The potential of breath analysis to improve outcome for patients with lung cancer

To cite this article before publication: Sophia Antoniou et al 2019 J. Breath Res. in press https://doi.org/10.1088/1752-7163/ab0bee

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2018 IOP Publishing Ltd.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.
Title: ‘The potential of breath analysis to improve outcome for patients with lung cancer’

Review Article

Authors

S X Antoniou*, E Gaude*, M Ruparel1, M P van der Schee2, S M Janes1 and R C Rintoul3,4, on behalf of LuCID research group*.

1 Lungs for Living Research Centre, UCL Respiratory, University College London, London UK

2 Owlstone Medical, Cambridge, UK

3 Papworth Trials Unit Collaboration, Royal Papworth Hospital, Cambridge UK

4 Department of Oncology, University of Cambridge

* Equal contribution

Corresponding author: M P van der Schee: marc.vanderschee@owlstone.co.uk

Conflict of Interest Statement

EG, and MvdS are fulltime employees of Owlstone Medical Ltd

Acknowledgements

We’d like to acknowledge the significant contribution all researchers, clinicians, nurses and patients make to our research. S.M.J. is a Wellcome Trust Senior Fellow in Clinical Science. S.M.J. is supported by the Rosetrees Trust, the Welton Trust, the Stonygate and the Garfield Weston Trust and UCLH Charitable Foundation. This work was partially undertaken at UCLH/UCL who received a proportion of funding from the Department of Health’s NIHR Biomedical Research Centre’s funding scheme (S.M.J.). S.M.J. and S.X.A. are funded by the Roy Castle Lung Cancer Foundation. RCR is funded by the NIHR Cambridge Biomedical Research Centre and Cancer Research UK Cambridge Cancer Centre.

Abstract

Lung cancer remains the most common cause of cancer related death in both the UK and USA. Development of diagnostic approaches that have the ability to detect lung cancer early are a research priority with potential to improve survival. Analysis of exhaled breath metabolites, or volatile organic compounds (VOCs) is an area of considerable interest as it could fulfil such requirements. Numerous studies have shown that VOC profiles are different in the breath of patients with lung cancer compared to healthy individuals or those with non-malignant lung diseases. This review provides a scientific and clinical assessment of the potential value of a breath test in lung cancer. It discusses the current understanding of metabolic pathways that contribute to exhaled VOC production in lung cancer and reviews the research conducted to date. Finally, we highlight important areas for future research and discuss how a breath test could be incorporated into various clinical pathways.
1. Introduction

Although the survival rate for lung cancer in the UK has doubled in the last fifteen years it remains the most common cause of cancer related death in both the UK and USA \(^1\). Unfortunately, around 75% of patients who present with lung cancer have advanced stage disease which, for the most part, is not curable. The development of diagnostic approaches that will have the ability to detect lung cancer early are a research priority. An ideal screening test should be sensitive, specific, minimally invasive and accessible to a large population as part of a national programme. Analysis of exhaled breath metabolites, or volatile organic compounds (VOCs) is an area of considerable interest as it could fulfil such requirements.

The basis of a lung cancer “Breath Biopsy” test is underpinned by the observation that alterations in key metabolic pathways are strongly linked to the transformation from healthy cells to malignant cells \(^2\). Some of the metabolites generated by these processes are volatile and excreted through the airways, making them measurable in exhaled breath.

The purpose of this review is to provide a scientific and clinical assessment of the potential value of a breath test in cancer, with a specific focus on lung cancer. We review the metabolic alterations occurring in lung cancer cells relevant to the production of VOCs. Subsequently, we will discuss research that has been conducted to date to validate the presence of volatile biomarkers in lung cancer. We will discuss the future steps that need to be taken in breath research, as despite decades of research looking at VOCs in disease, including lung cancer, translation into the clinical environment has been slow. This can, in part, be attributed to difficulties in standardising breath collection resulting in varying methods and small studies. The review summarises breath research to date and looks to the future; it is not an exhaustive review of this topic or breath collection processes, which have been discussed previously\(^3,4\). Finally, we will discuss how a breath test could be incorporated into various clinical pathways.
2. Association between altered metabolism in cancer and production of VOCs

2.1 Common metabolic traits of cancer

Together with well-established hallmarks of cancer transformation, such as uncontrolled cell proliferation, resistance to cell death or de-differentiation, the alteration of metabolism has been identified as among the most common aberrations in cancer. During the last few decades, a plethora of metabolic enzymes and pathways have been found to be altered in order to support the high energy requirements of cancer cells. It is now well-established that mutations of oncogenes and tumour suppressor genes are associated with activation of metabolic pathways that support anabolism, i.e. the construction of complex biosynthetic molecules to provide the building blocks for cell proliferation, a metabolic phenotype fundamental for survival and growth of cancer cells. In this section we will briefly introduce some examples of anabolic pathways activated by cancer cells. More extensive descriptions of specific pathways that support cancer growth have been reviewed elsewhere.

Activation of glycolysis is one of the hallmarks of cancer metabolism. Aberrant activation of glycolysis was first described by Otto Warburg in the 1920s, who observed increased production of lactic acid in slices of rat liver carcinoma. In recent years, Warburg’s findings have been widely confirmed by numerous studies, linking a plethora of cancer-associated mutations to activation of glycolysis. High glycolytic rates can support anabolism via the production of numerous intermediates that can initiate the generation of biosynthetic molecules. For instance, increased flux through glycolysis can provide intermediates of the pentose phosphate pathway (PPP), an important biosynthetic pathway leading to production of reduced nicotinamide adenine dinucleotide phosphate (NADPH), an important reducing intermediate utilised in the management of oxidative stress, among other functions. In addition, the PPP can provide the backbone sugar for RNA and DNA molecules, thus supporting cell proliferation via supply of intermediates for DNA replication. In line with this evidence, nucleotide biosynthesis has been suggested as a common driver of cancer cell proliferation in a wide screening of cancer cell lines, thus indicating the relevance of nucleotide supply for cancer cell growth. Indeed, inhibitors of nucleotide synthesis are among the most efficient chemotherapeutic drugs currently used in the clinic.

Activation of aerobic glycolysis in cancer cells has been hypothesised to be the bioenergetic consequence of dysfunction of mitochondrial metabolism. Mitochondria are the major site of oxygen consumption in the cell and they are responsible for the optimised production of energy and biosynthetic molecules in terminally differentiated cells. Although several studies have highlighted a strong association between mitochondrial metabolism and human cancers, the role of mitochondria during cancer transformation is highly heterogeneous. Several mutations of the mitochondrial genome have been observed in different cancers including colon, breast, lung, prostate, liver, pancreas, kidney, thyroid, brain, gastric carcinoma and ovarian cancer and have been implicated with metabolic alterations. In line with this evidence, when comparing the abundance of mtDNA between normal and cancer samples Reznik and colleagues found that mtDNA is widely depleted in several types of human cancer. A comprehensive study found similar mutational signatures in ~40 different types of cancer, with gene truncating mutations being enriched in kidney, colorectal and thyroid cancers. Interestingly, the authors also found that transfer of mitochondrial genes into the nucleus of cancer cells induced disruption of critical genes for cancer development, such as ErbB2.
addition, a recent study found that mitochondrial single nucleotide variants (mtSNVs) accumulate in
mutational hot-spots in the mtDNA of patients with prostate cancer \(^25\), suggesting positive selection
of specific mutations in the mitochondrial genome. Notably, mtSNVs were associated with copy
number alterations of the nuclear oncogene MYC and specific mtSNVs were linked to patient survival.
Together, this evidence suggests that mutation of mtDNA can not only affect metabolic enzymes, but
also mutation of nuclear genes associated to cancer transformation, providing a potential link
between mtDNA mutations and cancer formation. Genetic alteration of several enzymes of the
tricarboxylic acid (TCA) cycle, a central mitochondrial pathway, has been shown to drive formation of
an array of cancer types \(^{26-30}\). Classical examples are the TCA cycle enzymes isocitrate dehydrogenase
(IDH) 1-2, succinate dehydrogenase (SDH), and fumarate hydratase (FH), whose mutations are
associated with glioma and leukaemia, pheochromocytoma and paraganglioma, as well as hereditary
leiomyomatosis and renal cell cancer syndrome \(^9\). Furthermore, mutations affecting the mitochondrial
respiratory chain lead to mitochondrial dysfunction and have been associated with cancer formation
across a wide range of human tumours \(^9,23\). Overall, the metabolic changes caused by mitochondrial
dysfunction could drive the acquisition of proliferative metabolic programs which, together with a
permissive environment, can lead to cancer formation. These interesting concepts have been
extensively reviewed elsewhere \(^9\). Nevertheless, it is well established that complete mitochondrial
dysfunction is detrimental for the survival of cancer cells. Chemical ablation of mitochondrial DNA
inhibits cancer cells growth rate, proliferation in soft agar and tumour growth in nude mice \(^{31,32}\).
Moreover, intact mitochondrial metabolism has been shown to support proliferation of cancer cells
\(^{33,34}\). Together, these findings indicate that regulation of mitochondrial metabolism is an important
determinant of cancer transformation and likely depends on multiple factors, such as oncogenic
mutations and environmental conditions \(^21\).

The evidence collected over the last few decades indicates that a profound rewiring of
metabolism accompanies transformation of cancer cells. Mutated oncogenes and tumour suppressor
genes drive metabolic changes that, together with tissue-specific environmental cues, determine
cancer survival.

### 2.2 Lung cancer-specific metabolic traits

The metabolic landscape of cancer cells in the lung has been analysed in recent years and lung
cancer-specific metabolic traits have emerged. The laboratories of Ralph DeBerardinis and Theresa
Fan have conducted pioneering studies where they exploited carbon tracing in human patients to
investigate metabolic rewiring of human lung cancer \(^{35-38}\). Injection of lung cancer patients with \(^13\)C-
glucose showed that, compared to healthy lung tissue, cancer cells increase glucose metabolism
through glycolysis, as well as full oxidation of glucose in the mitochondria. In the same studies, activity
of the enzyme pyruvate carboxylase (PC) has been shown to support increased mitochondrial TCA
cycle \(^{35,36}\). In line with this evidence, Hensley and colleagues reported increased labelling of TCA cycle
intermediates from \(^13\)C-glucose in human lung tumours and argued that additional carbon sources
might contribute to the TCA cycle and increased mitochondrial metabolism \(^37\). Indeed, one year later,
a study from the same lab showed that lactate is a major carbon contributor to the TCA cycle in lung
tumours of human patients compared to healthy lung tissue \(^38\). This evidence was confirmed in an
independent study showing that lactate is the preferred anaplerotic substrate in tumours of the lung
specifically \(^39\). Finally, activation of mitochondrial metabolism has been confirmed by recent
bioinformatic studies where genes encoding for TCA cycle and oxidative phosphorylation pathways
were found increased in both lung adenocarcinomas and lung squamous carcinomas, compared to
healthy lung \(^21\). In addition, lung adenocarcinoma was found as the only human cancer displaying
increased mitochondrial DNA abundance \(^22\).
Other studies have highlighted the importance of redox regulation for lung cancer cells. Singh and colleagues reported that 20-40% of human lung tumours bear genetic inactivation of Kelch-like ECH-associated protein 1 (KEAP1), a negative regulator of Nuclear factor erythroid-2 related factor 2 (NRF2). This results in stabilisation of NRF2 and activation of a strong antioxidant response that includes increased activity of antioxidant enzymes and increased levels of glutathione. In line with this evidence, oncogenic mutations affecting KRAS, B-RAF, and c-Myc lead to upregulation of NRF2 and drive tumourigenesis in vivo via a mechanism that involves reactive oxygen species (ROS) regulation, demonstrating that multiple cancer mutations converge towards redox control as a common mechanism of cancer formation. Of note, A549 adenocarcinoma cell line is dependent on the oxidative arm of the PPP, a glucose-fed metabolic pathway for management of oxidative stress. In addition, A549 cells rewire the TCA cycle to support biosynthesis of glutathione and reduce ROS levels. Finally, KRAS-driven lung cancer in mice has been shown recently to be dependent on glutathione and glucose metabolism, confirming the role of glutathione biosynthesis in lung tumourigenesis. Excellent reviews provide a full picture of the current understanding of the molecular mechanisms governing metabolic adaptations in lung cancer.

The evidence presented here shows that cancer formation in the lung leads to a specific reprogramming of metabolism, characterised by increased glucose and lactate oxidation in the mitochondria, as well as by increased glutathione biosynthesis. Together with transcriptional control mechanisms via NRF2, the metabolic rewiring observed in lung cancer cells supports a strong antioxidant response that is fundamental for cancer survival. Increased reliance on oxidative metabolism and redox control might be dictated by the peculiar environmental conditions present in the lung, such as the relatively high concentration of oxygen.

2.3 Altered metabolism contributes to breath VOCs in lung cancer patients

Volatile organic compounds (VOCs) are molecules characterised by high volatility due to high vapour pressure at room temperature. When produced by the human body, VOCs are excreted via breath, skin, urine and faeces, among others, and are released into the external air. VOCs can be generated as part of physiological metabolic reactions and human breath is known to contain thousands of volatile small molecules. Metabolic changes occurring during cancer transformation have been shown to alter several metabolic pathways and can affect the production of small molecules, including VOCs. Indeed, numerous studies have reported altered levels of VOCs in the breath of patients with lung cancer, compared to healthy controls.

Despite this, the evidence collected so far on the association between lung cancer and altered VOCs remains correlative or anecdotal, and little is known about the mechanisms linking cancer metabolic rewiring and production of VOCs in lung cancer patients. In this section we review the current understanding of the metabolic pathways that could contribute to production of VOCs in the breath of patients with lung cancer.

Lipid peroxidation

Production of several volatile molecules, such as alkanes, alkenes and aldehydes, can arise from lipid peroxidation. Pentane and ethane, as well as malondialdehyde, are commonly found in
human breath and have been shown to correlate with oxidative stress in different clinical settings. Importantly, products of lipid peroxidation are among the most common breath biomarkers of lung cancer. The levels of decane, heptanial, octane, and undecane have consistently been found to be increased in the breath of patients with lung cancer, compared to controls. A study conducted on pulmonary nodules found that 1-octene, an alkene derivative of octane, could differentiate between benign and malignant nodules, thus indicating that markers of lipid peroxidation might be used for early detection of lung cancer. Moreover, Broza and colleagues analysed the breath of patients before and after surgical removal of lung tumours and found increased levels of the lipid peroxidation marker trimethylhexane, thus adding convincing evidence about the role of this pathway for breath diagnosis of lung cancer.

Lipid peroxidation is a process whereby free radicals and ROS oxidise membrane phospholipids, as well as free fatty acids. This process occurs primarily on polyunsaturated fatty acids, due to the presence of methylene groups residing between two double bonds (bisallylic groups) (Figure 1A). The hydrogen atoms of the methylene groups are attracted, and weakened, by the neighbouring double bonds, and can be easily abstracted by radicals, resulting in lipid peroxyl radicals and hydroperoxides. Membranes in all cellular and subcellular compartments can be attacked, with plasma and mitochondrial membrane being the main site of lipid peroxidation in the cell. While lipid peroxidation at the plasma membrane is mainly driven by NAD(P)H oxidases, mitochondrial lipid peroxidation is caused by accumulation of ROS (Figure 1A). Several lines of evidence indicate that abnormal function of mitochondrial metabolism is associated with increased production of ROS. Due to its electron transport activity, the respiratory chain is the main site of ROS production within the mitochondria. Reduced availability of final electron acceptors or excess of electron donors can cause the leakage of electrons from the respiratory chain, resulting in the one-electron reduction of molecular oxygen to oxygen radical. Moreover, the inner mitochondrial membrane is enriched in cardiolipin, a phospholipid particularly prone to attack by ROS due to its high content of polyunsaturated fatty acids. The combination of high yield of ROS in the mitochondria, together with the presence of proximal lipids that can act as electron acceptors, suggests that the mitochondria could be an important site of lipid peroxidation. In addition, mitochondria-derived ROS can diffuse into the cytosol and attack extra-mitochondrial lipids, as well as other molecules.

As mentioned above, detoxification of ROS is critical for lung cancers and impairment of antioxidant mechanisms have been shown to inhibit proliferation of lung cancer cells in several studies. In addition, defective mitochondrial redox control by polymorphisms of superoxide dismutase 2 (SOD2), one of the main mitochondrial antioxidant mechanisms, are linked to increased risk of lung cancer. Importantly, lung cancer cell lines and lung tumours select for molecular mechanisms that protect from ferroptosis, a form of regulated cell death induced by the accumulation of lipid peroxidation products via ROS- and iron-dependent reactions. Nitrogen fixation of S. cerevisiae homolog 1 (NFS1) is induced in well-differentiated, early-stage lung tumours, to provide Fe-S clusters to cancer cells and suppress iron-induced lipid peroxidation and ferroptosis. Suppression of NFS1 alone, or in combination with inhibition of antioxidant mechanisms, can significantly impair lung tumour formation in vivo. Glutathione peroxidase 4 (GPX4), a master regulator of ferroptosis (see figure 1A), is localised in the mitochondrial compartment and primarily acts on cardiolipin. Together with the evidence that ferroptosis induces changes of mitochondrial morphology this suggests that mitochondria could be important for the lipid peroxidation process that leads to
ferroptosis. Together with the important role of redox control in supporting proliferation of lung cancer cells, this evidence indicates that antioxidant mechanisms that control lipid peroxidation could be important for lung cancer cells.

This evidence indicates that detection of products of lipid peroxidation might be a valid strategy for diagnosis of lung cancer at different stages of disease progression. This might be justified, at least in part, by increased oxidative damage at different stages of lung tumourigenesis. Currently it is unknown what the net effect of the increased level of ROS and upregulation of detoxifying mechanisms are on the level of VOCs derived from lipid peroxidation.

Besides changes in the level of lipid peroxidation the nature of the products may also change; differences in membrane lipid composition are observed in lung cancer samples, compared to healthy lung \(^{77,78}\), suggesting that production of specific VOCs might arise from peroxidation of cancer-specific lipid species. Interestingly, a matrix of different alkanes and methylated alkanes originating from lipid peroxidation has been applied as a test for systemic oxidative stress \(^{79}\), suggesting that a similar approach might be tailored for detection of lung cancer biomarkers. Implementation of a panel of multiple biomarkers, including primary products of lipid peroxidation, as well as their derivatives, is likely to be necessary for associating biomarkers of lipid peroxidation with lung cancer specifically.

**Mevalonate biosynthesis pathway**

Synthesis of the mevalonate pathway intermediate dimethylallyl-diphosphate (IPP) is associated with production of the volatile molecule isoprene (Figure 1B), which has been linked to a plethora of pathophysiological conditions. Deneris and colleagues first reported the generation of isoprene from mevalonate in rat liver homogenates \(^{80}\). The authors proposed the explanation that an acid-catalysed elimination reaction caused the production of isoprene from DMPP \(^{80}\) and this reaction is thought to be catalysed under physiological conditions by isopentenyl-diphosphate isomerase \(^{81}\). However, the first association between breath isoprene levels and the mevalonate pathway derived from the observation that administration of lovastatin led to decreased breath isoprene, as well as blood cholesterol, in human healthy patients \(^{82}\). Numerous studies have reported the presence of increased levels of isoprene in the breath of lung cancer patients, compared to healthy controls or patients with other lung disorders \(^{51,52,83,84}\) and isoprene was among the top breath VOC biomarkers for lung cancer in a recent meta-analysis \(^{85}\). Although one study reported decreased levels of isoprene in the breath of cancer patients \(^{62}\), the evidence collected so far suggests that breath isoprene could be a biomarker for lung cancer.

The mevalonate pathway is a cytosolic chain of reactions that uses acetyl-CoA as a precursor for the generation of biosynthetic molecules important for cell membranes, protein modifications and mitochondrial respiration \(^{86}\). Acetyl-CoA can derive from the mitochondrial decarboxylation of pyruvate and \(\beta\)-oxidation of fatty acids, as well as from cytosolic metabolism of acetate \(^{87}\) and glutamine reductive carboxylation \(^{88}\); notably, deregulation of these metabolic pathways has been shown to support proliferation of cancer cells \(^{2}\). The pathway is initiated by the rate-limiting enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), which converts HMG-CoA to mevalonate, and continues with the production of several isoprenoids, such as dimethylallyl-, geranyl-, and farnesyl-diphosphate, important molecules for protein modification and function (Figure 1B) \(^{89}\). Farnesyl-
diphosphate (FPP) can then be subsequently metabolised to squalene, precursor of the fundamental cell membrane component cholesterol, or FPP can generate geranylgeranyl-diphosphate (GGPP), which can further lead to generation of dolichol, important for N-glycosylation, or ubiquinone, a fundamental co-factor for the mitochondrial electron transport chain. This is in line with research highlighting the link between exhaled isoprene and cholesterol metabolism.

Numerous studies have highlighted the association between activation of the mevalonate pathway and proliferation of cancer cells. Ectopic expression of HMGCR increases anchorage-independent growth in vitro, as well as growth in nude mice of liver and breast cancer cell lines. The same study also highlighted that increased mRNA expression of genes for mevalonate pathway is associated with poor survival of breast cancer patients. Furthermore, inhibition of HMGCR with different forms of statins was able to impair cancer cell proliferation in several animal models, including mouse colorectal cancer, rat liver carcinoma, mouse ovarian and prostate cancer. Finally, statins treatment has been shown to block proliferation of lung tumours in mice, as well as formation of lung metastasis from melanoma tumours. Together with the finding that injection of the metabolite mevalonate could increase the size of breast cancer xenografts in mice, this evidence indicates that the mevalonate pathway is an important determinant of cell proliferation in different cancer settings.

Beyond supporting cell proliferation by providing important anabolic molecules, the mevalonate pathway is intimately connected with cell metabolism via production of ubiquinone. Ubiquinone is a component of the mitochondrial respiratory chain, where it allows the shuffling of electrons through the mitochondrial complexes involved with oxidative phosphorylation. Inhibition of the mevalonate pathway with statins can reduce ubiquinone levels and is known to cause muscle disease via induction of mitochondrial dysfunction. The resulting leakage of electrons to oxygen can result in increased oxygen radicals, contributing to the aforementioned oxidative stress. In addition to its role in mitochondrial respiration, ubiquinone can function as an antioxidant molecule within cell membranes and it has been shown to reduce lipid peroxidation. In line with this hypothesis, statins can induce lipid peroxidation and sensitise therapy-resistant cancer cells to ferroptosis, indicating that mevalonate-derived ubiquinone can be an important mechanism for regulation of ferroptosis in cancer cells.

Despite the strong evidence, isoprene has not yet been applied as a biomarker for diagnosis of lung cancer in the clinic. This could be due to the multifactorial regulation of isoprene secretion, which includes age as well as with circadian rhythms. This is likely to increase the interpatient variability and hampers the association of breath isoprene to specific pathological conditions in real-life cross-sectional clinical settings. Targeted approaches aimed at eliciting isoprene secretion directly from cancer cells might help in the identification of cancer-specific isoprene levels.

Ketones

The term ketone bodies refer to three small molecules, namely acetoacetate, β-hydroxybutyrate and acetone, produced from the precursor acetyl-CoA. Although concentrations of the ketone body acetone are low compared to β-hydroxybutyrate or acetoacetate, acetone is readily
detected in human breath due to its high volatility. Acetone can be produced from the spontaneous decarboxylation of acetoacetate, and it has been used as a biomarker for activation of ketone metabolism. Indeed, breath acetone has been found increased in patients with type 1 diabetes and it is correlated with blood glucose levels, as well as being shown to be a valid readout of ketone bodies in epileptic patients undergoing a ketogenic diet. Moreover, several studies have found increased acetone in the breath of lung cancer patients, compared to controls, suggesting that metabolism of ketone bodies might be important for lung cancer cells, and that acetone might be a potential biomarker for diagnosis of lung cancer. Nevertheless, it is important to note that increased breath acetone has been associated with several other diseases, such as non-alcoholic fatty liver disease, alcohol hepatitis, and heart failure, indicating that acetone might not be a specific disease biomarker.

Under physiological, non-fasting conditions, mitochondrial acetyl-CoA produced via pyruvate decarboxylation, fatty acid β-oxidation, or amino acid catabolism, is condensed with oxaloacetate (OAA) to generate citrate, thus supporting energy production through the TCA cycle and mitochondrial respiration. Under glucose starvation or pathological conditions (e.g., diabetes), availability of oxaloacetate can become limiting, leading to accumulation of acetyl-CoA. Excess acetyl-CoA is condensed with acetoacetyl-CoA by mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase (HMGCS2), generating HMG-CoA, which is finally metabolised to acetoacetate and acetyl-CoA by HMG-CoA lyase (HMGCL). Acetoacetate can then be decarboxylated non-enzymatically to acetone or it can be converted to β-hydroxybutyrate by β-hydroxybutyrate dehydrogenase (BDH1). This set of reactions normally occurs in the liver, leading to secretion in the blood stream of β-hydroxybutyrate, which can be taken up by different tissues and converted back to acetoacetyl-CoA and acetyl-CoA, thus supporting energy production in the absence of glucose.

Existing evidence on the role of ketone bodies in cancer biology is somewhat controversial. A recent meta-analysis reported that ketogenic diet (KD), a low-carb high-fat dietary regimen designed to increase ketone body generation, reduces tumour growth in mice. Moreover, ketogenic diet (KD) has been proposed as an adjuvant strategy in conjunction with anti-cancer therapy, thus suggesting that increased levels of ketone bodies are detrimental for proliferation of cancer cells. Nevertheless, other studies have proposed a supportive role for ketone body metabolism on cancer cell growth. The work from Lisanti’s lab has proposed in recent years the concept of a two-compartment metabolism, whereby tumour stroma can provide several substrates that support proliferation of cancer cells. Among these, β-hydroxybutyrate has been shown to support tumour growth in a mouse xenograft model of breast cancer and, while ketogenic enzymes are increased in the stromal cells, epithelial cancer cells increase the expression of enzymes for ketones utilisation, suggesting that cancer cells could enhance degradation of ketone bodies to support energy generation. The controversy on the role of ketones in cancer biology could be explained by different cancer types being analysed, where different environmental conditions and/or genetic background could contribute to differential response of cancer cells to ketone bodies. A possible explanation could be linked to the role of ketone bodies in mitochondrial metabolism. Supplementation with the ketones β-hydroxybutyrate or butanediol induces mitochondrial biogenesis in a breast cancer cell line and treatment with β-hydroxybutyrate or acetoacetate increases respiration and diminishes production of mitochondrial ROS in primary neuron cultures. This evidence suggests that utilisation of ketones by cancer cells might depend on activation of mitochondrial metabolism, with ketone bodies degradation being...
increased in cancer types characterised by high oxidative metabolism, potentially to control ROS production. Nevertheless, further understanding of this pathway is necessary to elaborate specific hypotheses on the role of ketones in different cancer types.

Among ketones that have been detected in human breath of cancer patients, 2-butanone has been reported to have a strong association with lung cancer in a recent meta-analysis. Indeed, increased levels of 2-butanone were found in the breath of lung cancer patients compared to controls. Moreover, 2-butanone was increased in patients with stage II, III and IV lung cancers, compared to stage I, thus providing an association between disease progression and secretion of 2-butanone in the breath of lung cancer patients. Unlike other ketones, 2-butanone is not generated from acetoacetate through ketogenic metabolism, instead it is thought to derive from environmental pollutants such as paints and resins. Despite this, the strong association of exhaled 2-butanone and lung cancer might indicate an endogenous origin for this ketone. Interestingly, 2-butanone can be generated by non-enzymatic decarboxylation of methylacetoacetate, a catabolic intermediate in the degradation of the branched-chain amino acid isoleucine. Inborn defects of isoleucine metabolism can result in accumulation of 2-butanone in urine, which is utilised as a biomarker for these disorders. Of note, branched chain amino acid metabolism has recently been found to support tumour formation in the lung specifically, while this metabolic pathway was dispensable for tumourigenesis in the pancreas, indicating that activation of branched chain amino acid might be a specific metabolic rewiring of lung cancer cells. Based on this evidence, it is tempting to hypothesise that increased levels of 2-butanone in the breath of lung cancer patients might be associated with activation of branched chain amino acid in lung tumours. Dedicated research is needed to address this hypothesis.

Finally, detection of ketones has been facilitated by their high volatility, which leads to high secretion in human breath. Association of ketones, especially acetone, with different pathophysiological conditions is probably due to their intimate link with central carbon metabolism, a core metabolic pathway in diseases such as diabetes and cancer. Although this highlights the importance of ketones in detecting alterations of metabolism, the regulation of ketone secretion is likely to depend on multiple factors, as well as disease conditions, hampering the utilisation of these molecules as biomarkers for specific diseases. Nevertheless, triggering ketone metabolism through targeted strategies that rely on disease-specific metabolic alterations of this pathway might help the utilisation of ketones as specific breath biomarkers.

In conclusion, deep understanding of the metabolic pathways altered in lung cancer cells can help to shed light on the molecular pathways leading to production of breath VOCs from lung tumours specifically. The evidence reported here indicates that alteration of specific metabolic pathways in lung cancer can lead to production of particular VOCs. This opens the possibility to exploit cancer-specific metabolic alteration via design of strategies that specifically assess the extent of the metabolic aberration. An example of such strategies is offered by the breath test for infection of H. Pylori, where understanding of disease-specific metabolic alterations has led to the development of a specific diagnostic test based on breath biomarkers. Of note, such an approach might not only lead to the development of specific methods for early detection of lung cancer but could also drive the design of strategies for monitoring disease progression, as well as response to therapy. Despite the extensive evidence on metabolic alterations in lung cancer cells, as well as the numerous studies showing altered
levels of breath metabolites in lung cancer patients (see next sections), these two fields of research have seldom communicated. In fact, the evidence linking metabolic alterations in lung cancer cells and breath biomarkers remains highly correlative and very few studies have gathered mechanistic evidence in this regard. Combining the information from metabolic studies of lung cancer, together with robust analytical techniques of breath analysis, might lead to the development of effective breath tests for lung cancer biomarkers, potentially improving current clinical practice for diagnosis and management of lung cancer patients. In the next section we will review the existing evidence pertaining to VOC analysis for diagnosis of lung cancer with the aim of identifying advantages and disadvantages of different approaches in breath clinical research studies.

3. Clinical research to date

Table 1 summarises studies investigating the value of VOCs in lung cancer performed to date. There is significant variability in design and execution and the majority of studies are small. Therefore, whilst these studies provide a proof of concept and show that breath analysis can differentiate between lung cancer and health, they have several limitations which have prevented translation of this research into the clinical environment.

One of the largest studies is by Phillips and colleagues in 2008. They found that breath analysis could detect lung cancer. However, they found that an individual VOC or a small number (<5) did not have sufficient discriminatory power. They identified 30 VOCs with the greatest predictive power and combined these in an algorithm using weighted digital analysis (WDA). This distinguished lung cancer from controls with a sensitivity of 84.5% and a specificity of 81% without the stage of lung cancer or smoking history affecting the accuracy of the results. In a further study this group validated a similarly devised algorithm in a separate group of subjects, Table 1.

Similarly, other groups have found that a group of VOCs were able to differentiate lung cancer patients from controls. Bajtarevic and colleagues, found that the sensitivity to identify lung cancer compared to controls improved when a larger number of VOCs were included in the analysis. With 4 VOCs the sensitivity was only 52% but this increased to 80% with 21 VOCs. Wang and colleagues, separated patients with lung cancer from controls using a group of 23 VOCs with a sensitivity of 96.5% and sensitivity 97.5%. In studies which used eNose analysis certain ‘breath patterns’ were detected and able to distinguish between subjects with lung cancer and controls. Van der Goor and colleagues have conducted one of the largest studies looking at eNose analysis and were able to distinguish between lung cancer and control groups with a sensitivity of 88% and specificity 86% in their validation cohort. Machado and colleagues also performed breath analysis using an eNose. They initially generated a classifier that was tested in a validation group with 14 lung cancer patients. The authors reported a test sensitivity of 71.4%, specificity of 91.9%, and positive and negative predictive values of 66.6% and 93.4% respectively in the validation group. The histological type or stage of lung cancer or lung function did not impact these results. These studies as with others outlined in Table 1 support the proof of concept but interpretation of results and application is limited by the lack of robust validation as the varying statistical methods used limits their external validity.

Some groups have looked at headspace VOCs, which is defined as the air in the space around the lung cancer tissue or cell line. Wang and colleagues compared headspace VOCs to those from exhaled
breath. They found that exhaled breath contained a greater variety of VOCs compared to the cell line headspace. This finding supports the notion that VOCs for lung cancer likely originate both from the tumour itself and from systemic metabolic effects of tumourigenesis. Another group also analysed headspace of cells lines and compared the headspace of NSCLC and SCLC cell lines. This group identified 12 VOCs that differed between histological types and a further 9 VOCs that could differentiate subtypes of NSCLC and they subsequently developed gold nanoparticle sensors to detect these changes. They then showed, in vitro, that these sensors could differentiate NSCLC and SCLC, (sensitivity 100%, specificity 75%) as well as subtypes of NSCLC, (sensitivity 100%, specificity 67% and 90% accuracy). Translating these in vitro models to the in vivo population is the next step, and although these results and studies come from a small sample size they help provide further understanding of the origin and potential utility of VOCs in the diagnosis of lung cancer.

Despite these results, showing that analysis of VOCs can distinguish between lung cancer and controls, many of these studies have limitations which, in part, explains why after decades of research there is still no translation into routine clinical practice. They have often been conducted in small groups of patients, a common limitation of breath cancer research to date, and show significant variability in a number of areas including study design, collection methods and patient groups. This makes interpretation and pooling of data extremely difficult. Furthermore, an important factor limiting the translation of studies to date has been the lack of validation in independent cohorts, which limits both the reliability of study findings as well as the generalisability of individual findings to the overall population of lung cancer patients. Variations in study designs also impact the ability for findings to be generalised and translated to wider populations. A number of studies differ as to the stage of lung cancer in recruited patients. Some studies recruited patients undergoing curative treatment, with an early stage of lung cancer. While this population is important to investigate as they are the target for screening these studies are currently too small to allow for appropriate statistical validation.

Di Natale and colleagues studied 35 patients with early stage lung cancer and compared them to 18 controls. By utilising a quarts-micro-balance eNose they achieved 100% correct classification of patients and 94% correct classification of controls. Similarly, Dragonieri and colleagues compared 10 patients with NSCLC, 10 COPD patients and 10 healthy controls distinguishing them with 90% accuracy after cross-validation. In most other studies a significant proportion of the patients had stage III or IV disease. In one, the majority of patients (53.9%) had stage IV lung cancer, with only 11.5% having stages I & II, a limitation acknowledged by the authors. Studies differ in their interpretation of whether staging of disease has any impact on the sensitivity of detection. Phillips et al and Machado et al found that stage of lung cancer, smoking history or lung function had no impact on the results. Gasparri and colleagues, found they had a higher sensitivity in detecting stage I cancers, compared to stages II to IV, 92% and 58% respectively. However there were only 8 patients with stage III and IV lung cancer, which may explain the lack of observed differences. Patient conditioning also varied between studies. In Wang’s study participants did not eat or smoke for 12 hours prior to sampling, whereas Van der Goors’ study had no such limitations. These differences make combining data to apply conclusions to a large population very difficult.

There is also variability in the control groups used by different studies. Control group definitions range from healthy participants alone, a combination of healthy recruits and individuals with other lung diseases, and the latter group alone. Machado and colleagues’ control group combined healthy participants and those with non-cancerous respiratory conditions; a high number of subjects had alpha 1 antitrypsin deficiency or chronic granulomatous disease secondary to beryllium exposure. Whilst testing the performance of a breath biomarker test in the context of other respiratory conditions has value, these conditions are not commonly found in an average lung cancer diagnostic
or screening population. Grouping healthy participants and those with other respiratory diseases without understanding the effect of the latter on exhaled VOC profile may pose problems. This is especially relevant when using pattern-based techniques such as eNose which do not allow selection of biomarkers based on their independence from confounders. The variation between groups may explain the moderate sensitivity found by this group compared to other studies (Table 1)\(^{128}\). Dragonieri and colleagues, looked at comparing healthy controls and those with COPD independently to cancer. Using eNose analysis they found ‘breath prints’ from NSCLC patients were different when compared to patients with COPD\(^{132}\).

It is recognised that certain co-morbidities such as COPD, liver and kidney disease, can contribute to the exhaled breath VOC profile, potentially confounding the identification of VOC profiles specific for lung cancer if appropriate control groups have not been selected. For instance VOCs that have been associated with lung cancer such as acetone, isoprene and alkanes, have been linked to both other conditions and physiological processes, including cystic fibrosis\(^{133}\), asthma\(^{134}\), malaria\(^{135}\), influenza\(^{136}\), renal disease\(^{137}\), muscle activity\(^{138}\), oxidative stress\(^{139}\) and age\(^{79}\). In studies, researchers have managed the possible influences of co-morbidities and conditions differently. Phillips and colleagues assumed any potential effects would be cancelled out as they had matched their control and cancer groups well,\(^{124}\) Gasparrì conducted a relatively large study, with 70 lung cancer participants and 76 healthy controls and performed sub-analysis to evaluate possible confounding effects of co-morbid conditions. Interestingly, they did not find a significant difference between patients with metabolic comorbidities; e.g. diabetes, and obesity compared to healthy controls\(^{129}\).

Wang and colleagues performed an extensive study aimed at finding VOCs that originate specifically from lung cancer cells in vivo\(^{126}\). To this end, they identified a set of breath VOCs that were significantly different between lung cancer patients and controls and compared them to VOCs from cancer tissues or cancer cells in vitro. They found nonadecane and 2-pentadecanone in all specimens, suggesting that these VOCs could be bona fide biomarkers of cancer cell metabolism in vivo. Notably, nonadecane originates from lipid peroxidation, further confirming the link between oxidative stress and lung cancer metabolism. The finding that 2-pentadecanone is found increased in all cancer specimens is in line with the recent finding that 2-pentadecanone originates from glucose metabolism\(^{122}\), one of the major metabolic pathways altered in lung cancer cells. This evidence suggests that VOCs might be present in breath as a result of the altered metabolism of relatively small masses of cancer cells in the lungs.

The research performed to date, whilst identifying the ability of breath VOCs to distinguish lung cancer, has still not identified a clear group of VOCs or ‘breath pattern’ indicative of lung cancer. This may be due to the variation between studies as highlighted above. The need for breath collection to be standardised is well recognised, and the European Respiratory Society Task Force has published a technical standard to assist with future research. The report recognises the important questions that still need to be answered with regards to best practise in breath collection research and highlights important areas for future research to include a large multi-centre trial, before a technical standard for breath collection can be established,\(^{140}\). As part of the on-going efforts to standardise breath collection the RECIVA Breath sampler has been developed enabling multi-centre clinical trials regardless of analytical platform.
Table 1 (reorganised)

Summary of studies of exhaled breath in lung cancer\textsuperscript{141,142}.


<table>
<thead>
<tr>
<th>Author</th>
<th>Participants</th>
<th>Sensitivity</th>
<th>Sensitivity</th>
<th>Accuracy (PPV/NPV)</th>
<th>Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bajtarevic et al, 2009</td>
<td>65 LC 31 HC</td>
<td>52%, 71%, 80% (n=4, n=15, n=21 respectively)</td>
<td>100%</td>
<td>NR</td>
<td>SPME PTR-MS GC-MS</td>
</tr>
<tr>
<td>Broza et al, 2013</td>
<td>12 LC 10 other lung diseases</td>
<td>100%</td>
<td>80%</td>
<td>94.1%</td>
<td>NaNose SPME GC-MS</td>
</tr>
<tr>
<td>Buszewski et al, 2012</td>
<td>29 LC 44 HC</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>GC-MS</td>
</tr>
<tr>
<td>Capuano et al, 2015</td>
<td>20 LC 10 other lung diseases</td>
<td>NR</td>
<td>NR</td>
<td>90% NaNose 76% SPME/GC-MS</td>
<td>NaNose SPME/GC-MS</td>
</tr>
<tr>
<td>D’Amico et al, 2010</td>
<td>28 LC 36 HC</td>
<td>85%</td>
<td>100%</td>
<td>NR</td>
<td>eNose QMS sensors (GC-MS)</td>
</tr>
<tr>
<td>Di Natale et al, 2003</td>
<td>35 LC 18 controls 9 post surgery</td>
<td>NR</td>
<td>NR</td>
<td>100%</td>
<td>eNose LibraNose(GC-MS)</td>
</tr>
<tr>
<td>Dragonieri et al, 2009</td>
<td>10 LC 10 HC</td>
<td>NR</td>
<td>NR</td>
<td>90%</td>
<td>eNose Cyanose 320</td>
</tr>
<tr>
<td>Fuchs et al, 2010</td>
<td>12 LC 12 Healthy smokers 12 HC</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>GC-MS</td>
</tr>
<tr>
<td>Gaspar et al, 2009</td>
<td>LC Controls</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>GC-MS</td>
</tr>
<tr>
<td>Gasparri et al, 2016</td>
<td>70 LC 76 HC</td>
<td>81%</td>
<td>91%</td>
<td>NR</td>
<td>eNose QMS sensors</td>
</tr>
<tr>
<td>Gordon et al, 1985</td>
<td>12 LC 17 HC</td>
<td>NR</td>
<td>NR</td>
<td>93% (3 VOCs) 100% (22 VOCs)</td>
<td>TD GC-MS</td>
</tr>
<tr>
<td>Hubers et al, 2014</td>
<td>20 LC 31 COPD controls Validation 18 LC 8 HC</td>
<td>80%</td>
<td>48%</td>
<td>NR</td>
<td>eNose Cyanose 320</td>
</tr>
<tr>
<td>Study</td>
<td>Type</td>
<td>Controls</td>
<td>Cases</td>
<td>Healthy Smokers</td>
<td>HC</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
<td>----------</td>
<td>-------</td>
<td>-----------------</td>
<td>----</td>
</tr>
<tr>
<td>Kischkel et al, 2010</td>
<td>31 LC</td>
<td>31 Healthy smokers</td>
<td>31 HC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ligor et al, 2009</td>
<td>65 LC</td>
<td>31 HC</td>
<td></td>
<td>51%</td>
<td>100%</td>
</tr>
<tr>
<td>Machado et al, 2005</td>
<td>14 bronchogenic ca</td>
<td>45 HC</td>
<td>Validation</td>
<td>14 LC</td>
<td>62 HC</td>
</tr>
<tr>
<td>Mazzone et al, 2007</td>
<td>49 LC</td>
<td>63 other lung diseases</td>
<td>21 controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McWilliams et al, 2015</td>
<td>25 LC</td>
<td>166 controls</td>
<td></td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Nisreen et al, 2016</td>
<td>149 LC</td>
<td>56 controls (COPD/asthma)</td>
<td></td>
<td>87%</td>
<td>82%</td>
</tr>
<tr>
<td>Peled et al, 2012</td>
<td>53 LC</td>
<td>19 benign nodules</td>
<td></td>
<td>86%</td>
<td>96%</td>
</tr>
<tr>
<td>Phillips et al, 1999</td>
<td>60 LC</td>
<td>48 HC</td>
<td></td>
<td>71.7%</td>
<td>66.7%</td>
</tr>
<tr>
<td>Phillips et al, 2003</td>
<td>178 bronchoscopy patients (67 LC)</td>
<td>41 HC</td>
<td></td>
<td>85.1%</td>
<td>80.5%</td>
</tr>
<tr>
<td>Phillips et al, 2007</td>
<td>193 LC</td>
<td>211 HC</td>
<td></td>
<td>84.6%</td>
<td>80%</td>
</tr>
<tr>
<td>Phillips et al, 2008</td>
<td>193 LC</td>
<td>211 HC</td>
<td></td>
<td>84.5%</td>
<td>81%</td>
</tr>
<tr>
<td>Phillips et al, 2015</td>
<td>96 LC</td>
<td>205 controls</td>
<td>Validation</td>
<td>Site A</td>
<td>Site B</td>
</tr>
<tr>
<td>Poli et al, 2005</td>
<td>36 NSCLC</td>
<td>110 Controls (25 COPD, 35 smokers no COPD, 50 non-smokers)</td>
<td></td>
<td>72.2%</td>
<td>93.6%</td>
</tr>
<tr>
<td>Santonico et al, 2012</td>
<td>20 LC</td>
<td>10 controls</td>
<td></td>
<td>85%</td>
<td>85%</td>
</tr>
<tr>
<td>Schallschmidt et al, 2016</td>
<td>37 LC</td>
<td>23 HC</td>
<td></td>
<td>80%</td>
<td>90%</td>
</tr>
<tr>
<td>Van der Goor et al, 2018</td>
<td>52 LC</td>
<td>93 HC</td>
<td></td>
<td>83%</td>
<td>84%</td>
</tr>
</tbody>
</table>

Validation Site A Site B 68% 70.1% 68.4% 68%
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Controls (Benign Lung Disease, HC)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang et al, 2012</td>
<td>88 LC</td>
<td>155 controls (70 benign lung disease, 85 HC)</td>
<td>96.5%</td>
<td>97.5%</td>
<td>97.1%</td>
<td>SPME GC-MS</td>
</tr>
</tbody>
</table>
4. The next steps

As outlined in the previous sections there is a strong biological rationale for using breath biomarkers for the detection of lung cancer and a considerable amount of research has been performed in this area. Nonetheless, there is not a single clinically available breath test for cancer that has been validated or has regulatory approval. When surveying the literature to date, we identified some common elements which could help explain this apparent contradiction.

1) *Patients number:* most studies have been performed on relatively small group sizes (Table 1), limiting overall statistical power of the study and reducing the ability to discover robust disease biomarkers. This could partly be due to cost and time constraints, as recruitment of bigger sample sizes requires significant time and funds.

2) *Single site:* often breath analysis trials have been conducted in only one site (hospital or university), potentially introducing site-specific biases and limiting reproducibility of results.

3) *Validation in target population:* successful development of a test into a clinically deployable assay depends on validation of the assay in the intended use population. Often, biomarker sets are developed by comparing completely healthy individuals to subjects with advanced disease, and even sometimes to patients on treatment. The risk of biases, false discoveries and overfitting are many-fold and robust validation is crucial to enable translation of results.

4) *Multidisciplinary field:* Successful breath analysis requires integration of a wide range of expertise including engineering, chemistry, clinical and data science. The complexity of designing a medical device alongside breath collection algorithms combined with stringent quality control and robust statistical data analysis will likely require close collaboration between academia and industry.

5) *Analytical variability:* differences between study results could also be related to the use of different analytical platforms, with different characteristics in terms of their ability to quantitate and identify different subsets of VOCs characteristics. In addition, some studies have been conducted by using relatively immature breath sampling and analysis techniques which have not gone through proper analytical validation or lack calibration techniques. This latter challenge is especially profound for analytical techniques based on pattern recognition or so-called ‘breath prints’, such as the eNose, as they are unable to determine the chemical nature of the biomarkers in the assay. This makes disentangling biological and analytical variability nearly impossible and creates challenges for analytical verification of the assay as required for regulatory approval which have not been overcome to date.

As a consequence of these limitations data between different study sites cannot easily be pooled. Concerns may arise around stability of analytical platforms over time and therefore the full potential of VOCs for a specific application cannot be assessed reliably. Often this has resulted in a situation where biomarkers are not reproducible between studies and progression of a test beyond proof of concept is very challenging. A standardised and reproducible workflow would allow prospective validation and progression of an assay towards clinical practice. With the advent of tools such as the ReCIVA the hope is that we are entering a new era for breath VOC-analysis. Such tools have enabled the initiation of the current multi-centre prospective biomarker discovery trial, Lung Cancer Indicator Detection (LuCID), which has been designed to help address common issues associated with breath clinical studies by focusing on the intent to diagnose population and recruiting up to 4000 patients in an adaptive trial design.
5. Potential clinical utility of breath biomarkers

In the final section of this review we will describe potential uses for a VOC breath test in the clinic alongside relevant studies exploring these applications:

5.1 Lung Cancer Screening

Lung cancer screening using chest x-ray and sputum cytology has been tested in numerous studies over the past few decades but was not found to be an effective screening tool as it did not reduce lung cancer-specific mortality. Interest in lung cancer screening was revived following the striking mortality benefit demonstrated by low-dose computerised tomography (LDCT) in the National Lung Screen Trial, (NLST), 144, performed in the US. Subsequently the United Services Preventative Services Task Force (USPSTF), recommend LDCT screening in individuals who are between 55 and 80 years of age, with a 30 pack year history who still smoke or have quit within the last 15 years 145. However, LDCT is not without its limitations.

LDCT carries a number of potential harmful side effects, including exposure of individuals to ionising radiation. LDCT screening for lung cancer is therefore not recommended for those with a low risk of lung cancer. Whilst never-smokers have a low risk of lung cancer overall, 15% of lung cancers occur in never-smokers, and this sub-population of high-risk never-smokers may benefit from some form of lung cancer screening, however they are a difficult population to identify and current proposed screening programmes would not include them 146. A second concern is the frequent rate of indeterminate pulmonary nodules detected by CT, which require further interval CT surveillance.

Thirdly, increased detection by CT can lead to overdiagnosis. This occurs when a cancer is diagnosed, which will not cause them illness or death i.e. they will die with the cancer rather than from it. These factors affect the acceptability of the screening test to the public and policy makers and impact cost and resource considerations 147. The UK and many other European countries have thus far resisted national implementation of screening programmes. The long awaited Dutch-Belgium NELSON study has presented results showing that LDCT screening can result in an 25% reduction in lung cancer mortality in asymptomatic men at high risk148. These results are promising and build on previous research. The full results of this study are awaited and it is hoped they will assist with optimising the risks, benefits, costs and implementation of LDCT screening 149.

Lung cancer screening and case finding could be transformed by a successful biomarker. Thus far, no blood biomarkers have been clinically validated although blood based biomarkers, in particular circulating tumour DNA biomarkers are currently under development, 150,151. A breath biomarker could have value by being offered as an initial screening test to a wider population than currently recommended for LDCT. A breath test could pre-select individuals at a higher risk thereby enriching the population to undergo LDCT. This could help reduce the number of low risk individuals undergoing LDCT and increase the overall cost-effectiveness of a screening program by reducing the number of CT scans performed whilst increasing the number of cancers detected. Depending on the specific population of interest this test should have a high sensitivity and a moderate to high specificity. Ultimately, a breath test for lung cancer screening should be validated in a large screening trial of several thousand individuals in the intention to diagnose population after carefully considering the intended use population and screening interval.

5.2 Nodule Management

The management of pulmonary nodules remains a challenging area in clinical practice, despite a number of clinical guidelines 152,153. An indeterminate pulmonary nodule is a radiological abnormality observed on a CT-scan <30mm in diameter of unknown origin, 152. The most common way to exclude
lung cancer is to perform repeat CT-scans for several years. This results in considerable health care costs, potential psychological stress for patients and repeated radiation dosages. The issue of indeterminant nodules has been exacerbated in recent years by the increasing use of CT for body scanning and will further increase in the context of CT screening programmes.

A breath biomarker that could distinguish between benign and malignant pulmonary nodules could reduce unnecessary investigations and transform nodule follow-up and management. Peled and colleagues, in a pilot study in 72 patients, found that VOC analysis could differentiate between benign and malignant nodules with 86% sensitivity and 96% sensitivity using leave one out cross validation 153. Although this was a small samples size and no external validation was performed it gives some indication that breath VOC analysis has the potential to meet this clinical need.

To successfully validate a breath biomarker for accurately differentiating benign from malignant pulmonary nodules, a clinical trial would need to ensure at least a 2 year follow-up of nodules as this is what is currently recommended in terms of imaging follow up in clinical practice 152. A breath biomarker would need to show a high sensitivity and high negative predictive value (NPV) to ensure nodules are not inappropriately dismissed from surveillance as being negative for malignancy. The NPV and sensitivity could be enhanced by incorporating the breath test result into a risk score, combined with the demographic, clinical and radiological characteristics of the patient and their pulmonary nodule. This could lead to an improved risk of malignancy prediction model, compared to existing algorithms which are based on clinical and radiological characteristics alone, 152. Furthermore, it is extremely important to carefully think through the subset of nodules (size, radiological characteristics) for which such a test would be used. Likewise thinking about the test positioning in the pathway and test frequency are key drivers to success.

5.3 Following up recurrence and detecting further primaries

Individuals who have had lung cancer are both at risk of recurrence but also at risk of further new primary cancers. Current standard practice for surveillance is with serial CT or x-ray imaging, although the optimal imaging modality and follow-up intervals have not been elucidated. There is the potential risk that intervals between scans may be too long, missing rapidly growing nodules and hindering possible early intervention. Additionally, lung surgery and radical radiotherapy can induce anatomical changes and render radiological interpretation challenging, especially in the initial stages following treatment. The scanning intervals need to be balanced against added unnecessary radiation risks and also patient factors including time and psychological adverse effects after having already had treatment for an initial lung cancer, with effort to return to ‘normal’ life 154. Breath testing, in contrast, is non-invasive and could provide valuable information and reassurance in between follow-up scanning. It could be easily accessible in a follow-up clinic environment, carried out at regular intervals and could benefit care by reducing the need for imaging tests and expedite the detection of recurrent or new primary cancers, thus augmenting standard of care.

A number of groups have looked at VOCs post-surgery, however the studies have been small. Two groups found that VOC profiles indicative of lung cancer were reduced and breath profiles were subsequently more similar to non-cancer controls post-surgery 64,127. This is in contrast to some other groups who concluded there was no difference in VOC patterns post-surgical resection; Phillips and colleagues analysed a second breath sample after surgical resection and reported that the breath pattern was consistent with primary lung cancer. Interestingly this group concluded that they would not expect resection to alter the breath signal due to the fact that metabolic processes that had already been activated by the cancer, including the activation of cytochrome p450 pathway, would not be down-regulated 125. Phillips and colleagues. did not indicate the time between surgery and
breath test as this may be a limiting factor. It is likely that metabolic changes as a result of lung cancer would regress or change after the cancer has been resected but this may take time while VOCs wash out e.g. from fatty tissue. This concept is supported by a study evaluating the breath volatile limonene as a marker of liver cirrhosis. Limonene was found in the breath of individuals with liver cirrhosis at greater levels than controls. Following liver transplant, the levels of limonene reduced, however this occurred over weeks and not immediately after transplantation. This could similarly occur in lung cancer following resection and longitudinal measurements would be useful in clarifying this.

The contrasting conclusion of these small studies once again highlights the need for a larger study with well standardised methodology. A large prospective trial with follow-up of patients at different time intervals over 5 years, post treatment with repeated breath collection would be extremely informative about changes to VOC composition post curative treatment, to include both surgery and radiotherapy. The added benefit of such a study following individual patients with serial breath sampling is the reduction in patient variability. Through collecting breath sampling at regular intervals over the follow-up period, with and without contemporaneous CT scanning, an assessment could be made as to whether breath profiles change prior to CT changes, for example with recurrence. This could therefore prove useful in improving detection of early recurrence whilst minimising unnecessary radiation exposure or intervention.

5.4 Improving management of lung cancer

Some preliminary research has been performed evaluating the different ‘breath prints’ between different histological subtypes and genetic driver mutations such as epidermal growth factor receptor (EGFR) and receptor (PD1-PDL1) axis enabling prediction of therapeutic efficacy. Although the evidence base for the association between altered metabolism and breath VOCs is growing (see section 2.3) there is still much to be done. Besides its merits in personalisation of immunotherapy such research may also yield important candidate biomarkers for diagnostic purposes.

In recent years progress has been made by attempting to personalise the treatment of lung cancer with specific therapies targeting specific mutations such as EGFR and immune checkpoint inhibitors such as PD-L1. In their current form breath biomarkers are unlikely to ever reach the level of sensitivity and accuracy for characterisation of tumour mutations that the tissue targeted molecular assays can deliver and therefore it is unlikely that they will replace the current staging and molecular work-up of patients with confirmed lung cancer. However, advances in metabolomics approaches may be able to non-invasively characterise the molecular signature of lung cancer e.g. the PD1-PDL1 axis enabling prediction of therapeutic efficacy.

In addition to helping to make therapeutic decisions prior to initiation of therapy, VOC biomarkers may be able to assist with monitoring therapeutic response. Nardi-Agmon and colleagues, investigated the use of VOCs to determine treatment response. They collected a series of breath samples over time in patients with advanced lung cancer who were receiving treatment, including
Chemotherapy and tyrosine kinase inhibitors (directed at mutations to include EGFR and ALK). Through establishing sensor ‘breath patterns’ of disease and response they found evidence to suggest breath analysis could be used to assist with monitoring treatment effect. This was a small study and larger studies are needed to validate these findings. If validated this could potentially translate to the use of a breath biomarker alongside the Response Evaluation Criteria in Solid Tumours (RECIST) criteria currently used to assess tumour burden. It requires repeated imaging and labour-intensive reporting and at present there is a challenge between ensuring a timely diagnosis of progressive disease and allowing time for treatment to work. Combining a breath biomarker with current management has the potential to reduce the reliance on imaging and improve the timely diagnosis of progressive disease.

In this context another potential application of a breath biomarker could be in correctly identifying pseudoprogression. Immunotherapy works by targeting the immune system; activated immune cells invade tumour cells and cause an inflammatory response. Radiologically this is challenging to differentiate from true disease progression, however biopsies taken from enlarging lesions have shown the presence of inflammatory cells. The phenomenon of pseudoprogression has led to the development of a new immune related RECIST criteria. These criteria will continue to be updated as results of current research become available. Considering the substantial evidence around the utility of breath analysis in asthma and COPD, some of these biomarkers may hold relevance for quantification of this inflammatory response enabling differentiation between true progression and pseudoprogression.

To fully investigate whether breath analysis has potential for the management of patients with lung cancer, detailed prospective trials are required in patients undergoing different treatments for lung cancer. Analysis of any trial data would seek to assess whether treatment failure could have been predicted by breath profiles. Potential confounding effects of chemotherapy or immunotherapy, which may have their own effects on VOC profiles would also need to be studied and accounted for. Ultimately, a breath biomarker guided strategy will need to be compared to the current standard of care to assess clinical utility of such a breath test.

Overall these applications point towards the potential utility of breath biomarkers and their ability to help realise a much-needed personalised medicine approach to the treatment of lung cancer. This could generate considerable value for patients but may be challenging to realise as such a companion diagnostic is ultimately tied into the efficacy of the therapy itself. Careful health economic evaluation of the positioning of such a test should be done early during a research program to gain insights into the requirements to generate a clinically useful test.

Summary

There is a strong scientific rationale underpinning the production of breath biomarkers in lung cancer along with a wide range of pilot studies pointing toward clinical utility of such tests. Such tests have not yet made it beyond these initial proof of concept stages. In recent years considerable effort has gone into improving the standardisation of breath collection and VOC analysis. Therefore, we are now at a stage where interdisciplinary research programs can significantly progress our understanding of the utility of breath biomarkers for various applications in early detection and personalised medicine in lung cancer. Ultimately, this should help enable the common vision breath researchers have, namely saving lives by developing a non-invasive tool for the early detection of cancer.
References


17. Martinez-Outschoorn UE, Peiris-Pagés M, Pestell RG, Sotgia F, Lisanti MP. Cancer metabolism:


96. Kidera Y, Tsubaki M, Yamazoe Y, et al. Reduction of lung metastasis, cell invasion, and


Figure 1. Schematic representation of the main metabolic pathways leading to generation of breath VOCs. A) Mitochondrial respiratory chain complexes and cytosolic NADPH oxidase are depicted as examples of ROS generating mechanisms. Attack by hydroxyl radicals on polyunsaturated fatty acids (PUFAs), and subsequent generation of lipid radicals, is represented. B) Representation of the mevalonate pathway for synthesis of cholesterol, dolichol, and ubiquinone. Spontaneous generation of isoprene is highlighted. C) Generation of ketone bodies in the liver following lack of glucose availability due to fasting or diabetes. Ketone bodies generated in the liver are distributed through the bloodstream to extrahepatic tissues in order to support energy production. Acetone, the most volatile ketone body, is