M+P | Member of Müller-BBM group
The solution people

Wolfskamerweg 47
NL-5262 ES Vught
The Netherlands
www.mplusp.eu

To70 Aviation Consultants
Postbus 85818
NL-2508 CM Den Haag
The Netherlands
www.to70.nl

ANNEX

To	Dutch Ministry of Infrastructure and Water Management (IenW)
From	Bert Peeters – M+P Desley Kemper – To70
Report number	21.171.30 – Annex
Date	May 2022
Pages	18
Subject	WHO guidelines for aircraft noise - summary of recent debate

Summary

The World Health Organization Regional Office for Europe published their Environmental Noise Guidelines for the European Region in 2018 [1]. For aircraft noise, the WHO recommends reducing noise levels below 45 dB Lden and 40 dB Lnight to avoid significant health impacts, and to implement suitable reduction measures with a preference towards infrastructural changes. Since 2018, several publications have been made that critically evaluate the WHO's findings, both from the noise community and from the aircraft sector. To some extent, researchers involved in the Guidelines development have answered to this critique.

From the published discussions, several points of discussion emerge that WHO researchers agree on or where the critique is not or insufficiently disproven, as assessed by the authors of this document. Unresolved discussion points can be summarised as follows:

- the exposure-response function (ERF) found by the WHO for aircraft noise annoyance, based on evidence from 2000 – 2014, is higher than the earlier ERF commonly used in Europe since 2001. Researchers disagree as to what extent this is due to a real increase in annoyance at a given noise level or due to differences in evidence selection and analysis methods;
- the WHO evidence base for aircraft noise show considerable heterogeneity, possibly caused by severe differences in non-acoustic factors between the various studies. As the WHO report itself states ([1], p109), it may not be possible to determine the exact annoyance for each exposure level in a generalised situation, and evidence derived in a local context should be applied whenever possible. WHO's decision to formulate 'strong' recommendations based on 'moderate' quality evidence is questioned, also by the aircraft sector;
- one important non-acoustic factor is airport changes: annoyance by aircraft noise exposure is higher at airports that have recently undergone changes (e.g. expansions, abruptly increased traffic). As the WHO evidence base contains several such 'high rate change airports', the ERF may be somewhat biased towards higher annoyance, and consequently a lower recommended Lden threshold;
- there is agreement that the ERF depends strongly on the selection of survey studies underpinning it. Critics challenge WHO's selection: some studies should not have been included, for various reasons of incomparability, while others should have been. Researchers seem to agree that it

would be valuable to revisit this selection and to include a number of new studies that has become available in recent years;
- specifically for Dutch airports, ERFs found by RIVM from national data actually show annoyance and sleep disturbance similar to the new WHO ERF at low noise levels and higher annoyance at high noise levels. Significant differences are found between various regional Dutch airports, attributed also to non-acoustic factors;
- considerations of costs and benefits, as the WHO claims were included, have not been based on thorough research or objective evidence. The WHO recommendations may be regarded as being formulated primarily from a health perspective, and balancing these against economic and other societal interests is up to the policy makers.

Goal and sources

The WHO Environmental Noise Guidelines have been published by the WHO in October 2018. The implementation of these guidelines may have considerable implications. This requires due diligence and serious consideration of relevant external reviews and criticism.
In the paragraphs below, some important points of debate regarding the WHO's Environmental Noise Guidelines report, the separately published systematic reviews and the WHO's final recommendations are described and summarised. The information below is not our own, it is a summary of several public sources, being

1 some points of discussion brought forward by the WHO researchers themselves in their own report and publications;
2 a series of journal articles, reply's and comments initiated by Truls Gjestland in the International Journal of Environmental Research and Public Heath;
3 a second strain of discussion publications initiated by Gjestland in the Journal of the Acoustic Society of America;
4 a report published by the European region of the Aircrafts Council International.

Also included in the review are two Dutch reports by RIVM. These reports do not critically review the WHO findings, but focus on policy recommendations how the WHO Guidelines could be used in strengthening national policies about health improvement.

Most of the critique in sources 1-4 seems to focus on the relation between aircraft noise and annoyance. The systematic review on sleep disturbance and other effects are only briefly challenged in these sources. The RIVM reports have a wider scope, covering also road/rail noise and other health impacts; we have used only the information relating to aviation noise.

<u>Disclaimer</u>
We have not performed a wide literature survey to find more sources of comments supporting or challenging the WHO research and guidelines. These sources were known to us and we believe that several of the contacts that are interviewed for our project are also aware of these publications. None of the critique below is our own. We have tried to objectively summarise what the various authors have stated, and to draw conclusions on what relevant points remain in the end.

The focus of this research is on the research methodology of the WHO. The sources did not contain discussion on the quality of the exposure-response function per airport, although it is discussed if the methods used for each individual airport give comparable results. Delving deeper into the quality of the individual airport studies is out of scope for this research. Also not included is any discussion

regarding the other noise sources in the WHO guidelines (i.e. road, rail noise, wind turbines or leisure noise).

WHO Environmental Noise Guidelines

In October 2018, the World Health Organization Regional Office for Europe published their Environmental Noise Guidelines for the European Region [1] ('ENG'), following their 1999 community noise guidelines [2] and 2009 night noise guidelines for Europe [3]. Although the guidelines focus on the European Region, the evidence base included also research from Asia, Australia and the USA. The ENG provide separate recommendations for road, rail, aircraft, wind turbine and leisure noise. For each of these sources, a systematic review of existing evidence on noise and health was performed by the WHO researchers. Each review focuses on a particular health outcome (e.g. annoyance, sleep disturbance, cardiovascular disease) to find reliable indications of adverse health effects related to each noise source as well as the magnitude of such effects. If possible, an exposure-response function (ERF) that describes the prevalence or incidence of that health endpoint as a function of the long-term average noise level (Lden, Lnight) is established. The WHO reviewers followed a predefined and strict process (GRADE) to grade the quality of the evidence from **high**: "further evidence is very unlikely to change the estimate" to **very low**: "any effect estimate is uncertain". Following this process, the evidence quality for each specific health outcome was also graded as a whole: if different high-quality studies show contradictory results, the quality is downgraded.
A separate team of experts, the Guideline Development Group (GDG), has pre-set absolute or relative risk levels for each health outcome (e.g. ≤10% highly annoyed people). Based on the systematic review results and quality assessment, the GDG has then formulated recommendations, including maximum Lden and Lnight levels. Each recommendation is then rated **strong**, indicating it can be adopted in most situations, or **conditional**, indicating it may not apply to all circumstances and adoption should be considered in a policy-making process with stakeholders. Among other things, the recommendation strength depends on the evidence quality: setting a strong recommendation was only considered if the evidence was at least moderate quality.

<u>Recommendations for aircraft noise</u>
For aircraft noise, the WHO recommendations are formulated as follows:

For average noise exposure, the GDG strongly recommends reducing noise levels produced by aircraft below **45 dB Lden**, as aircraft noise above this level is associated with adverse health effects.
For night noise exposure, the GDG strongly recommends reducing noise levels produced by aircraft during night time below **40 dB Lnight**, as aircraft noise above this level is associated with adverse effects on sleep.
To reduce health effects, the GDG strongly recommends that policy-makers implement suitable measures to reduce noise exposure from aircraft in the population exposed to levels above the guideline values for average and night noise exposure. For specific interventions the GDG recommends implementing suitable changes in infrastructure.

The WHO report that contains these recommendations gives further explanation and guidance. It is stated, for instance, that "data and exposure–response curves derived in a local context should be applied whenever possible to assess the specific relationship between noise and annoyance in a given situation". It is acknowledged that the response is different from airport to airport due to a

different fleet mix, night closure, frequency etc. The WHO nevertheless recommends the specific noise levels mentioned above. It is not stated that local exposure-response functions should be used to redetermine local noise limits.

The report does not recommend explicit methods to determine the exposure. The systematic reviewers mention that in their quality selection of studies to include, they assessed if the method to determine the Lden or Lnight was reported and reliable or standardised. It is expected that not all studies have used the same method (i.e. Doc29) and that results using different methods have been mixed.

Annoyance

The systematic review on environmental noise and annoyance [4] revealed that for aircraft noise there is high quality evidence of a correlation between the Lden noise level and the occurrence of annoyance. There is also evidence that the occurrence of high annoyance (HA) increases with an increase of the Lden. In numeric terms, the odds ratio (OR) is 4.7 for a 10 dB increase of Lden: at 60 dB Lden, the chance of being highly annoyed is 4.7 times higher than at 50 dB. The reviewers indicate high quality evidence of the OR being positive, but moderate quality evidence for its magnitude.

Finally, they establish an ERF from various available studies, see figure 1. The evidence for this ERF was graded as moderate quality. The quality of each individual study is regarded high, but as there is a wide scatter between the data of the various studies, as is clear from the figure, the overall evidence quality is downgraded to moderate.

Based on the ERF resulting from the full dataset (the black line in figure 1) and the pre-set benchmark level of maximum 10% HA, the GDG comes to the recommendation to reduce Lden levels for aircraft noise below 45 dB. The GDG rates this as a strong recommendation.

Other health impacts related to the Lden were also identified, such as ischaemic heart disease or cognitive impairment of school children, but with lower quality evidence and/or starting at higher levels.

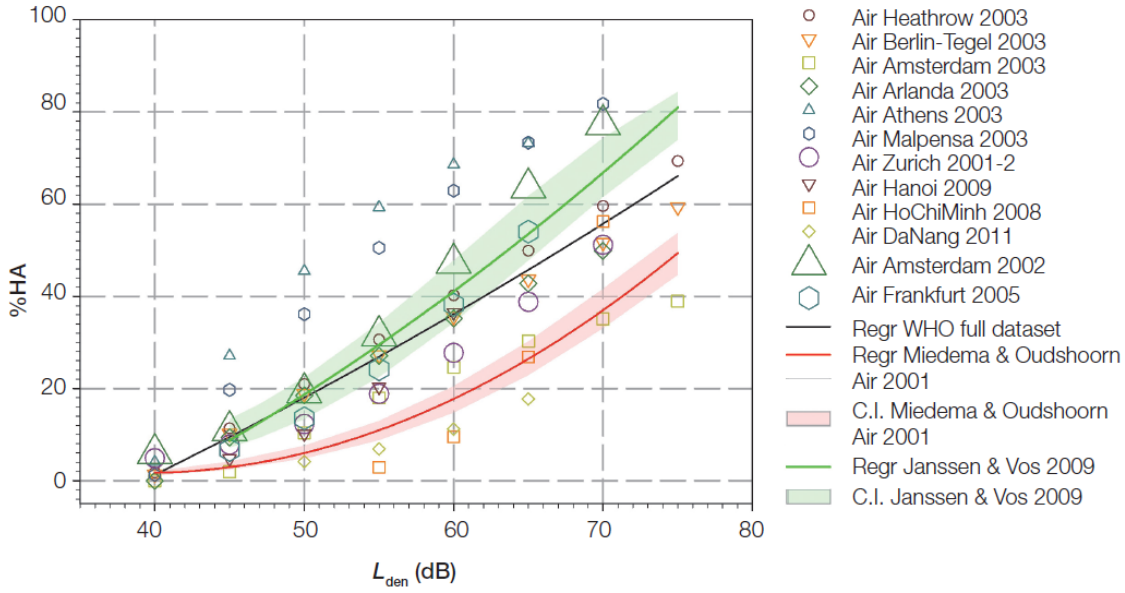

figure 1 Exposure-response function for aircraft noise: % highly annoyed population vs. Lden; black line shows the ERF from the 2018 WHO guidelines, red line shows the earlier EU standard curve (Miedema & Oudshoorn), green line shows an ERF by Janssen & Vos in 2009 (figure from [1])

Sleep disturbance

A systematic review was performed for environmental noise and sleep disturbance [5]. The review regarded studies of measured sleep disturbance, using polysomnography, but this did not lead to sufficient evidence to establish reliable health impacts. Rather, the WHO findings for sleep disturbance have been based on studies using 'self-reporting': surveys with questions about conscious problems with falling asleep, awakening, or being able to sleep continuously.

The reviewers conclude that there is evidence of an increased chance of sleep disturbance, combined from all sleep questions, with increasing Lnight values, with an OR of 1.94. Also, the percentage of highly sleep disturbed (HSD) population was established as a function of the Lnight in 5 dB classes. A second order polynomial was fitted through these data points, see figure 2. The evidence was rated moderate quality. It should be noted that these results refer to the surveys that questioned specifically how noise affects sleep. For other surveys that did not specifically mention noise, the evidence quality was rated very low. It should also be noted that the Lnight levels in all studies are outdoor noise levels at the most exposed building façade, which is not necessarily the bedroom façade.

Based on the polynomial function (black line in figure 2), the GDG recommends that Lnight is reduced below 40 dB. At this level, there is still 11% of HSD, which is considerably higher than the pre-set benchmark of maximum 3%. However, calculated noise levels below 40 dB Lnight were considered to be too inaccurate for any recommendation.

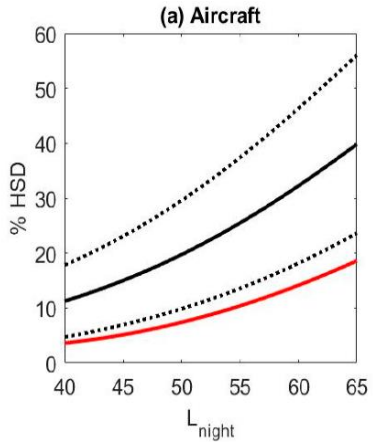

figure 2 Percentage of highly sleep disturbed based on self-reporting, for aircraft noise vs. the Lnight in dB. Black dashed lines show the 95% confidence intervals; red line shows the ERF by Miedema and Vos in 2007 (figure from [5])

Discussion

Self-criticism

Most important for the WHO recommended L_{den} level of max. 45 dB is the exposure-response function for annoyance (figure 1). The ERF is based on 12 recent studies, from the period 2001 – 2014. The authors of the WHO systematic review on annoyance mention a few points of discussion in their own publication.

- study heterogeneity: The variance between studies ('heterogeneity') is very large. As the selection criteria for studies to be included in the WHO's meta-analysis are sound and objective, it is assumed that the variance represents 'true' differences in the relation between annoyance and L_{den} in the various situations. Unfortunately, due to the analysis methods chosen, confidence intervals around the ERF could not be established, so there is no indication of the accuracy of the ERF.
- airport changes: Several causes for the heterogeneity have been explored, such as the different ranges of noise levels (large vs. smaller airports) in each study, the response rate, and the survey method, but these did not clearly explain the variance. One factor mentioned by the authors is the influence of recent abrupt airport changes that may, at least for some period, bias the opinion of the local community towards more annoyance. Examples are the opening a new runway or an abrupt increase in the number of flights, or public discussion about intentions to do so. The WHO reviewers regard five of the 12 airport studies to be 'high rate change' airports where such a cause of negative bias exists, and another five studies are regarded as 'low rate change' airports. For two airports it was not clear if they should be regarded as 'high rate change'. Separate ERFs have been established for the high and low rate change airport studies, see figure 3. The analysis shows that high rate change airports indeed show higher %HA at the same L_{den} level than low rate change airports. It is also clear that the low rate change ERF is closer to the 2001 Miedema/Oudshoorn standard curve, but still higher.

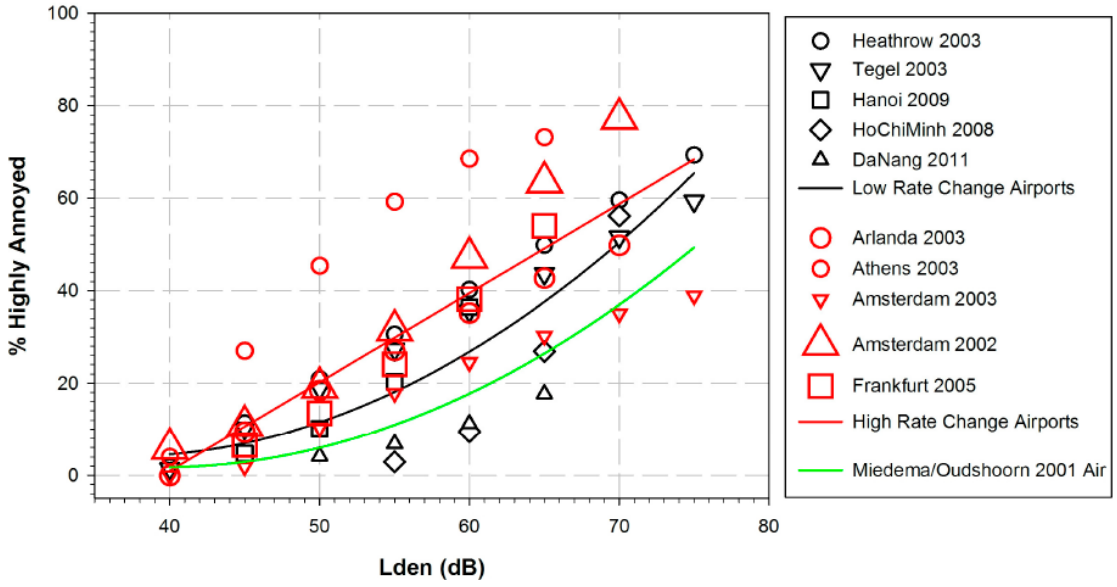

figure 3 Re-analysis of the studies used for the new ERF for annoyance, now separated between low rate change (black) and high rate change (red) airports; two studies not in either category are excluded

Gjestland vs. Guski et al.

Shortly after the publication of the WHO Guidelines in October 2018, Truls Gjestland published a paper [6] in the International Journal of Environmental Research and Public Health (IJERPH), the same journal in which the WHO researchers published their systematic reviews. Gjestland's paper critically reviews the WHO's systematic review on annoyance by Guski et al. [4], specifically for their work and findings for aircraft noise.

Gjestland is an experienced researcher at SINTEF in Norway, one of Europe's largest independent research organisations. He has over 50 years of research experience and has no direct affiliation with the aircraft industry, nor with the health sector. He was part of the WHO External Review Team within the development process of the ENG.

The critical review by Gjestland (Dec 2018) was commented in a IJERPH publication by Guski et al. [7] (Mar 2019), which Gjestland replied to in another IJERPH publication [8] (Mar 2019).

In summary, the following points have been discussed:

- non-acoustic factors: Gjestland claims that the importance of non-acoustic factors has been underestimated: only 1/3 of annoyance can be explained by acoustic factors relating to the L_{den} (maximum levels, number of flights, …) and non-acoustic factors such as population age, education, or the general attitude towards flying are more important than the noise levels. This is confirmed by Guski et al., claiming that another 1/3 of annoyance is explained from non-acoustic factors and the last 1/3 is unexplained variance. An example provided by Gjestland is that the Milan Malpensa airport study may have been influenced by a plane crash at the nearby Milan Linate two years before, leading to increased fear among the Milanese population. Guski et al. claim that scientific evidence for an influence of fear of accidents on perceived annoyance is scarce or just anecdotal, and that it would not be so local (such accidents would lead to fear of aircraft also at other airports).

- HYENA studies – population age: of the 12 airport studies included in the ERF evidence, six stem from a single research project HYENA[1]. The data used for these six studies only contain participants aged 45 – 70 years and did not include the younger age group (18 – 45 years). Gjestland references other research showing that sensitivity towards noise annoyance is highest for people around 45 years and much lower among people aged 20 years. The HYENA studies may thus have biased the ERF of the full dataset towards higher %HA. Guski et al. quote other research sources showing very small or insignificant effects of age on sensitivity. Gjestland replies by quoting Guski et al. themselves stating in their original review that a certain bias can be assumed for these studies, but was not or could not be tested. He also quotes the HYENA researchers stating the due to the limited age range their results may not be fully comparable to the earlier EU standard curves, i.e. 2001 the Miedema & Oudshoorn ERF.
- HYENA studies – non-standard questions: Gjestland, being involved with the revision of the ISO15666 standard on noise annoyance surveys, points to the fact that the six HYENA questionnaires deviated from the standardised questions as they asked separate questions about annoyance during the day and during the night. Guski et al. have then used only the daytime questions in their analysis of annoyance, together with date from other studies that did use the standardised questions not mentioning a specific time period. Daytime-only questions could lead to lower annoyance, assuming nighttime noise is more annoying. This critique is not commented by Guski et al.
  Another point made by Gjestland is that for the HYENA study around Heathrow, a local noise interest group (HACAN) advised its members by letter to participate in the study. As HACAN members will generally be opposed against Heatthrow airport, this may have led to self-selection bias and higher %HA. As this point was made in his last reply, there has been no reaction from Guski et al.
- study weighting by response count: The data used for the WHO's ERF (figure 1) are weighted, as indicated by the size of the datapoints. The weighting factor for each study is according to the number of study participants: smaller studies have lower weights. Guski comments that this could be valid for small studies up to a certain number of participants, but at some point adding more participants will no longer influence the ERF for a particular airport. The Schiphol study, for instance, has ca. 6000 participants, which is 40% of the total, and weighs particularly heavy in the final ERF. Guski et al. in their reply do not dispute this critique as such, but do show that an unweighted version of their ERF is very close to the original weighted version, see figure 4.
- high / low rate change airports: Gjestland confirms what the original systematic review states, namely that high rate change airports show a higher ERF than low rate change airports, quoting also other research showing this effect, albeit slightly larger. Gjestland claims that not five, but eight of the twelve airports in the WHO dataset, representing 83% of the total number of participants, should be regarded as high rate change airports, for various reasons. Guski et al. do not explicitly deny this, but claim that a typical or representative EU airport is not defined and may not exist. Gjestland quotes another study based on a different (older, but larger) dataset showing only 35% of study participants living near high rate change airports. Also, looking at the dataset underpinning the Miedema & Vos ERF, only two out of 20 airports could be classified as high rate change.
- study selection: Gjestland presents his own selection of 18 aircraft noise annoyance studies between 2001 – 2015 fulfilling the WHO requirements, of which six overlap with the WHO dataset. In his analysis of these data, Gjestland follows the Community Tolerance Level (CTL)

[1] HYENA (HYpertension and Exposure to Noise near Airports) was a research project by a European consortium on health impacts from aircraft noise, finished around 2009 with study data ranging from the period 2003 – 2005.

approach which differs from the WHO methods. Yet, applying the CTL approach to his own dataset as well as the earlier dataset used by Miedema and Vos, he shows that 'his' exposure-response curve is closer to the Miedema & Vos curve than the new ERF found by the WHO. Guski et al. comment that they see no value of the CTL approach over their own analysis method and show that CTL curves also exhibit large heterogeneity. They do not comment on the different results found from a different dataset. However, as several of the new studies selected by Gjestland were not available earlier, they have not been overlooked by the WHO reviewers, but were simply not yet available. The ratio of high rate change vs. low rate change airports in Gjestland's dataset is ca. 50/50.

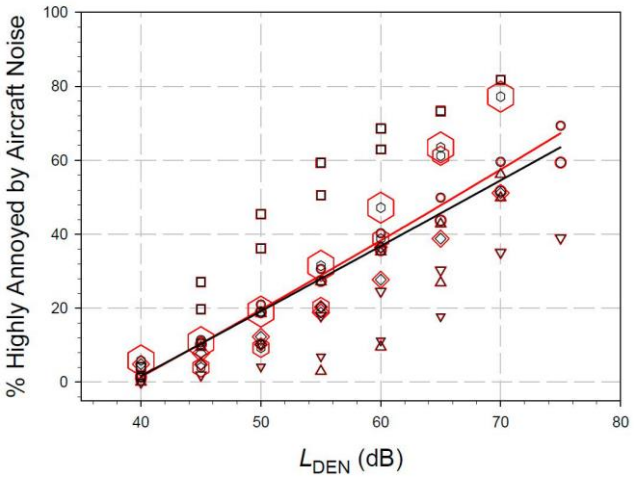

figure 4 WHO exposure-response function for %HA vs. L_{den} in dB for aircraft noise; red line is weighted according to (square root of the) sample size, black line is unweighted

In conclusion, Guski et al. state that 'there were no specific flaws, faults, or inaccuracies in the analysis of the available evidence' and that the GDG 'did not come to false conclusions and that their recommended guideline value for aircraft noise is not unjustifiably low'. Gjestland retains in his second reply that 'WHO's new recommendations for limiting aircraft noise are based on questionable evidence'.

Gjestland vs. Brink

In August 2020, Truls Gjestland started a second strain of discussion on the same topic with an article [9] in the Journal of the Acoustic Society of America (JASA), an international scientific journal well-known in the acoustics community. Like the previous discussion, Gjestland's critique is aimed at the WHO's recommendations regarding aircraft noise annoyance. As a reaction on this paper, a Letter to the Editor was published ([10], Dec 2020) from Mark Brink of the Swiss environmental agency, who was a member of the WHO Guideline Development Group. Gjestland replies by sending another Letter to the Editor ([11], Mar 2021).

Some of Gjestland critique overlaps with the previous IJERPH discussion, but the main focus is on the question whether or not there is a temporal trend in the level of annoyance experienced from aircraft noise exposure. Does the fact that the new ERF found by WHO is significantly higher than the earlier curves (Miedema & Oudshoorn, 2001) mean that the public's attitude towards aircraft

noise has become more negative? Or does the difference stem from difference in research methods and selection of underlying studies, and is there not such a temporal trend?
Gjestland presents a list of 65 noise annoyance studies regarding different airport, from the period 1961 to 2014. Figure 5 shows two ERFs[2] calculated from these data representing different periods: 1961-2000 and 2001-2015. Gjestland states that the 'younger' ERF is somewhat higher than the 'older' ERF, indicating an increase in noise annoyance with time. However, as the 95% confidence bounds overlap he claims this difference is not significant, and the conclusion is that 'in contrast with the findings of Guski et al. (2017), the prevalence of high annoyance with aircraft noise has not significantly changed over the last half century'.
With figure 6, Gjestland shows another way to identify a possible temporal trend: The GDG has formulated their recommendation of max. 45 dB L_{den} as this is the level that corresponds in their ERF with 10% HA. Gjestland has calculated the L_{dn} corresponding to 10% HA for each of the 65 studies he collected. These levels are shows in figure 6, as well as the average of these levels over successive 5-years period. An increase of noise annoyance over time should reveal a clear downward trend of these levels over time, which he claims this graph does not show.
In conclusions, Gjestland states that the 2001 ERF by Miedema and Oudshoorn 'can still be considered the best estimate for prevalence of annoyance with aircraft noise. According to their exposure-response curve, an annoyance prevalence rate of 10% HA corresponding to the limit to avoid adverse health effects should be set at L_{dn} = 54 dB, not 45 dB, as recommended by WHO'.

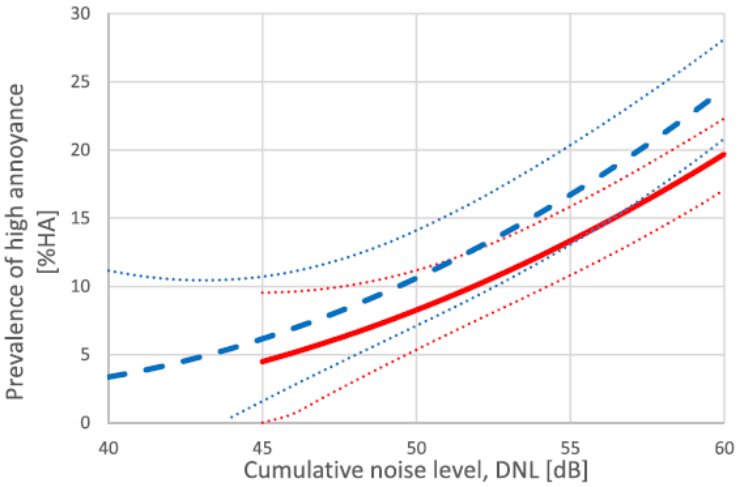

figure 5 ERF for aircraft noise annoyance, as %HA vs. the L_{dn} / Day-Night Level, combined from 43 surveys before the year 2000 (red line) and 22 post-2000 surveys (blue dashed line); dotted lines indicate 95 confidence intervals, but are calculated from the aggregated data from individual studies and do not represent confidence directly related to survey response (figure from [9])

[2] In Gjestland's paper, day-night levels (or Ldn) are combined with day-evening-night levels (Lden) without conversion, as the difference between these values is typically less than 0.5 dB. This is not disputed by Brink in his response.

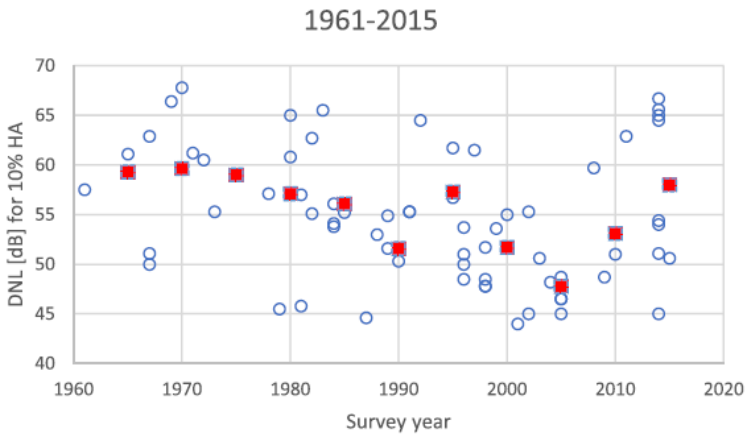

figure 6 The Day-Night Level (Ldn) corresponding to 10% HA from 65 studies vs. the study year 1961 – 2015; red squares indicate the average of these 10% HA noise levels over successive 5-year periods (from [9])

Brink responds in his Letter [10] that Gjestland's analysis lacks a clear study objective, does not provide clear selection criteria for his collection of studies and also lacks any description of how the data were extracted from these studies, some of which are decades old. Brink confirms that the selection of which studies to include in such a meta-analysis or not may severely affect the results, stressing that objective criteria to select only good quality and comparable studies are indispensable. Without these, 'we will not know (1) if aircraft noise annoyance has increased or remained stable over the last decades, and (2) if the WHO guideline value for aircraft noise is appropriate or not.' Gjestland replies by using a similar argument, stating that the study selection by Guski et al. was too restrictive ('cherry-picking') and that their results depend on the specific characteristics of the limited number of studies used. Also, he points out that his results (figure 6) show that of the 65 studies he found, only 5 studies report a level corresponding to 10% HA that is at or slightly below 45 dB L_{dn}, and the majority of studies shows (much) higher corresponding levels.
Gjestland also responds to Brink's point that older aircraft noise studies are based on lower quality noise assessment models, by stating that this would lead to more uncertainty, but not to a systematic trend towards a higher ERF.

Airports Council International

The Airports Council International, European region (ACI Europe) published an analysis paper in 2018, primarily as a response to the 2018 WHO Guidelines and express their own views and concerns on the topic. ACI Europe is the professional association of airport operators, representing 500 airports in 45 European countries. The names of the report authors and their capacity are not given.

In their introduction, ACI express that the modern world is dynamic and full of noise, especially in the busy cities. They also point to the fact that many people expose themselves to sound by talking on the phone and using personal entertainment systems with headphones, partially isolating themselves from environmental noise. This fact is clearly highlighted by the graphical presence of earphones on the report's cover (see figure 7). ACI also stresses the importance of mobility in modern life, being intrinsic to our existence ('living is moving'). And they state that they welcome WHO's work done to collect all the evidence on noise impacts, but that 'for air transport the resulting Environmental Noise Guidelines are not as constructive as we had initially hoped'. Further in their report, they describe

their critique in more detail, and show the current and past efforts of the airport sector to decrease its noise impact.
It is clear and fully understandable that the ACI report is written in defence of their own interests, as this is their role. In our document, we will limit our contemplations on the ACI report to a few valid points of critique towards the WHO guidelines. We are not aware of any public or official reaction from the WHO research team on the ACI report.

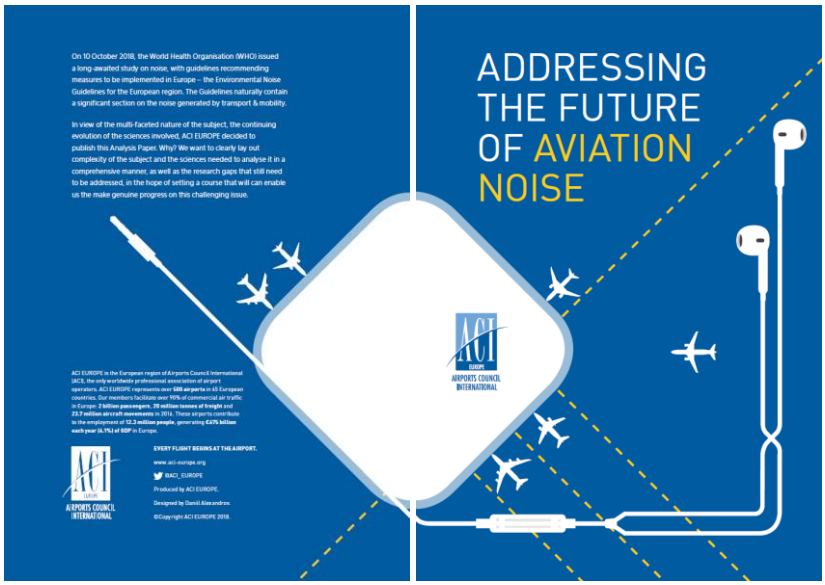

figure 7 Back and front cover of the ACI Europe report [12]

The main points brought forward by ACI are the following:

- ACI quotes results of the NORAH[3] research project. Regarding annoyance it was concluded that for equal noise levels, annoyance has increased over time since 2011, but it was stressed that there was no increase in the noise levels (i.e. L_{max} of flight events or the number of flights above a certain noise level). It was recognised that non-acoustic factors contribute significantly to annoyance.
- ACI stresses the importance of these non-acoustic factors and the fact that it is not clear how WHO considered them in their definition of the recommendations. ACI also points to the influence of airport changes and highlights a possible discrepancy with the third WHO recommendation to implement infrastructure changes: wouldn't such changes imply a risk of actually increasing annoyance, at least temporarily?
- ACI does not understand the recommendation to implement suitable changes in infrastructure, and they quote the WHO report stating elsewhere that source reduction measures are most effective, and that progress is actually made in research and development of quieter aircraft.
- The WHO rates the evidence quality for annoyance and sleep disturbance from aircraft noise as **moderate**, meaning 'further research is likely to have an important impact on the certainty of the effect estimate and is likely to change the estimate'. Yet, they recommend maximum levels of

[3] NORAH (Noise Related Annoyance, Cognition and Health) was a multidisciplinary research project from 2011 – 2014 regarding specifically the German Rhine/Main area around Frankfurt airport, plus interviews around other German airports. Some of the researchers were also involved in the development of the WHO guidelines.

45 dB L_{den} and 40 dB L_{night} and grade this as a **strong** recommendation. ACI disagrees with the decision of WHO to base strong recommendations of these levels on moderate quality evidence and finds this approach 'worrying'.

- ACI disputes the decision to base L_{night} recommendations on outdoor noise levels. Indoor levels corresponding to a 40 dB L_{night} will be very low, ranging from 15 to 30 dB depending on windows being open or closed, according to the WHO. ACI questions why indoor noise levels were not considered, as people sleep inside dwellings.
- Although WHO claims in the ENG report that it also considered other factors when developing recommendations, including the costs and benefits of noise mitigation, they also state that 'no comprehensive cost-benefit analysis for the WHO European Region yet exists'. WHO has later admitted that such an analysis has not been performed and these considerations have been based on expert judgment, see also the RIVM reports mentioned below. As examples, ACI quotes from the ENG that changes in flight paths 'in principle … do not involve any direct costs' and more generally that some noise mitigation interventions can be implemented 'at very low cost'. ACI claims that the WHO underestimates or overlooks the costs that are actually involved with noise interventions, such as extra fuel and extra costs for air traffic control and radar systems, for instance.
- The WHO guidelines, although aiming to improve public health, are based on a noise perspective only. Some noise mitigation measures lead to increased fuel burn, CO_2 and other polluting air emissions, for example, if flight paths are diverted to longer routes over less populated areas, or if a night flight ban leads to the necessity to use longer routes during the busy daytime.
- ACI explains that implementation of 45 dB L_{den} / 40 dB L_{night} would have unrealistic consequences for mobility around some airports. As an example, they state that 45 dB L_{den} at 3 km from the airport corresponds to only 10 daily flight movements of a Boeing B737-800. Also, Frankfurt airport would have to reduce the number of flights by 98% to reach the recommended levels in all surrounding housing areas. Also, the 45 dB L_{den} / 40 dB L_{night} contours for Frankfurt go as far as 70 km from the airport, or 40 km in the case of Madrid Barajas Airport.

<u>RIVM reports on WHO Guidelines</u>

In June 2020, RIVM published a report on the consequences of the new WHO Guidelines for Dutch policy for environmental noise [13], including aviation noise. A supplementary report was published in 2021 with further background information, including a comparison of current and previous national and international exposure-response functions for annoyance and sleep disturbance [14].

With regards to the WHO guidelines and systematic reviews for aviation noise, the RIVM reports show the following:

- Several researchers in the past have suggested that aircraft noise annoyance, at the same noise levels, has increased over time. Janssen and Vos in 2009 [15] have derived a new ERF based on seven then recent studies, which showed a significant increase with respect to the reference Miedema and Oudshoorn ERF, supporting the suggested increase[4].
- Several ERFs for annoyance are available for Schiphol and regional Dutch airports, showing significant differences. For Schiphol, the ERF from 2002 at low levels, up to 53 dB L_{den}, is quite similar to the 2018 WHO ERF by Guski et al. and to the 2009 Janssen and Vos ERF, see figure 8. The 2016 Dutch curve is based on the 'GGD Gezondheidsmonitor' for all Dutch airports, using the same survey questions and noise levels range as used for the WHO systematic review evidence base. There is a difference in age range: the GGD study does not include people over

[4] This 2009 ERF is also referenced in the WHO guidelines, see the green line with confidence bounds in figure 1.

65 years, whereas some of the WHO studies include all ages above 18, and some mention specifically that 90+ ages are included. Above 53 dB, all Dutch ERFs show higher annoyance than the international curves.

- RIVM mentions methodological differences as well as non-acoustic factors as explanations for the observed differences. Regarding methodological differences, it is stated that the increased use of 11-point numeric scales instead of 5-point label scales in recent annoyance surveys has been associated with higher annoyance ratings. The effect was investigated by Guski et al. in their systematic review. It was found to be not significant and results were inconsistent, e.g. for road and rail noise the effect was reversed. RIVM also mentions the limited age range for several surveys used by WHO, as was mentioned by Gjestland. Regarding non-acoustic factors RIVM mentions the influence of airport changes, as highlighted by others. They also state that it is unknown which non-acoustic factors best explain changes between Dutch airports.
- The RIVM report briefly mentions the critical reviews published by Gjestland and the ERF he proposes based on additional surveys. The fact that Gjestland only provides his ERF in a graph, not in a formula, is considered unfortunate.
- The reports reference a recent scoping review by Van Kamp et al., RIVM researchers also involved in the WHO's work, confirming that 13 eligible new studies on aircraft noise annoyance have been published since 2015 [17]. They state that 'differences in effect due to the in- or exclusion of different types of study should also be discussed in more detail' [18], referencing the Gjestland vs. Guski et al. discussion described above, and that 'an update of the review and its consequences for the current Guideline values for air traffic noise and annoyance need close examination of which studies should be included […]' [17].
- For sleep disturbance, RIVM explains that standardised methods, similar to the ICBEN and ISO surveys for annoyance, are not available for self-reported sleep disturbance. It is highlighted that the association between night noise levels and sleep disturbance was only significant if the survey questions specifically mentioned noise. RIVM shows that the 2018 ERF for sleep disturbance derived by WHO is quite comparable to the ERF found for Schiphol in 2002; more comparable than the earlier EU standard curve.
- Regarding cardiovascular disease, RIVM highlights that the recent WHO work shows indications that such bodily effects are onset from 53 dB L_{den}, whereas earlier these were only associated with much higher levels around 70 dB. This is only mentioned for road and rail noise, however, not for aviation.
- RIVM confirms that a good cost-benefit analysis for measures to reduce exposure to noise and the consequential negative impacts for the European region is not available [13]. Nevertheless, several measures are considered cost-effective by the WHO, including low noise tyres and pavements, noise barriers and façade insulation.

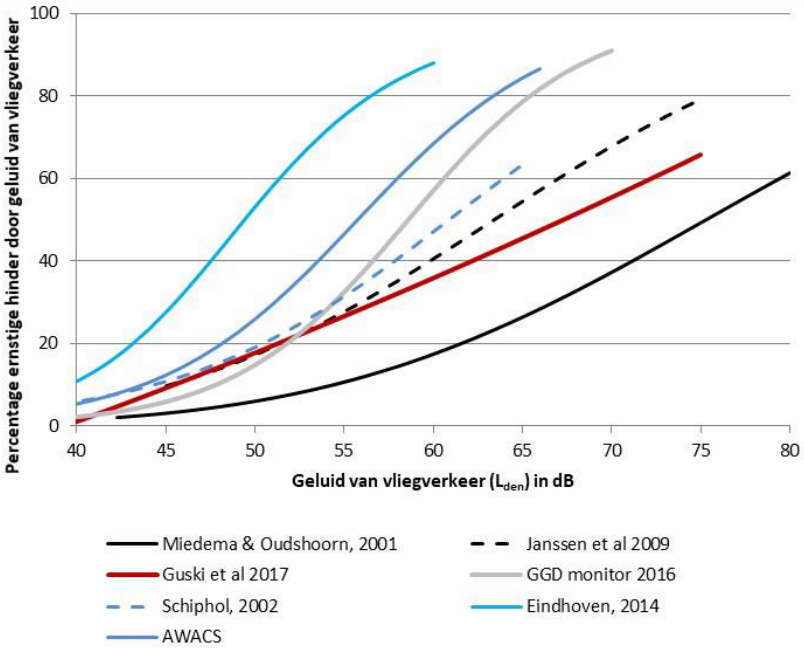

figure 8 Exposure-response functions for aircraft noise annoyance from international studies (Miedema, Janssen, Guski) and several Dutch studies: the national GGD monitor 2016, regional curves for Schiphol and Eindhoven, and the AWACS aircraft around Geilenkirchen

Conclusions

In this report, we have summarised some recent evaluations and criticism towards the WHO environmental noise guidelines and their recommendations for aircraft noise. Most of the criticism focuses on the exposure-response functions (ERF) found for annoyance, and consequently the recommended value of maximum 45 dB Lden. There is no judge to decide if the scientific claim of this maximum is right or wrong. However, like in court, the likelihood that a (scientific) claim is true should be based on evidence. Both quantity and quality of evidence, either pro or contra, determines the strength of any claim. So, 'judging' from the various publications, we conclude that there are points of criticism towards the WHO's guidelines that policy makers should be aware of and should take into account when considering to implement the recommendations. More research may be needed to clarify these points and the consequences for the WHO-recommended noise levels.

- There is agreement that annoyance due to aircraft noise exposure is higher at airports that have recently undergone changes (expansions, abruptly increased traffic), so-called 'high rate change' airports, with respect to other 'low rate change' airports. The new WHO ERF is based on a mix of both types of airports, but the amount of high rate change airports seems to be relatively high. This has a rising effect on the ERF, leading to a lower value for the recommended maximum Lden. Future work may be needed on this issue: In a recent White Paper [13] by a collaboration of several authors including Gjestland and Guski, it is concluded that 'In future work, existing exposure-response functions should be updated and diversified to account for various acoustic and non-acoustic factors. The difference between a high rate change and a low rate change

situation seems to be particularly important.' Other factors such as the age range and range of noise levels should also be regarded.

- Researchers agree that the ERF strongly depends on the exact selection of studies included to calculate it. On one hand, critics claim that some of the studies included by the WHO should not have been, due to various reasons, and that the actual selection by the WHO may have biased the ERF in the upward direction. Also, claims are that other studies have not been included that should have. And, in the meantime, several more studies have become available that would also change the ERF. As suggested by RIVM, this deserves an update of the WHO systematic review for aircraft noise annoyance and other impacts.
- All in all, the new ERF by the WHO is considerably higher than the earlier 'EU standard' ERF established by Miedema & Oudshoorn (2001), and consequently the recommended 45 dB Lden is much lower than it would be according to the older curve. We conclude that researchers do not agree whether this represents the fact that people are more annoyed by aircraft noise at the same level than they used to be, or that it is caused by different research methods and more high rate change airports being considered. We do observe that at least to some extent annoyance seems to have increased, as even for low rate change airports more recent data reveal a higher ERF than older data. The 2009 ERF by Janssen and Vos [15], also referenced in the WHO Guidelines and in the RIVM reports, supports such a significant increase but is not given any attention by the critics. We also observe that the shift in ERFs is challenged by questioning the validity of the new WHO ERF, while there seems to be little or no critique towards the earlier Miedema & Oudshoorn results.
- Generally, the WHO evidence base underpinning their aircraft noise recommendations show considerable heterogeneity, possibly caused by severe differences in non-acoustic factors between the various studies. Critique is that the resulting quality of the total dataset is only 'moderate' and that this should not lead to 'strong' recommendations.
- Specifically for Dutch airports, RIVM concludes that the ERFs found from national data are actually show quite similar annoyance as the new WHO ERF at low levels, up to 53 dB L_{den}, and even higher annoyance at high levels. Significant differences are found between various regional Dutch airports, attributed also to non-acoustic factors. The ERF for sleep disturbance around Schiphol is also shown to be comparable to the new WHO ERF. Also, RIVM highlights that cardiovascular diseases, at least for road noise, are now associated with lower levels than before, and this should be investigated further also for aircraft noise.
- The WHO claim that the recommendation of 45 dB Lden and 40 dB Lnight is formulated including considerations of costs and benefits for interventions seems unsubstantiated, as any cost/benefit considerations are based on limited amounts of expert judgment, not substantiated by thorough research or objective evidence. We conclude that these recommendations should be regarded as being formulated from a health perspective only, and that balancing these against economic and other societal interests is left up to the policy makers.

References

[1] WHO/Europe (2018). *Environmental Noise Guidelines for the European Region*, World Health Organization Regional Office for Europe, ISBN 978 92 890 5356 3 - link

[2] Berglund, Birgitta, Lindvall, Thomas, Schwela, Dietrich H & World Health Organization. Occupational and Environmental Health Team. (1999). *Guidelines for Community Noise.* World Health Organization - link

[3] WHO/Europe (2009). *Night Noise Guidelines for Europe*, World Health Organization Regional Office for Europe, ISBN 978 92 890 4173 7 - link

[4] Guski R, Schreckenberg D, Schuemer R. (2017). WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Annoyance. *International Journal of Environmental Research and Public Health*. 2017; 14(12):1539 - link, with Supplementary Materials - link

[5] Basner M, McGuire S (2018). WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Effects on Sleep. *International Journal of Environmental Research and Public Health*. 2018; 15(3):519 - link, with Supplementary Materials - link

[6] Gjestland T (2018). A Systematic Review of the Basis for WHO's New Recommendation for Limiting Aircraft Noise Annoyance. *International Journal of Environmental Research and Public Health*. 2018; 15(12):2717 - link

[7] Guski R, Schreckenberg D, Schuemer R, Brink M, Stansfeld SA (2019). 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. *International Journal of Environmental Research and Public Health*. 2019; 16(7):1088 - link

[8] Gjestland T (2019). Reply 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. *International Journal of Environmental Research and Public Health*. 2019; 16(7):1105 - link

[9] Gjestland T (2020). Recent World Health Organization regulatory recommendations are not supported by existing evidence, *The Journal of the Acoustical Society of America*. 2020; 148(2):511-517 - link

[10] Brink M (2020). Comment on "Recent World Health Organization regulatory recommendations are not supported by existing evidence" [J. Acoust. Soc. Am. 148, 511–517 (2020)], *The Journal of the Acoustical Society of America*. 2020; 148(6):3397-3398 - link

[11] Gjestland T (2021). Response to "Comment on 'Recent World Health Organization regulatory recommendations are not supported by existing evidence' " [J. Acoust. Soc. Am. 148, 3397–3398 (2020)], *The Journal of the Acoustical Society of America*. 2021; 149(3):1967-1968 - link

[12] ACI Europe (2018). *Addressing the Future of Aviation Noise*, Airports Council International Europe, retrieved from https://www.aeroport.fr/uploads/documents/voir-l%27etude-sur-le-bruit.pdf?v12

[13] Welkers D, Van Kempen E, Helder R, Verheijen E, Van Poll R (2020). *Schonis motion and WHO guidelines for environmental noise (2018) : The end justifies the means*, RIVM Report 2019-0227 (in Dutch) - link

[14] Van Kempen E (2021). *New WHO environmental noise guidelines. Additional health-based analyses*, RIVM Report 2020-0148 (in Dutch) - link

[15] Janssen SA, Vos H (2009). *A comparison of recent surveys to aircraft noise exposure-response relationships*, TNO report TNO-034-DTM-2009-01799

[16] Sparrow V, Gjestland T, Guski R, et al. (2019). *Aviation Noise Impacts White Paper*, prepared for ICAO - link

[17] Van Kamp I, Van Kempen E, Simon S, Baliatsas C (2019). *Review of Evidence Relating to Environmental Noise Exposure and Annoyance, Sleep Disturbance, Cardio-Vascular and Metabolic Health Outcomes in the Context of ICGB(N)*, RIVM Report 2019-0088, prepared for DEFRA UK - link

[18] Van Kamp I, Simon S, Notley H, Baliatsas C, van Kempen E (2020). Evidence Relating to Environmental Noise Exposure and Annoyance, Sleep Disturbance, Cardio-Vascular and Metabolic Health Outcomes in the Context of IGCB (N): A Scoping Review of New

Evidence. *International Journal of Environmental Research and Public Health*. 2020; 17(9):3016 - link