving 4158 potentially eligible patients; 1169 (28%) patients were excluded as GPs could not confirm them as SLS participants or due to incapacity. Of 2989 patients invited, 1189 (40%) consented.

Conclusions: Developing an EHR-based trial extension was achieved, with reasonable consent rates amongst invited patients. The resulting Ext-SLS is a unique and valuable research

Affiliations for Wilhelmine Meeraus, Jeanne M. Pimenta and Rupert Jones were correct at the time of study.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12325-021-01827-2.

W. Meeraus · M. Fry · J. M. Pimenta · J. Astrom · D. Leather (✉)
GlaxoSmithKline plc, Brentford, London, UK
e-mail: david.a.leather@gsk.com

R. Yeatman
IgniteData, Reading, UK

A. Barth
Graphnet Health Ltd, Milton Keynes, UK

S. McCorkindale
NIHR Clinical Research Network Greater Manchester, Manchester, UK

R. Jones
University of Plymouth, Plymouth, UK

resource. Leveraging EHRs and technology reduced GP burden, facilitating participation. Initiation of extension studies prior to study close-out may help increase GP and patient participation.

Keywords: Asthma; COPD; EHR studies; Primary care research; Real-world evidence

Key Summary Points

Why carry out this study?

The Salford Lung Studies (SLS) in asthma and COPD were unique phase IIIb randomised controlled trials into the effectiveness and safety of initiating fluticasone furoate/vilanterol versus continuing usual care; however, the SLS only covered a short period, limiting the data available from the participants.

In order to broaden the capture of SLS patients' data, we undertook an extension study (the Extended SLS [Ext-SLS]) to capture retrospective and prospective data from SLS participants.

The Ext-SLS collected retrospective and prospective data from SLS participants' electronic health records (EHR) and questionnaires to better understand the patient disease journey and the effects of treatment in a real-world setting.

What was learned from the study?

Developing an EHR-based trial extension is achievable, with reasonable consent rates.

Significant challenges were identifying patients, addressing new data protection legislation and allaying the concerns of general practitioners about increased workload; these challenges could be addressed in future by initiation of extension studies prior to study close-out.

INTRODUCTION

Real-world evidence is becoming increasingly important in healthcare decision-making as a complement to randomised controlled trials (RCTs) [1–4]. Real-world data, such as information contained in electronic health records (EHRs), disease registries, and insurance claims, can be used to assess treatment effectiveness and the long-term safety of pharmaceutical products in a larger, more representative population than is typically found in RCTs [1–3]. In this way, the value of interventional controlled trial data may be enhanced by long-term follow-up studies in real-world settings [5].

Pragmatic randomised trials conducted under normal clinical care conditions (effectiveness trials) are increasingly turning to routinely collected healthcare data (RCHD) for patient follow-up [5]. As an example, the Salford Lung Studies (SLS) were two phase III RCTs that assessed the effectiveness and safety of initiating once-daily fluticasone furoate/vilanterol (FF/VI) versus continuing usual care (UC) in patients with chronic obstructive pulmonary disease (COPD; SLS COPD [6]) or asthma (SLS asthma [7]). These studies were the first in the world to assess the real-world effectiveness of a pre-licence medicine [8]. The SLS were designed to recruit a broad cohort, representative of patients with asthma and patients with COPD in the real-world, and to minimise interference in participants' routine clinical management through the use of EHRs, thus providing novel, detailed safety and effectiveness data [8]. The patient experience was as close to routine care as possible to preserve the real-world nature of the study, with key exposure and outcome data captured remotely via EHRs.

Studies using RCHD such as those in EHRs may reduce the cost of clinical trials, enabling a greater number of large definitive trials to be conducted and facilitating efficient long-term assessment of interventions used in clinical practice [9]. Furthermore, RCHD studies may reduce the burden of participation (and broaden the included population) by removing the need for repeated study visits and regimens of data collection. Follow-up of RCTs through extension studies utilising RCHD

has confirmed the results of the RCTs, demonstrated changes in effectiveness over time, and revealed long-term safety profiles [10]. Thus, extension studies of this type can provide valuable information. This is increasingly recognised in the UK, with one study finding that approximately half of the publicly funded studies examined planned to collect information on outcomes using RCHD [11].

The SLS capitalised on the EHR infrastructure in the Salford region which linked primary and secondary care data with patient-level prescription information in real time [12]. As a result, SLS patients represent a population whose disease experience and management are extremely well characterised. However, EHR data collected in the SLS were limited to 3 years prior to randomisation and the 12-month interventional treatment period. This finite period of data coverage limited the potential to address scientific questions of clinical interest related to early life events and exposures, or long-term COPD/asthma disease progression and associated outcomes.

In order to broaden the capture of SLS patients' data through the additional collection of patient-level data encompassing past and periodic future, we undertook an extension study, the Extended SLS (Ext-SLS), using RCHD. The Ext-SLS was designed to extend the utility of the SLS patient population through provision of an enriched patient dataset by collecting historic and prospective primary and secondary care data to investigate and better understand the patient disease journey, factors affecting disease progression and the effects of treatment in a real-world setting. In this paper, we present study design information for the Ext-SLS and outline the challenges we have encountered to date and the learnings we have gathered from the creation of this real-world extension study.

METHODS

Study Design

The SLS in COPD and Asthma

The SLS methods have been extensively reported elsewhere [6, 7]. Briefly, patients were included if they had a diagnosis of COPD with a history of exacerbations or of symptomatic asthma and were taking regular maintenance therapy; there were minimal exclusion criteria. Patients were recruited from primary care (general practice) clinics in Salford and South Manchester, UK, between March 2012 and October 2014 for SLS COPD and November 2012 and December 2016 for SLS asthma. Patients were randomised 1:1 to initiate FF/VI or continue UC. Primary endpoints were mean annual rate of moderate/severe exacerbations (SLS COPD), percentage of patients with an Asthma Control Test (ACT) score of at least 20 and/or an increase in ACT score of at least 3 at week 24 (SLS asthma). Patients consented to relevant data being collected from up to 3 years prior to the study and for the 12-month trial period. The decision to constrain the duration of data collection was based on the assumptions that limiting the availability of patient data would make the study more acceptable to patients, improve recruitment and be more acceptable to the ethics committee. Overall, 75 general practices and 7039 patients were randomised into the SLS.

The Ext-SLS

The Ext-SLS is a retrospective and longitudinal prospective observational study of patients who completed the SLS (Fig. 1). Data were collected from patients' EHRs and questionnaires were used to capture detailed information about patients' history with their disease, the impact of the disease on everyday life and their level of disease control.

Patient consent for the original SLS did not cover an extension study; therefore, ethical approval was sought and granted by the North West–Greater Manchester East Research Ethics Committee (REC Number 17/NW/0122). Patients enrolled in the Ext-SLS have provided consent for the collection of data via their EHRs both retrospectively to the earliest available record and prospectively for up to 10 years from consent into the study.

A key consideration for the Ext-SLS was to create a study with minimal burden on participating general practitioners (GPs) and patients. To this end, we employed innovative methods of patient identification and data collection,

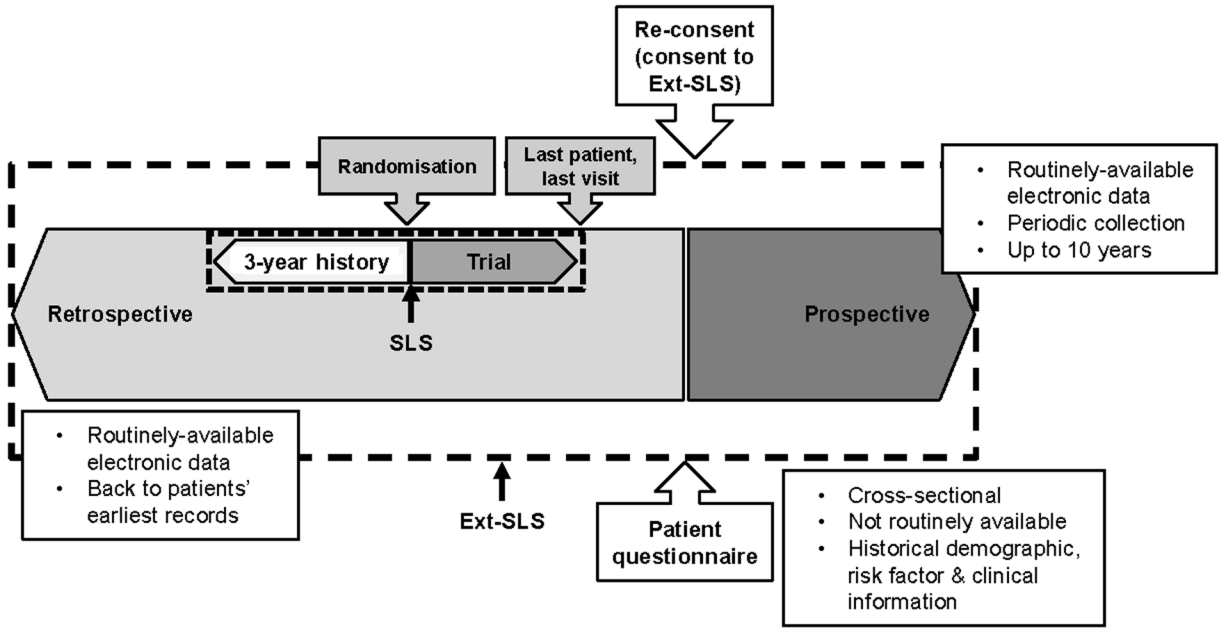

Fig. 1 Extended SLS study design. Ext-SLS, Extended Salford Lung Study; SLS, Salford Lung Studies

creating a 'light-touch' approach from the perspective of the GPs.

GP and Patient Recruitment

GP sites that were included in the original SLS were invited by the sponsor to participate in the Ext-SLS. More information was provided to GPs by a third-party company (IgniteData, Reading, UK) who worked with the National Institute for Health Research Clinical Research Network to follow up with GPs and facilitate their participation. GPs who agreed to participate were asked to triage lists of patients, excluding those who they deemed to be too unwell to participate or provide informed consent. The final patient lists were then used to mail informed consent packages to patients (Table S1 in the supplementary material) along with COPD- or asthma-specific questionnaires.

As records of which patients were in the original SLS were not easily accessible by GPs, and the SLS sponsor did not hold patient-identifiable information on SLS patients, a novel approach was taken to assist GPs to identify patients suitable for the Ext-SLS. De-identified trial data were used to create lists of patients with at least four dates of primary care contacts (PCC) during the SLS trial period, patient year of birth and GP practice. The lists were used to match patients to their SLS identifier; positive identification was defined as a match of year of birth and four or more known PCC dates between the patient and the identifier. In this way, GP-specific lists of potentially eligible patients were compiled on behalf of GPs, allowing them to easily triage and consent patients. Piloting of GP triage and automated mail-out of consent packs was conducted at two sites to inform measures to enhance patient consent rates.

Development of Patient Questionnaires

Disease-specific questionnaires were developed to gather detailed patient-reported information about disease management (via the Asthma Control Questionnaire [13], ACT [14], COPD and Asthma Sleep Impact Scale [15] and COPD Assessment Test [16]) and patient history that might not be captured in the EHR; for example, the questionnaires contained questions on living and working environments. Prior to patient recruitment, and alongside ethical approval, the questionnaires were validated in a pilot study including 80 patients (40 asthma, 40 COPD).

Questionnaires were re-formatted after the pilot to optimise self-completion; for example, larger font was used in the COPD questionnaire because of the older age of the patients.

The Ext-SLS Data Framework

Primary care data from source EHRs were extracted and collated by two trusted third-party companies (IgniteData, Reading, UK; and Graphnet Health, Milton Keynes, UK). Graphnet Health managed data arriving nightly from GP sites as part of the local Integrated Digital Care Records. As Data Controllers, GP sites agreed to have their data reviewed and collated. At the agreed data-cut time points, the latest 'cut' of the primary care (GP) data was transferred to a dedicated virtual network for the study. Here the data underwent quality control checks (including removal of patient-identifiable information) by IgniteData before being made available to the study researchers.

Data from patient questionnaires were similarly processed by IgniteData who transcribed patient responses into a bespoke database. Final pseudonymised data were secured with an encryption key. Only authorised study personnel could access data for analysis.

Linked secondary care EHR data were requested from National Health Service (NHS) Digital, including hospital episode statistic (HES) data on hospital admissions, outpatient appointments and emergency hospital attendances. The application to NHS Digital began around the time ethics approval was sought in order to allow the NHS Digital Independent Group Advising on the Release of Data to review the consent form, ensuring the language relating to NHS Digital secondary care data was sufficiently detailed. The data available from NHS Digital were reviewed as part of the application so that any elements of the patient records that were deemed as sensitive or irrelevant to the study were not requested.

To date, historical data encompassing all routinely available primary care electronic demographic and health-related data have been collected from the patients' EHRs. Additional historical demographic data, COPD/asthma risk factor information and clinical data not routinely available have also been collected via patient-completed questionnaires. Prospective data collection from primary care EHRs of RCHD is ongoing. The original intention was to collect full retrospective data and prospective data for up to 10 years after consent; however, ultimately only approximately 10 years of patient data including retrospective and some prospective data were obtained and are available. Similarly, at the time of writing, the secondary care data are not yet available.

Proof-of-Concept Study

Given the novelty and complexity of the Ext-SLS, we conducted a proof-of-concept (PoC) study which allowed for testing of the processes for matching patients and extracting primary care data using a small subset of GP sites and patients participating in the Ext-SLS. Selection of sites for the PoC was intended to represent the geographical diversity of GP sites and to include sites using both Egton Medical Information Systems and Vision GP software. A primary objective of the study was to assess whether extracted primary care data for Ext-SLS patients met specific objective criteria (metrics) relating to participation in the SLS, e.g. an asthma or COPD diagnosis, a prescription for respiratory maintenance therapy. For each metric, the proportion of patients with a valid value above, below or at a pre-specified threshold was assessed. Threshold values were based on previous experience with primary care EHR data and respiratory clinical data. As a further test to determine whether primary care EHR data were extracted correctly, prescribing records from the 12-month SLS study period were reviewed.

RESULTS

Recruitment of SLS GP Sites

A total of 75 GP sites were included in the original SLS (of these 74 SLS asthma, 75 SLS COPD) and all were invited to participate in the

Ext-SLS. Of these, 40 sites (53%) participated in the Ext-SLS (48 initially agreed to participate but eight declined participation at a later stage). Collectively, these sites provided a pool of 4158 potentially eligible patients (64% of the SLS cohort) who completed the original SLS (Fig. 2).

PoC Study

A sample of 329 patients (192 asthma, 137 COPD) were included in the PoC study. These patients were registered at six of the first GP sites that provided informed consent to participate in the Ext-SLS. For both patients with asthma and patients with COPD, all metrics exceeded the threshold values. It was therefore concluded that the primary care data had been appropriately extracted (Table S2).

Table S3 shows that patients with asthma and patients with COPD received multiple prescriptions for respiratory medications during the SLS. Moreover, only patients randomised to FF/VI were prescribed FF/VI during the trial. A point to note is that the mean number of FF/VI prescriptions per patient during the SLS study period was low (1 [range 0–2] for both asthma and COPD). This likely reflects the fact that FF/VI was not approved for use until the end of the SLS trial period and therefore was not on the formulary for much of the SLS study period. Before approval, prescriptions of FF/VI were recorded in the patient record as free text, which was excluded from Ext-SLS data collection because of the likelihood of it containing sensitive information.

Recruitment of SLS Patients

As part of the pilot, consent packs were mailed to 209 eligible patients at the two participating GP sites; 30% of patients consented after receiving one mailing. These figures were extrapolated to provide overall estimates of the expected number of patients agreeing to participate in the Ext-SLS and reinforced the decision to send eligible patients a follow-up consent pack mailing 4 weeks after the initial pack was mailed.

Fig. 2 Flow of GP site and patient recruitment. *Proportion of patients who completed the SLS. COPD, chronic obstructive pulmonary disease; EHR, electronic health record; FF/VI, fluticasone furoate/vilanterol; GP, general practice; SLS, Salford Lung Studies

From the patients potentially eligible for inclusion in the Ext-SLS, 1055 could not be positively identified as living SLS participants who were still registered with the same GP site, and were therefore unable to participate in the Ext-SLS (Fig. 2). Following exclusions made by GPs during triage, consent packages were sent to 2989 eligible patients; of these, 1183 (40%; 813 asthma, 370 COPD) consented. A small number (36) of consented patients did not complete the questionnaires or did not have valid GP data (e.g. GP records unavailable or missing crucial information such as diagnosis codes), resulting in a final cohort of 1147 (38%; 798 asthma, 349 COPD) (Fig. 2) with a mean time between completion of the SLS and consent to the Ext-SLS of 3.2 years (2.7 years asthma, 4.2 years COPD). The FF/VI to UC ratio of 1:1 achieved in the original SLS was mostly maintained in the Ext-SLS.

DISCUSSION

The Ext-SLS study promised to deliver a unique dataset for the understanding of chronic respiratory disease. Ultimately, 53% of GP sites from the original SLS agreed to participate, representing 64% (4158) of the patients who completed the original SLS. The final cohort included 1147 patients (798 asthma, 349 COPD)—approximately 40% of invited patients, but only 18% of patients completing the SLS.

Although the Ext-SLS research cohort comprises only 18% of the patients who completed the SLS, the longitudinal dataset is unique, with historic and ongoing data collection. The dataset combines trial data, primary and secondary care EHRs with self-reported questionnaire data, to provide researchers with a tool to better understand patient disease journeys with

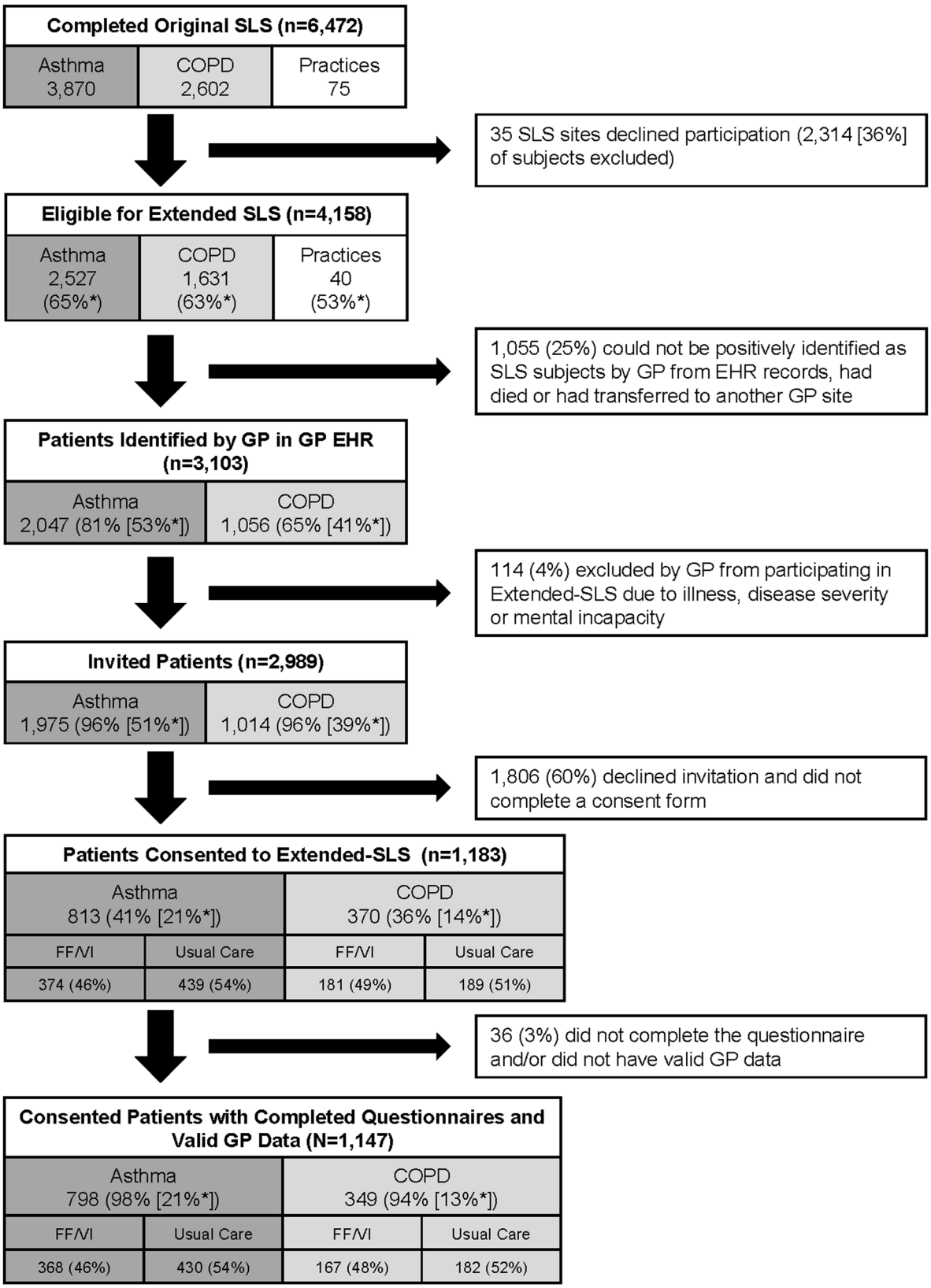
Completed Original SLS (n=6,472)
Asthma 3,870
COPD 2,602
Practices 75
35 SLS sites declined participation (2,314 [36%] of subjects excluded)
Eligible for Extended SLS (n=4,158)
Asthma 2,527 (65%*)
COPD 1,631 (63%*)
Practices 40 (53%*)
1,055 (25%) could not be positively identified as SLS subjects by GP from EHR records, had died or had transferred to another GP site
Patients Identified by GP in GP EHR (n=3,103)
Asthma 2,047 (81% [53%*])
COPD 1,056 (65% [41%*])
114 (4%) excluded by GP from participating in Extended-SLS due to illness, disease severity or mental incapacity
Invited Patients (n=2,989)
Asthma 1,975 (96% [51%*])
COPD 1,014 (96% [39%*])
1,806 (60%) declined invitation and did not complete a consent form
Patients Consented to Extended-SLS (n=1,183)
Asthma 813 (41% [21%*])
COPD 370 (36% [14%*])
FF/VI
Usual Care
FF/VI
Usual Care
374 (46%)
439 (54%)
181 (49%)
189 (51%)
36 (3%) did not complete the questionnaire and/or did not have valid GP data
Consented Patients with Completed Questionnaires and Valid GP Data (N=1,147)
Asthma 798 (98% [21%*])
COPD 349 (94% [13%*])
FF/VI
Usual Care
FF/VI
Usual Care
368 (46%)
430 (54%)
167 (48%)
182 (52%)

Table 1 Challenges of setting up an extension study

Challenge	Description/effect on the Ext-SLS	Potential solution for future extension studies
Changes to data protection legislation	The introduction of new data protection legislation (EU GDPR) concerned GPs, making them less likely to participate in the study. GPs were advised against participating in research while there was uncertainty about GDPR	Awareness of upcoming changes to relevant laws and better education of GPs on what this means for them and for the use of patient data may allay fears
Time between parent and extension study	A long period of time between the studies reduced the likelihood of GPs agreeing to participate as their knowledge of the study had reduced and their circumstances had changed	Preparing for an extension study during or even prior to the parent study may reduce the intervening time and ensure that sites are expecting to continue their involvement
Identification of patients	Following completion of the parent study, flags in the EHR marking patients as participants were removed, making direct identification impossible. It was also not possible to identify patients who had moved to a different GP site or who were deceased	Preparing for the extension study during or even prior to the parent study by ensuring that accessible flags are maintained on patient records should help them to be identified later
GP workload	GPs have high workloads without the inclusion of research-related work, making their participation in research unmanageable	Design the extension study with clinical workload in mind and attempt to reduce the burden on healthcare professionals at every opportunity. This can be achieved through automation and the use of third parties where possible. Being able to demonstrate an available level of support during the study is crucial to encouraging the participation of GPs
Quality of data	Real-world data are challenging to work with and this is compounded by discrepancies in recording methods between GP sites and differing software (EMIS vs VISION) used by GP sites, especially when a data dictionary is unavailable	PoC studies allow for the testing and refining of data management approaches and are invaluable for the conduct of an extension study. New initiatives in the UK should improve the consistency of data recording for future studies

EHR electronic health record, *EU* European Union, *GDPR* General Data Protection Regulation, *GP* general practitioner, *PoC* proof-of-concept, *UK* United Kingdom

asthma and COPD and the effects of treatment in a real-world setting.

Here, we discuss the key challenges we encountered during the set-up of the Ext-SLS and share learnings for future pragmatic trials and extension studies (Table 1).

One of the biggest obstacles we faced was the introduction of new data protection legislation during study set-up; ultimately affecting GP site recruitment, delaying patient recruitment by months and complicating the application for secondary care EHRs from NHS Digital. The European General Data Protection Regulation (GDPR) was widely publicised in the time leading up to its introduction and GPs were concerned about the implications that this had for their participation in studies using patient data. There was uncertainty around how to comply

with the new regulations and for many GPs, the safest option was not to take part in this study. Although not all of the 35 GP sites declined to participate in the Ext-SLS for this reason, future extension studies should consider upcoming legislation that might affect participation in a trial during the design of the study.

An unanticipated consequence of the new legislation was the need to alter key study documents, including patient consent forms and information leaflets, once GDPR had come into effect. This led to significant delays negotiating the required changes with the sponsor, the Ethics Committee and the Health Research Authority. Compounded by the fact that the extension study had not been planned as part of the original SLS, these delays meant there was at least 12 months between the SLS ending and Ext-SLS site recruitment, and 18–24 months between GPs receiving invitations to participate in the Ext-SLS and consent packs being mailed out. These extended periods of time may have contributed to the attrition of GP practices. Had an extension study been planned as part of the original SLS, some of these delays could have been avoided; however, GDPR was a notable factor hampering the distribution of invitations.

The arrival of GDPR also complicated the already extensive process of applying for data from NHS Digital. Obtaining linked secondary care data was a lengthy process, further hampered by the prioritisation of COVID-19-related data requests above others. HES data were approved for use in the Ext-SLS in January 2021 and are expected to be available in May 2021, following the application process begun in 2017. In general, we would not recommend direct data procurement of secondary care data in this way under the current application process.

A second challenge we faced with site recruitment potentially stemmed from perceptions of the GP workload associated with the Ext-SLS. During the original SLS, GPs received support from the research nurse network which was crucial in allowing them to participate while minimising the impact on their usual work [17]. However, by the time invitations for the Ext-SLS were sent, the SLS support team had been largely de-mobilised and this may have impacted whether GPs decided to participate. GPs are known to have a high workload of NHS tasks [18] and the additional work of research may have been too much for many practices. To mitigate this, the Ext-SLS was designed to be 'light-touch' from the perspective of GPs, utilising third-party assistance and EHRs for data collection and patient identification; successfully demonstrating this approach was key to recruitment.

Another hurdle during the set-up of the Ext-SLS was in providing lists of potentially eligible patients for GPs to triage. The primary difficulty in identifying patients was that after completion of the SLS, flags in patient EHRs held by GP sites that identified individuals as SLS participants were removed, making it impossible to generate patient lists from the local integrated Digital Care Records alone. A total of 1055 patients (16% of the original SLS cohort) were excluded from the lists because they could not be positively identified as SLS patients, either because the patients did not have sufficient trial data available for matching or because they had died or moved to a new GP practice. Whilst eligible patients could have been identified if GPs were asked to search their archived paper records, this would have been very labour-intensive and against the 'light-touch' nature of the study. To overcome this hurdle, we developed a 'key' that could link SLS identifiers to individual patients using non-sensitive data from the SLS, an approach that was validated with the pilot study conducted at two GP sites. Nonetheless, failure of patient identification (including for death) remained a point of attrition and this highlights the need to consider extension studies when planning RCTs. This is particularly relevant when the subjects in the 'parent' study are elderly and at increased risk of death, e.g. mortality rate among patients with COPD is high (with estimates from one Dutch population-based study ranging from 41.9 to 249.9 per 1000 patient years depending on disease severity [19]). Movement of patients from their original GP practice also presented a challenge to patient identification as their record did not match that of the SLS.

An additional technical difficulty, related to the investigation of an initially unlicensed product in the SLS, was that the patient records did not necessarily contain details of the prescription of FF/VI. These prescriptions were entered as free text which was not included in data collection as it may have contained sensitive information. As such, patients from the FF/VI arm of the SLS could not be identified by their prescription record.

Beyond the challenges of recruiting participants to the study, the collection and handling of data from EHRs also proved difficult. By their very nature, real-world data are designed to record routine care and are hard to standardise for the purposes of research, even with schemes in place such as the Quality Outcomes Framework, which incentivises GPs to provide improved care, measured by consistent recording of clinical data in EHRs [20]. Additionally, GP practices are not research-ready organisations and data are not always entered with research in mind. Using a small sample in a PoC study, we were able to assess and refine our approach and this streamlined the remainder of the study. We strongly recommend including a PoC step in future studies. New systems are being implemented that will make data entry easier for GPs through the use of artificial intelligence to assign coding structures to free-text inputs, which should help improve the availability of data to research [21].

At first glance, extending a community-based pragmatic trial seems a relatively easy and straightforward task; however, setting up the Ext-SLS was challenging and technically difficult. The early decision to limit the access of data in the SLS to 4 years was a decision taken to drive success in the SLS. Future investigators undertaking pragmatic randomised trials using routinely available EHR data should carefully assess the benefits and risks of continued data collection beyond the interventional period. Nonetheless, developing an EHR-based trial extension to the SLS was achieved and consent rates were reasonable. We found that implementing EHR technology to reduce the burden on busy GPs may have helped to facilitate their participation. Direct collection of primary care data to avoid electronic case report forms was possible, but secondary care data could not be accessed in a timely manner. In future, initiation of extension studies prior to 'parent' study close-out may help in reducing patient attrition.

In summary, studies like the Ext-SLS are not without their challenges. However, with careful design, they can be a valuable source of patient-level, disease-specific data. The Ext-SLS comprises an extremely well-characterised cohort which includes data from a randomised clinical trial in its timeline. Furthermore, the Ext-SLS captured data not routinely collected in healthcare records alongside data from primary care; the coming addition of prospective data will further enhance the dataset and cement the Ext-SLS as an important source of patient-level information on chronic respiratory disease.

ACKNOWLEDGEMENTS

The authors would like to thank Loretta Jacques for her leadership at GSK on the SLS studies. The authors would like to thank the patients, HCPs and practice staff who enabled the SLS studies and the extension to happen.

Trademarks are owned by or licensed to their respective owners (the GlaxoSmithKline group of companies, Graphnet Health Ltd, IgniteData or QualityMetric Inc).

Funding The extended Salford lung study was funded by GlaxoSmithKline plc. (study 207531; ClinicalTrials.gov ID NCT03152669). Funding for the journal's Rapid Service and Open Access Fees was also provided by GlaxoSmithKline plc.

Authorship All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author's Contributions All authors contributed to the analysis and interpretation of the data; WM, MF and RJ were involved in the study concept and design; WM, AB, JA, RY, JMP and MF were involved in data acquisition.

Medical Writing and/or Editorial Assistance Editorial support (in the form of writing assistance, assembling tables and figures, collating author comments, grammatical editing, and referencing) was provided by Fraser Shearer, MSc, of Ashfield MedComms, an Ashfield Health company (Glasgow, UK) and was funded by GlaxoSmithKline plc.

Compliance with Ethics Guidelines Ethical approval for this study was granted by the North West–Greater Manchester East Research Ethics Committee (REC Number 17/NW/0122).

Disclosures **Wilhelmine Meeraus:** Employment with and stock/share ownership in GlaxoSmithKline plc. **Alan Barth:** Employment with Graphnet Health Ltd, which has received funding from GlaxoSmithKline plc. **Jamila Astrom:** Employment with and stock/share ownership in GlaxoSmithKline plc. **Richard Yeatman:** Director of and major shareholder in IgniteData, which has received funding from GlaxoSmithKline plc. **Sheila McCorkindale:** Membership of the study governance board for the Salford Lung Study and has been an advisory board member for GlaxoSmithKline plc. **David Leather:** Employment with and stock/share ownership in GlaxoSmithKline plc. **Jeanne Pimenta:** Current employment with BioMarin Pharmaceutical Inc., employment with GlaxoSmithKline plc. at time of study and stock/share ownership in GlaxoSmithKline plc. **Mark Fry:** Employment with and stock/share ownership in GlaxoSmithKline plc. **Rupert Jones:** Current employment with Plymouth Marjon University. Grants from AstraZeneca, Teva Pharmaceutical Industries Ltd. and GlaxoSmithKline plc., and personal fees from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline plc., Novartis, Nutricia and OPRI outside the submitted work.

Data Availability The datasets generated and/or analysed during the current study are not publicly available due to ongoing analyses preventing their wider release. Additionally, the secondary care data associated with this study will be under license from NHS Digital and cannot be released as such.

REFERENCES

1. McDonald L, Lambrelli D, Wasiak R, Ramagopalan SV. Real-world data in the United Kingdom: opportunities and challenges. BMC Med. 2016;14:97.

2. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and impact of real-world clinical data for the practicing clinician. Adv Ther. 2018;35:1763–74.

3. Breckenridge AM, Breckenridge RA, Peck CC. Report on the current status of the use of real-world data (RWD) and real-world evidence (RWE) in drug development and regulation. Br J Clin Pharmacol. 2019;85:1874–7.

4. US Food and Drug Administration. Real-world evidence. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence. Accessed 14 Aug 2020.

5. 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.

6. Vestbo J, Leather D, Bakerly ND, et al. Effectiveness of fluticasone furoate-vilanterol for COPD in clinical practice. N Engl J Med. 2016;375(13):1253–60.

7. Woodcock A, Vestbo J, Bakerly ND, et al. Effectiveness of fluticasone furoate plus vilanterol on asthma control in clinical practice: an open-label, parallel group, randomised controlled trial. Lancet. 2017;390:2247–55.

8. New JP, Bakerly ND, Leather D, Woodcock A. Obtaining real-world evidence: the Salford lung study. Thorax. 2014;69:1152–4.

9. Lensen S, Macnair A, Love SB, et al. Access to routinely collected health data for clinical trials—review of successful data requests to UK registries. Trials. 2020;21:398.

10. Fitzpatrick T, Perrier L, Shakik S, et al. Assessment of long-term follow-up of randomized trial participants by linkage to routinely collected data: a scoping review and analysis. JAMA Netw Open. 2018;21(1):e186019. https://doi.org/10.1186/s13063-020-04329-8.

11. McKay AJ, Jones AP, Gamble CL, Farmer AJ, Williamson PR. Use of routinely collected data in a UK cohort of publicly funded randomised clinical trials. F1000Res. 2020;9:323.

12. Leather DA, Jones R, Woodcock A, Vestbo J, Jacques L, Thomas M. Real-world data and randomised controlled trials: the Salford lung study. Adv Ther. 2020;37:977–87.

13. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J. 1999;14:902–7.

14. Nathan RA, Sorkness CA, Kosinski M, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol. 2004;113(1):59–65.

15. Pokrzywinski RF, Meads DM, McKenna SP, Glendenning GA, Revicki DA. Development and psychometric assessment of the COPD and asthma sleep impact scale (CASIS). Health Qual Life Outcomes. 2009;7:98.

16. Jones PW, Harding G, Berry P, Wiklund I, Chen W-H, Kline LN. Development and first validation of the COPD assessement test. Eur Respir J. 2009;34:648–54.

17. Gemzoe K, Crawford R, Caress A, et al. Patient and healthcare professional experiences of the Salford lung studies: qualitative insights for future effectiveness trials. Trials. 2020;21:789.

18. Gibson J, Sutton M, Spooner S, Checkland K. Ninth national GP worklife survey. Manchester: University of Manchester; 2017.

19. Afonso ASM, Verhamme KMC, Sturkenboom MCJM, Brusselle GGO. COPD in the general population: prevalence, incidence and survival. Respir Med. 2011;105:1872–84.

20. NHS Digital. Quality and Outcomes Framework (QOF), enhanced services and core contract extraction specifications (business rules). https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof. Accessed 13 Nov 2020.

21. Royal College of General Practitioners. Artificial intelligence and primary care. https://www.rcgp.org.uk/-/media/Files/CIRC/artificial-intelligence-and-primary-care-jan-2019.ashx?la=en. Accessed 31 Mar 2021.