Exploring the 175-year history of spirometry and the vital lessons it can teach us today

Andrew Kouri1, Ronald J. Dandurand2,3,4, Omar S. Usmani5 and Chung-Wai Chow6,7

1Division of Respirology, Dept of Medicine, St. Michael’s Hospital, Unity Health Toronto, Ontario, Canada. 2Lakeshore General Hospital, Quebec, Canada. 3Dept of Medicine, Respiratory Division, McGill University, Montreal, Quebec, Canada. 4Montreal Chest Institute, Meakins-Christie Labs and Oscillometry Unit of the Centre for Innovative Medicine, McGill University Health Centre and Research Institute, Montreal, Canada. 5National Heart and Lung Institute, Imperial College London and Royal Brompton Hospital, London, UK. 6Dept of Medicine, University of Toronto, Toronto, Canada. 7Division of Respirology and Multi-Organ Transplant Programme, Dept of Medicine, Toronto General Hospital, University Health Network, Toronto, Canada.

Corresponding author: Omar Usmani (o.usmani@imperial.ac.uk)

Abstract

175 years have elapsed since John Hutchinson introduced the world to his version of an apparatus that had been in development for nearly two centuries, the spirometer. Though he was not the first to build a device that sought to measure breathing and quantify the impact of disease and occupation on lung function, Hutchinson coined the terms spirometer and vital capacity that are still in use today, securing his place in medical history. As Hutchinson envisioned, spirometry would become crucial to our growing knowledge of respiratory pathophysiology, from Tiffeneau and Pinelli’s work on forced expiratory volumes, to Fry and Hyatt’s description of the flow–volume curve. In the 20th century, standardization of spirometry further broadened its reach and prognostic potential. Today, spirometry is recognized as essential to respiratory disease diagnosis, management and research. However, controversy exists in some of its applications, uptake in primary care remains sub-optimal and there are concerns related to the way in which race is factored into interpretation. Moving forward, these failings must be addressed, and innovations like Internet-enabled portable spirometers may present novel opportunities. We must also consider the physiologic and practical limitations inherent to spirometry and further investigate complementary technologies such as respiratory oscillometry and other emerging technologies that assess lung function. Through an exploration of the storied history of spirometry, we can better contextualize its current landscape and appreciate the trends that have repeatedly arisen over time. This may help to improve our current use of spirometry and may allow us to anticipate the obstacles confronting emerging pulmonary function technologies.

Respiration from antiquity to the modern era

For thousands of years, humanity struggled to understand the functions and mechanics of respiration. Hippocrates and the Greek philosophers reasoned that the purpose of the lungs was to temper and cool blood in the heart [1]. Aristotle posited we only have one lung, and though Galen did perform some rudimentary experiments on capturing exhaled air into an artificial bladder, he believed the role of the lungs was to carry off vapours, like smoke from a fire [1, 2]. Our knowledge of the lungs remained stagnant until the scientific revolution of the 17th century, when new thinkers further elucidated the nature of breathing. Examples include William Harvey, who in 1628 first detailed pulmonary circulation, Evangelista Toricelli (a student of Galileo), who discovered the initial laws of atmospheric pressure in 1643, and Robert Hooke in 1667, who kept a dog alive by artificial respiration with bellows [1]. The Italian mathematician and physiologist Giovanni Borelli (1681) was the first to try to precisely measure the volume of air inspired in one breath, though his apparatus, which involved sucking liquid up a cylindrical
tube, did not account for the effects of negative pressure and provided inaccurate results [2]. In the following two centuries, many others worked on improving Borelli’s measurements. In 1798, Thomas Beddoes, with the help of his colleagues Humphry Davy, James Watt and William Clayfield at the Pneumatic Institution in England, created a “mercurial airholder and breathing machine” capable of accurately measuring lung capacity (figure 1). 16 years later, Edward Kentish would be the first to use a similar apparatus he named the “pulmometer” in respiratory disease diagnosis, and soon thereafter Charles Turner Thackrah performed a health survey across England with the help of such a device, establishing early insights into the relationship between occupational history and lung health [4]. These years of accelerating incremental innovation finally culminated in 1846, when John Hutchinson, a British surgeon, published his magnus opus on the newly coined “spirometer” [1]. Hutchinson had refined the pulmometer to optimize its usability, used it to systematically collect a large sample of respiratory measurements, and applied modern statistical methods establishing “normal” spirometric values and illuminating the previously unrecognized relationships between age, height and vital capacity [5]. His work, inspired by Thackrah and others before him, breathed new life into the field of respiratory mechanics and revealed the full potential of spirometry [6].

As 175 years have now passed since Hutchinson’s landmark publication, a look back through the illustrious history of spirometry and towards its potential future is merited.

The Hutchinson spirometer

Hutchinson’s spirometer consisted of a counterbalanced bell inverted in water, whereby breathing into a connected pneumatic tube resulted in measurable vertical motion of the bell (figure 2) [6, 8]. Using this device, he performed spirometry on 2130 individuals, anachronistically categorized as sailors, paupers, artisans, pugilists, compositors and draymen, to name a few (figure 3) [1]. From these measurements, he described and named the fundamental respiratory volumes: residual air, now known as the residual volume; reserve air, now expiratory reserve volume; breathing air, now tidal volume; complemental air, now inspiratory reserve volume; and, finally, the vital capacity, which to this day still represents “the greatest voluntary expiration, following the deepest inspiration” — the capacity for life [1, 8]. Hutchinson also correctly identified the relationship between vital capacity, height and age, and recognized that diseases like tuberculosis could negatively affect vital capacity [2, 8]. Though Hutchinson was not the first to invent the spirometer or recognize its importance in respiratory disease — in fact even his identification

FIGURE 1 Mercurial air holder and breathing machine developed by Humphry Davy and colleagues at the Pneumatic Institution in England, circa 1798, 48 years prior to Hutchinson’s spirometer. Reproduced from [3] courtesy of the Wellcome Collection.
of the respiratory volumes was preceded by Julius Jeffreys in 1843 – his synthesis and expansion of the work of those before him and extensive publications and promotion of the value of spirometry as a clinical tool made him a central figure in its history [4, 5]. Unfortunately, despite its initial positive reception by contemporary scientists and physiologists, and Hutchinson’s own efforts to promote his spirometer, its use was relegated predominantly to research for more than 50 years after its invention [6, 8]. This may be partly due to the fact that Hutchinson’s background selling life insurance prompted him to promote the spirometer first for actuarial estimations, that its correct use required a degree of patient education and coordination, or that medicine’s understanding of clinical respiratory pathology was still developing [2, 6]. Whatever the reason, it would not be the last time in the history of pulmonary function testing that the clinical value of new discoveries and techniques was initially overlooked.

The golden age of spirometry

The 1920s marked the next leap forward for pulmonary function testing. Physicians like Georges Dreyer pioneered the use of spirometry in pre-operative thoracic surgery assessments, evaluation of occupational lung disease and the respiratory problems experienced by pilots during World War I, and documenting the increasingly recognized complications of asthma and emphysema [2, 6]. In the 1930s, Alvin Barach quantified the response of expiratory flow rates to nebulized epinephrine in asthma and emphysema, the first published report of the effectiveness of bronchodilation. Soon after, Courand and Richards of Columbia University proposed that respiratory disease could be divided into problems relating to ventilatory obstruction and restriction, establishing the basis for the classification system that remains in use today [6]. The next significant advancement came in 1947, when two French physicians, Robert Tiffeneau and André Pinelli, proposed the capacité pulmonaire utilisable à l’effort (CPUE), defined as the largest forced volume that could be expired in one second following a maximal inspiration, known in English as the forced expiratory volume in one second (FEV1) [9]. Tiffeneau and Pinelli recognized that FEV1 more sensitively predicted ventilatory limitations during exertion than the vital capacity, was easier to perform than other popular manoeuvres such as the maximum breathing capacity (maximal voluntary ventilation) and was relatively constant between measurements [9, 10]. They also introduced the Tiffeneau–Pinelli index, representing the ratio of CPUE and vital capacity, now known as the FEV1/forced vital capacity (FVC) ratio [6]. Independently of his French colleagues, Edward Gaensler at Boston University also recognized in 1951 that the FEV1 was an important spirometric measurement that most closely correlated to the maximum breathing capacity [11]. However, despite the significance of Tiffeneau
and Pinelli’s initial discovery, and possibly because it was published solely in French, it took nearly 5 years and the parallel work of Gaensler for the FEV1 to be globally recognized as an important pulmonary function measurement [6].

The flow–volume curve

Over the years that followed, different approaches to measuring ventilatory function proliferated, from Fowler’s work on the mean flow over the middle-part of expiration, known now as the forced expiratory flow at 25–75% of forced vital capacity (or FEF25–75%), to Wright’s creation of the peak flowmeter in 1959, which is still in use today [6]. While these and traditional forced-breathing methods were indeed useful indices of chronic respiratory disease, they failed to capture the complete nature of respiratory behaviour and were often highly susceptible to patient effort [12]. Pioneering work from Fry and Hyatt and Permutt and Mead in the late 1950s and early 1960s illuminated this knowledge gap, providing a rich and novel picture of the relationship between pulmonary pressures, volumes and flows [12, 13]. Fry and Hyatt’s description of the “flow–volume curve” robustly defined the intrinsic characteristics of the intrathoracic pulmonary system in motion, was highly reproducible across efforts and more accurately characterized perturbations in expiratory flow due to pathology or inhalants [13]. Permutt and Mead’s work complemented this representation by providing an explanation of the dynamic flow limitation observed in flow–volume curves through what is now known as the “equal pressure point theory” [6, 13]. In fact, these giants of pulmonary physiology and their colleagues would occasionally meet informally over dinner to debate and discuss their ideas, starting a tradition of scientific collaboration and camaraderie that continues to this day and known as the Flow-Volume Underworld [14]. Though the work of Fry, Hyatt, Macklem, Mead, Permutt and other members of the Flow-Volume Underworld is now recognized to have revolutionized our understanding of pulmonary physiology, once again broad clinical uptake of their findings was slow, and Fry’s early reports were said to have “languished in obscurity” for several years. This may be attributed partially to the complexity of the computations involved in creating flow–volume curves, and to a lack of standardization of “normal” values and technical methodology [13].

Standardization of spirometry

With increased understanding of the theoretical basis of respiration, the popularity of spirometers grew. Throughout the 1960s and 1970s, many groups sought to better describe populations using spirometry, recognizing that FEV1 and FVC could help delineate between and characterize restrictive and obstructive disease, and that the flow–volume loop could detect intra- and extra-thoracic obstruction [15, 16]. However, the wide variety of available spirometers and the lack of standard testing methods resulted in considerable measurement variability, limiting spirometry’s broader potential [6]. This led the American Thoracic Society (ATS) to convene a meeting in Snowbird, Utah, where 22 American scientists met and produced the first standardization guidelines for spirometry in 1979 [6, 17]. Their recommendations were the
culmination of a 2-year effort and represented the most thoroughly reviewed statement ever published by the ATS. They included sections on instrument specifications, validation, quality control, manoeuvre performance, acceptability and reproducibility standards, and reference values [18]. A similar effort led to European guidelines in 1983, and though later updates of both these guidelines were broadly similar in their recommendations, it took 22 years for the two societies to officially harmonize their standardization efforts in the 2005 “ATS/ERS Task Force: Standardization of Lung Function Testing” document [19]. This document was only recently updated in 2019, a testament to the depth and breadth of the original effort [20].

In concert with standardization, accurately interpreting spirometry required the development of reference equations to account for the observed differences in expected values across age, height, sex and ethnicity [21]. The first broadly adopted reference equations came from the European Coal and Steel Community in 1983, which arose out of work to better characterise occupational lung disease [22, 23]. Though these equations were in use for many years, they were criticized for their lack of longitudinal equations covering childhood to old age, their underestimation of predicted FEV₁ and FVC, and their ethnic homogeneity, leading the ATS to initially develop and use their own reference equations [24–27]. As with standardization, it was only in 2012, nearly 30 years later, that the Global Lung Function Initiative (GLI) published their seminal work addressing many of the previous limitations in reference equation research, pooling spirometry data from 74,187 asymptomatic non-smokers across 26 countries to develop international spirometry reference equations that are still in use today [21].

The prognostic potential of spirometry
Standardization of spirometry led to an abundance of research to explore its potential prognostic significance. These efforts were bolstered by findings from the Framingham study in the late 1960s that showed a strong independent relationship between reduced FVC and the risk of coronary heart disease [8, 28]. The 1970s and 1980s saw several population level investigations, with publications from Friedman et al. finding that FVC was also a risk factor for myocardial infarction [29]. Beaty et al. extended these findings from the Baltimore Longitudinal Study of Aging, showing that FEV₁ predicted all-cause mortality even when adjusted for age and smoking status [30]. Lange et al. confirmed this further in 662 male and 2048 female never-smokers enrolled in the Copenhagen Heart Study, concluding that decreases in both FEV₁ and FVC were significant risk factors for mortality [31]. This work continued into the 1990s, with large population cohorts such as the National Health and Nutrition Examination Survey (NHANES) corroborating the evidence for spirometry as a risk factor for overall mortality, lung cancer and cardiovascular disease [32–34]. Despite this growing evidence and many calling for spirometry to be widely adopted as a screening tool in general health assessments, only 20–30% of primary care physicians in the US had a spirometer in their offices or used one regularly by 2002 [8, 34]. This failure once again to translate the potential of spirometry into clinical practice led Macklem and Permutt two revered pulmonologists to wonder if the “absence of the use of pulmonary function ... must be borne by the expert in pulmonary medicine and especially in their] relation to the non-specialist” [8].

The evidence for spirometry
After the results of the Framingham study validated the prognostic potential of spirometry in cardiorespiratory disease, interest in spirometry as a tool for managing chronic respiratory disease grew [35]. In the subsequent years that overlapped with the birth of the “evidence-based medicine” movement of the 1990s, the clinical utility of spirometry was increasingly studied. These efforts resulted in spirometry becoming the gold standard tool for respiratory disease diagnosis, monitoring disease status, assessing response to therapies, prognostication, pre-operative assessment, determining eligibility for specialized therapies and as an important outcome in pulmonary research [36]. However, some controversy still exists around the optimal role for spirometry in several important respiratory conditions [35, 37–42].

The work of Fletcher and Peto in the late 1970s cemented the role of spirometry in the diagnosis and management of chronic obstructive lung disease (COPD), establishing the basis for our current understanding of COPD as a disease of progressive lung function loss related to cigarette smoking [43, 44]. They also introduced the now famous “Fletcher–Peto diagram” of FEV₁ decline over time [45, 46]. Building on their discoveries, Burrows and colleagues showed that FEV₁ decline was also a strong predictor of 10-year mortality in COPD [35, 47]. The value of spirometry in diagnosing COPD was further supported by the NHANES III study, which found that 63% of asymptomatic smokers with “low lung function” (FEV₁/FVC ratio of less than 0.70 and FEV₁ less than 80% predicted) denied a previous or current history of obstructive lung disease [35, 48]. Current international COPD guidelines now require spirometry to make the diagnosis of COPD in patients with appropriate symptoms and/or risk factors and endorse the use of spirometry in COPD prognostication using tools such as the BODE (Body-mass index, airflow Obstruction, Dyspnea and Exercise) index [49]. However, the value of spirometry for following disease status is more contentious. One underappreciated aspect of Fletcher and Peto’s early work is that they demonstrated
significant heterogeneity in FEV$_1$ decline over time in cigarette smokers, suggesting that not all patients with COPD follow the Fletcher–Peto curve [44]. Contemporary research has expanded on these findings, revealing that only roughly half of adults with COPD demonstrate enhanced FEV$_1$ decline over time, that clinical symptoms often do not correlate with spirometry and that there remains no strong evidence to support the use of routine periodic spirometry in the management of stable COPD [37, 41, 46, 50, 51].

Comparably to COPD, the importance of spirometry in the diagnosis and management of asthma was poorly understood for most of the 20th century. In the early 1900s, physicians considered asthma to be an acute disorder of episodic bronchospastic exacerbations due to dysregulated airway neural control rather than a chronic inflammatory process [52]. Over time, understanding of the inflammatory and immunologic nature of asthma grew, as did the awareness that chronic asthma could lead to airway remodelling identifiable with spirometry [53]. Current international asthma guidelines now strongly endorse the importance of spirometry with bronchodilator testing in the initial diagnosis of asthma, and ongoing monitoring with regular spirometry, as FEV$_1$ is a strong independent predictor of asthma exacerbation risk [54]. Without spirometry, clinicians also risk over-diagnosing and treating patients who present with respiratory symptoms, as a recent Canadian study showed that 33.1% of patients diagnosed with asthma in the community had no objective evidence of bronchodilator reversibility or airway hyperreactivity when rigorously tested [55]. However, the value of spirometry in following patients with asthma has also been contested by the work of Abramson et al., that did not demonstrate any differences in asthma-related quality of life, asthma control, or exacerbation rates with regular spirometry versus usual care in children and adults with asthma followed in primary care practices [40, 56].

The line between asthma and COPD is also an area where interpreting the results of spirometry can be challenging. Spirometry with bronchodilator testing was initially explored in the 1950s in an attempt to clearly differentiate airway diseases into reversible (i.e. asthma) and irreversible (i.e. COPD) forms [57]. However, research since the 1980s called this into question, showing no clear cut-off between asthma and COPD in bronchodilator response, and sometimes significant post-bronchodilator changes in patients with severe emphysema [57–60]. The more recent concept of asthma COPD overlap (ACO) further complicates this binary approach to interpreting spirometry [61, 62]. Given the persistent lack of a universally accepted definition for ACO, uncertain epidemiology, and significant clinical heterogeneity, the effectiveness of spirometry in diagnosing and managing patients with ACO is still being debated [62, 63].

There are a multitude of other conditions in which spirometry is commonly used in modern clinical care, though evidence often remains imperfect. For example, in cystic fibrosis (CF), spirometry is considered essential to monitoring disease progression and diagnosing exacerbations, to grading the severity of CF lung disease and to decision making regarding response to therapies and timing of transplant referral [64–67]. Spirometric measurements are also frequently included as important outcomes in CF clinical trials, predicting *Pseudomonas* acquisition, exacerbation rate and mortality [64, 66, 68]. However, spirometry is harder to obtain in younger children with CF and measurements are more variable [65]. Alternative techniques to assess airway function in CF like the lung clearance index derived from multiple breath inert gas washout have also been shown to be more sensitive than FEV$_1$ in detecting early structural changes in the small airways, in assessing treatment-related changes in ventilation across ages and are easier to perform in children [64, 65, 69–71]. Another example is idiopathic pulmonary fibrosis (IPF), where FVC has been explored as a prognostic indicator, with declines found to be associated with decreased survival time and all-cause mortality [72, 73]. IPF guidelines recommend regular spirometry with particular attention to decreases in FVC greater than 10% of a patient’s baseline, as this holds particular prognostic value [74, 75]. Change in FVC has also become the standard accepted primary endpoint in all major trials and is used as a surrogate for mortality [76]. However, this metric is controversial as critics argue that FVC has not been sufficiently validated as a surrogate endpoint in IPF, citing a lack of evidence to support that treatment-related changes in FVC predict clinically meaningful changes, and that without an effective therapy that significantly improves mortality, there is no benchmark against which FVC can be validated [39, 77]. Spirometry is also critical in the management and monitoring of lung transplant patients for identification of potential treatable causes of graft injury such as acute rejection and for prognostication in chronic lung allograft dysfunction (CLAD) [78, 79]. Through the various developments in our understanding of CLAD and the different CLAD phenotypes, spirometry remained the primary means by which these pathophysiologic changes in allograft function are diagnosed and categorized [78, 80]. However, the recognition that acute rejection and CLAD initially manifest in the small airways has raised concern that spirometry – which assesses the large and medium-sized airways – may be missing disease in this “silent zone”, preventing earlier appreciation and treatment of clinically important dysfunction [81, 82]. For example, spirometry has been shown to only have 57% sensitivity for detecting clinically significant acute allograft rejection [83]. The routine use of spirometry in the preoperative assessment of lung cancer resection surgery has also been challenged, as much of the evidence is based on data from the 1980s [84–86].
More recent analysis has questioned its ability to predict postoperative complications depending on the surgical technique, particularly in patients with co-existing COPD where the “lobar volume reduction effect” may paradoxically improve respiratory mechanics post-resection [87, 88].

**Moving spirometry out of the laboratory and into the office**

Despite standardization efforts from the 1990s onward, and growing evidence supporting the utility of spirometry in the routine care of respiratory patients, adoption of spirometry in primary care remained suboptimal by the turn of the 21st century [89, 90]. One of the barriers contributing to this problem was the reliance on conventional hospital-based in-laboratory spirometry, which could create delays in care and was often inconvenient for patients [91]. To address these problems, electronic spirometers designed for office use were developed and marketed to primary care practices [92]. Based on results from NHANES III and the multicentre Lung Health Study, the National Lung Health Education Program (NLHEP) also strongly promoted office spirometry in a consensus statement in 2000, recognizing its potential for earlier identification of patients at risk for COPD before the onset of clinical symptoms and the possibility of modifying the natural history of disease [90, 93]. However, despite the availability of low-cost office spirometers and the well-publicized NLHEP motto to “Test your lungs. Know your numbers”, increasing spirometry uptake remained challenging [94]. Surveys in primary care in the early 2000s found that lack of training, uncertainty about test impact and low confidence in the use and quality of spirometry were important obstacles [94–96].

In the last 10 years, newer generations of electronic portable spirometers have addressed some of the technical issues of early models, and recent guidelines support the notion that high quality testing is possible with only limited staff training [97–101]. Nonetheless, the 2018 update from the NLHEP found that the full potential of office spirometry has still not been realized, though adoption in primary care has improved over time [37].

**The intersection of spirometry and race**

Pulmonary function testing is one of the few broadly applied clinical tests that incorporates self-identified race or ethnicity into its reference definitions [102]. The view that genetic ancestry and racial differences intrinsically influence pulmonary capacity is widely accepted in pulmonary medicine, but many have challenged the scientific validity of these claims [103–107]. Race/ethnicity is a social construct that is undeniably associated with variance in health and disease outcomes, but the notion that these differences are principally related to innate genetic factors rather than environmental ones has a disturbing and racist history [105]. Former President Thomas Jefferson’s 1785 *Notes of the State of Virginia* are one of the first recorded instances of proposed differences in pulmonary function based on race, suggesting a “difference of structure in the pulmonary apparatus” between black slaves and white colonists [104, 105]. Plantation owner, physician and slaveholder Samuel Cartwright, expanded on Jefferson’s ideas in the 19th century, promoting the idea of “pulmonary deficiency” related to black people [104]. The popularization of Hutchinson’s spirometer helped further this bias, exemplified by Benjamin Gould’s 1864 US Sanitary Commission survey of 21,752 Civil War soldiers, sailors, prisoners, and students using an adapted field-spirometer, that reported decreased lung capacity measurements in non-white subjects without adjusting for height, age, occupation or socioeconomic status (figure 4) [4, 104, 108]. This prejudice

---

**TABLE 1. Average Capacity of Lungs.**

<table>
<thead>
<tr>
<th></th>
<th>In usual Vigor</th>
<th>Not in usual Vigor</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. Men</td>
<td>Cubic Inches</td>
<td>No. Men</td>
</tr>
<tr>
<td>White Soldiers, Earlier Series</td>
<td>4,857</td>
<td>175.655</td>
<td>1,915</td>
</tr>
<tr>
<td>White Soldiers, Later Series</td>
<td>8,890</td>
<td>187.898</td>
<td>1,041</td>
</tr>
<tr>
<td>Sailors</td>
<td>1,104</td>
<td>179.217</td>
<td>-</td>
</tr>
<tr>
<td>Students</td>
<td>286</td>
<td>204.382</td>
<td>-</td>
</tr>
<tr>
<td>Fall Blacks</td>
<td>1,631</td>
<td>163.313</td>
<td>221</td>
</tr>
<tr>
<td>Mulattoes</td>
<td>671</td>
<td>161.615</td>
<td>138</td>
</tr>
<tr>
<td>Indians</td>
<td>504</td>
<td>185.006</td>
<td>7</td>
</tr>
</tbody>
</table>

**FIGURE 4** Gould’s unadjusted breakdown of lung capacity in cubic inches by race and health status. Reproduced from [108].

https://doi.org/10.1183/16000617.0081-2021
spirometers are mainly being developed and marketed for home patient use. These spirometers include in-laboratory testing into patient homes, as the current generation of lower-cost portable electronic

One trend that will likely meaningfully influence the future of spirometry is the continued shift from addressing the inefficiencies and inequities in our current use of spirometry are needed.

discussed above, it is clear that further research, improved public policy and advocacy efforts devoted to race and pulmonary function testing [116].

African American intellectuals like W.E.B. DuBois and Kelly Miller, who contested that inequalities in socioeconomic status and access to medical care were not accounted for in survey data [103, 104].

freedom for African Americans [103, 104]. His work gave broad credibility to scientific racism, despite the dissenting work of leading African American intellectuals like W.E.B. DuBois and Kelly Miller, who contested that inequalities in socioeconomic status and access to medical care were not accounted for in survey data [103, 104].

Over the next century, numerous studies were carried out that attempted to document differences in spirometry related to race or ethnicity, however a 2013 systematic review by Braun et al. found that among 226 articles published between 1922 and 2008, the majority neither defined race/ethnicity, nor accounted for socioeconomic status, despite nearly all claiming that non-white racial/ethnic groups had lower lung function [109]. These limitations and the contentious history of spirometry has gained more attention in the research developing reference equations or race/ethnicity “correction factors” in the last 10 years. The most important example is the GLI 2012 publication discussed above, which provided reference equations for white people, African Americans, North East Asians and South East Asians, and acknowledged the limitations of self-reported race/ethnicity and the potential influence of socioeconomic and environmental factors [21, 110]. More recently, the Canadian Health Measures Survey also provided correction factors for multiple races and indigenous peoples, who have previously largely been excluded from research [111].

While the principle objective of these equations is to reduce misinterpretation of spirometry in populations that have been observed to have varying normal measurement ranges, basing these distinctions on self-identified race/ethnicity remains problematic [103–105, 110]. For one, the concept of “race correction” is linguistically prejudicial, as it implies that “correct” values are those corresponding to white populations. This idea of race correction was originally discussed in the 2005 ATS/European Respiratory Society (ERS) guidelines for when race or ethnicity-specific equations were not available but has since been removed in the 2017 reporting guidelines and 2019 standardization of spirometry update, which favour instead the use of race-specific reference equations like those provided by the GLI [19, 20, 112]. Secondly, current racial classifications often fail to account for the variation that exists within the same categories of self-identified race/ethnicity. For example, North American-born individuals of Chinese, Japanese and Indian ancestry have been shown to have different reference spirometry values than those born outside the US and Canada, and various Chinese ethnic subgroups have been shown to have different normal ranges [110, 112, 113]. Though there are studies that have used genetic techniques to show that sitting height ratio (a known predictor or lung size) is heritable [114], and that genetic markers of ancestry can accurately predict lung function [102], the influence of unmeasured confounders such as premature birth, socioeconomic status and other environmental factors associated with structural racism cannot be fully accounted for in these studies [115]. Finally, even assuming some degree of intrinsic genetic difference, growing admixture between populations is likely to render existing equations obsolete over time [110].

Finding solutions to these issues that prioritise both patient equity and scientific rigour will remain challenging and likely require much trial and error. In 2013, Dr. Lindy Braun, Professor of Medical Science and Africana Studies at Brown University, supported by the late globally respected pulmonologist Dr. Philip H. Quanjer, called for an international workshop where pulmonologists, historians, anthropologists and sociologists could meet to discuss and develop methodologies to guide future research into racial and ethnic disparities in pulmonary function testing [103, 109]. This year, such a conference was convened by the ATS, and a report is expected by the end of 2021. The 2021 American Thoracic Society Conference in May also held several symposia on racism and health disparities, with sessions devoted to race and pulmonary function testing [116].

The future of spirometry

Appreciating the complex and sometimes controversial 175-year history of spirometry, we must now ask what the future holds for this important tool. For one, optimal uptake and use of spirometry is still not ideal, and there remains significant health equity disparities preventing broader access to spirometry [91, 110, 117, 118]. In primary care, for example, the percentage of spirometry performed according to ATS/ERS criteria is often less than 50%, and agreement between general practitioners and pulmonologists on diagnosis is often low [117]. Many patients are still diagnosed with asthma and COPD without appropriate spirometric testing, and underserved populations, such as older adults and those with lower socioeconomic status, suffer disproportionately from these care gaps [55, 110, 119, 120]. Considering also the issues related to race discussed above, it is clear that further research, improved public policy and advocacy efforts devoted to addressing the inefficiencies and inequities in our current use of spirometry are needed.

One trend that will likely meaningfully influence the future of spirometry is the continued shift from in-laboratory testing into patient homes, as the current generation of lower-cost portable electronic spirometers are mainly being developed and marketed for home patient use. These spirometers include
features such as linked smartphone apps that provide usage guidance and performance feedback, and data storage in cloud-based services that allow patients to share results securely with their physicians [101]. Home spirometry has been studied in asthma and COPD, IPF, CF and post stem cell and lung transplantation [121–126]. It has been shown to be usable, to provide comparable results to in-laboratory testing for FEV\(_1\) and FVC [101, 126, 127], and to be less variable and less susceptible to placebo improvements than clinic-based spirometry [126]. There are still important limitations to home spirometry, as not all devices meet ATS/ERS quality standards nor are they suitable for all patient populations without proper education and training. Furthermore, studies that show a clear relationship between home spirometry use, clinical important outcomes and cost-effectiveness are still lacking [101]. However, given the severe limitations on access to conventional laboratory and clinic spirometry during the current coronavirus disease 2019 pandemic due to concerns regarding infection control, home alternatives may become a more prominent part of respiratory care in the near future [128].

In addition to the continued miniaturization and mobility of spirometry, novel analyses of existing spirometric data may also lead to improved disease classification and new standards. For example, analysis of the “peak index” – the number of peaks adjusted for lung size on the expiratory flow-volume curve of spirometry – has recently been explored as a new measure associated with radiographic disease, lung function decline, and mortality in COPD [129]. Machine learning techniques have also recently been used to analyse the “area under the expiratory flow–volume curve” and develop a robust classification algorithm that can accurately differentiate between normal, obstructive disease, restrictive disease and mixed impairments using spirometry alone [130]. There is also a growing body of research investigating how smartphone microphones could be used to generate precise spirometric measurements, replacing the conventional spirometer entirely [131]. These smartphone spirometers have so far only been tested in pilot studies and will require much more validation prior to being clinically viable, but the ubiquity of modern smartphones and capacity for broad dissemination make them a promising future direction [132–136].

**Thinking beyond spirometry**

Considering what the history of spirometry has taught us, it is also necessary to contemplate the ways in which modern spirometry does not satisfy our physiologic and clinical needs, and how other approaches to understanding respiratory function may complement or replace spirometry in the coming years. One crucial limitation of spirometry that has been present since the work of Tiffeneau and Pinelli is its reliance on forced expiratory manoeuvres. As the lungs function not in forced expiration, but with tidal breathing, our dependency on spirometry to describe and classify lung disease will always be fraught with variability and reproducibility concerns [82, 137]. Due to the effort dependent nature of spirometry, it is also often sub-optimally performed by those unable to fully cooperate such as children or those with cognitive and physical impairment, creating challenges in the diagnosis and management of these frequently underserved populations [138, 139]. Spirometry is equally unable to fully capture pathology until more than 50% of small airways have been affected [140], often resulting in a disconnect between patient symptoms and spirometric measurements, and the inability to accurately identify early respiratory disease affecting the small airways [37, 81, 141, 142]. There also remains considerable debate around one of the most fundamental uses of spirometry – the diagnosis of obstruction [143–146].

One emerging lung function testing technology that addresses some of these limitations is respiratory oscillometry. While first described by duBois et al. in 1951, the advent of modern computing technology was necessary to realise its full potential [147–149]. Performed while tidal breathing, oscillometry directly assesses the mechanical properties of the entire respiratory system through the application of multi-frequency acoustic waves at the mouth, resulting in measurements of respiratory impedance [81, 150]. It can provide accurate measurements across age ranges and in patients currently underserved by spirometry such as infants, the elderly and those with physical or cognitive impairments [151, 152]. Oscillometry is also more sensitive than spirometry in the early diagnosis of COPD and acute cellular rejection post lung transplant, and correlates better with respiratory symptoms, airway hyper-responsiveness and overall control in asthma [81, 152–157]. Though there is still a paucity of normal reference values for oscillometry, and variation between different devices can be problematic, recently published technical and interpretation guidelines and efforts to harmonise measurements from different commercial oscillometers are important developments [150, 158, 159]. Additionally, collection and improvement of oscillometry normative data by the GLI is ongoing, and efforts to standardize interpretation will facilitate its translation into routine clinical use [160–162]. With continued research and clinical validation, oscillometry may serve as an important adjunct to spirometry, and may even replace it in many situations.

Though we highlight oscillometry due to the many historical parallels between it and spirometry in their path to validation and clinical use, other technologies to assess lung health such as the fraction of exhaled...
nitric oxide (FENO) also deserve mention. FENO, like oscillometry, does not require forced expiratory manoeuvres, and is more sensitive than spirometry at identifying type 2 airway inflammation in asthma [163, 164]. As portable and cheaper FENO devices become more broadly available, they may take on a larger role in assessing lung function [165]. Similarly, the lung clearance index from multiple breath inert gas washout discussed above is a promising alternative to spirometry, and other technologies like functional pulmonary magnetic resonance imaging and electrical impedance tomography are providing novel insights to understanding lung function [166–170].

Conclusion

The history of spirometry from conception to present provides valuable insights into the overall progression of our understanding of the respiratory system and helps contextualize many of the issues preventing the ideal use of this powerful tool today. Contemplating this history also allows us to anticipate the obstacles we are likely to face with novel technologies like oscillometry providing a roadmap to avoid the same historical pitfalls of effectively translating new pulmonary function testing discoveries into clinical practice. First, a deep understanding of the involved underlying respiratory physiology is needed. Next, technical, reference, and interpretative standards and guidelines must be developed and refined. Then, the evidence for routine clinical use must be rigorously established and implementation strategies must be created to address the barriers of appropriate dissemination. Crucially, equitable use must be strived for, through studying and addressing how the social determinants of health influence our patients’ healthcare trajectories. These steps all depend on strong collaboration and cooperation between international scientists and clinicians, between generalists and specialists, and between public and professional bodies. Finally, we must remember and consider the original goal of Hutchinson 175 years ago and continue like him to seek a greater understanding of respiratory function. By doing so, we may discover the next techniques that will illuminate the complex mechanics of respiration and reveal new vital measurements that describe our capacity for life.

Provenance: Submitted article, peer-reviewed.

Conflict of interest: A. Kouri has nothing to disclose. R.J. Dandurand has a patent WO 2020/191499 licensed to Thorasys Thoracic Medical Systems Inc. O.S. Usmani reports personal fees from Thorasys, outside the submitted work. C-W. Chow has nothing to disclose.

References


Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical practice.


Quanjer PH. Lung function, race and ethnicity: a conundrum. 

Braun L. Race, ethnicity and lung function: a brief history. 


Coates AL, Tamari IE, Graham BL. Role of spirometry in primary care. 


Quanjer PH. Lung function, race and ethnicity: a conundrum. 

Braun L. Race, ethnicity and lung function: a brief history.

Lujan HL, DiCarlo SE. Science reflects history as society influences science: brief history of “race”, “race correction”, and the spirometer. 

Agrawal A, Aggarwal M, Sonnappa S, et al. Ethnicity and spirometric indices: hostage to tunnel vision?

Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms.


Braun L, Wolfgang M, Dickersin K. Defining race/ethnicity and explaining difference in research studies on lung function.


Enright P, Brusasco V. Counterpoint: should we abandon FEV1/FVC<0.70 To detect airway obstruction? yes. *Chest* 2010; 138: 1040–1042.


Thompson BR. Dysanapsis-once believed to be a physiological curiosity-is now clinically important. *Am J Respir Crit Care Med* 2017; 195: 277–278.


