Editorial Title:
Road Map to improve the quality of Lung Cancer risk data.

Professor John K Field
The University of Liverpool
Department of Molecular and Clinical Cancer Medicine
Institute of Systems, Molecular & Integrative Biology
The William Duncan Building
6 West Derby Street
L7 8TX
LIVERPOOL
UK.

Email: J.K.Field@liverpool.ac.uk

The success of the UKLS trial and subsequent implementation projects in Liverpool, Manchester and London provided the springboard for the NHS Eng. Targeted Lung Health Check (TLHC) programme (1-6), which started in 2019 and has been recently pledged by the Prime Minister (Rishi Sunak) (7) to become the National Lung Cancer screening programme in England and also in time, in the devolved nations. Thus, the UK maybe considered one of the global leaders implementing lung cancer screening.

One has to be cognisant of the fact that despite the excellent protocol and quality assurance programme underlining the TLHC (8), (9), there is a need for continued implementation research (10). Indeed, Goodley and co-authors (11) have identified one of the fundamental issues which faces the TLHC; identification of the risk population from primary care patient records and furthermore, the issue that certain general medical practices are not providing this data. This has been exemplified by SUMMIT's targeted approach, which found discordance between the primary care code and self-reported smoking status in a quarter of invitees (12). Indeed, the authors point out, that there is uncertainty regarding the accuracy of existing primary care records and whether practices, not contributing to research databases have higher levels of missing or inaccurate data (11).

Goodley and co-authors have utilised their own Manchester Lung Health Check data and undertook a novel approach, which could improve the targeted selection of high-risk individuals. In brief, data extracted from primary care was categorised as 'live data', whilst all previous smoking status dates and codes were called 'historical data'. Each participant had a 'live' and 'historical' ever-smoking status, but if no data was available, was referred to as 'absent'. The authors modelled targeted invitation strategies into four subgroups, based on the above combinations of data availability (11).

The authors' analysis provides some disquieting findings; their study identified inaccuracies in primary care recorded smoking status, the main one being that one-quarter of those with a 'live' never-smoking code, had historic data indicating the patients had a previous smoking history. In addition, there was a significant discrepancy between self-reported smoking status and that recorded in primary care. The pivotal message was that, self-reported smoking status was taken as the ground truth for the purposes of eligibility for the Lung Health Check, but

this is ‘fallible’(11). Overall, this study provides an impetus to reconsider the best way to identify high-risk populations, especially in the lower socio-economic groups who will gain the most from such a lung cancer screening invitation (13). Machine learning / AI have been demonstrated to improve risk prediction applied to incomplete datasets available. However, this is best suited to existing, large scale data collections, where overall prediction is judged en masse. It is less clear how these techniques can be easily and justifiably applied to reliably supply an individual's risk score, where they might be denied life-saving diagnosis on the basis of the imputation. Especially as it is preferable to ask the individual to provide their own risk data.

The TLHC Lung cancer screening programme in the UK, is unique as it incorporates two validated risk models (Liverpool Lung Project (LLP_{v2}) & $PLCO_{m2012}$) in the selection of high-risk participants. However, there is a need to ensure we have updated clinical data and this will require an integrated approach to ensure accurate risk data is collected and incorporated into the risk models. Many of the proposed steps are already in place but are not fully functional or poorly implemented.

- The smoking histories need to be regularly updated when patients visit a GP Practice.
- GP Practices need to optimise their automated messaging systems (txt, email, SMS) to update patients smoking data and other risk factors. We need to learn from the Swedish patient communication model utilising emails, which has a high level of acceptability (14).
- The Introduction of the NHS App/ PATCHS-AI (15) needs to be updated to collect risk data from patients.
- Regions about to launch the TLHC programme for the first time need to have intense communications management systems (COMS) and advise patients to update this data on the Primary Care systems.
- Recommendations to integrate and share data across UK cancer screening databases would enhance all the cancer screening services.
- Inclusion of both ‘Live’ & ‘Historical smoking data’ needs to be available on primary care databases, as recommended by Goodley et al.

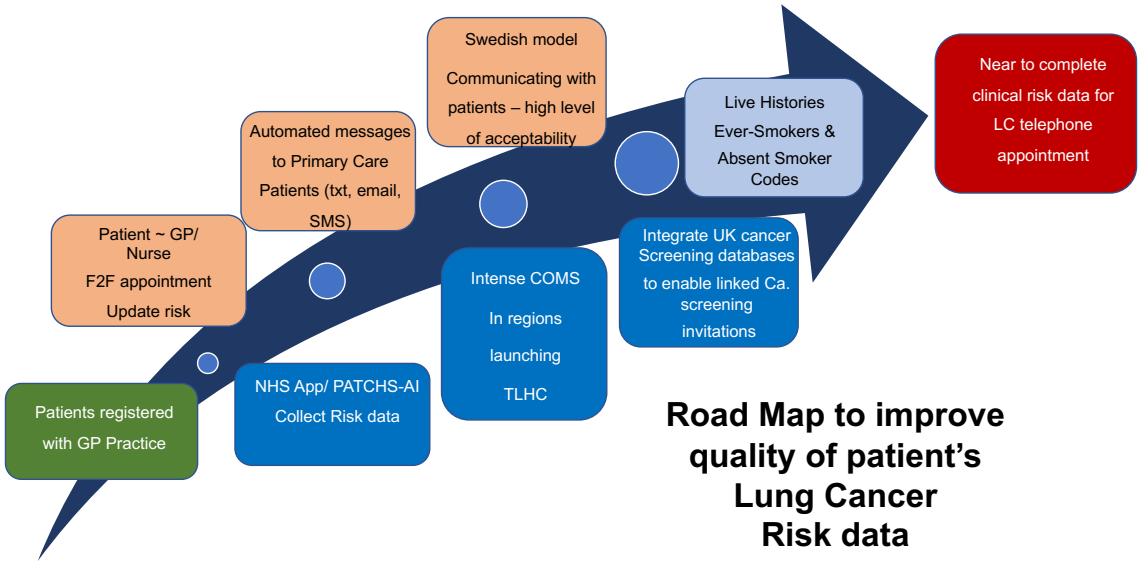
The long-term success of the TLHC depends on the availability of accurate patient risk-data in primary care; in theory these systems are in place but not always in practice. This Editorial proposes a Road Map to improve quality of patient’s Lung Cancer risk data (Figure 1), which summarises the link between the systems required to improve lung cancer screening invitation strategies and the goal to have ‘near to complete’ clinical risk data available for the Lung Cancer telephone screening appointment system.

The lessons learnt from the TLHC are applicable to all international lung cancer LDCT screening implementation programmes; accurate smoking data is one of the key parameters for identifying high risk populations, which in the UK is from the primary care setting whilst registry data is used is some countries. However, inconsistent reporting of smoking data within these electronic health care records will seriously impact on the targeted recruitment of individuals at risk of developing lung cancer.

1. Field JK, Duffy SW, Baldwin DR, Brain KE, Devaraj A, Eisen T, Green BA, Holemans JA, Kavanagh T, Kerr KM, Ledson M, Lifford KJ, McRonald FE, Nair A, Page RD, Parmar MK, Rintoul RC, Screaton N, Wald NJ, Weller D, Whynes DK, Williamson PR, Yadegarfar G, Hansell DM. The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. *Health Technol Assess* 2016; 20: 1-146.
2. Ghimire B, Maroni R, Vulkan D, Shah Z, Gaynor E, Timoney M, Jones L, Arvanitis R, Ledson M, Lukehirst L, Rutherford P, Clarke F, Gardner K, Marcus MW, Hill S, Fidoe D, Mason S, Smith SG, Quaife SL, Fitzgerald K, Poirier V, Duffy SW, Field JK. Evaluation of a health service adopting proactive approach to reduce high risk of lung cancer: The Liverpool Healthy Lung Programme. *Lung Cancer* 2019; 134: 66-71.
3. Crosbie PA, Balata H, Evison M, Atack M, Bayliss-Brideaux V, Colligan D, Duerden R, Eaglesfield J, Edwards T, Elton P, Foster J, Greaves M, Hayler G, Higgins C, Howells J, Irion K, Karunaratne D, Kelly J, King Z, Manson S, Mellor S, Miller D, Myerscough A, Newton T, O'Leary M, Pearson R, Pickford J, Sawyer R, Screaton NJ, Sharman A, Simmons M, Smith E, Taylor B, Taylor S, Walsham A, Watts A, Whittaker J, Yarnell L, Threlfall A, Barber PV, Tonge J, Booton R. Implementing lung cancer screening: baseline results from a community-based 'Lung Health Check' pilot in deprived areas of Manchester. *Thorax* 2019; 74: 405-409.
4. Bartlett EC, Kemp SV, Ridge CA, Desai SR, Mirsadraee S, Morjaria JB, Shah PL, Popat S, Nicholson AG, Rice AJ, Jordan S, Begum S, Mani A, Derbyshire J, Morris K, Chen M, Peacock C, Addis J, Martins M, Kaye SB, Padley SPG, Devaraj A, West London Lung Screening Collaboration G. Baseline Results of the West London lung cancer screening pilot study - Impact of mobile scanners and dual risk model utilisation. *Lung Cancer* 2020; 148: 12-19.
5. Crosbie PA, Gabe R, Simmonds I, Kennedy M, Rogerson S, Ahmed N, Baldwin DR, Booton R, Cochrane A, Darby M, Franks K, Hinde S, Janes SM, Macleod U, Messenger M, Moller H, Murray RL, Neal RD, Quaife SL, Sculpher M, Tharmanathan P, Torgerson D, Callister ME. Yorkshire Lung Screening Trial (YLST): protocol for a randomised controlled trial to evaluate invitation to community-based low-dose CT screening for lung cancer versus usual care in a targeted population at risk. *BMJ open* 2020; 10: e037075.
6. Field JK, Vulkan D, Davies MPA, Baldwin DR, Brain KE, Devaraj A, Eisen T, Gosney J, Green BA, Holemans JA, Kavanagh T, Kerr KM, Ledson M, Lifford KJ, McRonald FE, Nair A, Page RD, Parmar MKB, Rassl DM, Rintoul RC, Screaton NJ, Wald NJ, Weller D, Whynes DK, Williamson PR, Yadegarfar G, Gabe R, Duffy SW. Lung cancer mortality reduction by LDCT screening: UKLS randomised trial results and international meta-analysis. *Lancet Reg Health Eur* 2021; 10: 100179.
7. GOV.UK. New lung cancer screening roll out to detect cancer sooner. https://www.gov.uk/government/news/new-lung-cancer-screening-roll-out-to-detect-cancer-sooner: Department of Health and Social Care; 2023.
8. O'Dowd EL, Lee RW, Akram AR, Bartlett EC, Bradley SH, Brain K, Callister MEJ, Chen Y, Devaraj A, Eccles SR, Field JK, Fox J, Grundy S, Janes SM, Ledson M, MacKean M, Mackie A, McManus KG, Murray RL, Nair A, Quaife SL, Rintoul R, Stevenson A, Summers Y, Wilkinson LS, Booton R, Baldwin DR, Crosbie P. Defining the road map to a UK national lung cancer screening programme. *Lancet Oncol* 2023; 24: e207-e218.

9. NHS-Eng-National-Cancer-Programme. Targeted Screening for Lung Cancer with Low Radiation Dose Computed Tomography; Standard Protocol prepared for the Targeted Lung Health Checks Programme. 2019 [Access date13-01-2023]. Available from: https://www.england.nhs.uk/publication/targeted-screening-for-lung-cancer/.
10. Oudkerk M, Liu S, Heuvelmans MA, Walter JE, Field JK. Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives. *Nat Rev Clin Oncol* 2021; 18: 135-151.
11. Goodley P. Invitation strategies and participation in a community based lung cancer screening programme located in areas of high socio-economic deprivation. *Thorax* 2023; (in Press).
12. Dickson JL, Hall H, Horst C, Tisi S, Verghese P, Worboys S, Perugia A, Rusius J, Mullin AM, Teague J, Farrelly L, Bowyer V, Gyertson K, Bojang F, Levermore C, Anastasiadis T, McCabe J, Devaraj A, Nair A, Navani N, Hackshaw A, Quaife SL, Janes SM, consortium S. Utilisation of primary care electronic patient records for identification and targeted invitation of individuals to a lung cancer screening programme. *Lung Cancer* 2022; 173: 94-100.
13. Davies MPA, Vulkan D, Gabe R, Duffy SW, Field JK. Impact of Low Dose CT Screening on Cause of Death in Different Socio-Economic Groups. *Journal of Thoracic Oncology* 2022; 17: S190.
14. Belfrage S, Helgesson G, Lynoe N. Trust and digital privacy in healthcare: a cross-sectional descriptive study of trust and attitudes towards uses of electronic health data among the general public in Sweden. *BMC Med Ethics* 2022; 23: 19.
15. NHS. PATCHS- Contacting your practice via the NHS App. 2023 [cited 2023. Available from: https://help.patchs.ai/hc/en-gb/articles/10287963513111-Contacting-your-practice-via-the-NHS-App.

Figure 1


Patients registered with GP Practice
Patient ~ GP/ Nurse F2F appointment Update risk
NHS App/ PATCHS-AI Collect Risk data
Automated messages to Primary Care Patients (txt, email, SMS)
Swedish model Communicating with patients – high level of acceptability
Intense COMS In regions launching TLHC
Live Histories Ever-Smokers & Absent Smoker Codes
Integrate UK cancer Screening databases to enable linked Ca. screening invitations
Near to complete clinical risk data for LC telephone appointment
Road Map to improve quality of patient's Lung Cancer Risk data