PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

A critical review of the basis for WHO's new recommendations for limiting annoyance due to environmental noise

Truls GJESTLAND

[1] SINTEF DIGITAL, Norway

truls.gjestland@sintef.no

ABSTRACT

The new WHO Environmental Noise Guidelines for the European Region include recommendations for limiting noise exposure to levels without adverse health effects. This paper reviews WHO's evidence with respect to aircraft and road traffic noise annoyance. It demonstrates that WHO's new guidelines are based on a non-representative selection of existing studies whose findings cannot be generalized to the residential population at large. The new WHO recommendations are unwarranted and unsupported by the reported evidence.

Keywords: Aircraft noise, road traffic noise, WHO recommendations

1. INTRODUCTION

The European Regional Office of the World Health Organization (1) recently dramatically lowered its former (2) recommendations for cumulative aircraft noise exposure levels associated with risks of adverse public health effects. Similarly, the recommendations for road traffic noise were also reduced to a lesser degree.

WHO's recommendations, although lacking the force of law, are nonetheless of interest to national regulatory bodies and to the public at large. It is therefore important that WHO's recent recommendations receive and withstand careful scrutiny.

WHO (2018) strongly recommends decreasing permissible noise exposure levels produced by aircraft to below L_{den} = 45 dB to prevent adverse public health consequences. WHO's newly identified noise exposure levels are an order of magnitude lower than those identified by WHO in 2000. WHO's 2000 recommendations were not source-specific recommendations, but suggested a limit of $L_{\mathrm{pA,16h}} =$ 55 dB to avoid health effects mediated by serious annoyance. The corresponding L_{den} value would have been higher for a full 24- (rather than 16-hour) day period. A source-specific correction for aircraft, however, would have moved WHO's recommendation in the opposite direction. WHO's new (2018) recommendation thus represents a shift of about 10 dB from its former recommendation. A 10 dB reduction in the duration-corrected loudness of cumulative noise exposure (to L_{den} = 45 dB from $L_{\mathrm{pA,16h}} =$ 55 dB) represents a halving of loudness as WHO's prior recommendation. This is a dramatic shift in the recommended "safe" limit on aircraft noise exposure, for which strong and reliable evidence is essential.

WHO also recommends reducing exposure to road traffic noise to less than L_{den} = 53 dB to avoid adverse health effects. This is a less drastic reduction, but still represents a lowering of the "safe" exposure level by 3 to 5 dB compared with WHO's former 2000 recommendation.

2. BASIS FOR NEW RECOMMENDATIONS

WHO commissioned different groups of noise experts to provide a basis for the new recommendations. Three researchers, R Guski, D Schreckenberg and R Schuemer, were given the task of reviewing the connection between environmental noise and annoyance. Their review was published as a separate report (3).

2.1 Road traffic noise

Guski *et al.* identified 25 studies of road traffic noise annoyance for their analysis. A scatterplot of the calculated percentage highly annoyed (%HA) for these studies is shown in Figure 1, together with the exposure-response function (ERF) for the complete dataset (black curve), and the ERF presented by Miedema and Oudshoorn (red curve) that is currently being used as a reference curve by EU.

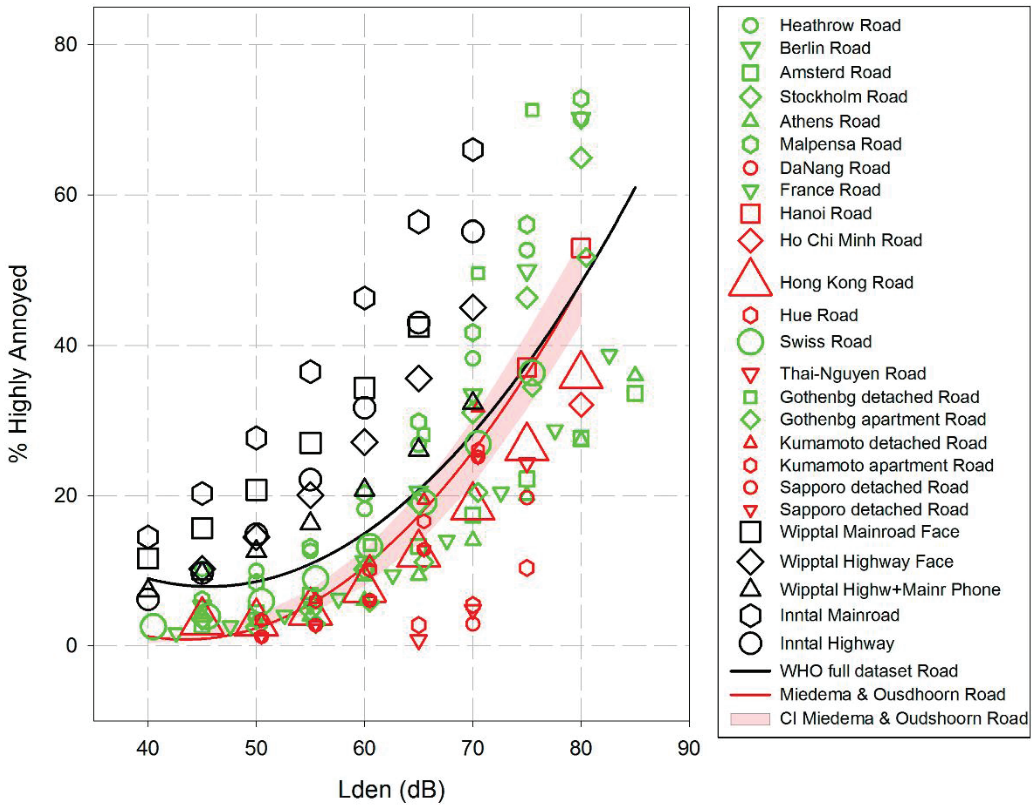

Figure 1. Scatterplot and quadratic regression functions for the relation between Lden and the calculated %HA for 25 road traffic noise studies reported by Guski *et al.*(3).

The most striking feature of the proposed ERF is the U-shape with a minimum at L_{den} 45.5 dB and increasing values for %HA for decreasing exposure levels below this point. This result is presented despite the observation made by Guski *et al.* : *It should be noted that 20 of 21 road traffic noise studies reporting a correlation show a statistically significant correlation between noise levels and raw scores, and they all show a clear increase of %HA with increasing noise levels, too* (3) (Supplement S14). The shape of the derived ERF is, of course, just an artifact of the curve-fitting methods that Guski *et al.* employed. Response data from only five studies are available at L_{den} 40 dB, of which four were conducted in the Austrian Alpine region. The latter studies yielded exceptionally high annoyance scores.

2.1.1 Alpine surveys

The dataset compiled by Guski *et al.* included five surveys from the Austrian Alpine region. According to Guski *et al.* three of these surveys used a cut-off at 60 % for "highly annoyed" while the other two used a cut-off at 73 %. This is incorrect, however. A closer reading of the report by Lercher *et al.* (4) reveals that they used the upper four categories of the 11-point scale (not the usual three) to classify respondents as highly annoyed in two of the studies. This corresponds to a cut-off at 64 %. The main reason for the high annoyance scores for the Alpine surveys (black plotting symbols in Figure 1) is thus different definitions of *highly annoyed*. Lercher *et al.* (4) also offer response data for the Alpine studies based on the usual cut-off at 73 %. The ERF curve for these data is shifted about 5 dB compared to the cut-off used by Guski *et al.* It is, however, not possible to validly compare data that has been derived using different definitions of highly annoyed. The five Alpine studies should therefore not have been included in the WHO dataset with a cut-off at 60-64 %.

In addition, Lercher *et al.* comment in their report that the noise prediction model they used may not be applicable to the special topography found in the narrow Alpine valleys. This is yet another reason for not including the results in the WHO dataset.

2.1.2 The Hyena study - road

Six of the surveys in the WHO dataset were part of the HYENA study (5). Respondents in the HYENA study were limited to residents aged 45 – 70 because it was designed primarily to study prevalence of hypertension. Noise annoyance has been shown to vary with age (6), with a maximum noise sensitivity around 45 years of age. The prevalence of highly annoyed residents in the HYENA study thus over-estimates the prevalence of annoyance in the general population in a more representative group aged 18 – 80 (as is commonplace in such surveys). Guski *et al.* comment on this fact but cite other studies that show only weak links between annoyance and age.

Social surveys on noise annoyance nowadays are usually carried out according to the ICBEN recommendations that specifies standard survey questions and response scales (7). This facilitates inter-survey comparisons. The HYENA study used a modified annoyance question. Instead of asking about annoyance in general, as is recommended by ICBEN, the participants were asked about the annoyance during the day and during the night. Guski *et al.* assumed, without documentation that annoyance during the day is no different from annoyance in general.

The assumptions that modification of the annoyance question and limitation of the age range make no difference to the response can be tested by comparing the average ERF for the HYENA study with the average ERF for the other surveys. The HYENA response is shifted 3.5 dB toward higher annoyance. The results of the HYENA study cannot therefore be compared directly with results from surveys conducted strictly according to the ICBEN recommendations and should not have been included in the WHO dataset.

2.1.3 Discussion - road traffic noise

The response data for all of the road traffic noise studies shown in Figure 1 have been analyzed by the CTL method defined in ISO 1996-1 (8). Original data reported from the surveys (pairs of noise exposure levels and % HA values) were used, rather than a secondary dataset calculated from individual, regression-derived exposure-response functions. Each survey was analyzed separately with no weighting of sample size differences. As Guski *et al.* (9) note, the influence of sample-size weighting is relatively small. Three ERFs were constructed: results from the five Alpine surveys, results from the six HYENA surveys, and the results from the remaining surveys. These curves are shown in Figure 2 together with the Miedema & Oudshorn curve which is currently used by the EU as a reference for road traffic noise annoyance.

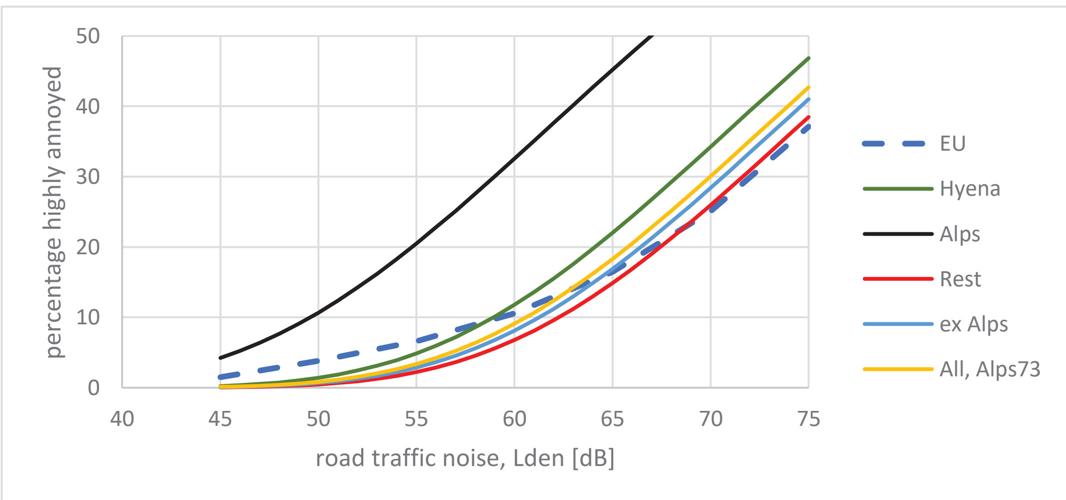

Figure 2. Exposure-response curves for road traffic noise annoyance calculated for the surveys presented in Figure 1.

The ERF based on the surveys that were conducted according to the ICBEN recommendations and using a cut-off of 73 % for highly annoyed (red curve) lies below the EU reference curve for most of the exposure range. The ERF for the HYENA surveys (green curve) displays higher annoyance equivalent to a shift in exposure level of 3.5 dB. It may be argued that the difference between the HYENA surveys and the rest is not severe enough to warrant exclusion of the whole study. The ERF "ex Alps" (light blue curve) shows slightly higher annoyance but still lies below the EU reference for exposure levels below L_{den} 65 dB.

The ERF for the Alpine studies (black curve) indicates a much higher prevalence of annoyance. This is due to a different definition of *highly annoyed.* It is evident from Figure 2 that the results from the Alpine surveys should not have been included in the manner elected by Guski *et al.* However, Lercher *et al.* (4) have provided alternative results for the Alpine studies based on the standard cut-off definition 73 %. An ERF based on all the surveys in the WHO dataset but with a comparable cut-off for the Alpine surveys is also shown in Figure 2 (yellow curve). Although based on some surveys of questionable quality, this curve also lies below the EU reference at low exposure levels but displays annoyance prevalence rates above the EU reference at higher levels.

An analysis of 47 surveys on annoyance from road traffic noise conducted between 1969 and 2015 yields an average CTL value of 78.7 dB (8). The ERF associated with this value is very similar to the EU reference curve. If the selection is limited to studies conducted after 1996, similar to the WHO dataset, the average CTL value for 19 post-1996 studies is 1 dB higher, L_{ct} 79.7 dB, indicating that people are slightly less annoyed.

A survey on noise annoyance in Switzerland has recently been reported by Brink *et al.*(9). This was a large-scale survey comprising more than 5400 respondents. The results reported by Brink *et al.* for road traffic noise have a CTL value of 78.3 dB (r^2=0.98), which is practically identical to the average CTL value for the 47 surveys analyzed by Gjestland (8). Actual response data reported by Brink *et al.* have been plotted in Figure 3 together with the ERF associated with L_{ct} 78.7 dB, the ERF proposed by Guski *et al.* (3), and the EU reference curve.

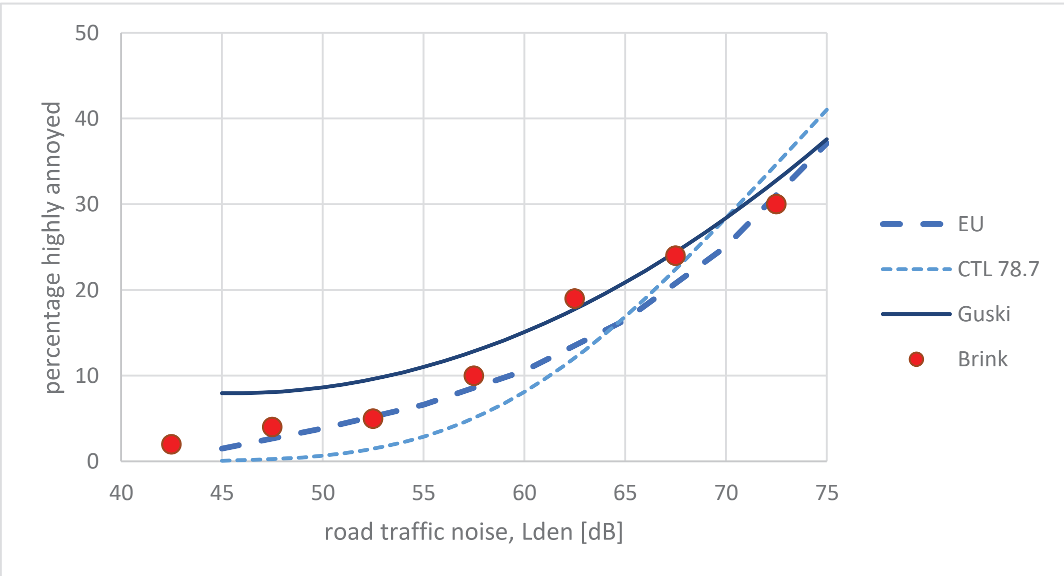

Figure 3. Exposure-response curves for road traffic noise annoyance as proposed by Guski *et al.*(3), the average of 47 surveys analyzed by Gjestland (8) (CTL 78.7 dB), and the EU reference curve. Red dots are results reported by Brink *et al.*

2.1.4. Conclusions – road traffic noise

The WHO recommendation for limiting road traffic noise annoyance deals with exposure at low levels. The yellow curve in Figure 2 shows that an exposure-response curve based on comparable results from available post-2000 surveys on road traffic noise annoyance displays annoyance prevalence rates below the current EU reference curve at exposure levels below L_{den} = 60 dB. Similarly, a comprehensive study, recently conducted in Switzerland with more than 5400 respondents, reported by Brink *et al.* shows results comparable to the current EU reference curve for low exposure levels. The lowering of the recommended guideline value to L_{den} = 53 dB to avoid adverse health effects therefore seems unwarranted and unsupported by existing evidence. At this noise level the average prevalence of highly annoyance is about 2 percent, a prevalence rate that cannot plausibly be considered evidence of a serious public health hazard.

2.2 Aircraft noise

Guski *et al.* included 12 studies of aircraft noise annoyance in their analysis (3). Figure 4 is a scatterplot of the calculated %HA for these studies, together with the exposure-response function for the complete dataset (black curve) and the ERF presented by Miedema and Oudshoorn that currently serves as a reference curve by EU (red curve).

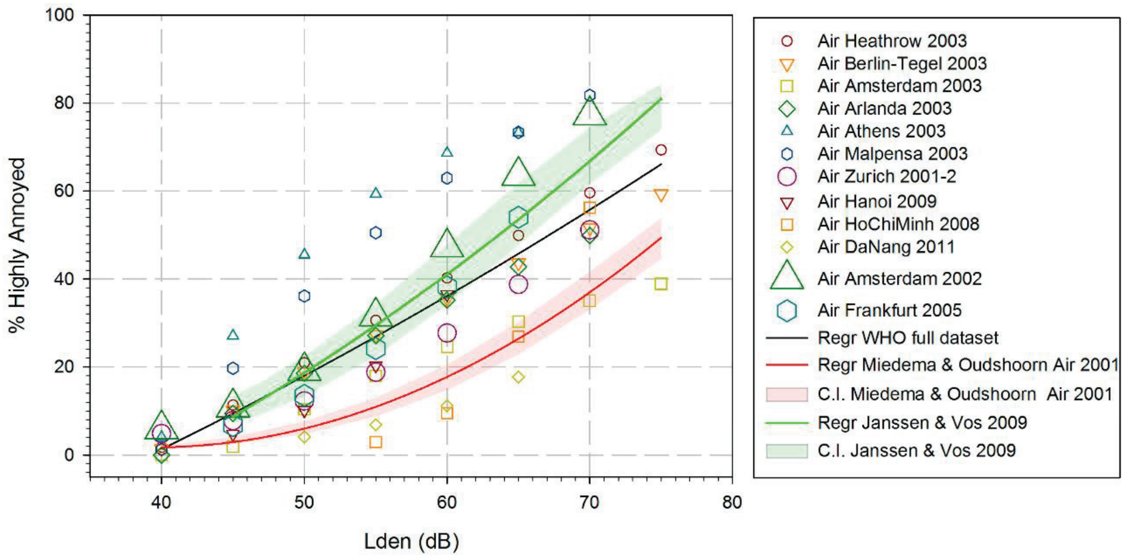

Figure 4. Scatterplot and quadratic regression functions for the relation between L_{den} and the calculated %HA for 12 aircraft noise studies reported by Guski *et al.*(3).

The data points in Figure 4 do not represent aggregated empirical observations as is usual in such plots. Instead, they represent predicted values estimated from regression curves developed for each of the studies. Inconsistent regression methods have been used in the various studies, and the regressions have been developed over different exposure ranges. Further, the results for the WHO full dataset were derived using a quadratic regression model that weighted studies by sample size.

A procedure based on combining all responses from different surveys in this manner does not take into account the fact that only about one third of the variance in the response data is explained by the cumulative noise exposure (12) while the rest of the variance is accounted for by so-called non-acoustic factors. As is evident from the very large spread in response data, such non-acoustic factors are especially prominent in aircraft noise surveys. Simple visual inspection of the data in Figure 4 shows that for the noise exposure range of most practical interest for regulatory purposes, L_{den} = 50 dB to L_{den} = 60 dB, the prevalence of highly annoyed residents varies between about 5 % and 70 %. This enormous spread is not necessarily fully attributable to personal or situational attitudes toward the cumulative noise exposure. A more plausible explanation would be that there must be other factors that also play an important role.

2.2.1 The HYENA study - aircraft

As in section 2.1.2 above, the HYENA study was designed to investigate the prevalence of hypertension among residents of neighborhoods near airports. The questionnaire did not follow the ICBEN recommendations. The selection of residents was limited to people aged 45 – 70 years. According to Guski *et al.* the respondents were selected at random from available registries, but at least at one airport, (Heathrow), some respondents may have been self-selected. A letter was distributed to members of a noise interests group urging them to participate in the survey and telling them that if you "would be interested in taking part, please telephone NN" at Imperial College, even if they did not belong to the randomly selected group.

Likewise, there is reason to question the adequacy of the exposure estimates at several of the HYENA airports. Babisch *et al.* (5) cite aircraft noise levels as low as $L_{\mathrm{Aeq\ 24h}}$ = 11 dB (Stockholm) and $L_{\mathrm{Aeq\ 24h}}$ = 22 dB (Milan). Such aircraft noise levels are not credible as modeling estimates, much less as acoustic measurements. Referring to Brink (13), Guski *et al.* (3) convert these estimates to L_{den} values by applying a correction factor of 2.6414 dB. These values suggest that the researchers relied heavily on software modeling to describe the aircraft noise exposure, with little evident concern for the limitations of aircraft noise modeling software, nor for the reasonableness of predicted exposure levels. No existing aircraft noise prediction programs can yield credible exposure predictions at such low levels (indicating a very long distance from the source). Applying an adjustment with a precision of "one ten-thousandth of a decibel", to a number that has been predicted by a model that yields output values with a standard error of at least 1-2 dB (14) , makes no sense.

Visual inspection of the annoyance data from the HYENA study reveals that two airports, Athens (ATH) and Milan (MXP) have an exceptionally high prevalence of highly annoyed neighbors, (see upper six surveys in Figure 4). The field work for the Athens study was conducted in 2003, but this airport was not opened until March 2001. Someone who has endured a noisy construction period of perhaps 3-4 years and then suddenly has been exposed to unfamiliar aircraft noise for two years, can not be considered a typical airport neighbor.

In a report on the results of the HYENA study its authors comment on the very high annoyance scoring of the Athens and Milan airports. They discuss several reasons for this including a bias in the selection of respondents and conclude that the data from these two airports is not representative for airports in general. They therefore exclude the data from their subsequent pooled analyses (5) (p.1175). Guski *et al.* nevertheless included both airports in the WHO full dataset.

Likewise, Babisch *et al.* state that “Our results may not be fully comparable to the EU curve, because the HYENA study annoyance was assessed in the limited age range of 45-70-year-old subjects" (5) (page 1175). The original researchers’ concerns were not shared by Guski *et al.* (11).

All in all, the HYENA study design deviates so much from standard survey procedures that the results should not have been included in the WHO dataset.

2.2.2 High-rate and low-rate airport change situation

Most airports experience growth in traffic over time. This increase usually occurs gradually over many years. Other airports are characterized by large abrupt changes due to the opening of a new runway, introduction of new flight paths, an abrupt increase in number of aircraft movements, *etc*.

Janssen and Guski (15) classify airports low-rate change airports if there is no indication of a sustained abrupt change of aircraft movements, or the intention of the airport to change the number of movements within three years before and after the annoyance study. They offer the following definition: "An abrupt change is defined here as a significant deviation in the trend of aircraft movements from the trend typical for the airport. If the typical trend is disrupted significantly and permanent, we call this a ‘high-rate change airport’. We also classify this airport in the latter category if there has been public discussion about operational plans within [three] years before and after the study". Low-rate change is the default characterization.

Gelderblom et al. (16) have applied this "high-rate/low-rate" classification to 62 aircraft noise annoyance studies conducted over the past half century. They show that there is a difference in the annoyance response between the two types amounting to about 9 dB. To express a certain degree of annoyance people at a high-rate change (HRC) airport on average "tolerate" 9 dB less noise than people at a low-rate change (LRC) airport Any attempt to develop an average dose-response curve from at set of studies will therefore be highly sensitive to the types of airports that are included. A high percentage of HRC airports will increase the average prevalence of highly annoyed people.

The WHO dataset includes 8 out of 12 HRC airports and about 83 % of the respondents. In contrast, in the dataset presented by Gelderblom et al., 17 out of 62 airports and about 35 % of the respondents are HRC airports. In the original dataset used by Miedema and Vos for their dose-response curve (17), only 2 out of 20 airports or about 10 % of the respondents were categorized as HRC.

2.2.3 Discussion - aircraft noise

The response data for all of the aircraft noise studies shown in Figure 4 were analyzed according to the CTL method as defined by ISO 1996-1 (10). Original data reported from the surveys (pairs of noise exposure and % HA) were used, rather than a secondary dataset calculated from individual exposure-response curves, except for the Amsterdam-2002 study for which original data could not be obtained. Each survey was analyzed separately, without any sample size weighting. The red curve in Figure 5 shows an average exposure-response function based on CTL analysis together with the ERF derived by Guski *et al.* (black curve) and the current EU reference curve for aircraft noise annoyance (dashed blue curve). As noted above some of the surveys included in WHO dataset are of a dubious quality (conducted in a non-standardized manner). These results cannot therefore be considered representative for the noise situation around airports in general.

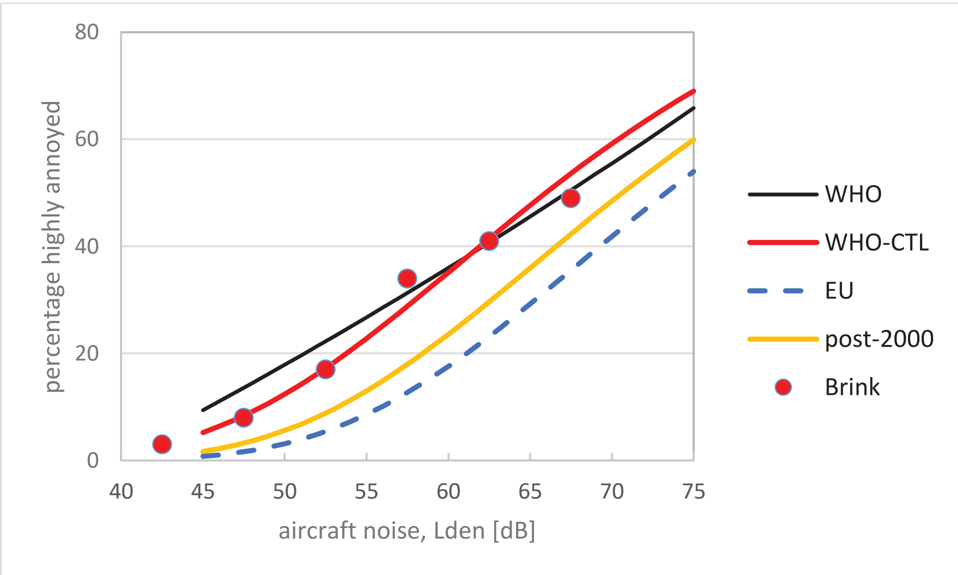

Figure 5. Exposure-response curves for aircraft noise annoyance

Gjestland (18) has compiled a different selection of post-2000 aircraft noise surveys that adhere to the inclusion protocol defined by Guski et al. (3) and for which sufficient data were available to do a comparative analysis. His dataset includes 18 surveys with 16,047 individual respondents. Half of the airports were characterized as HRC airports. The average exposure-response curve for these airports has also been plotted in Figure 5 (yellow curve). This curve lies above the EU reference curve, indicating a higher prevalence of highly annoyed people, but lies substantially below the ERF based on the WHO dataset. The difference between the EU reference and the ERF based on the alternative dataset by Gjestland is less than 1σ (one standard deviation) so they cannot be considered significantly different. This ERF shows 10 % highly annoyed at an exposure level L_{den} = 53 dB.

Brink *et al.*(9) have recently published results from a large Swiss study on annoyance from environmental noise. Data for aircraft noise annoyance have been compiled on the basis of the response from 2664 participants. Data from this study has also been plotted in Figure 5 (red dots). The results presented by Brink *et al.* confirm that the ERF suggested by Guski *et al.* overestimates the annoyance at low exposure levels. According to Brink *et al.* a prevalence rate of 10 % highly annoyed can be found for an exposure level L_{den} = 51 dB.

2.2.4 Conclusions aircraft noise

Even if all the surveys analyzed by Guski *et al.* (3), the WHO dataset, had been conducted by irreproachable methods, the fact that a similar analysis of a different and larger set of survey data (18) yields a very different result clearly indicates that the findings of Guski *et al.* (3) are not representative, and cannot be generalized to community response to aircraft noise around airports.

As has been shown by Gjestland (18) in his original critical remarks on the new WHO recommendations, and further documented in (19) and in the current paper, WHO's new recommendations for limiting aircraft noise are based on questionable evidence.

References

1. World Health Organization. *Environmental noise guidelines for the European region*. 2018, Copenhagen, Denmark: WHO Regional Office for Europe

2. Berglund, B, Lindvall, T, Schwela, T, Goh, K T, *Guidelines for community noise*, World Health Organization, 2000, ISBN 9971-88-770-3

3. Guski, R., Schreckenberg, D., & Schuemer, R. *WHO Environmental Noise Guidelines for the European Region. A systematic review on environmental noise and annoyance*. Int J of Environmental Research and Public Health, 2017, doi:10.3390/ijerph14121539

4. Lercher, P.; de Greve, B.; Botteldooren, D.; Rüdisser, J. *A comparison of regional noise-annoyance-curves in alpine areas with the European standard curves*. In Proceedings of the 9th International Conference on Noise as a Public Health Problem (ICBEN 2008), Foxwoods, CT, USA, 21–25 July 2008.

5. Babisch, W; Houthuijs, D; Cadum, E; Katsouyanni, K; Velonakis, M; Dudley, M.L; Swart, W; Breugelmans, O; Bluhm, G; Selander, J; Vigna-Taglianti, F; Pisani, S; Haralabidis, A; Dimakopoulou, K; Järup, L. *Annoyance due to aircraft noise has increased over the years - results of the HYENA study*. Environment International, 2009, 1169 - 1176.

6. Van Gerven, P., Vos, H., Van Boxtel, M., Janssen, S., & Miedema, H. *Annoyance from environmental noise across the lifespan*. J Acoust Soc Am, 2009, 187-194.

7. Fields; Jong, de; Gjestland; Flindell; Job; Kurra; Lercher; Yano; Guski; Felscher-Suhr. *Standardized noise-reaction questions for community noise surveys: research and a recommendation*. J Sound Vib, 2001, 641-679

8. Gjestland, T. *Forty-five years of surveys on annoyance from road traffic noise,* Proc. 23rd ICA, Aachen, Germany

9. Brink, M., Schäffer, B., Vienneau, D., Foraster, M., Pieren, R., Eze. I. C., Cajochen, C., Probst-Hensch, N., Röösli, M. Wunderli, J-M., *A survey on exposure-response relationships for road, rail, and aircraft noise annoyance: Differences between continuous and intermittent noise.* Environ. Int. Feb 2019, doi: 10.1016/j.envint.2019.01.043

10. ISO 1996-1. *Acoustics - Description, measurement and assessment of environmental noise - part 1: Basic quantities and assessment procedures.* 2016, International Standardization Organization.

11 .Guski, R., Schreckenberg, D., Schuemer, R., Brink, M., Stansfeld, S. A. *Comment on Gjestland, T. A systematic Review of the Basis for WHO's New Recommendation for limiting Aircraft Noise Annoyance. Int. J. Env. Res. Pub. Health 2018, 15, 2717, 1-9,* Int. J. Env. Res. Pub. Health 2019, 16

12. Basner, M., Clark, C., Hansell, A., Hileman, J. I., Janssen, S., Shepherd, K., & Sparrow, V. *Aviation noise impacts: State of the Science*. Noise & Health, 2017, 19: 41-50

13. Brink, M. (2015). *Conversion of transportation noise exposure metrics, Working paper*. Bern, Switzerland: Federal Office for the Environment.

14. Fidell, S., & Schomer, P. (2007). *Uncertainties in measuring aircraft noise and predicting community response to it.* Noise Control Engineering Journal, 82-88.

15. Janssen, S., & Guski, R. Chapter 7: *Aircraft noise annoyance. In Evidence Review on Aircraft Noise and Health.* S. A. Stansfeld, Eds, DG-JRC and DG-Env, European Union, Brussels, Belgium, 2017, (in press)

16. Gelderblom, F. B., Gjestland, T., Fidell, S., & Berry, B. *On the stability of community tolerance for aircraft noise.* Acta Acustica united with Acustica, 17-27, 2017

17. Miedema, H. M., & Vos, H. *Exposure-response relationships for transportation noise*. J Acoust Soc Am, 1998, 3432-3445.

18. Gjestland, T. (2018). *A systematic review of the basis for WHO's new recommendation for limiting aircraft noise annoyance*. Int J Environ Res Public Health, doi:10.3390/ijerph15122717

19. Gjestland, T. (2019) *Rebuttal to Guski, Schreckenberg, Schuemer, Brink and Stansfeld: Comment on Gjestland, T. A Systematic Review of the Basis for WHO's New Recommendation for Limiting Aircraft Noise Annoyance. Int. J. Env. Res. Pub. Health 2018, 15, 2717,1-9*. Int. J. Env. Res. Pub. Health 2019, 16