The Effect of Using the Lower Limit of Normal 2.5 in Pulmonary Aeromedical Assessments

Yara Q. Wingelaar-Jagt; Thijs T. Wingelaar; Metin Bülbül; Pijke P. vd Bergh; Erik Frijters; Erik Staudt

INTRODUCTION: Many regulations for aeromedical assessments state that a ratio between forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) of < 0.7 should be evaluated by a pulmonary specialist. The Global Lung Initiative (GLI) reference values introduced the lower limit of normal (LLN 2.5), in which the lowest 2.5% of the population is regarded as abnormal, instead of a fixed ratio. This study assesses the impact of adopting GLI reference values on aeromedical evaluation and referrals.

METHODS: The Royal Netherlands Air Force performed 7492 aeromedical assessments between February 2012 and April 2017. Cases with FEV_1/FVC < 0.7 from three groups were selected: 1) men < 25 yr; 2) men > 40 yr; and 3) women, with twice as many matched controls. Pearson's Chi-squared and Fisher's exact tests were used to analyze the data.

RESULTS: From the database, 23 (group 1), 62 (group 2), and 7 (group 3) cases were selected, with 184 controls. Respectively, 17%, 84%, and 29% would not be referred using the GLI. In the controls, this would lead to one additional referral (group 1). Qualitative analysis of the cases who would not be referred using the GLI showed that no significant diagnoses would have been missed.

DISCUSSION: Using the GLI LLN 2.5 reference values for pulmonary function tests leads to significantly fewer referrals to a pulmonary specialist without missing relevant pulmonary pathology in our aircrew. This would reduce resources spent on the assessment of aircrew without compromising flight safety.

KEYWORDS: pulmonary function tests, reference values, flight safety, aircrew, fitness to fly.

Wingelaar-Jagt YQ, Wingelaar TT, Bülbül M, vd Bergh PP, Frijters E, Staudt E. *The effect of using the lower limit of normal 2.5 in pulmonary aeromedical assessments*. Aerosp Med Hum Perform. 2020; 91(8):636–640.

Regular aeromedical examinations are performed to prevent—to the highest degree possible—in-flight medical events.[22] An in-flight medical event, when severe enough, may lead to pilot incapacitation and jeopardize flight safety.[11] Studies show annual incapacitation rates from 0.25% in general commercial aviation to 1.5% in fatal aviation accidents.[7,17] The most common categories of in-flight medical incapacitating events are of cardiovascular and neurological origin.[3,17,24] Other causes of in-flight medical events are gastro-intestinal complaints, psychiatry, and pulmonary problems.[4,7]

All military aeromedical examinations in the Royal Netherlands Air Force (RNLAF) are performed at the Center for Man in Aviation (CMA). The Military Aviation Regulations (MAR-FCL) are based upon the regulations imposed by the European Aviation Safety Agency, although they are more stringent in some respects.[5,14] According to both sets of regulations, applicants with significant impairment of pulmonary function shall be assessed as unfit. They state that a pulmonary function test (PFT) must be performed and an FEV_1/FVC ratio < 0.7 requires evaluation by a pulmonary specialist.[6,14] This commonly used fixed cutoff value for the FEV_1/FVC ratio was introduced by the Global Initiative for Chronic Obstructive Lung Disease guidelines in the early nineties as a "rule of thumb."[8] However, many studies have shown that defining the lower limit of normal (LLN) as this fixed ratio leads to extensive under- and over diagnosing in, respectively, younger and older adults.[2,10,23] This implies that using the fixed FEV_1/FVC ratio can lead to

From the Center for Man in Aviation, Royal Netherlands Air Force, Soesterberg, The Netherlands; the Diving Medical Center, Royal Netherlands Navy, Den Helder, The Netherlands; and the Central Military Hospital, Utrecht, The Netherlands.

This manuscript was received for review in December 2019. It was accepted for publication in April 2020.

Address correspondence to: Maj. Y. Q. Wingelaar-Jagt, Center for Man in Aviation, Kampweg 53, 3769 DE Soesterberg, The Netherlands; yq.wingelaar.jagt@mindef.nl.

DOI: https://doi.org/10.3357/AMHP.5566.2020

false-negative and false-positive conclusions when assessing PFTs in pilots.

Aside from this limitation of the fixed FEV_1/FVC ratio, the reference set used can also be of great influence. Since 1960, more than 70 different spirometry reference sets have been published, which all have significant variability in the definition of "normal," statistical approaches used, and included population.[18] This led to much debate worldwide and the Global Lung Initiative (GLI) was formed to generate a valid new dataset which could be used for all ages, both sexes, and included multiple ethnicities.[21] In 2012, the task force introduced a new dataset: the GLI 2012. This database is based on 74,187 records from healthy nonsmoking males and females between 3 and 95 yr of age from 26 countries across 5 continents.

Implementing this new dataset increases the validity for pulmonary function testing as it represents a wider population. Additionally, the fixed cutoff point for the FEV_1/FVC ratio was replaced by the Z-score, which indicates how many standard deviations a measurement lies from its predicted value.[21] This Z-score is independent of age, height, sex, and ethnicity. The American Thoracic Society and the European Respiratory Society (ERS) both recommend the use of the 5^{th} percentile to define the LLN in a clinical setting. This equals a Z-score of -1.64.[18] To reduce the number of false-positives in asymptomatic individuals, the authors of the GLI 2012 recommend the use of the LLN 2.5 (Z-score -1.96) as the cutoff point for screening for pulmonary disease in a healthy population.

Before implementing this new cutoff point in our aeromedical examinations, an impact assessment on the number of referrals and false-positive and false-negative results was needed. The general objective of this study was to illustrate what the effect will be of using the GLI 2012 on the number of referrals for evaluation by a pulmonary specialist when assessing past aeromedical examinations. Primarily, for patients who were referred using the ERS 1993, we wanted to compare if we would still refer them using the GLI 2012. For this specific group, we also wanted to study the clinical findings from the pulmonary specialist, so as to see what possible pulmonary diagnoses we might miss using the GLI 2012. Lastly, we wanted to investigate how many extra patients we would have referred had we used the LLN 2.5. Our hypothesis was that using the GLI 2012 would not lead to underdiagnosis of pulmonary disease. As a secondary outcome, we expected that use of the GLI 2012 would reduce the number of referrals for pulmonary evaluation in older applicants, but would slightly increase the number of referrals in young applicants.

METHODS

This study is a retrospective case-control analysis of routinely obtained data from annual medical assessments of pilots, aircrew, and air traffic controllers in the CMA of the RNLAF. In this study, the number of referrals for pulmonary evaluation using the GLI 2012 and the European Community for Steel and Coal/ERS 1993 dataset with their respective cutoff values were compared quantitatively.[19,21] In addition, in order to assess the pulmonary diagnoses that may be missed, we qualitatively analyzed the cases that would not be referred using the LLN 2.5, but were referred using the 0.7 FEV_1/FVC ratio.

The Medical Ethics Committee affiliated with the University Medical Center Utrecht (Utrecht, Netherlands) decided that our study was exempt from ethical review as the data were collected during regular aeromedical examinations and did not involve invasive interventions. Data were collected between February 2012 and April 2017. All data were obtained from the digital medical database which the CMA started to use in February 2012. All aeromedical examinations entered were eligible for inclusion. All aeromedical examinations were performed according to the MAR-FCL 3 standards.[9] All PFTs were conducted in compliance with American Thoracic Society/ERS recommendations.[15,18]

To avoid confounding when analyzing the entire database, we selected which groups would be most valuable to assess. Previous studies have shown that using the FEV_1/FVC ratio < 0.7 leads to significantly more misclassification in both older and younger patients.[16] Additionally, to remove bias caused by gender, we have separated both sexes. As the number of women in our database proved to be too small to stratify appropriately, we have included all female applicants. To summarize, we have included all individuals with an FEV_1/FVC ratio < 0.7 from 1) men under 25 yr, 2) men over 40 yr, and 3) all women.

As a rule of thumb, twice as many subjects were selected, matched for age, sex, and ethnicity, for the control groups. When several matched options for controls were available, the controls were selected randomly.

Baseline characteristics, results of PFT, pulmonary history, result of the aeromedical examination, and (if applicable) the results of the evaluation by a pulmonary specialist were obtained. All spirometry data used were baseline status. The Z-scores were calculated using the equations provided by the GLI.[9] These Z-scores were compared to the LLN 2.5 (a Z-score below -1.96).

Statistical analysis was performed using SPSS statistical software (IBM Corp; Armonk, NY: 2015, version 23.0). Differences between groups and subgroups were tested using Pearson's Chi-squared test. In the case of null values, Fisher's exact test was used. Statistical significance will be assumed with an $\alpha < 0.05$.

RESULTS

Between February 2012 and April 2017, 7492 records were entered in the CMA database. This corresponds to 1873 individuals, i.e., the majority of the individuals were tested more than once. Of these 7492 records, 366 (4.8%) had an FEV_1/FVC ratio below 0.7. Of these 366 records, 9 (2.5%) were women, the 9 records representing 7 different subjects. Of the 357 male records, 40 (23 persons) were younger than 25 yr of age and 183 records (62 individuals) were over 40 yr of age. See **Fig. 1** for the database search results.

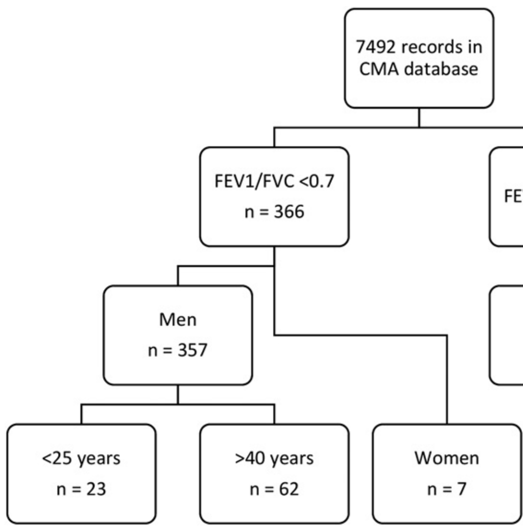

Fig. 1. Database search results.

For these 92 cases, 184 controls were selected, resulting in a total of 276 spirometry results for analysis. Of these 276 persons, 7 (2.5%) were declared unfit for duty for reasons other than pulmonary pathology. Baseline characteristics are displayed in **Table I**.

Table II shows the results of the comparison of using the fixed value of 0.7 or the LLN 2.5 as the cutoff point on the number of referrals. Referrals are significantly reduced (with $\chi^2 = 0.001$ in a Pearson's Chi-squared test) when using the LLN 2.5.

In a subgroup analysis of group 1 (men < 25) and group 2 (men > 40), similar values were found (both $\chi^2 = 0.001$). In the female analysis, a Fisher's exact test was used (due to the null values) and a P-value of 0.001 was found. This leads to the conclusion that referrals were significantly reduced in the three subgroups when using the LLN 2.5 instead of the fixed cutoff value of 0.7.

Adopting the LLN 2.5 would lead to one additional referral in the control group to the pulmonary specialist (0.5%): a man < 25 yr old. As the FEV_1/FVC ratio was > 0.7, he was not referred to a pulmonary specialist, so we do not know whether he had pulmonary pathology.

Using the LLN 2.5 as a cutoff point instead of using the fixed ratio of 0.7 will reduce the number of referrals by 58. Of these referrals, 52 (90%) were men in the group > 40 yr of age (group 2). Detailed analysis of their medical files shows that of these 58 persons, 34 (59%) had no pulmonary pathology according to the pulmonary specialist. There were 14 (24%) who were not referred because of mitigating circumstances such as poor technique or recent upper respiratory tract infection. Only 10 (17%) subjects were diagnosed with mild pulmonary pathology. The pulmonary diagnoses can be found in **Table III**.

Of the 10 cases with mild pulmonary pathology, 3 used pulmonary medication: corticosteroids with or without a long-acting beta-agonist. They had no pulmonary complaints and the medication did not seem to improve their pulmonary function. One of the three had already stopped using inhalation medication and in another case the pulmonologist considered stopping the medication. Two of the three had pulmonary complaints in childhood, so they could have been identified by a questionnaire.

DISCUSSION

Switching from the fixed cutoff value of 0.7 for the FEV_1/FVC ratio to the LLN 2.5 leads to significantly fewer referrals for pulmonary evaluation without missing relevant diagnoses, thus increasing validity and reducing resources for a fitness-to-fly assessment. This retrospective study was performed as part of an ongoing process to keep the aeromedical examinations at the CMA evidence based. This study met the objective, which was to illustrate what the effect would be on the number of referrals to a pulmonary specialist of using the GLI 2012 and the LLN 2.5 when looking at past aeromedical examinations. The results confirmed our hypotheses that using the GLI 2012 did not lead to underdiagnosis of pulmonary disease and did reduce the number of referrals in older applicants. To the best of our knowledge, no studies have been published studying the effect of switching from one reference set to another for pulmonary screening in aeromedical examinations. Due to international collaboration in NATO working groups, we know that there is no consensus as to what pulmonary screening for aircrew should include. Some countries only need a medical history, while others perform PFT examinations annually.

Both the Dutch MAR-FCL and the European Aviation Safety Agency regulations use the fixed cutoff value of 0.7 for the FEV_1/FVC ratio as an indication for referral. The present study shows that using a fixed cutoff value leads to over-diagnosing in aircrew > 40 yr. Using the LLN 2.5 would have reduced the number of referrals by 58, of which 90% were over the age of 40. We found limited evidence for the underdiagnosing in younger individuals when using the fixed cutoff value, as we only found one extra subject for referral.

Several studies have compared the consequences of adopting the GLI 2012 reference values to the reference values commonly used.[13,20] Most studies show that the GLI 2012 equations adequately describe lung function in most populations,

Table I. Baseline Characteristics.

GROUP	NO. OF SUBJECTS: CASES	NO. OF SUBJECTS: CONTROLS	MEAN AGE	ETHNICITY: CAUCASIAN	ETHNICITY: OTHER	SEX: MALE	SEX: FEMALE
1. Men < 25 yr	23	46	22.4 (18–24)	67	2	69	0
2. Men > 40 yr	62	124	49.0 (40–57)	186	0	186	0
3. Women	7	14	36.6 (26–51)	21	0	0	21

Table II. Number of Referrals.

	<LLN 2.5	>LLN 2.5	TOTAL
Cases, FEV_1/FVC < 0.7			
Subgroup 1	19 (83%)	4* (17%)	23
Subgroup 2	10 (16%)	52* (84%)	62
Subgroup 3	5 (71%)	2* (29%)	7
Subtotal	34	58*	92
Controls, FEV_1/FVC > 0.7			
Subgroup 1	1† (2%)	45 (98%)	46
Subgroup 2	0 (0%)	124 (100%)	124
Subgroup 3	0 (0%)	14 (100%)	14
Subtotal	1†	183	184
Total	35	241	276

* Cases that would not be referred using the LLN 2.5; †Controls that would be referred additionally using the LLN 2.5.

although some populations might have different values.[1,12] Furthermore, the interpretation of results using the Z-score describes how much the measurement deviates from the mean predicted value and makes assessment less biased by age, sex, and ethnicity. As we are generally assessing healthy aircrew, we adopted the 2.5th percentile as the LLN. This would result in fewer referrals without missing relevant diagnoses. When using the LLN 5 instead of the LLN 2.5, this would lead to the additional referral of 10 persons instead of 1 (almost all of them men younger than 25). Also, using the LLN 5 would lead to 29 fewer referrals instead of the 58 fewer with the LLN 2.5. So, using the LLN 5 instead of the LLN 2.5 would mean more referrals to a pulmonary specialist without generating new, relevant pulmonary diagnoses.

Decreasing the number of (unnecessary) referrals as part of the aeromedical examinations has many advantages. For one, a referral slows down the process, which in the case of a renewal may lead to an applicant being grounded until the pulmonary analysis is completed. Secondly, an investigation by a pulmonary specialist affects resources, both time and money, thus increasing medical cost.

The main limitation of this study is the small number of cases, particularly of female subjects. In the RNLAF, women are still a minority and therefore underrepresented in this study. Another limitation is that not all cases with an FEV_1/FVC ratio < 0.7 were referred to a pulmonary specialist (despite regulations saying that they should be referred). However, none of them declared pulmonary complaints and none were unfit to fly. Most of them had a one-time inadequate result due to recent upper respiratory tract infection.

Table III. Pulmonary Diagnoses of Subjects Who Would Not Be Referred Using the LLN 2.5.

DIAGNOSIS	*N*
Allergic asthma	2
Allergic asthma, medication	2
Nonallergic asthma	1
Nonallergic asthma, medication	1
Exercise-induced asthma	1
COPD GOLD I	3
Total	10

The main strength of this study is that it looks at the pulmonary standards used in aeromedical examinations and the consequences of implementing the GLI 2012 reference equations. There is no international consensus as to which pulmonary tests should be used, or what the criteria are. Previous studies have shown that the fixed cutoff value of 0.7 has no association with respiratory disease, but it is part of many aviation regulations.[5,14,16] As an aeromedical center, we always remain critical of the medical screening we perform and the standards used without making concessions to the safety of aircrew. We believe that herein lies the opportunity to keep improving the field of aviation medicine.

In conclusion, using the GLI 2012 reference values with the LLN 2.5 for PFT analysis would have led to significantly fewer referrals to a pulmonary specialist without missing relevant pulmonary pathology in our aircrew. This study has provided the impact assessment required to implement the GLI 2012 reference values and the LLN 2.5 in the Dutch Military Aviation Regulations. From an aeromedical perspective, less referrals to a pulmonary specialist reduces resources (time and money) spent on the assessment of aircrew without compromising flight safety. We would recommend other aeromedical centers to remain critical of the reference set and cutoff value used in their aeromedical examinations.

ACKNOWLEDGMENTS

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and affiliations: Yara Q. Wingelaar-Jagt, M.D., Erik Frijters, M.D., M.Sc., D.Av.Med., and Erik Staudt, M.D., Department of Aerospace Medicine, Center for Man in Aviation, Royal Netherlands Air Force, Soesterberg, The Netherlands; Thijs T. Wingelaar, M.D., Ph.D., Diving Medical Center, Royal Netherlands Navy, Den Helder, The Netherlands; and Metin Bülbül, M.D., and Pijke P. vd Bergh, M.D., Department of Pulmonology, Central Military Hospital, Utrecht, The Netherlands.

REFERENCES

1. Ben Saad H, El Attar MN, Hadj Mabrouk K, Ben Abdelaziz A, Abdelghani A, et al. The recent multi-ethnic global lung initiative 2012 (GLI2012) reference values don't reflect contemporary adult's North African spirometry. Respir Med. 2013; 107(12):2000–2008.
2. Celli BR, Halbert RJ, Isonaka S, Schau B. Population impact of different definitions of airway obstruction. Eur Respir J. 2003; 22(2):268–273.
3. DeJohn CA, Wolbrink AM, Larcher JG. In-flight medical incapacitation and impairment of airline pilots. Aviat Space Environ Med. 2006; 77(10):1077–1079.
4. DeJohn CA, Wolbrink AM, Larcher JG. In-flight medical incapacitation and impairment of U.S. airline pilots: 1993 to 1998. Washington (DC): Federal Aviation Administration, Office of Aerospace Medicine; 2004. Report No.: DOT/FAA/AM-04/16.
5. European Union Aviation Safety Agency (EASA). Commission Regulation No. 1178/2011: Aircrew Regulation - Annexes I to IV - Flight Crew Licensing and Medical Requirements. 2011. [Accessed 2019 Nov. 26]. Available from: https://www.easa.europa.eu/document-library/regulations/commission-regulation-eu-no-11782011.

6. European Union Aviation Safety Agency (EASA). Acceptable Means of Compliance and Guidance Material to Commission Regulation No 1178/2011. 2011. [Accessed 2019 Nov. 26]. Available from: https://www.easa.europa.eu/document-library/acceptable-means-of-compliance-and-guidance-materials.
7. Evans S, Radcliffe SA. The annual incapacitation rate of commercial pilots. Aviat Space Environ Med. 2012; 83(1):42–49.
8. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Fontana (WI): Global Initiative for Chronic Obstructive Lung Disease Inc; 2019.
9. Global Lung Initiative. E-learning resources, spirometry tools. 2012. [Accessed: 2019 Nov. 26]. Available from: https://www.ers-education.org/guidelines/global-lung-function-initiative/spirometry-tools.aspx.
10. Hansen JE, Sun XG, Wasserman K. Spirometric criteria for airway obstruction: Use percentage of FEV1/FVC ratio below the fifth percentile, not <70%. Chest. 2007; 131(2):349–355. Comment in: Conflicting definitions of airways obstruction: Drawing the line between normal and abnormal. [Chest. 2007]; and Shifting the goalpost to stay younger: normal spirometry but short of breath! [Chest. 2007].
11. International Civil Aviation Organization (ICAO). Manual of civil aviation medicine. 2012. [Accessed 2019 Nov. 26]. Available from: https://www.icao.int/publications/documents/8984_cons_en.pdf.
12. Kainu A, Timonen KL, Toikka J, Qaiser B, Pitkaniemi J, et al. Reference values of spirometry for Finnish adults. Clin Physiol Funct Imaging. 2016; 36(5):346–358.
13. Ketfi A, Gharnaout M, Bougrida M, Ben Saad H. The multi-ethnic global lung initiative 2012 (GLI-2012) norms reflect contemporary adult's Algerian spirometry. PLoS One. 2018; 13(9):e0203023.
14. Military Aviation Authority - the Netherlands (MAA-NLD). Military Aviation Requirements - Flight Crew Licensing 3, Part 3 (Medical). 2018. [Accessed 2019 Nov. 26]. Available from: https://english.defensie.nl/downloads/publications/2016/08/26/mar-fcl-3-flight-crew-licensing-medical.
15. Miller MR, Hankinson J, Brusasco V, Burgos R, Casaburi R, et al. Standardisation of spirometry. Eur Respir J. 2005; 26(2):319–338.
16. Miller MR, Quanjer PH, Swanney MP, Ruppel G, Enright PL. Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. Chest. 2011; 139(1):52–59. Erratum in: Chest. 2011; 139(3):733. Comment in: Lower limit of normal is better than 70% or 80%. [Chest. 2011].
17. Mitchell SJ, Evans AD. Flight safety and medical incapacitation risk of airline pilots. Aviat Space Environ Med. 2004; 75(3):260–268.
18. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, et al. Interpretative strategies for lung function tests. Eur Respir J. 2005; 26(5):948–968.
19. Quanjer PH. Standardized lung function testing. Bull Eur Physiopathol Respir. 1983; 19(Suppl. 5):45–51.
20. Quanjer PH, Brazzale DJ, Boros PW, Pretto JJ. Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. Eur Respir J. 2013; 42(4):1046–1054.
21. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, et al. 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. Comment in: New spirometric reference values. [Arch Bronconeumol. 2013].
22. Siedenburg J. Legal background of aviation medicine in Europe and its future development. Hippokratia. 2008; 12(Suppl. 1):64–73.
23. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Miller MR, et al. Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. Thorax. 2008; 63(12):1046–1051.
24. Taneja N, Wiegmann DA. An analysis of in-flight impairment and incapacitation in fatal general aviation accidents (1990–1998). Proc Hum Factors Ergon Soc Annu Meet. 2002; 46(1):155–159.