Age and body mass index affect fit of spirometry GLI references in schoolchildren

Rebeca Mozun^{1,2}, Cristina Ardura-Garcia¹, Eva S. L. Pedersen¹, Jakob Usemann^{4,5}, Florian Singer^{3,6}, Philipp Latzin³, Alexander Moeller^{4*}, Claudia E. Kuehni^{1,3*} on behalf of the LUIS Study Group⁷

*These authors contributed equally.

1. Institute of Social and Preventive Medicine, University of Bern, Switzerland
2. Graduate School for Cellular and Biomedical Sciences, University of Bern, Switzerland
3. Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland.
4. Department of Respiratory Medicine, University Children's Hospital Zurich and Childhood Research Center, University of Zurich, Switzerland
5. University Children's Hospital Basel (UKBB), Basel, Switzerland
6. PEDNET, paediatric clinical trial unit, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland.
7. For a full list of members of the LUIS Study Group, please refer to the Acknowledgements section

Correspondence to:

Prof. Claudia E. Kuehni
Institute of Social and Preventive Medicine, University of Bern
Mittelstrasse 43, 3012, Bern
Email: claudia.kuehni@ispm.unibe.ch

Take home message (250/256 characters):

Our study suggests GLI-based FEV₁, FVC, and FEF₂₅₋₇₅ z-scores over detect abnormal lung function in Swiss adolescents, and more so among slimmer adolescents, which has important implications for clinical care, research, and international comparisons.

Key words: spirometry, child, adolescent, reference values, Switzerland, Global Lung function Initiative.

Word count: 2832 / 3000

Abstract (250/250)

Background: References from the Global Lung Function Initiative (GLI) are widely used to interpret children's spirometry results. We assessed fit for healthy schoolchildren.

Methods: LuftiBus in the school (LUIS) is a population-based cross-sectional study done from 2013-2016 in the canton of Zurich, Switzerland. Parents and children aged 6-17 years answered questionnaires about respiratory symptoms and lifestyle. Children underwent spirometry in a mobile lung function lab. We calculated GLI-based z-scores for FEV₁, FVC, FEV₁/FVC, and FEF₂₅₋₇₅ for healthy White participants. We defined appropriate fit to GLI references by mean values ± 0.5 z-scores. We assessed if fit varied by age, body mass index, height, and sex using linear regression models.

Results: We analysed data from 2036 children with valid FEV₁ measurements of which 1762 also had valid FVC measurements. The median age was 12.2 years. Fit was appropriate for children aged 6-11 years for all indices. In adolescents aged 12-17 years, fit was appropriate for FEV₁/FVC (mean: -0.09; SD: 1.02) z-scores, but not for FEV₁ (mean: -0.62; SD: 0.98), FVC (mean: -0.60; SD: 0.98), and FEF₂₅₋₇₅ (mean: -0.54; SD: 1.02). FEV₁, FVC, and FEF₂₅₋₇₅ z-scores fitted better in children considered overweight (means: -0.25, -0.13, -0.38) than normal weight (means: -0.55, -0.50, -0.55; *p*-trend: <0.001, 0.014, <0.001). FEV₁, FVC, and FEF₂₅₋₇₅ z-scores depended on both age and height (*p* interaction: 0.034, 0.019, <0.01).

Conclusion: GLI-based FEV₁, FVC, and FEF₂₅₋₇₅ z-scores do not fit White Swiss adolescents well. This should be considered when using reference equations for clinical decision making, research and international comparison.

Introduction

Standardized global spirometry reference equations are important for interpreting spirometry results, supporting clinical decision-making, and allowing comparison among research findings. To adequately interpret spirometry results and avoid misclassification in clinical care and research, lung function references must fit the local population [1]. The Global Lung Function Initiative (GLI) collected large amounts of spirometry data from healthy non-smoking populations around the world and developed these reference equations [2]. GLI references are now widely used in clinics and studies worldwide. An ideal fit for GLI references with a local population is defined by mean z-scores of spirometry indices equal to 0 with a standard deviation of 1; however, a shift of up to ± 0.5 z-scores of the mean is considered acceptable due to sampling variations [3-5]. GLI references account for ages 3-95, ethnicities, height, and sex [2]. Thus, z-scores of spirometry indices should not be affected by these factors. GLI references do not adjust for body composition assessed by weight or body mass index (BMI), though these factors may explain some variability in z-score results in children [6, 7] and excessive weight gain is reported to adversely affect lung growth [8, 9].

Standardizing lung function references during adolescence is particularly complex [10] because physiological changes in lung growth and body structure may occur at different ages for males and females of the same ethnicity [11-13]. European studies showed that GLI references fitted children from the France, Norway, Spain, and the United Kingdom (UK) [14-17], yet only partially fitted children from Italy and Germany [18, 19]. At present, we know little about the factors that contribute to the poor fit for some populations. The fit of GLI references had not been assessed for Swiss

children. In this study, we assessed GLI reference equation fit for healthy White schoolchildren in Switzerland and investigated factors that may affect fit.

Methods

Study design and setting

LuftiBus in the school (LUIS) is a population-based cross-sectional study of children aged 6-17 years and their respiratory health from 2013-16 in the canton of Zurich, Switzerland (ClinicalTrials.gov: NCT03659838). Zurich is the most populated canton in Switzerland, and it includes many persons born in other countries. The methodology of LUIS has been described [20]. All schools in the canton of Zurich were invited to participate. If the head of a school agreed, trained lung function technicians visited the school in a bus with equipment to measure lung function. Parents completed a detailed questionnaire at home; children answered a questionnaire via interview with technicians at school and underwent lung function testing. The ethics committee of the canton of Zurich approved the study (KEK-ZH-Nr: 2014-0491) and informed consent was obtained from parents and children prior to participation.

Selection of the study population

We included children with available parent and child questionnaires. To obtain a sample of healthy children, we excluded those with parent-reported wheeze in the past year, use of inhaled corticosteroids (ICS) in the past year or lifetime doctor's diagnosis of asthma (Figure 1). We excluded children who reported smoking at least once per week and those who reported cough or a cold on the day of the measurement. These are standard criteria to define healthy participants (with some variations between studies) [6, 19, 21, 22]. We excluded children with ethnicities other

than White because there were too few to assess fit [1]. We also excluded children with implausible height or body mass index (BMI) z-scores (> 4 or < -4 z-scores), which probably was the result of data recording errors during height and weight measurements. We did quality control checks of lung function measurements, as previously described [20]. We excluded children with invalid lung function measurements, whose flow-volume curves had signs of hesitation at the start of expiration, submaximal effort, early termination of expiration, or cough or glottis closure before the first second of expiration for FEV₁ and after the first second of expiration for FVC.

Study procedures

The questionnaire for parents was paper-based and completed at home [20]. It asked about diagnoses, family and household characteristics and lifestyle, medication, and respiratory symptoms. The interview questionnaire for children was shorter and focused on respiratory symptoms, presence of cough or a cold on the day of the measurement, and active smoking. Questions were based on the International Study of Asthma and Allergies in Childhood and the Leicester Respiratory Cohort studies questionnaires [23, 24].

For spirometry we used Masterlab (Jaeger, Würzburg, Germany). Trained technicians conducted spirometry according to ERS/ATS standards [25]. As main outcomes, we used forced expiratory volume in 1 second (FEV₁), forced vital capacity (FVC), FEV₁/FVC, and forced expiratory flow between 25 and 75% of the FVC (FEF₂₅₋₇₅). We applied GLI ethnicity-specific reference equations for Whites to produce z-scores for spirometry indices using the GLI desktop software [26]. Children's standing height and weight were measured without shoes by the technicians at school according to

standard [27]. We calculated BMI z-scores and height-for-age z-scores based on WHO references [27]. We classified BMI z-scores into 4 categories (underweight: <-2 z-scores; normal weight: ≥ -2 to <1 z-scores; overweight: ≥ 1 to <2 z-scores; obese: ≥ 2 z-scores) [27].

Statistical analysis

We calculated the mean and 95% confidence interval (CI) for z-scores of spirometry indices. Deviations of the mean outside the range of -0.5 to $+0.5$ z-scores are generally considered physiologically meaningful [3, 5, 19]. We described the proportion of the population below lower limits of normal ([LLN]; i.e., <-1.645 z-scores) and its 95% CI. Per definition, 5% of the population should fall below the LLN. To address systematic deviations in the fit of GLI references by age, we stratified by 2 age groups: children younger than 12 years and adolescents aged 12 years or older [19]. We chose this age cut-off because puberty begins at different ages for males and females, yet it has usually begun in both sexes by age 12 [28]. To allow further comparisons, the online supplement also shows results stratified by 4 age groups: 6-9, 10-11, 12-13, and 14-17 years [19].

We studied associations of GLI-based z-scores for FEV_1 , FVC, FEV_1/FVC , and FEF_{25-75} with age, BMI z-scores, height, and sex using scatterplots and univariable linear regression models. We tested for age-height, age-sex, and height-sex interactions using multivariable regression models for each spirometry parameter and compared models with and without interaction terms using likelihood ratio tests.

We performed 3 sensitivity analyses. In the first, we used stricter criteria to define healthy children; in the second, we used looser criteria; and in the third, we excluded children with migration backgrounds (i.e., ethnically White children who

neither they nor their parents were born in Switzerland [see online supplement, Table S2]).

We used STATA (version 16, Stata Corp, College Station, TX, USA) for statistical analysis and graphs. We followed STROBE reporting guidelines [29].

Results

Study population and sample selection

The LUIS study included 3870 children from 37 schools in the canton of Zurich. From these 3870 children, we analysed data from 2213 children who had parent completed questionnaires and child answered questionnaires, spirometry, and anthropometric measurements, were ethnically White, and healthy (Figure 1). After quality control of flow-volume curves, we included data from 2036 children. All had valid FEV₁ and 1762 had valid FVC measurements. The median age was 12.2 years (range 6.3-17.0, interquartile range 9.6-14.0) and 49% were males (Table 1). Most children were born in Switzerland (n=1814, 90%) but only two-thirds of their parents were (table 1). Compared to WHO references, our population had a similar BMI distribution (mean BMI z-scores: 0.04) but was slightly taller (mean height-for-age z-scores: 0.56).

Table 1: Characteristics of the whole study population and the study sample.

	Whole population N=2858 n (%)	Sample included for analysis N=2036 n (%)
Sex, male	1409 (49%)	995 (49%)
Sex, female	1449 (51%)	1041 (51%)
Age of the child in years, median (range)	12.6 (6.0, 17.2)	12.2 (6.3, 17.0)
Age groups		
6–11 years	1268 (44%)	975 (48%)
12–17 years	1590 (56%)	1061 (52%)
Height in cm, median (range)	156 (105-195)	154 (105, 194)
Height-for-age z-scores, mean (SD)	0.56 (1.01)	0.56 (0.98)
BMI z-scores, mean (SD)	0.06 (1.14)	0.04 (1.10)
BMI groups		
Underweight (z-scores <-2)	99 (3%)	61 (3%)
Normal weight (z-scores >-2, <1)	2168 (76%)	1573 (77%)
Overweight (z-scores >1, <2)	449 (16%)	310 (15%)
Obesity (z-scores >2)	142 (5%)	92 (5%)
Child born in Switzerland*	2522 (88%)	1815 (90%)
Mother originally from Switzerland*	1833 (64%)	1338 (67%)
Father originally from Switzerland*	1876 (66%)	1362 (68%)
Swiss SEP, median (range)	70.2 (28.0, 98.5)	70.4 (28.0, 98.5)
Parent history of asthma*		
No	2404 (84%)	1772 (87%)
Mother	236 (8%)	130 (6%)
Father	164 (6%)	102 (5%)
Both	18 (1%)	9 (<1%)
Wheeze, past 12m*	230 (8%)	-
Use of inhaled corticosteroids, past 12m*	146 (5%)	-
Physician diagnosed asthma ever*	254 (9%)	-

Abbreviations: SEP = socioeconomic position index; BMI = body mass index. Height-for-age and BMI z-scores based on WHO references. *Information obtained from the parent questionnaires.

Fit of GLI references

For children aged 6–11 years, fit was appropriate for FEV₁ (mean: -0.37; SD: 0.94), FVC (-0.26; 0.98), and FEF₂₅₋₇₅ (-0.49; 0.94). For adolescents aged 12–17 years, the fit for all 3 outcomes was poorer: mean FEV₁ was -0.62 (SD: 0.98), FVC was -0.60 (0.98), and FEF₂₅₋₇₅ -0.54 (1.02) (Figures 2 and 3, Table S1). Fit of FEV₁/FVC z-scores was appropriate for both age groups (mean z-scores: 6-11 years: -0.14; SD: 1.06; 12-17 years: -0.09; 1.02). The proportion of participants with mean z-scores below the LLN

was above the expected 5% for the 6–11 year age group for FEV₁ (9%; 95% CI: 7-11), FVC (8%; 6-10), and FEF₂₅₋₇₅ (11%; 9-13). Among adolescents, this was even more pronounced for FEV₁ (15%; 13-17), FVC (14%; 12-16), and FEF₂₅₋₇₅ (14%; 11-16). For FEV₁/FVC, proportions below LLN were close to 5% for males aged 6–11 years (4%; 3-7) and females (7%; 5-10) and males (7%; 5-10) aged 12-17 years, yet not for females aged 6–11 years (9%; 7-12). Findings were very similar in the 3 sensitivity analyses (Tables S2 and S3).

Associations of GLI-based spirometry z-scores by age, BMI, height, and sex

FEV₁, FVC, and FEF₂₅₋₇₅ z-scores were lower for older than for younger children (Figure 4). For ages 10-11 years, FEV₁, FVC, and FEF₂₅₋₇₅ z-scores were lower for females than males (Figure 4, Table S1). FEV₁/FVC z-scores did not vary with age (Figure 4) or height (Figure 5).

There was an inverse association of FEV₁, FVC, and FEF₂₅₋₇₅ z-scores by age in multivariable linear models (Table S4) and this association with age varied by height (*p* values for interaction: <0.034, 0.019, <0.001) (Figure 6). For those aged 6-11 years, FEV₁ and FEF₂₅₋₇₅ z-scores decreased with height, while z-scores increased with height for those aged 12-17 years. FVC z-scores increased with height in all age groups, but more so for adolescents. We found no interactions with sex (data not shown).

Children considered overweight or obese had higher FEV₁, FVC, and FEF₂₅₋₇₅ z-scores than children considered normal weight or underweight. Fit was poor for children classified as underweight (e.g., mean FEV₁: -1.13; SD: 0.92) or normal weight (-0.55; 0.96), but good for children considered overweight (-0.25; 0.92) or obese (-0.08; 0.93) (Figure 7, Table S5). FEV₁, FVC, and FEF₂₅₋₇₅ z-scores also increased with BMI in linear regression models (Table S4).

Discussion

Our large population-based study of healthy schoolchildren found that FEV₁, FVC, and FEF₂₅₋₇₅ z-scores were lower than expected from GLI reference values for Swiss schoolchildren, particularly adolescents. The association of FEV₁, FVC, and FEF₂₅₋₇₅ z-scores with height differed by age. Although FEV₁ z-scores decreased with height for younger children, they increased in older children. The fit of references for FEV₁, FVC, and FEF₂₅₋₇₅ z-scores was much better for children with higher BMI. This suggests that it is worthwhile considering whether weight or BMI data could be used to improve future GLI reference equations, resulting in a potentially better fit for samples who are slimmer or more overweight than the average.

Strengths and limitations

Our study had several strengths. The large sample size allowed us to assess the fit of GLI-based lung function z-scores in 2 age groups of White schoolchildren each with well over 150 males and 150 females, according to GLI recommendations [1]. We performed spirometry using a state of the art set-up instead of hand-held portable spirometers and did careful quality control of spirometry flow-volume curves, which increases the reliability of our results and makes measurement bias due to poor quality unlikely [20]. We obtained information about respiratory symptoms from both parents and children which allowed us to perform sensitivity analyses to define a healthy population, adding robustness to our findings. However, our study also had some limitations. First, we cannot exclude sampling bias because not all schools took part in the study. However, participating schools were similar to the whole canton [20]. Second, we lacked information about the onset of puberty; this information would

have allowed us to better explore whether pubertal stage influences fit of GLI references in adolescents. Last, although the canton of Zurich is the most populated and diverse region of Switzerland, we cannot assess if our findings apply to all Swiss schoolchildren.

Comparison with other studies

Other studies have assessed the fit of national data with GLI reference values. The fit was good in studies from the France, Norway, Spain, and the UK [14-17]; however, fit was poorer in studies from Germany and Italy [18, 19]. Our findings are comparable to those of Hüls et al. [19]. They assessed the applicability of GLI references for White children from several German cities using a school-based study of 1943 children aged 4-19 years and a population-based birth cohort of 1042 adolescents aged 15 years. They found sufficient fit for children younger than 10 years but systematically lower mean z-scores for FEV₁ and FVC for children older than 10 years. In the German study, mean FEV₁ and FVC z-scores were higher for those aged 6-9 years (e.g., mean FEV₁ z-scores = 0.01 for females and 0.02 for males), than in our study (-0.41 for females; -0.42 for males). In contrast to our study where females had lower z-scores than males (14-17 years = -0.70 for females; -0.53 for males), older males had lower FEV₁ and FVC z-scores than older females in the 2 German cohorts (e.g., mean FEV₁ z-scores for those aged 15-18-years = -0.24 for females; for males -0.41). In their study and in ours, fit for FEV₁/FVC was appropriate for all age groups. Fasola et al. found that GLI references produced mean z-scores >0.5 for FEV₁/FVC for 1243 males (mean: 0.75) and females (0.81) aged 7-16 years from southern Italy, which led to an underestimation of children with reduced FEV₁/FVC

[18]. In contrast, we found a good fit of FEV₁/FVC z-scores. Because mean z-scores were reduced both for FEV₁ and FVC, the ratio of these 2 remained normal.

The poor fit of GLI references for FEV₁, FVC, and FEF₂₅₋₇₅ for adolescents from our study may be related to differences in body size or timing of pubertal growth spurts that affect lung growth. Age of puberty onset, height, and timing of growth spurt vary across European countries [30, 31]. Children in northern countries generally have higher median height and later onset of puberty compared to southern countries [30, 31]. The height range in our study population was 113-194 cm for males and 105-180 cm for females aged 6-17 years, whereas the height range in the GLI reference population was 96-199 cm for males and 100-188 cm for females aged 6-18 years [32]. The mean height-for-age of our participants was higher (mean z-score: 0.56, SD: 1.01) than the WHO references [27]. This confirms that children from the canton of Zurich are on average taller than the WHO reference population [33]. Previous studies reported a slightly higher age of peak height growth in Switzerland (13.9 years for males; 12.2 years for females) than the UK (13.6 years for males; 11.7 years for females), Canada (13.4 years for males; 11.8 years for females), and the United States (13 years for males, 11 years for females) [34-36].

Another determinant of lung function in children is body composition [6]. Children considered overweight or obese have higher values for both FEV₁ and FVC, particularly for FVC, so that the FEV₁/FVC ratio is lower than for normal weight children [37, 38]. This aligns with our findings and may be due to asymmetric lung growth (i.e., increased dysanapsis) for children considered overweight and obese [9].

Implications

GLI references are too high for the lung function of healthy adolescents from our study. The GLI-based LLN underestimates dynamic lung volumes and forced flows (i.e., over detects abnormal results). This misclassification affects interpretations of lung function data in research and clinical practice [39]. In research, this affects cross-sectional studies when the lung function of children with a disease is compared to normal values. It also affects longitudinal analyses when using GLI references may make lung function in older children at follow-up appear poorer than at baseline. Clinicians commonly use spirometry to diagnose respiratory disease in children. Over detection of pulmonary restriction and obstruction may induce unnecessary treatments and surveillance burdens on patients, as well as unnecessary costs on healthcare systems [39]. For example, underestimated FEV₁ in adolescents may lead to misdiagnosis of asthma [40]. Many clinical trials use FEV₁ to enrol or assign individuals to an intervention. Use of GLI-based FEV₁ z-scores would classify more children below LLN and may lead to inclusion of children with lower severity. People with chronic respiratory disease are often included in national registries [41-43]. GLI-based FEV₁ z-scores of Swiss registry participants may wrongly seem to decline from early school-age through adolescence only because of the age-related change in fit with GLI z-score, not because their disease becomes worse. This can affect international comparisons of lung health. Clinicians and researchers interpreting GLI-based FEV₁, FVC, and FEF₂₅₋₇₅ z-scores should consider the possibility of finding results below LLN in slimmer, yet otherwise healthy children. Future studies should assess ways to compensate for BMI on spirometry indices.

In conclusion, we found evidence that GLI-based spirometry reference equations fit data for White children aged 6-11 years, but not for adolescents aged 12-17 years in Switzerland. BMI is a possible driver of deviations in FEV₁, FVC, and FEF₂₅₋₇₅ z-scores and future longitudinal studies should assess the impact of timing of pubertal growth spurt on these indices. We also advise caution when interpreting borderline spirometry results in slimmer adolescents.

Acknowledgements

We thank the staff from the schools, the children, and their families for taking part in the study, as well as LUIS study fieldworkers for their technical support during the study. We thank Marcel Zwahlen for his statistical advice. We thank Johanna M. Kurz, Andras Soti, Marc-Alexander Oestreich, Corin Willers (Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland), Léonie Hüsler, Eugénie Collaud, and Carmen C. M. de Jong (Institute of Social and Preventive Medicine, University of Bern, Switzerland) for their help in the assessment of the quality of the spirometry flow-volume curves. We thank Kristin Marie Bivens for her editorial contributions.

Author's contributions

Claudia E. Kuehni, Alexander Moeller, and Phillip Latzin conceptualised and designed the study. Alexander Moeller supervised data collection. Rebeca Mozun analysed the data and drafted the manuscript. Cristina Ardura-Garcia and Eva S.L. Pedersen supported the statistical analysis. All authors gave input for interpretation of the data. All authors critically revised and approved the manuscript.

The LuftiBus in the school (LUIS) Study Group

Alexander Moeller, Jakob Usemann (Division of Respiratory Medicine, University Children's Hospital Zurich and Childhood Research Centre, University of Zurich, Switzerland); Philipp Latzin, Florian Singer, and Johanna M. Kurz (Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland); Claudia E. Kuehni, Rebeca Mozun, Cristina Ardura-Garcia, Myrofora Goutaki, Eva S.L. Pedersen, and Maria Christina Mallet

(Institute of Social and Preventive Medicine, University of Bern, Switzerland); Kees de Hoogh (Swiss Tropical and Public Health Institute, Basel, Switzerland).

Funding

Lunge Zürich, Switzerland, funded the study set-up, development, and data collection with a grant to Alexander Moeller. *Lunge Zürich*, Switzerland, and University Children's Hospital Zurich and Children's Research Center, University of Zurich, Switzerland, funds LUIS data management, data analysis, and publications. Analysis was supported by a grant from the Swiss National Science Foundation (320030_182628) to Claudia Kuehni. Jakob Usemann and Florian Singer received grants from the Swiss lung foundation and the Bern lung foundation.

Conflicts of interest

Dr. Moeller reports personal fees from Vertex and OM Pharma—outside the present work. Dr. Latzin reports personal fees from Gilead, Novartis, OM pharma, Polyphor, Roche, Santhera, Schwabe, Vertex, Vifor, Zambon and grants from Vertex—all outside the present work. Dr. Singer reports personal fees from Vertex and Novartis—outside the present work. Dr. Usemann reports personal fees from Vertex—outside the present work.

References

1. Quanjer PH, Stanojevic S. Do the Global Lung Function Initiative 2012 equations fit my population? *Eur Respir J* 2016; 48(6): 1782-1785.
2. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, Initiative ERSGLF. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40(6): 1324-1343.
3. Hall GL, Thompson BR, Stanojevic S, Abramson MJ, Beasley R, Coates A, Dent A, Eckert B, James A, Filsell S, Musk AW, Nolan G, Dixon B, O'Dea C, Savage J, Stocks J, Swanney MP. The Global Lung Initiative 2012 reference values reflect contemporary Australasian spirometry. *Respirology* 2012; 17(7): 1150-1151.
4. Quanjer PH, Stocks J, Cole TJ, Hall GL, Stanojevic S, Global Lungs I. Influence of secular trends and sample size on reference equations for lung function tests. *Eur Respir J* 2011; 37(3): 658-664.
5. Thompson BR, Stanojevic S, Abramson MJ, Beasley R, Coates A, Dent A, Eckert B, James A, Filsell S, Musk AW, Nolan G, Dixon B, O'Dea C, Savage J, Stocks J, Swanney MP, Hall GL, Party ARRW. The all-age spirometry reference ranges reflect contemporary Australasian spirometry. *Respirology* 2011; 16(6): 912-917.
6. Chang SM, Tsai HJ, Tzeng JY, Yeh KW, Chen LC, Lai SH, Liao SL, Hua MC, Tsai MH, Huang JL, Yao TC. Reference equations for spirometry in healthy Asian children aged 5 to 18 years in Taiwan. *World Allergy Organ J* 2019; 12(11): 100074.
7. Arigliani M, Canciani MC, Mottini G, Altomare M, Magnolato A, Loa Clemente SV, Tshilolo L, Cogo P, Quanjer PH. Evaluation of the Global Lung Initiative 2012 Reference Values for Spirometry in African Children. *Am J Respir Crit Care Med* 2017; 195(2): 229-236.
8. den Dekker HT, Sonnenschein-van der Voort AMM, de Jongste JC, Anessi-Maesano I, Arshad SH, Barros H, Beardsmore CS, Bisgaard H, Phar SC, Craig L, Devereux G, van der Ent CK, Esplugues A, Fantini MP, Flexeder C, Frey U, Forastiere F, Gehring U, Gori D, van der Gugten AC, Henderson AJ, Heude B, Ibarluzea J, Inskip HM, Keil T, Kogevinas M, Kreiner-Møller E, Kuehni CE, Lau S, Mélen E, Mommers M, Morales E, Penders J, Pike KC, Porta D, Reiss IK, Roberts G, Schmidt A, Schultz ES, Schulz H, Sunyer J, Torrent M, Vassilaki M, Wijga AH, Zabaleta C, Jaddoe VWV, Duijts L. Early growth characteristics and the risk of reduced lung function and asthma: A meta-analysis of 25,000 children. *J Allergy Clin Immunol* 2016; 137(4): 1026-1035.
9. Forno E, Weiner DJ, Mullen J, Sawicki G, Kurland G, Han YY, Cloutier MM, Canino G, Weiss ST, Litonjua AA, Celedón JC. Obesity and Airway Dysanapsis in Children with and without Asthma. *Am J Respir Crit Care Med* 2017; 195(3): 314-323.
10. Quanjer PH, Stanojevic S, Stocks J, Hall GL, Prasad KV, Cole TJ, Rosenthal M, Perez-Padilla R, Hankinson JL, Falaschetti E, Golshan M, Brunekreef B, Al-Rawas O, Kühr J, Trabelsi Y, Ip MS. Changes in the FEV₁/FVC ratio during childhood and adolescence: an intercontinental study. *Eur Respir J* 2010; 36(6): 1391-1399.
11. Neve V, Girard F, Flahault A, Boule M. Lung and thorax development during adolescence: relationship with pubertal status. *Eur Respir J* 2002; 20(5): 1292-1298.
12. Lebowitz MD, Sherrill DL. The assessment and interpretation of spirometry during the transition from childhood to adulthood. *Pediatr Pulmonol* 1995; 19(2): 143-149.
13. Wang X, Dockery DW, Wypij D, Fay ME, Ferris BG, Jr. Pulmonary function between 6 and 18 years of age. *Pediatr Pulmonol* 1993; 15(2): 75-88.

14. Bonner R, Lum S, Stocks J, Kirkby J, Wade A, Sonnappa S. Applicability of the global lung function spirometry equations in contemporary multiethnic children. *Am J Respir Crit Care Med* 2013; 188(4): 515-516.
15. Neve V, Machuron F, Behal H, Howsam M, Methlin CM, Delille C, Baquet G, Matran R. Global Lung Initiative spirometry references in healthy 3-15-year-old French children. *ERJ Open Res* 2019; 5(3).
16. Martín de Vicente C, de Mir Messa I, Rovira Amigo S, Torrent Vernetta A, Gartner S, Iglesias Serrano I, Carrascosa Lezcano A, Moreno Galdó A. Validation of Global Lung Function Initiative and All Ages Reference Equations for Forced Spirometry in Healthy Spanish Preschoolers. *Arch Bronconeumol* 2018; 54(1): 24-30.
17. Langhammer A, Johannessen A, Holmen TL, Melbye H, Stanojevic S, Lund MB, Melsom MN, Bakke P, Quanjer PH. Global Lung Function Initiative 2012 reference equations for spirometry in the Norwegian population. *Eur Respir J* 2016; 48(6): 1602-1611.
18. Fasola S, La Grutta S, Cibella F, Cilluffo G, Viegi G. Global Lung Function Initiative 2012 reference values for spirometry in South Italian children. *Respir Med* 2017; 131: 11-17.
19. Hüls A, Krämer U, Gappa M, Müller-Brandes C, Schikowski T, von Berg A, Hoffmann B, Schuster A, Wisbauer M, Flexeder C, Heinrich J, Schulz H, Berdel D. Age Dependency of GLI Reference Values Compared with Paediatric Lung Function Data in Two German Studies (GINIplus and LUNOKID). *PLoS One* 2016; 11(7): e0159678.
20. Mozun R, Kuehni CE, Pedersen ESL, Goutaki M, Kurz JM, de Hoogh K, Usemann J, Singer F, Latzin P, Moeller A, On Behalf Of The Luis Study G. LuftiBus in the school (LUIS): a population-based study on respiratory health in schoolchildren. *Swiss Med Wkly* 2021; 151: w20544.
21. Madanhire T, Ferrand RA, Attia EF, Sibanda EN, Rusakaniko S, Rehman AM. Validation of the global lung initiative 2012 multi-ethnic spirometric reference equations in healthy urban Zimbabwean 7-13 year-old school children: a cross-sectional observational study. *BMC Pulm Med* 2020; 20(1): 56.
22. Al-Qerem WA. How applicable are GLI 2012 equations to a sample of Middle Eastern school-age children? *Pediatr Pulmonol* 2020; 55(4): 986-993.
23. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, Mitchell EA, Pearce N, Sibbald B, Stewart AW, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995; 8(3): 483-491.
24. Kuehni CE, Brooke AM, Strippoli MP, Spycher BD, Davis A, Silverman M. Cohort profile: the Leicester respiratory cohorts. *Int J Epidemiol* 2007; 36(5): 977-985.
25. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J, Force AET. Standardisation of spirometry. *Eur Respir J* 2005; 26(2): 319-338.
26. Quanjer P, Stanojevic S, Cole TJ, Stocks J. Global Lungs Initiative 2012 desktop calculator. 2012 [cited 2021; Available from: <https://www.ers-education.org/guidelines/global-lung-function-initiative/spirometry-tools/desktop-sheet-calculator/>]
27. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; 85(9): 660-667.

28. Brix N, Ernst A, Lauridsen LLB, Parner E, Støvring H, Olsen J, Henriksen TB, Ramlau-Hansen CH. Timing of puberty in boys and girls: A population-based study. *Paediatr Perinat Epidemiol* 2019; 33(1): 70-78.
29. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61(4): 344-349.
30. de Muinich Keizer SM, Mul D. Trends in pubertal development in Europe. *Hum Reprod Update* 2001; 7(3): 287-291.
31. Bonthuis M, van Stralen KJ, Verrina E, Edefonti A, Molchanova EA, Hokken-Koelega AC, Schaefer F, Jager KJ. Use of national and international growth charts for studying height in European children: development of up-to-date European height-for-age charts. *PLoS One* 2012; 7(8): e42506.
32. Quanjer PH, Hall GL, Stanojevic S, Cole TJ, Stocks J. Age- and height-based prediction bias in spirometry reference equations. *Eur Respir J* 2012; 40(1): 190-197.
33. Eiholzer U, Fritz C, Katschnig C, Dinkelmann R, Stephan A. Contemporary height, weight and body mass index references for children aged 0 to adulthood in Switzerland compared to the Prader reference, WHO and neighbouring countries. *Ann Hum Biol* 2019; 46(6): 437-447.
34. Largo RH, Gasser T, Prader A, Stuetzle W, Huber PJ. Analysis of the adolescent growth spurt using smoothing spline functions. *Ann Hum Biol* 1978; 5(5): 421-434.
35. Abbassi V. Growth and normal puberty. *Pediatrics* 1998; 102(2 Pt 3): 507-511.
36. Frysz M, Howe LD, Tobias JH, Paternoster L. Using SITAR (SuperImposition by Translation and Rotation) to estimate age at peak height velocity in Avon Longitudinal Study of Parents and Children. *Wellcome Open Res* 2018; 3: 90.
37. Forno E, Han YY, Mullen J, Celedón JC. Overweight, Obesity, and Lung Function in Children and Adults-A Meta-analysis. *J Allergy Clin Immunol Pract* 2018; 6(2): 570-581.e510.
38. Köchli S, Endes K, Bartenstein T, Usemann J, Schmidt-Trucksäss A, Frey U, Zahner L, Hanssen H. Lung function, obesity and physical fitness in young children: The EXAMIN YOUTH study. *Respir Med* 2019; 159: 105813.
39. Gaillard EA, Kuehni CE, Turner S, Goutaki M, Holden KA, de Jong CCM, Lex C, Lo DKH, Lucas JS, Midulla F, Mozun R, Piacentini G, Rigau D, Rottier B, Thomas M, Tonia T, Usemann J, Yilmaz O, Zacharasiewicz A, Moeller A. European Respiratory Society clinical practice guidelines for the diagnosis of asthma in children aged 5-16 years. *Eur Respir J* 2021.
40. Global Initiative for Asthma. Global strategy for asthma management and prevention, 2021. [cited 2021 14. May]; Available from: www.ginasthma.org
41. Kuehni CE, Rueegg CS, Michel G, Rebholz CE, Strippoli MP, Niggli FK, Egger M, von der Weid NX. Cohort profile: the Swiss childhood cancer survivor study. *Int J Epidemiol* 2012; 41(6): 1553-1564.
42. Korten I, Kieninger E, Yammine S, Regamey N, Nyilas S, Ramsey K, Casaulta C, Latzin P, For The Scild Study G. The Swiss Cystic Fibrosis Infant Lung Development (SCILD) cohort. *Swiss Med Wkly* 2018; 148: w14618.
43. Goutaki M, Eich MO, Halbeisen FS, Barben J, Casaulta C, Clarenbach C, Hafen G, Latzin P, Regamey N, Lazor R, Tschanz S, Zanolari M, Maurer E, Kuehni CE. The Swiss Primary Ciliary Dyskinesia registry: objectives, methods and first results. *Swiss Med Wkly* 2019; 149.

Figure 1: Flow-chart showing the selection of the study population.

Wheeze or asthma: parent reported wheeze in the past 12 months; use of inhaled corticosteroids in the past 12 months; or asthma diagnosed by a doctor ever in life. Active smoking: child reported smoking cigarettes, shishas, or e-cigarettes at least once per week. Sick the day of the measurement: child reported strong cough or cold the day of the measurement. Spirometry of inappropriate quality: We also excluded results for FVC-related parameters when we identified signs of early termination of expiration or cough or glottis closure after the first second of expiration in flow-volume curves. Abbreviations: FEV₁ = forced expiratory volume in 1 second. FVC = forced vital capacity. BMI = body mass index.

Figure 2: Kernel density plots showing the distribution of z-scores for spirometry parameters based on Global Lung Function Initiative (GLI) reference equations stratified by sex (A) and age groups (B) among healthy schoolchildren from the LUIS study.

Dashed vertical lines mark upper and lower limits of normal (ULN, LLN) at 5% (equivalent to +1.64, -1.64 z-scores). Abbreviations: FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity. FEV₁/FVC = forced expiratory volume in 1 second over forced vital capacity; FEF₂₅₋₇₅ = forced expiratory flow between 25-75% of the forced vital capacity.

Figure 3: Mean spirometry GLI-based z-scores of healthy schoolchildren from the LUIS study and percentage below lower limits of normal (LLN) stratified by age and sex.

LLN corresponds to z-scores <1.645. Abbreviations: GLI = Global Lung Function Initiative; FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity; FEV₁/FVC = forced expiratory volume in 1 second over forced vital capacity; FEF₂₅₋₇₅ = forced expiratory flow between 25-75% of forced vital capacity.

Figure 4: Scatter plots of GLI-based z-scores for spirometry parameters by age and by sex among healthy schoolchildren from the LUIS study.

Thick black lines show median age splines for z-scores of lung function parameters. Dashed thin horizontal lines mark upper and lower limits of normal (ULN, LLN) at 5% (+1.64, -1.64 z-scores). Abbreviations: GLI = Global Lung Function Initiative; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity; FEV_1/FVC = forced expiratory volume in 1 second over forced vital capacity; FEF_{25-75} = forced expiratory flow between 25-75% of forced vital capacity.

Figure 5: Scatter plots of GLI-based z-scores for spirometry parameters by height and by sex among healthy schoolchildren from the LUIS study.

Thick black lines show median height splines for z-scores of lung function parameters. Dashed thin horizontal lines mark upper and lower limits of normal (ULN, LLN) at 5% (+1.64, -1.64 z-scores). Abbreviations: GLI = Global Lung Function Initiative; FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity; FEV₁/FVC = forced expiratory volume in 1 second over forced vital capacity; FEF₂₅₋₇₅ = forced expiratory flow between 25-75% of forced vital capacity.

Figure 6: Marginal plot for predicted GLI-based z-scores for FEV₁, FVC, and FEF₂₅₋₇₅ at different heights in 4 age groups of healthy schoolchildren from the LUIS study.

P values for interaction were produced by likelihood ratio tests comparing the models with and without the interaction term= 0.033 for FEV₁ z-scores, 0.019 for FVC z-scores, <0.001 for FEF₂₅₋₇₅ z-scores. Abbreviations: GLI = Global Lung Function Initiative; FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity; FEF₂₅₋₇₅ = forced expiratory flow between 25-75% of forced vital capacity.

Figure 7: Distribution of GLI-based z-scores stratified by body mass index (BMI) categories among healthy schoolchildren from the LUIS study.

P values for a trend in z-scores across ordered categories of BMI were <0.001 for FEV₁, 0.014 for FVC, <0.001 for FEV₁/FVC, and <0.001 for FEF₂₅₋₇₅. The top lines of the boxes represent 75th percentiles, middle lines show the median, and bottom lines the 25th percentiles of GLI-based z-scores of spirometry parameters. BMI z-scores were calculated based on WHO references. BMI categories were based on WHO recommendations (underweight: <-2 z-scores; normal weight: ≥ -2, <1 z-scores; overweight: ≥ 1 z-scores; obesity: ≥2 z-scores). Abbreviations: GLI = Global Lung Function Initiative; FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity; FEV₁/FVC = forced expiratory volume in 1 second over forced vital capacity; FEF₂₅₋₇₅ = forced expiratory flow between 25-75% of forced vital capacity. Table S4 shows numeric results that correspond to this figure.