**Environmental noise and health in Low-Middle-Income-Countries: a systematic review of epidemiological evidence**

Yingxin Chen1, 2, Anna L Hansell1, 2, Sierra Clark3, Yutong Samuel Cai1, 2,

1: Centre for Environmental Health and Sustainability, University of Leicester, Leicester, UK

2: The National Institute of Health Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposure and Health at the University of Leicester, Leicester, UK

3: Noise and Public Health, Radiation Chemical and Environmental Hazards, Science Group, UK Health Security Agency

Correspondence to: Ms. Yingxin Chen, Centre for Environmental Health and Sustainability, University of Leicester, Leicester, University Road, LE1 7RH, UK. Email: [yc310@leicester.ac.uk](mailto:yc310@leicester.ac.uk)

# Supplementary 1 List of countries

Countries in alphabetical order:

|  |  |  |  |
| --- | --- | --- | --- |
| Afghanistan | Dominican Republic | Liberia | Senegal |
| Albania | Ecuador | Libya | Serbia |
| Algeria | Egypt, Arab Rep. | Madagascar | Sierra Leone |
| American Samoa | El Salvador | Malawi | Solomon Islands |
| Angola | Equatorial Guinea | Malaysia | Somalia |
| Argentina | Eritrea | Maldives | South Africa |
| Armenia | Eswatini | Mali | South Sudan |
| Azerbaijan | Ethiopia | Marshall Islands | Sri Lanka |
| Bangladesh | Fiji | Mauritania | St. Lucia |
| Belarus | Gabon | Mauritius | St. Vincent and the Grenadines |
| Belize | Gambia, The | Mexico | Sudan |
| Benin | Georgia | Micronesia, Fed. Sts. | Suriname |
| Bhutan | Ghana | Moldova | ﻿﻿Syrian Arab Republic |
| Bolivia | Grenada | Mongolia | Tajikistan |
| Bosnia and Herzegovina | Guatemala | Montenegro | Tanzania |
| Botswana | Guinea | Morocco | Thailand |
| Brazil | Guinea-Bissau | Mozambique | Timor-Leste |
| Bulgaria | Guyana | Myanmar | Togo |
| Burkina Faso | Haiti | Namibia | Tonga |
| Burundi | Honduras | Nepal | Tunisia |
| Cabo Verde | India | Nicaragua | Turkey |
| Cambodia | Indonesia | Niger | Turkmenistan |
| Cameroon | Iran, Islamic Rep | Nigeria | Tuvalu |
| Central African Republic | Iraq | North Macedonia | Uganda |
| Chad | ﻿﻿Jamaica | Pakistan | Ukraine |
| China | Jordan | Panama | Uzbekistan |
| Colombia | Kazakhstan | Papua New Guinea | Vanuatu |
| Comoros | Kenya | Paraguay | Vietnam |
| Congo, Dem. Rep | Kiribati | Peru | West Bank and Gaza |
| Congo, Rep. | Korea, Dem. People's Rep | Philippines | Yemen, Rep. |
| Costa Rica | Kosovo | Romania | Zambia |
| Côte d'Ivoire | Kyrgyz Republic | Russian Federation | Zimbabwe |
| Cuba | Lao PDR | Rwanda |  |
| Djibouti | Lebanon | Samoa |  |
| Dominica | Lesotho | São Tomé and Principe |  |

Source: World Bank (2022) World Bank Country and Lending Groups [online] [available at: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (assessed at 10 November 2021)

# Supplementary 2

## 2.1 Medline search

Database: All Ovid journals (the University does not have subscriptions to all these), University of Leicester subscribed Ovid journals, Books@Ovid <November 01, 2021>, Health and Psychosocial Instruments <1985 to October 2021>, Medline 1946-date (includes epub ahead of print and in process citations. Updated daily), Maternity & Infant Care Database (MIDIRS) <1971 to November 02, 2021>

Search Strategy:

--------------------------------------------------------------------------------

1 "epidemiologic".af. (408616)

2 "epidemiological".af. (666827)

3 "epidemiology".af. (3374782)

4 "health".af. (10069443)

5 1 or 2 or 3 or 4 (11676875)

6 "South America".af. (74126)

7 "Oceania".af. (22633)

8 "North America".af. (367864)

9 "Europe".af. (715879)

10 "Central America".af. (20836)

11 "Caribbean".af. (61376)

12 "Asia".af. (376597)

13 "Africa".af. (503352)

14 "low income country".af. (3687)

15 "LMIC".af. (7686)

16 "low middle income country".af. (2526)

17 "poverty countries".af. (106)

18 "poverty cities".af. (35)

19 "poverty regions".af. (45)

20 "poverty areas".af. (7419)

21 "developing regions".af. (3711)

22 "developing areas".af. (2314)

23 "developing countries".af. (282231)

24 "Afghanistan".af. (28493)

25 "Albania".af. (6914)

26 "Algeria".af. (19437)

27 "Angola".af. (5703)

28 "American Samoa".af. (1510)

29 "Argentina".af. (217011)

30 "Armenia".af. (10399)

31 "Azerbaijan".af. (6840)

32 "Bangladesh".af. (60090)

33 "Belarus".af. (17217)

34 "Belize".af. (3277)

35 "Benin".af. (14439)

36 "Bhutan".af. (3228)

37 "Bolivia".af. (13742)

38 "Bosnia and Herzegovina".af. (16429)

39 "Botswana".af. (13867)

40 "Brazil".af. (774945)

41 "Burkina Faso".af. (14963)

42 "Burundi".af. (3792)

43 "Bulgaria".af. (72475)

44 "Cabo Verde".af. (580)

45 "Cambodia".af. (16303)

46 "Cameroon".af. (25431)

47 "Central African Republic".af. (3813)

48 "Chad".af. (28230)

49 "China".af. (2922308)

50 "Colombia".af. (77485)

51 "Comoros".af. (1281)

52 "Congo, Dem. Rep.".af. (11)

53 "Congo,Rep.".af. (5)

54 "Costa Rica".af. (25612)

55 "Côte d'Ivoire".af. (13)

56 "Cuba".af. (31724)

57 "Djibouti".af. (1531)

58 "Dominica".af. (1950)

59 "Dominican Republic".af. (9788)

60 "Ecuador".af. (23295)

61 "Egypt".af. (207466)

62 "Egypt, Arab Rep.".af. (12)

63 "El Salvador".af. (6794)

64 "Equatorial Guinea".af. (1458)

65 "Eritrea".af. (3861)

66 "Eswatini".af. (1062)

67 "Ethiopia".af. (52400)

68 "Fiji".af. (12742)

69 "Gabon".af. (7095)

70 "Gambia".af. (13761)

71 "Georgia".af. (325821)

72 "Ghana".af. (40837)

73 "Grenada".af. (3778)

74 "Guatemala".af. (16682)

75 "Guinea".af. (299217)

76 "Guinea-Bissau".af. (4699)

77 "Guyana".af. (4259)

78 "Haiti".af. (13803)

79 "Honduras".af. (8427)

80 "India".af. (1147092)

81 "Indonesia".af. (66474)

82 "Iran, Islamic Rep".af. (9)

83 "Iraq".af. (43011)

84 "Jamaica".af. (24308)

85 "Jordan".af. (197277)

86 "Kazakhstan".af. (12257)

87 "Kenya".af. (86886)

88 "Kiribati".af. (776)

89 "Kosovo".af. (4633)

90 "Korea, Dem. People's Rep".af. (4)

91 "Kyrgyz Republic".af. (525)

92 "Lao PDR".af. (2490)

93 "Lebanon".af. (76476)

94 "Lesotho".af. (3710)

95 "Liberia".af. (6218)

96 "Libya".af. (7086)

97 "Madagascar".af. (13764)

98 "Malawi".af. (32643)

99 "Malaysia".af. (123872)

100 "Maldives".af. (1631)

101 "Mali".af. (23176)

102 "Marshall Islands".af. (1230)

103 "Mauritania".af. (2271)

104 "Mauritius".af. (5390)

105 "Mexico".af. (382156)

106 "Micronesia, Fed. Sts.".af. (4)

107 "Moldova".af. (4975)

108 "Mongolia".af. (23572)

109 "Montenegro".af. (17623)

110 "Morocco".af. (35924)

111 "Mozambique".af. (14857)

112 "Myanmar".af. (11736)

113 "Namibia".af. (7425)

114 "Nepal".af. (42495)

115 "Nicaragua".af. (8484)

116 "Niger".af. (33540)

117 "Nigeria".af. (107978)

118 "North Macedonia".af. (4327)

119 "Pakistan".af. (127577)

120 "Panama".af. (20279)

121 "Papua New Guinea".af. (15552)

122 "Paraguay".af. (9066)

123 "Peru".af. (53234)

124 "Philippines".af. (52739)

125 "Romania".af. (113960)

126 "Russian Federation".af. (46571)

127 "Rwanda".af. (15734)

128 "Samoa".af. (4368)

129 "São Tomé and Príncipe".af. (0)

130 "Senegal".af. (24465)

131 "Serbia".af. (76581)

132 "Sierra Leone".af. (9001)

133 "Solomon Islands".af. (2879)

134 "Somalia".af. (7725)

135 "South Africa".af. (278313)

136 "South Sudan".af. (2129)

137 "Sri Lanka".af. (30909)

138 "St. Lucia".af. (22315)

139 "St. Vincent and the Grenadines".af. (231)

140 "Sudan".af. (33187)

141 "Suriname".af. (2920)

142 "Syrian Arab Republic".af. (816)

143 "Tajikistan".af. (2790)

144 "Tanzania".af. (52405)

145 "Thailand".af. (197157)

146 "Timor-Leste".af. (1357)

147 "Togo".af. (9513)

148 "Thailand".af. (197157)

149 "Tonga".af. (2353)

150 "Tunisia".af. (47820)

151 "Turkey".af. (490471)

152 "Turkmenistan".af. (1656)

153 "Tuvalu".af. (363)

154 "Uganda".af. (71334)

155 "Ukraine".af. (126815)

156 "Uzbekistan".af. (8255)

157 "Vanuatu".af. (2211)

158 "Vietnam".af. (75734)

159 "West Bank and Gaza".af. (650)

160 "Yemen, Rep.".af. (5)

161 "Zambia".af. (24435)

162 "Zimbabwe".af. (30755)

163 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114 or 115 or 116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124 or 125 or 126 or 127 or 128 or 129 or 130 or 131 or 132 or 133 or 134 or 135 or 136 or 137 or 138 or 139 or 140 or 141 or 142 or 143 or 144 or 145 or 146 or 147 or 148 or 149 or 150 or 151 or 152 or 153 or 154 or 155 or 156 or 157 or 158 or 159 or 160 or 161 or 162 (9275312)

164 ((rail\* or aircraft or airport\* or road\* or traffic\* or vehicle\* or automobile\* or turbine) adj5 noise).tw. (5273)

165 exp \*Transportation/ (48682)

166 Aircraft/ or Airports/ or Railway/ or motor vehicles/ or turbine.mp. [mp=ti, ab, tx, ct, bt, ac, de, md, sd, so, ot, nm, hw, fx, kf, ox, px, rx, ui, sy] (17321)

167 \*noise/ (12947)

168 Noise, transportation/ (1555)

169 (164 or 165 or 166) and (164 or 167 or 168) (6026)

170 exp Mortality/ (409498)

171 mortality.tw. (2268332)

172 (admission adj3 hospital\*).tw. (222980)

173 hospital\*.tw. (4670032)

174 emergency.af. (1252798)

175 morbidity.af. (1427727)

176 lung diseases, obstructive/ (18244)

177 exp pulmonary disease, chronic obstructive/ (61059)

178 (copd or coad or cobd or aecb).tw. (123985)

179 asthma.af. (503770)

180 cancer.af. (4729079)

181 carcino\*.af. (1988047)

182 malignan\*.af. (1545140)

183 neoplas\*.af. (3615769)

184 tumor.af. (3065427)

185 depressi\*.af. (1335442)

186 exp Alzheimer Disease/ (104116)

187 exp Dementia, Vascular/ (7032)

188 exp Dementia, Multi-Infarct/ (1089)

189 exp Cognition Disorders/ or exp Dementia/ (253096)

190 dement\*.mp. (372210)

191 alzheimer\*.mp. (372329)

192 exp cardiovascular disease/ or exp vascular disease/ or exp isch?emic heart disease/ (2549123)

193 (isch?emic heart disease\* or coronary heart disease\* or angina pectoris or myocard\* infarct\* or heart disease).mp. [mp=ti, ab, tx, ct, bt, ac, de, md, sd, so, ot, nm, hw, fx, kf, ox, px, rx, ui, sy] (1140369)

194 exp stroke/ (151128)

195 exp hypertension/ or exp blood pressure/ (517585)

196 exp diabetes mellitus/ (460396)

197 exp obesity/ or exp overweight/ or exp body mass index/ (323060)

198 (diabetes or obesit\* or overweight or bmi or body mass index).tw. (2353323)

199 ((Pre term or preterm or premature or early or immature) adj5 (labo?r or birth$ or childbirth$ or deliver$ or partu$ or ruptur$)).tw. (270960)

200 exp sleep/ (87334)

201 (insomnia or arousal or awak\* or tired$ or fatigue).tw. (750641)

202 (perceived well-being or quality of life or mental health or mental well-being or anxiety or psychological symptom or emotional disorder or cortisol).tw. (1928226)

203 exp depression/ (134561)

204 exp anxiety/ (97544)

205 (annoyance or noise annoyance).tw. (8057)

206 (attention or concentrat\* or intellig\* or memor\* or comprehens\*).tw. (7025021)

207 (cognitive impairment or cognitive fucntion or cognitive development).tw. (241022)

208 (executive function or reasoning or task flexibility or problem solving or hyperactivity).tw. (409948)

209 or/170-208 (21488657)

210 5 and 163 and 169 and 209 (1454)

211 noise.ti. (38565)

212 210 and 211 (535)

213 212 not occupat\*.ti. (514)

214 limit 213 to humans [Limit not valid in Journals@Ovid,Your Journals@Ovid,Books@Ovid,HAPI,MWIC; records were retained] (473)

215 limit 214 to yr="2009 -Current" (399)

## 2.2 Web of Science search

|  |  |  |
| --- | --- | --- |
| search | query | Result |
| #1 | ALL=(epidemiologic or epidemiological or epidemiology or health) | 9,292,983 |
| #2 | AB=(South America or Oceania or North America or Europe or Central America or Caribbean or Asia or Africa or low income country or LMIC or low middle income country or poverty countries or poverty cities or poverty regions or poverty areas or developing regions or developing areas or developing countries or Afghanistan or Albania or Algeria or Angola or American Samoa or Argentina or Armenia or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bhutan or Bolivia or Bosnia and Herzegovina or Botswana or Brazil or Burkina Faso or Burundi or Bulgaria or Cabo Verde or Cambodia or Cameroon or Central African Republic or Chad or China or Colombia or Comoros or Congo, Dem. Rep. or Congo,Rep. or Costa Rica or Côte d'Ivoire or Cuba or Djibouti or Dominica or Dominican Republic or Ecuador or Egypt or Egypt, Arab Rep. or El Salvador or Equatorial Guinea or Eritrea or Eswatini or Ethiopia or Fiji or Gabon or Gambia or Georgia or Ghana or Grenada or Guatemala or Guinea or Guinea-Bissau or Guyana or Haiti or Honduras or India or Indonesia or Iran, Islamic Rep or Iraq or Jamaica or Jordan or Kazakhstan or Kenya or Kiribati or Kosovo or Korea, Dem. People's Rep or Kyrgyz Republic or Lao PDR or Lebanon or Lesotho or Liberia or Libya or Madagascar or Malawi or Malaysia or Maldives or Mali or Marshall Islands or Mauritania or Mauritius or Mexico or Micronesia, Fed. Sts. or Moldova or Mongolia or Montenegro or Morocco or Mozambique or Myanmar or Namibia or Nepal or Nicaragua or Niger or Nigeria or North Macedonia or Pakistan or Panama or Papua New Guinea or Paraguay or Peru or Philippines or Romania or Russian Federation or Rwanda or Samoa or São Tomé and Príncipe or Senegal or Serbia or Sierra Leone or Solomon Islands or Somalia or South Africa or South Sudan or Sri Lanka or St. Lucia or St. Vincent and the Grenadines or Sudan or Suriname or Syrian Arab Republic or Tajikistan or Tanzania or Thailand or Timor-Leste or Togo or Thailand or Tonga or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Vanuatu or Vietnam or West Bank and Gaza or Yemen, Rep. or Zambia or Zimbabwe) | 3,321,372 |
| #20 | AB=(executive function or reasoning or task flexibility or problem solving or hyperactivity) | 1,504,289 |
| #3 | TI=(South America or Oceania or North America or Europe or Central America or Caribbean or Asia or Africa or low income country or LMIC or low middle income country or poverty countries or poverty cities or poverty regions or poverty areas or developing regions or developing areas or developing countries or Afghanistan or Albania or Algeria or Angola or American Samoa or Argentina or Armenia or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bhutan or Bolivia or Bosnia and Herzegovina or Botswana or Brazil or Burkina Faso or Burundi or Bulgaria or Cabo Verde or Cambodia or Cameroon or Central African Republic or Chad or China or Colombia or Comoros or Congo, Dem. Rep. or Congo,Rep. or Costa Rica or Côte d'Ivoire or Cuba or Djibouti or Dominica or Dominican Republic or Ecuador or Egypt or Egypt, Arab Rep. or El Salvador or Equatorial Guinea or Eritrea or Eswatini or Ethiopia or Fiji or Gabon or Gambia or Georgia or Ghana or Grenada or Guatemala or Guinea or Guinea-Bissau or Guyana or Haiti or Honduras or India or Indonesia or Iran, Islamic Rep or Iraq or Jamaica or Jordan or Kazakhstan or Kenya or Kiribati or Kosovo or Korea, Dem. People's Rep or Kyrgyz Republic or Lao PDR or Lebanon or Lesotho or Liberia or Libya or Madagascar or Malawi or Malaysia or Maldives or Mali or Marshall Islands or Mauritania or Mauritius or Mexico or Micronesia, Fed. Sts. or Moldova or Mongolia or Montenegro or Morocco or Mozambique or Myanmar or Namibia or Nepal or Nicaragua or Niger or Nigeria or North Macedonia or Pakistan or Panama or Papua New Guinea or Paraguay or Peru or Philippines or Romania or Russian Federation or Rwanda or Samoa or São Tomé and Príncipe or Senegal or Serbia or Sierra Leone or Solomon Islands or Somalia or South Africa or South Sudan or Sri Lanka or St. Lucia or St. Vincent and the Grenadines or Sudan or Suriname or Syrian Arab Republic or Tajikistan or Tanzania or Thailand or Timor-Leste or Togo or Thailand or Tonga or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Vanuatu or Vietnam or West Bank and Gaza or Yemen, Rep. or Zambia or Zimbabwe) | 2,060,961 |
| #4 | (#2) OR #3 | 4,307,823 |
| #5 | ALL=((rail\* or aircraft or airport\* or road\* or traffic\* or vehicle\* or automobile\* or turbine) adj noise) | 50,140 |
| #6 | ALL=(Transportation or Aircraft or Airports or Railway or motor vehicles or turbine) | 699,064 |
| #7 | (#5) OR #6 | 728,543 |
| #8 | AB=(mortality or (admission adj hospital\*) or hospital\* or emergency or morbidity) | 1,898,047 |
| #9 | AB=(lung diseases or (obstructive pulmonary disease) or (chronic obstructive) or copd or coad or cobd or aecb orasthma or cancer or carcino\* or malignan\* or neoplas\* or tumor) | 2,731,558 |
| #10 | AB=((exp Alzheimer Disease) or (exp Dementia adj Vascular) or (exp Dementia Multi-Infarct) or (exp Cognition Disorders) or (exp Dementia) or dement\* or alzheimer\*) | 203,361 |
| #11 | AB=((exp cardiovascular disease) or (exp vascular disease) or (exp isch?emic heart disease) or (isch?emic heart disease\*) or (coronary heart disease\*) or (angina pectoris) or (myocard\* infarct\*) or (heart disease) or (exp stroke) or (exp hypertension) or (exp blood pressure)) | 334,300 |
| #12 | AB=((exp diabetes mellitus) or (exp obesity) or (exp overweight) or (exp body mass index) or diabetes or obesit\* or overweight or bmi or body mass index) | 781,376 |
| #13 | AB=((Pre term or preterm or premature or early or immature) adj (labo?r or birth$ or childbirth$ or deliver$ or partu$ or ruptur$)) | 180,108 |
| #14 | AB=((exp sleep) or insomnia or arousal or awak\* or tired$ or fatigue) | 298,860 |
| #15 | AB=(perceived well-being or quality of life or mental health or mental well-being or anxiety or psychological symptom or emotional disorder or cortisol) | 833,390 |
| #16 | AB=((exp depression) or exp anxiety) | 240 |
| #17 | AB=(annoyance or (noise annoyance)) | 3,467 |
| #18 | AB=(attention or concentrat\* or intellig\* or memor\* or comprehens\*) | 5,541,402 |
| #19 | AB=(cognitive adj (impairment or fucntion or development)) | 161,002 |
| #20 | AB=(executive function or reasoning or task flexibility or problem solving or hyperactivity) | 1,504,289 |
| #21 | ((((((((((((#8) OR #9) OR #10) OR #11) OR #12) OR #13) OR #14) OR #15) OR #16) OR #17) OR #18) OR #19) OR #20 | 12,117,993 |
| #22 | (((#1) AND #4) AND #7) AND #21 | 5,274 |
| #23 | TI =(Noise not occupational) | 160,592 |
| #24 | ((TI =(Noise not occupational)) AND #22) AND #23 | 202 |
| #25 | (((TI =(Noise not occupational)) AND #22) AND #23) AND PY=(2009-2022) | 189 |

# Supplementary 3 Risk of bias assessment criteria

Most of the text from these instructions and criteria for judging risk of bias has been adopted verbatim or adapted from one of the latest *Navigation Guide* (Woodruff and Sutton 2014) systematic reviews (Lamet al. 2016).

Instructions:

* + *Please evaluate each individual study for the following nine risk of bias domains. Please answer “low risk,” “probably low risk,” “probably high risk,” “high risk,” or “not applicable” and provide details/justification for your rating. If there is empirical evidence or other knowledge that informs the direction of bias, please include this in your answer as well; however, if there is not enough information to do so please do not guess at the direction of bias.*
  + *Additionally, please note that some internal validity issues could potentially be appropriately captured in several different risk of bias considerations. In this situation, please select the single most appropriate domain to evaluate this potential bias, to avoid double-counting the same internal validity concern.*

## Are the study groups at risk of not representing their source populations in a manner that might introduce selection bias?

The source population is viewed as the population for which study investigators are targeting their study question of interest. Examples of considerations for this risk of bias domain include: 1) level of detail reported for participant inclusion/exclusion (including details from previously published papers referenced in the article for an existing cohort); 2) participation rates and whether this differed by exposure or outcome group; 3) attrition rates and reasons; and 4) comparisons of study characteristics between the study population and full cohort.

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”): EITHER:

* 1. The descriptions of the source population, inclusion/exclusion criteria, recruitment and enrollment procedures, participation and follow-up rates were sufficiently detailed, and adequate data were supplied on the distribution of relevant study sample and population characteristics to support the assertion that risk of selection effects was minimal.

OR

* 1. Although the descriptions and/or data as indicated in “a” above suggested the potential for selection effects, adequate support was given indicating that potential selection effects were *not* differential across both exposure and outcome.

OR

* 1. Although the descriptions and/or data as indicated in “a” above suggested the potential for selection effects and there was no support indicating that potential selection effects were *not* differential across both exposure and outcome, selection factors appeared to be well-understood, were measured in the data set, and appropriate adjustment post hoc techniques were used to control for selection bias.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about participant selection to permit a judgment of low risk of bias, but there is indirect evidence which suggests that inclusion/exclusion criteria, recruitment and enrollment procedures, and participation and follow-up rates were consistent across groups as described by the criteria for a judgment of low risk of bias.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about participant selection to permit a judgment of high risk of bias, but there is indirect evidence which suggests that inclusion/exclusion criteria, recruitment and enrollment procedures, and participation and follow-up rates were inconsistent across groups, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

1. There were indications from descriptions of the source population, inclusion/exclusion criteria, recruitment and enrolment procedures, participation and follow-up rates, or data on the distribution of relevant study sample and population characteristics that risk of selection effects were substantial; and
2. There was no support to indicate that potential selection effects were *not* differential across both exposure and outcome; and
3. Adjustment post hoc techniques were not used to control for selection bias.

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that participant selection is not an element of study design capable of introducing risk of bias in the study.

## Was knowledge of the group assignments inadequately prevented (i.e. blinded or masked) during the study, potentially leading to subjective measurement of either exposure or outcome?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”): Any of the following:

* No blinding, but the review authors judge that the outcome measures as well as the exposure measures are not likely to be influenced by lack of blinding (such as differential outcome assessment where the outcome is assessed using different measurement or estimation metrics across the exposure groups, or differential exposure assessment where exposure is assessed using different measurement or estimation metrics across the diagnostic or outcome groups); or
* Blinding of key study personnel was ensured, and it is unlikely that the blinding could have been broken; or
* Some key study personnel were not blinded, but exposure and outcome assessment was blinded and the non-blinding of others is unlikely to introduce bias.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about blinding to permit a judgment of low risk of bias, but there is indirect evidence which suggests the study was adequately blinded, as described by the criteria for a judgment of low risk of bias. For example, investigators were effectively blinded to the exposure and/or outcome groups if the exposure was measured by a separate entity and the outcome was obtained from a hospital record.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about blinding to permit a judgment of high risk of bias, but there is indirect evidence which suggests the study was not adequately blinded, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”): Any of the following:

* No blinding or incomplete blinding, and the outcome measures or exposure measures is likely to be influenced by lack of blinding (i.e. differential outcome or exposure assessment); or
* Blinding of key study personnel attempted, but likely that the blinding could have been broken so as to introduce bias; or
* Some key study personnel were not blinded, and the non-blinding of others was likely to introduce bias.

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that blinding is not an element of study design capable of introducing risk of bias in the study.

## Were exposure assessment methods lacking accuracy?

*The following list of considerations represents a collection of factors proposed by experts in various fields that may potentially influence the internal validity of the exposure assessment in a systematic manner (not those that may randomly affect overall study results).* ***These should be interpreted only as suggested considerations, and should not be viewed as scoring or a checklist.***

**List of Considerations**:

*Possible sources of exposure assessment metrics:*

1. *Official Records (Ministry of Health, Ministry of Labour or other official sources)*
2. *Organization*
3. *Self-reported*
4. *Combination of the above options*

*For each, overall considerations include:*

1. *What is the quality of the source of the metric being used?*
2. *Is the exposure measured in the study a surrogate for the exposure?*
3. *What was the temporal coverage (i.e. short or long-term exposure)?*
4. *Did the analysis account for prediction uncertainty?*
5. *How was missing data accounted for, and any data imputations incorporated?*
6. *Were sensitivity analyses performed?*

*In particular, for exposure assessment models:*

1. *Were the input data in the study suspected to systematically under- or over-estimate exposure?*
2. *What type of model was used?*
3. *What was geographic/spatial accuracy (county, census tract, organization, individual residence)?*
4. *What was the temporal specificity and variation?*
5. *What was the space-time coverage of the model?*
6. *Were time-activity patterns accounted for?*

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

The reviewers judge that there is low risk of exposure misclassification, i.e.:

* + There is high confidence in the accuracy of the exposure assessment methods, such as methods that have been tested for validity and reliability in measuring the targeted exposure.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about the exposure assessment methods to permit a judgment of low risk of bias, but there is indirect evidence that suggests that methods were robust, as described by the criteria for a judgment of low risk of bias.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

The reviewers judge that there is high risk of exposure misclassification and any one of the following:

* + - There is low confidence in the accuracy of the exposure assessment methods; or
    - Less-established or less direct exposure measurements are not validated and are suspected to introduce bias that impacts the outcome assessment; or
    - Uncertain how exposure information was obtained.

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that exposure assessment methods are not capable of introducing risk of bias in the study.

## Were outcome assessment methods lacking accuracy?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

The reviewers judge that there is low risk of outcome misclassification, i.e.:

* Outcomes were assessed and defined consistently across all study participants, using valid and reliable measures (all non-fatal or fatal stroke events with solid medical records). Note that all outcome assessment measures captured in the PECO statement are considered beforehand to be valid and reliable, unless other information provided within the study warrants a consideration otherwise; or
* Less-established or less direct outcome measurements are validated against well- established or direct methods; or
* Appropriate sensitivity analyses were conducted that suggest the influence of outcome misclassification would be minimal
* AND, if applicable, appropriate QA/QC for methods is described and is satisfactory.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about the outcome assessment methods to permit a judgment of low risk of bias, but there is indirect evidence which suggests that methods were robust, as described by the criteria for a judgment of low risk of bias. Appropriate QA/QC for methods are not described but the review authors judge that the outcome and the outcome assessment are objective and uniform across study groups.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about the outcome assessment methods to permit a judgment of high risk of bias, but there is indirect evidence which suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

The reviewers judge that there is high risk of outcome misclassification and any one of the following:

* There is low confidence in the accuracy of the outcome assessment methods; or
* Less-established or less direct outcome measurements are not validated and are suspected to introduce bias that impacts the outcome assessment
* Uncertain how outcome information was obtained

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that outcome assessment methods are not capable of introducing risk of bias in the study.

## Was potential confounding inadequately incorporated?

List of important potential confounders, collectively generated by review authors prior to the initiation of screening for studies based on expert opinion and knowledge gathered from the literature:

Tier I: Important confounders:

* Age, sex and socioeconomic position.

Tier II: Other potentially important confounders:

* None identified.

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

* The study appropriately assessed and accounted for (i.e. matched, stratified, excluded certain populations or statistically controlled for) all important confounders (Tier I) using appropriate statistical techniques, or reported that important confounders were evaluated and omitted because inclusion did not substantially affect the results. The determination of specific confounders may also be informed by, but not limited to, the studies included in the overall review,
* AND the study appropriately assessed and accounted for (i.e. matched, stratified, or statistically controlled for) other potentially important confounders relevant (Tier II) using appropriate statistical techniques, or reported that these confounders were evaluated and omitted because inclusion did not substantially affect the results,
* AND the important potential confounders were measured consistently across study groups using valid and reliable methods, or the influence of covariate measurement error was determined, through sensitivity analysis, to be minimal.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

* The study appropriately accounted for most but not all of the important confounders (Tier I) or used appropriate statistical techniques;
* AND some of the other potentially important confounders relevant (Tier II) using appropriate statistical techniques,
* OR reported that these confounders were evaluated and omitted because inclusion did not substantially affect the results;
* AND this is not expected to introduce substantial bias.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

* The study evaluated some but not all of the important confounders (Tier I),
* AND some but not all of the other potentially important confounders relevant (Tier II), OR used questionable statistical techniques for confounder adjustment;
* AND this is expected to introduce substantial bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

* The study did not account for or evaluate multiple important confounders (Tier I),
* AND did not account for or evaluate multiple other potentially important confounders relevant (Tier II),
* OR the important potential confounders were inappropriately measured and/or inappropriately analyzed across study groups.

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that outcome assessment methods are not capable of introducing risk of bias in the study.

## Were incomplete outcome data inadequately addressed?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

Participants were followed long enough to obtain outcome measurements

OR any one of the following:

* No missing outcome data; or
* Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to introduce bias); or
* Attrition or missing outcome data balanced in numbers across exposure groups, with similar reasons for missing data across groups; or
* For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a relevant impact on the exposure effect estimate; or
* For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a relevant impact on the observed effect size; or
* Missing data have been imputed using appropriate methods

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about incomplete outcome data to permit a judgment of low risk of bias, but there is indirect evidence which suggests incomplete outcome data was adequately addressed, as described by the criteria for a judgment of low risk of bias.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about incomplete outcome data to permit a judgment of high risk of bias, but there is indirect evidence which suggests incomplete outcome data was not adequately addressed, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

Participants were not followed long enough to obtain outcome measurements OR any one of the following:

* Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across exposure groups; or
* For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce biologically relevant bias in intervention effect estimate; or
* For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce biologically relevant bias in observed effect size; or
* Potentially inappropriate application of imputation.

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that incomplete outcome data is not capable of introducing risk of bias in the study.

## Does the study report appear to have selective outcome reporting?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

All of the study’s pre-specified (primary and secondary) outcomes outlined in the pre- published protocol or the published manuscript’s methods, abstract, and/or introduction section that are of interest in the review have been reported in the pre-specified way.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information about selective outcome reporting to permit a judgment of low risk of bias, but there is indirect evidence which suggests the study was free of selective reporting, as described by the criteria for a judgment of low risk of bias. This includes if a pre-published protocol is not available but the study’s pre-specified (primary and secondary) outcomes outlined in the published manuscript’s methods, abstract, and/or introduction section that are of interest in the review have been reported in the pre-specified way.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information about selective outcome reporting to permit a judgment of high risk of bias, but there is indirect evidence which suggests the study was not free of selective reporting, as described by the criteria for a judgment of high risk of bias. This includes if a pre-published protocol is not available and the study’s pre-specified (primary and secondary) outcomes outlined in the published manuscript’s methods, abstract, and/or introduction section that are of interest in the review have not been reported in the pre-specified way.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”): Any one of the following:

* Not all of the study’s pre-specified primary outcomes (as outlined in the pre-published protocol or published manuscript’s methods, abstract, and/or introduction) have been reported; or
* One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; or
* One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected effect); or
* One or more outcomes of interest are reported incompletely

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that selective outcome reporting is not capable of introducing risk of bias in the study.

## Did the study receive any support from a company, study author, or other entity having a financial interest in any of the exposures studied?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples include the following:

* Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations;
* Chemicals or other treatment used in study were purchased from a supplier;
* Company affiliated staff are not mentioned in the acknowledgements section;
* Authors were not employees of a company with a financial interest in the outcome of the study;
* Company with a financial interest in the outcome of the study was not involved in the design, conduct, analysis, or reporting of the study and authors had complete access to the data;
* Study authors make a claim denying conflicts of interest;
* Study authors are unaffiliated with companies with financial interest, and there is no reason to believe a conflict of interest exists;
* All study authors are affiliated with a government agency (are prohibited from involvement in projects for which there is a conflict of interest or an appearance of conflict of interest).

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information to permit a judgment of low risk of bias, but there is indirect evidence which suggests the study was free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of low risk of bias.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information to permit a judgment of high risk of bias, but there is indirect evidence which suggests the study was not free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

The study received support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples of support include:

* Research funds;
* Chemicals, equipment or testing provided at no cost;
* Writing services;
* Author/staff from study was employee or otherwise affiliated with company with financial interest;
* Company limited author access to the data;
* Company was involved in the design, conduct, analysis, or reporting of the study;
* Study authors claim a conflict of interest

Criteria for the judgment of NOT APPLICABLE (risk of bias domain is not applicable to study):

There is evidence that conflicts of interest are not capable of introducing risk of bias in the study.

## Did the study appear to have other problems that could put it at a risk of bias?

Criteria for a judgment of LOW risk of bias (i.e. answer: “No”):

The study appears to be free of other sources of bias.

Criteria for the judgment of PROBABLY LOW risk of bias (i.e. answer: “Probably No”):

There is insufficient information to permit a judgment of low risk of bias, but there is indirect evidence which suggests the study was free of other threats to validity.

Criteria for the judgment of PROBABLY HIGH risk of bias (i.e. answer: “Probably Yes”):

There is insufficient information to permit a judgment of high risk of bias, but there is indirect evidence which suggests the study was not free of other threats to validity, as described by the criteria for a judgment of high risk of bias.

Criteria for the judgment of HIGH risk of bias (i.e. answer: “Yes”):

There is at least one important risk of bias. For example, the study:

* Had a potential source of bias related to the specific study design used; or
* Stopped early due to some data-dependent process (including a formal-stopping rule); or
* The conduct of the study is affected by interim results (e.g. recruiting additional participants from a subgroup showing greater or lesser effect); or
* Has been claimed to have been fraudulent; or
* Had some other problem.

# Supplementary 4 Instructions for grading the quality of evidence

Most of the text from these instructions and criteria for judging risk of bias has been adopted verbatim or adapted from one of the latest Navigation Guide systematic reviews (Lamet al. 2016).

## 3.1 Grading Quality

Each of the categories to consider in downgrading or upgrading the evidence is described in detail below. Please record your results on the chart at the end of each category, including a brief explanation for your ratings.

## Downgrade Categories

### Category 1. Quality of Study Limitations (Risk of Bias)

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

The evidence from studies can be rated down if most of the relevant evidence comes from studies that suffer from a high risk of bias. Risk of bias is rated by outcome across studies. Study limitations for each outcome for individual studies and across studies are summarized in the heat maps. GRADE outlines the following principles for moving from risk of bias in individual studies to rating quality of evidence across studies.

1. In deciding on the overall quality of evidence, one does not average across studies (for instance if some studies have no serious limitations, some serious limitations, and some very serious limitations, one does not automatically rate quality down by one level because of an average rating of serious limitations). Rather, judicious consideration of the contribution of each study, with a general guide to focus on the high-quality studies is warranted.[[1]](#footnote-2)
2. This judicious consideration requires evaluating the extent to which each study contributes toward the estimate of magnitude of effect. The contribution that each study makes will usually reflect study sample size and number of outcome events. Larger studies with many events will contribute more, much larger studies with many more events will contribute much more.
3. One should be conservative in the judgment of rating down. That is, one should be confident that there is substantial risk of bias across most of the body of available evidence before one rates down for risk of bias.
4. The risk of bias should be considered in the context of other limitations. If, for instance, reviewers find themselves in a close-call situation with respect to two quality issues (risk of bias and, say, precision), GRADE suggests rating down for at least one of the two.
5. Notwithstanding the first four principles, reviewers will face close-call situations. You should acknowledge that you are in such a situation, make it explicit why you think this is the case, and make the reasons for your ultimate judgment apparent.

|  |  |  |
| --- | --- | --- |
| **Rating for Risk of Bias (Study Limitations)**  0 no change  -1 decrease quality 1 level  -2 decrease quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 2. Indirectness of Evidence

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

Quality of evidence (your confidence in estimates of effect) may decrease when substantial differences exist between the population, exposure, or outcomes measured in the research studies under consideration in the review.

Evidence is direct when it directly compares the exposures in which we are interested in the populations in which we are interested and measures outcomes important to the study question (in GRADE the outcomes must be important to patients).

Based on GRADE (Guyattet al. 2011), evidence can be indirect in one of three ways.[[2]](#footnote-3)

* 1. The population studied differs from the population of interest (the term applicability is often used for this form of indirectness). GRADE states that in general, one should not rate down for population differences unless one has compelling reason to think that the biology in the population of interest is so different than the population tested that the magnitude of effect will differ substantially. According to GRADE, most often, this will not be the case.
  2. The intervention (exposure) tested may differ from the exposure of interest, i.e., a difference in the chemical, route and/or dose. Decisions regarding indirectness of populations and exposure depend on an understanding of whether biological or social factors are sufficiently different that one might expect substantial differences in the magnitude of effect. GRADE also states, “As with all other aspects of rating quality of evidence, there is a continuum of similarity of the intervention that will require judgment. It is rare, and usually unnecessary, for the intended populations and interventions to be identical to those in the studies, and we should only rate down if the differences are considered sufficient to make a difference in outcome likely.”
  3. Outcomes may differ from those of primary interest; for instance, surrogate outcomes that are not themselves important, but measured in the presumption that changes in the surrogate reflect changes in an important outcome. The difference between desired and measured outcomes may relate to time frame. When there is a discrepancy between the time frame of measurement and that of interest, whether to rate down by one or two levels will depend on the magnitude of the discrepancy. Another source of indirectness related to measurement of outcomes is the use of substitute or surrogate endpoints in place of the exposed population’s important outcome of interest. In general, the use of a surrogate outcome requires rating down the quality of evidence by one, or even two, levels. Consideration of the biology, mechanism, and natural history of the disease can be helpful in making a decision about indirectness. Surrogates that are closer in the putative causal pathway to the adverse outcomes warrant rating down by only one level for indirectness. GRADE states that rarely, surrogates are sufficiently well established that one should choose not to rate down quality of evidence for indirectness. In general, evidence based on surrogate outcomes should usually trigger rating down, whereas the other types of indirectness will require a more considered judgment.

|  |  |  |
| --- | --- | --- |
| **Rating for Indirectness**  0 no change  -1 decrease quality 1 level  -2 decrease quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 3. Inconsistency of Evidence

Possible ratings: 0 = no change; -1 or -2 downgrade 1 or 2 levels

According to Cochrane, “when studies yield widely differing estimates of effect (heterogeneity or variability in results) investigators should look for robust explanations for that heterogeneity.

…When heterogeneity exists and affects the interpretation of results, but authors fail to identify a plausible explanation, the quality of the evidence decreases.”

Based on GRADE (Guyattet al. 2011), **a body of evidence is not rated up in quality if studies yield consistent results**, **but may be rated down in quality if inconsistent**. Their stated reason is that a consistent bias will lead to consistent, spurious findings.

GRADE suggests rating down the quality of evidence if large inconsistency (heterogeneity) in study results remains after exploration of a priori hypotheses that might explain heterogeneity. Judgment of the extent of heterogeneity is based on similarity of point estimates, extent of overlap of confidence intervals, and statistical criteria. GRADE’s recommendations refer to inconsistencies in effect size, specifically to relative measures (risk ratios and hazard ratios or odds ratios), not absolute measures.

Based on GRADE, reviewers should consider rating down for inconsistency when:

1. Point estimates vary widely across studies;
2. Confidence intervals (Andersonet al. 2011) show minimal or no overlap;
3. The statistical test for heterogeneity-which tests the null hypothesis that all studies in a meta- analysis have the same underlying magnitude of effect- shows a low P-value;
4. The I*2* -which quantifies the proportion of the variation in point estimates due to among-study differences-is large. (I.e., the I2 index quantifies the degree of heterogeneity in a meta-analysis).

GRADE states that inconsistency is important only when it reduces confidence in results in relation to a particular decision. Even when inconsistency is large, it may not reduce confidence in results regarding a particular decision. For example, studies that are inconsistent related to the magnitude of a beneficial or harmful effect (but are in the same direction) would not be rated down; in instances when results are inconsistent as to whether there is a benefit or harm of treatment, GRADE would rate down the quality of evidence as a result of variability in results, because the meaning of the inconsistency is so relevant to the decision to treat or not to treat.

|  |  |  |
| --- | --- | --- |
| **Rating for Inconsistency**  0 no change  -1 decrease quality 1 level  -2 decrease quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 4. Imprecision of Evidence

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

Cochrane states that when studies have few participants and few events, and thus have wide confidence intervals (Andersonet al. 2011), authors can lower their rating of the quality of evidence. These ratings of precision are made as judgments by review authors. The ratings are made by looking across studies, or, if available, on the results of a meta-analysis.

GRADE defines evidence quality differently for systematic reviews and guidelines. For systematic reviews, quality refers to confidence in the estimates of effect. For guidelines, quality refers to the extent to which confidence in the effect estimate is adequate to support a particular decision (Guyattet al. 2011). For the purpose of step 3 of Navigation Guide, we will use the systematic review definition, because the decision phase does not occur until step 4 when recommendations for prevention are made. Thus, when reviewing the data for imprecision, evaluate your confidence in the estimate of the effect.

According to GRADE, to a large extent, CIs inform the impact of random error on evidence quality. Thus, when considering imprecision, the issue is whether the CI around the estimate of exposure effect is sufficiently narrow. If it is not, GRADE rates down the evidence quality by one level (for instance, from high to moderate). If the CI is very wide, GRADE might rate down by two levels.

|  |  |  |
| --- | --- | --- |
| **Rating for Imprecision**  0 no change  -1 decrease quality 1 level  -2 decrease quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 5. Publication Bias

Possible ratings: 0 = no change; -1 or -2 downgrade 1 or 2 levels

GRADE (Guyattet al. 2011) and Cochrane (Higgins and Green 2011) assess publication bias in a similar manner. Whereas “selective outcome reporting” is assessed for each study included in the review as part of the risk of bias assessment, “publication bias” is assessed on the body of evidence. GRADE states that “when an entire study remains unreported and the results relate to the size of the effect- publication bias- one can assess the likelihood of publication bias only by looking at a group of studies.”

Cochrane’s definition of publication bias is “the *publication* or *non-publication* of research findings depending on the nature and direction of the results.” Cochrane and GRADE are primarily concerned with *overestimates* of true effects of treatments or pharmaceuticals, especially related to “small studies effects”, i.e., the tendency for estimates of an intervention to be more beneficial in smaller studies. There is empirical evidence in the clinical sciences that publication and other reporting biases result in over estimating the effects of interventions (Higgins and Green 2011).

In contrast, in environmental health, we are primarily concerned with *underestimating* the true effects of a chemical exposure, since in many cases population wide exposure has already occurred. We are also concerned that studies finding no association are less likely to be published because journals are less likely to publish “negative” findings.

Applying this inverted concern to GRADE’s assessment for publication bias, leads to these considerations when rating publication bias:

* + Early *negative* studies, particularly if small in size, are suspect. (GRADE is concerned with early *positive* studies).
  + Authors of systematic reviews should suspect publication bias when studies are uniformly small, particularly when sponsored by the industry. (Same as GRADE)
  + Empirical examination of patterns of results (e.g., funnel plots) may suggest publication bias but should be interpreted with caution. (Same as GRADE)
  + More compelling than any of these theoretical exercises is authors’ success in obtaining the results of some unpublished studies and demonstrating that the published and unpublished data show different results. (Same as GRADE)
  + Comprehensive searches of the literature including unpublished studies, i.e., the grey literature, and a search for research in other languages are important to addressing publication bias. Note that Cochrane also states “comprehensive searching is not sufficient to prevent some substantial potential biases.”

|  |  |  |
| --- | --- | --- |
| **Rating for Publication Bias**  0 no change  -1 decrease quality 1 level  -2 decrease quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

## Upgrade Categories

GRADE states that the circumstances for upgrading likely occur infrequently and are primarily relevant to observational and other non-randomized studies. Although it is possible to rate up results from randomized controlled trials, GRADE has yet to find a compelling circumstance for doing so (Guyattet al. 2011). GRADE specifies 3 categories for increasing the quality of evidence (Guyattet al. 2011).

### Category 6. Large Magnitude of Effect

Possible ratings: 0 = no change; +1 or +2 upgrade 1 or 2 levels

Modeling studies suggests that confounding (from non-random allocation) alone is unlikely to explain associations with a relative risk (RR) greater than 2 (or less than 0.5), and very unlikely to explain associations with an RR greater than 5 (or less than 0.2). Thus, these are the definitions of “large magnitude of effect” used by GRADE to upgrade 1 or 2 levels, respectively. Also, GRADE is more likely to rate up if the effect is rapid and out of keeping with prior trajectory; usually supported by indirect evidence. GRADE presents empirical evidence to support these conclusions, and states that “although further research is warranted, both modeling and empirical work suggest the size of bias from confounding is unpredictable in direction but bounded in size.

Hence, the GRADE group has previously suggested guidelines for rating quality of evidence up by one category (typically from low to moderate) for associations greater than 2, and up by two categories for associations greater than 5.”

Applying the GRADE definitions of large magnitude of effect i.e., RR greater than 2 or 5 is problematic in environmental health because for dichotomous outcomes RR is a function of the exposure comparator; these definitions also are not applicable to results from continuous variables. At present, we do not have an empirically defined “large magnitude of effect.” Therefore, for the purpose of this case study, review authors should assess whether the results indicate a large magnitude of effect using their expert judgment of “large effects” in environmental health and state their definition for discussion by the group.

|  |  |  |
| --- | --- | --- |
| **Rating for Large Magnitude of Effect**  0 no change  +1 increase quality 1 level  +2 increase quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 7. Dose-response

Possible ratings: 0 = no change; +1 or +2 upgrade 1 or 2 levels

Possible considerations include consistent dose response gradients in one or multiple studies, and/or dose response across studies, depending on the overall relevance to the body of evidence.

|  |  |  |
| --- | --- | --- |
| **Rating for Dose-Response**  0 no change  +1 increase quality 1 level  +2 increase quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

### Category 8. Residual Confounding Increases Confidence

Possible ratings: 0=no change; +1 or +2 upgrade 1 or 2 levels

Upgrade if consideration of all plausible residual confounders, biases, or effect modification would underestimate the effect or suggest a spurious effect when results show no effect. If a study reports an association despite the presence of residual confounding, biases or effect modification that would diminish the association, confidence in the association is increased. GRADE provides an illustrative example related to bias: rating up observational evidence finding lack of association between vaccination and autism, which occurred despite empirically confirmed bias that parents of autistic children may be more likely to remember their vaccine experience. The negative findings despite this form of recall bias suggest rating up the quality of evidence (Guyattet al. 2011).

|  |  |  |
| --- | --- | --- |
| **Rating for Residual Confounding Increases Confidence**  0 no change  +1 increase quality 1 level  +2 increase quality 2 levels | | **Rationale for your judgment** |
| Human |  |  |

The results of the reviewers’ ratings by population will be compiled and discussed leading to a final decision on overall quality of human evidence. The rationale for the decision will be fully documented.

## Final decision on overall quality of human evidence:

(Example: Moderate quality is upgraded 1 step to high for XYZ reason(s))

---- High

---- Moderate

---- Low

---- Very Low

# Supplementary 5 Rate the strength of evidence

The strength of evidence was rated based on a combination of four criteria: (1) Quality of the entire body of evidence; (2) Direction of the effect estimate; (3) Confidence in the effect estimate; and (4) Other compelling attributes of the evidence that may influence certainty. The strength of evidence ratings are summarized below, where their meaning is further defined.

|  |  |
| --- | --- |
| Sufficient evidence of harmfulness | The available evidence usually includes consistent results from well-designed, well-conducted studies, and the conclusion is unlikely to be strongly affected by the results of future studies. For human evidence a positive relationship is observed between exposure and outcome where chance, bias, and confounding, can be ruled out with reasonable confidence. |
| Limited evidence of harmfulness | The available evidence is sufficient to determine the effects of the exposure, but confidence in the estimate is constrained by such factors as: the number, size, or quality of individual studies, the confidence in the effect, or inconsistency of findings across individual studies. As more information becomes available, the observed effect could change, and this change may be large enough to alter the conclusion. For human evidence a positive relationship is observed between exposure and outcome where chance, bias, and confounding cannot be ruled out with reasonable confidence. |
| Inadequate evidence of harmfulness | Studies permit no conclusion about a toxic effect. The available evidence is insufficient to assess effects of the exposure. Evidence is insufficient because of: the limited number or size of studies, low quality of individual studies, or inconsistency of findings across individual studies. More information may allow an estimation of effects. |
| Evidence of lack of harmfulness | The available evidence includes consistent results from well-designed, well-conducted studies, and the conclusion is unlikely to be strongly affected by the results of future studies. For human evidence more than one study showed no effect on the outcome of interest at the full range of exposure levels that humans are known to encounter, where bias and confounding can be ruled out with reasonable confidence. The conclusion is limited to the age at exposure and/or other conditions and levels of exposure studied. |

# Supplementary 6 Studies excluded from full-text assessment with reasons

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Title | Authors | Published Year | Journal | Study | Notes |
| Aircraft Noise Exposure and Its Effects on Quality of Life and Cognitive Function Among Thai Residents. | Padungtod, Chantana; Ekpanyaskul, Chatchai; Nuchpongsai, Piyarat; Laemun, Nattapong; Matsui, Toshihito; Hiramatsu, Kozo | 2011 | Epidemiology | Padungtod 2011 | Exclusion reason: Conference abstract, full text not available; |
| Estimation of airport noise impacts on public health. A case study of Izmir Adnan Menderes Airport | Ozkurt, N; Hamamci, SF; Sari, D | 2015 | Transportation Research Part D-Transport And Environment | Ozkurt 2015 | Exclusion reason: Wrong study design; |
| Estimation of the relationship between noise and annoyance caused by vehicular traffic: an application in the city of Medellin, Colombia | Restrepo, FJC; Munera, JDO; Campo, CAC | 2018 | Revista De Estudios Regionales | Restrepo 2018 | Exclusion reason: Conference abstract, full text not available; |
| Is Community Noise Associated with Metabolic Control in Patients with Cardiovascular Disease? | Dzhambov, AM; Tokmakova, MP; Gatseva, PD; Vladeva, SV; Zdravkov, NG; Vasileva, EV; Gencheva, DG; Ivanova, NG; Karastanev, KI; Donchev, AT | 2017 | Acoustics Australia | Dzhambov 2017 | Exclusion reason: Wrong outcomes; |
| Study of physio-psychological effects on traffic wardens due to traffic noise pollution; exposure-effect relation | Tabraiz, S; Ahmad, S; Shehzadi, I; Asif, MB | 2015 | Journal Of Environmental Health Science And Engineering | Tabraiz 2015 | Exclusion reason: Wrong patient population; |
| Association and moderation of self-reported hypotension with traffic noise exposure: a neglected relationship. | Lercher, Peter; Widmann, Ulrich | 2013 | Noise & Health | Lercher 2013 | Exclusion reason: High income country; |
| Evaluating the impact of wind turbine noise on health-related quality of life. | Shepherd, Daniel; McBride, David; Welch, David; Dirks, Kim N; Hill, Erin M | 2011 | Noise & Health | Shepherd 2011 | Exclusion reason: High income country; |
| Cardiovascular effects of environmental noise: research in Serbia. | Belojevic, Goran; Paunovic, Katarina; Jakovljevic, Branko; Stojanov, Vesna; Ilic, Jelena; Slepcevic, Vesna; Saric-Tanaskovic, Mica | 2011 | Noise & Health | Belojevic 2011 | Exclusion reason: Wrong study design; |
| A Health-Based Metric for Evaluating the Effectiveness of Noise Barrier Mitigation Associated With Transport Infrastructure Noise. | Prendergast, Geoffrey P; Staff, Michael | 2017 | Noise & Health | Prendergast 2017 | Exclusion reason: High income country; |
| The short-term association of road traffic noise with cardiovascular, respiratory, and diabetes-related mortality. | Recio, Alberto; Linares, Cristina; Banegas, Jose; Diaz, Julio | 2016 | Environmental Research | Recio 2016 | Exclusion reason: High income country; |
| Road traffic noise and hypertension - Accounting for the location of rooms. | Babisch, Wolfgang; Wolke, Gabriele; Heinrich, Joachim; Straff, Wolfgang | 2014 | Environmental Research | Babisch 2014 | Exclusion reason: High income country; |
| Noise levels and cardiovascular mortality: A case-crossover analysis. | Tobias, A; Recio, A; Diaz, J; Linares, C | 2015 | European Journal Of Preventive Cardiology | Tobias 2015 | Exclusion reason: High income country; |
| Cardiovascular and stress responses to short-term noise exposures-A panel study in healthy males. | Walker, Erica; Brammer, Anthony; Cherniack, Martin; Laden, Francine; Cavallari, Jennifer | 2016 | Environmental Research | Walker 2016 | Exclusion reason: Wrong patient population; |
| Community Noise Exposure and its Effect on Blood Pressure and Renal Function in Patients with Hypertension and Cardiovascular Disease. | Dzhambov, Angel M; Tokmakova, Mariya P; Gatseva, Penka D; Zdravkov, Nikolai G; Gencheva, Dolina G; Ivanova, Nevena G; Karastanev, Krasimir I; Vladeva, Stefka V; Donchev, Aleksandar T; Dermendzhiev, Svetlan M | 2017 | Folia Medica | Dzhambov 2017 | Exclusion reason: Wrong patient population; |
| Association between community noise and adiposity in patients with cardiovascular disease. | Dzhambov, Angel M; Gatseva, Penka D; Tokmakova, Mariya P; Zdravkov, Nikolai G; Vladeva, Stefka V; Gencheva, Dolina G; Ivanova, Nevena G; Karastanev, Krasimir I; Vasileva, Emanuela V; Donchev, Aleksandar T | 2017 | Noise & Health | Dzhambov 2017 | Exclusion reason: Wrong patient population; |
| Cardiovascular benefits of reducing personal exposure to traffic-related noise and particulate air pollution: A randomized crossover study in the Beijing subway system. | Yang, X.; Jia, X.; Dong, W.; Wu, S.; Miller, M.; Hu, D.; Li, H.; Pan, L.; Deng, F.; Guo, X. | 2018 | Indoor Air | Yang 2018 | Exclusion reason: Wrong study design; |
| Residential greenspace modifies the effect of road traffic noise exposure on mental health in youth: Donka Dimitrova. | Dzhambov, A; Markevych, I; Tilov, B; Dimitrova, D | 2017 | European Journal Of Public Health | Dzhambov 2017 | Exclusion reason: Conference abstract, full text not available; |
| The Influence of Daytime Road-Traffic Noise on Blood Pressure of School Children: Pp. 19.250 | Paunovic, K; Belojevic, G; Jakovljevic, B; Stojanov, V | 2010 | Journal Of Hypertension | Paunovic 2010 | Exclusion reason: Full text not available (Poster); |
| Evaluation of the social and economic burden of road traffic noise-attributed myocardial infarction in Bulgarian urban population. | Dzhambov, Angel M; Dimitrova, Donka D | 2015 | Arhiv Za Higijenu Rada I Toksikologiju | Dzhambov 2015 | Exclusion reason: Wrong outcomes; |
| Relationship among long-term aircraft noise exposure, blood pressure profile, and arterial stiffness. | Rojek, Marta; Rajzer, Marek; Wojciechowska, Wiktoria; Drozdz, Tomasz; Skalski, Pawel; Pizon, Tomasz; Januszewicz, Andrzej; Czarnecka, Danuta | 2019 | Journal Of Hypertension | Rojek 2019 | Exclusion reason: High income country; |
| Effect of nighttime aircraft noise exposure on endothelial function and stress hormone release in healthy adults. | Schmidt, Frank; Basner, Mathias; Kroger, Gunnar; Weck, Stefanie; Schnorbus, Boris; Muttray, Axel; Sariyar, Murat; Binder, Harald; Gori, Tommaso; Warnholtz, Ascan; Munzel, Thomas | 2013 | European Heart Journal | Schmidt 2013 | Exclusion reason: Wrong outcomes; |
| Community response to a step change in the aircraft noise exposure around Hanoi Noi Bai International Airport | Nguyen, T. L., Nguyen, T. L., Morinaga, M., Yokoshima, S., Yano, T., Sato, T., & Yamada, I | 2018 | The Journal Of The Acoustical Society Of America | Nguyen 2018 | Exclusion reason: replaced by most recent study (Nguyen 2020) |
| Effects of step change in aircraft noise exposure on activity disturbances: Socio-acoustic surveys around Hanoi Noi Bai International Airport. | Morinaga, M., Nguyen, T. L., Shimoyama, K., Yokoshima, S., & Yano, T. | 2020 | Acoustical Science And Technology | Morinaga 2019 | Exclusion reason: replaced by most recent study (Nguyen 2020) |
| The associations between noise annoyance and psychological distress with blood pressure in children and adolescents: The CASPIAN-V Study | Badihian, N., Riahi, R., Qorbani, M., Motlagh, M.E., Heshmat, R. and Kelishadi, R. | 2020 | The Journal of Clinical Hypertension | Badihian, 2020 | Exclusion reason: Wrong exposure (noise annoyance as exposure) |

# Supplementary 7 Detail information about exposure assessment for each study

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Exposure assessment** (type) | **Equipment** | **Number of sites** (with reason if applicable) | **Noise metrics** | **Other information about exposure data** | **Exposure range (mean+SD)** |
| Tao 2021 | Personal SLM monitoring | Wearable devices included a portable air pollutant sensor, a portable noise sensor and a GPS-equipped smartphone | N/A as PE | LAeq | During the 48-h survey (including a weekday and weekend.), participants were required to carry a device package. The noise and air pollutant sensor were exposed to the open air and adjacent to participants' breathing zone. | Median: 58.96 IQR: 7.98 |
| Huang 2013 | Personal SLM monitoring | Personal noise exposure meter (Model AWA5610B; Aihua Instrument, Hangzhou, China) | N/A as PE | LAeq | Personal exposure to noise were measured continuously, and simultaneously with the ambulatory HRV monitoring. The specially designed backpack with monitoring equipment was placed beside the volunteers throughout the exposure period. | Median+IQR: 72.54+3.02dBA |
| Ma 2020 | Personal SLM monitoring | Portable SLM and a GPS-equipped smartphone | N/A as PE | LAeq | During the 48-h survey (including a weekday and weekend.), participants were required to carry a device package. The noise and air pollutant sensor were exposed to the open air and adjacent to participants' breathing zone. | 36 to 97 dBA |
| Tao 2020 | Personal SLM monitoring | Portable SLM-25 | N/A as PE | LAeq | Data was collected from Dec. 2017 to Feb. 2018 in 6 waves. Each wave spanned over a continuous 48-h period that covered a workday and a weekend day. Every participant was required to carry a portable noise sensor and a GPS-equipped smartphone, a health questionnaire, and an activity-travel diary. | 37.7 to 96.7 dBA, mean 60.1 db(A) |
| Song 2016 | Fixed site SLM measurement & propagation-based model | AIHUA type AWA6291 SLM | Background noise: 10 sites in quiet areas within the wind farm (wind speed at the hub height was <3 m/s of wind turbines), and outside of 65 houses (1 m from facade) around wind farm. | LAeq, Lden | Randomly measured when the wind turbines (WT) were under normal operation during daytime (DT) in 5 days. Measuring points: 1.5 m higher above the ground and the measuring time was 15 min. Apart from the A-weighted SPLs (LAeq), frequency analysis in 1/3-octave bands from 10 Hz to 20 kHz was carried out. Used noise model predictions from CadnA software to assign exposures to LAeq based on a wind speed of 8 m/s and at a height of 10m in neutral atmosphere | Aerodynamic noise: 47.6+2.8 dBA; background noise: 32.5+1.5 dBA; |
| Trieu 2021 | Fixed site SLM measurement & propagation-based model | NL-21 and NL-22, RION, Tokyo, Japan | 10 | Lden | 2008 (measurement): ACN measured every 1 s from Sep 22-29, 2008. Microphones covered with all-weather wind screens were positioned on the rooftops (1.5 m above the roofs & 1 m away from facade). Flight arrival/departure times obtained from the airport office.  2019 (noise model): August 4–11. Lden and Lnight were estimated from noise contour maps calculated using the Integrated Noise Model 7.0 (INM). Flight route data were collected from 5–9 August 2019. Predictive calculations were performed for three days and average value was used of Lden and Lnight. RMSE between measured and model estimates was 1.9 Lden and 2.4 Lnight dB. | Lden: 53–71 dB in 2008, 63–81 dB in 2019. |
| Nguyen 2020 | Fixed site SLM measurement & propagation-based model | RION NL-42, NL-21, NL-22) | 13 | Lden , Lnight | 2014-2015 (measurement): microphones were set at the height of 1.5 m above the roofs &1 m from facade. Sampled at 1 s were recorded continuously over 7 days. The noise data of each day for each site was compared with flight logs to identify the aircraft events and then calculate the Lden.  2017 and 2018 (noise model): Lden and Lnight estimated from noise contour maps calculated using the Integrated Noise Model (INM). The model combines information on all aircrafts operating in 2017 and 2018: flight logs and weather conditions during the surveys, and compared with the seasonal average traffic Compared model estimated and measured noise levels RMSE between 2.4 and 3.9 dB | Lden ranged 45-66 dB in 2014; 44-73 dB in 2018 |
| Gilani 2021 | Propagation-based model | Validation with Casella CEL-633C class 1 SLM | Validation: 32 measurement sites, 24-h noise measurements during August- October 2020. | Lden | The noise levels were obtained by using TNM 2.5 (Traffic Noise Model) model developed by Federal highway administration. Penalties +5 dB(A) for evening and +10 dB(A) for NT to derive Lden. Data inputs: GPS coordinates of the survey area; data obtained from videographic measurements (Road traffic volume, traffic speed) and field visits (Pavement surface type, height of the residential buildings, lane width). Validity: high correlation of r = 0.921, p (two-tailed) < 0.01. Exposure dichotomized as low <60 vs high >60 exposure. | Mean: 66.82, SD: 4.91 |
| Roca-Barcelo 2021 | Propagation-based model | N/A | N/A | Ldn | Modelled using the integrated noise model (INM) developed by the US Federal Aviation Administration (FAA) and approved by the Brazilian National Civil Aviation Agency (ANAC). Model combines information on all aircrafts operating in 2017: flight patterns, times of arrival/departure, height, speed, and engine power. The modelled ground noise levels represented the LDN; with a penalty of +10 dB for NT noise (22:00-7:00). INFRAERO data were available as LDN contours categorized as: < 65, 65–70, 71–75, 76–80, 81–85, and > 85 dB. | ≤50 dB to > 65 dB |
| Dzhambov 2017 | Propagation-based model | NA | N/A | Lden | Combined Lden data from all traffic sources were elicited from the official strategic noise maps (10 m × 10 m grid, 4 m height) of Plovdiv, Bulgaria created in 2009 in compliance with the Environmental Noise Directive 2002/49/EC. Lden categorized as follows: 20-55, 55-60, 60-65, 65–70, 71–75 | 50-80 dBA |
| Dzhambov 2016a | Propagation-based model | NA | N/A | Lden | Combined Lden data from all traffic sources were elicited from the official strategic noise maps (10 m × 10 m grid, 4 m height) of Plovdiv, Bulgaria created in 2009 in compliance with the Environmental Noise Directive 2002/49/EC | Not reported |
| Paiva 2019 | Propagation-based model | Validation with Brüel & Kjaer BK 2237 SLM, calibrated using a B&K 4231 sound level calibrator | 75, selected exposed & not exposed to traffic noise areas based on number of vehicles at 75 selected points on the streets of the neighbourhood. | Leq, Lden, Ln | Measurement: procedure followed the Brazilian standard NBR 10151:2000. Time interval: 15 min.a total of about 20 h of measurement. Noise mapping: using B&K 7810 version 8.11 Predictor software (Bruel and Kjaer, 2010), into which were inserted physical characteristics (topographical) of the study area and parameters of the noise source (traffic-number of vehicles). The ISO 9613-1/2 Road model was chosen, and the calculation parameters were a 5 × 5 grid with a height of 4 m, as recommended by EC Directive 2002/49. Validation: The greatest difference between the measured and simulated values was 3.5 dB(A). | Range from 66.5 -78.3 dB |
| Banerjee 2014b | Propagation-based model | N/A | N/A | Lden | CORTN model using the SoundPLAN Essential 2 from Braunstein & Berndt GmbH, Backnang, Germany. Data input: GPS coordinates of sampling sites, elevation, building geometry, road geometry and details, hourly classified vehicle count, vehicle flow rate, vehicle speed, and percentage of heavy vehicles, were collected using strategic traffic sampling stations. Traffic data from field, digitized base map, noise levels for selected locations where participants resided to generate noise maps. The participants' dwellings close to one of the sampling stations were grouped under a similar noise exposure level. Those living in areas where L den was <60 dB(A) were marked as the reference group (quieter areas). | 63.5 ± 6.6 dBA |
| Dzhambov 2015 | Propagation-based model | NA | N/A | Lden | Combined Lden data from all traffic sources were elicited from the official strategic noise maps (10 m × 10 m grid, 4 m height) of Plovdiv, Bulgaria created in 2009 in compliance with the Environmental Noise Directive 2002/49/EC. | Not reported |
| Banerjee 2014a | Propagation-based model | N/A | N/A | Lden | CORTN model using the SoundPLAN Essential 2 from Braunstein & Berndt GmbH, Backnang, Germany. Data input: GPS coordinates of sampling sites, elevation, building geometry, road geometry and details, hourly classified vehicle count, vehicle flow rate, vehicle speed, and percentage of heavy vehicles, were collected using strategic traffic sampling stations. Traffic data from field, digitized base map, noise levels for selected locations where participants resided to generate noise maps. The participants' dwellings close to one of the sampling stations were grouped under a similar noise exposure level. Those living in areas where L den was <60 dB(A) were marked as the reference group (quieter areas). | Range 55–80 dB |
| Dzhambov 2016b | Propagation-based model | NA | N/A | Lden | Extracted from official noise map of Plovdiv. Lden was dichotomized (≥65 dB vs.<65 dB). | Lden ≥ 65 dB, n (%) 184 (64.79%) |
| Dzhambov 2018a | Propagation-based model | NA | N/A | Lden | Combined Lden data from all traffic sources were elicited from the official strategic noise maps (10 m × 10 m grid, 4 m height) of Plovdiv, Bulgaria created in 2009 in compliance with the Environmental Noise Directive 2002/49/EC. Contains noise levels in the 5-dB contours from 50 dB to 80 dB. | 50-80 dBA |
| Sieber 2018 | LUR model | Type-II SLM Data Logger Noise Sentry RT | 134 | LAeq/Lden | Measurements were collected over 1-week periods in an appropriately sampled way at 134 locations. Penalties +5 dB(A) for evening and +10 dB(A) for NT to derive Lden, used measurements to build a LUR model, which had predictive accuracy of r2=0.13. | Not reported |
| Dzhambov 2018b | LUR model | Not mentioned | 40 locations (traffic sites, industrial sites, sites in residential and recreational areas) | LAeq - daytime | The LUR was developed based on noise measurements in 2016. Measurements were conducted over the 12-h period from 07.00 to 19.00 h, according to the ISO 1996–2:1987. The final LUR has an adjusted R2 of 0.72 and leave-one-out cross validation R2 of 0.65. Exposure assigned as continuous variable. Stated in the Discussion that noise exposures likely overestimated at addresses on minor roads and in suburbs. | 62.4–73.5 dB |
| Dzhambov 2019 | LUR model | Not mentioned | 40 locations (traffic sites, industrial sites, sites in residential and recreational areas) | LAeq - daytime | The LUR was developed based on noise measurements in 2016 (range: 62.4–73.5 dB(A)). Measurements were conducted over the 12-h period from 07.00 to 19.00 h, according to the ISO 1996–2:1987. The final LUR has an adjusted R2 of 0.72 and leave-one-out cross validation R2 of 0.65. Exposure assigned as continuous variables. Stated in the Discussion that noise exposures likely overestimated at addresses on minor roads and in suburbs. | Depression No (N=351): 67.10 ± 1.59 dB; Yes (N=86): 67.12 ± 1.51 dB |
| Banerjee 2013 | Fixed site SLM measurement | Not mentioned | Close proximity to nearest outdoor façade of subject’s dwelling | Lden | Sample from those residing within 50 m of roadways who had lived for a minimum of ten yrs; measurements performed manually for 24h in 3 shifts of (7:00 -19:00, 19:00-23:00, and 23:00-07:00) | Not reported, |
| Phan 2010 | Fixed site SLM measurement | RION NL-06 and NL-22 | 8, based on traffic volume conditions & the typical characteristic of combined residential and commercial areas | LAeq, Lden, Lday, Lnight | For row house: measure for participants lived within approximately 1km; measured every second for 24 h. Used measurement values to form a distance reduction equation (From road shoulders to houses). For apartments, values were assigned to all apartment respondents, as the level averaged by weighting the number of dwellers on each floor. Traffic density: measured by counting the number of vehicles through video camera recordings | Day range:71-78(Hanoi)/74-80(Ho Chi Minh City);  LAeq night:66-76(Hanoi))/70-76(Ho Chi Minh City) |
| Stosic 2009 | Fixed site SLM measurement | Brüel & Kjær noise level analyzer type 4426. | 6, 3 in busiest streets and 3 in quiet streets | Leq | Total number of samples: 9000, with a period of sampling 0.1 s in a fast dynamic range during two NT intervals (22:00 - 01:00, 02:00-05:00). Residents had to live at their address for over a year and have their bedroom windows face the street. | Leq in noisy street 61.33+2.31; in quiet street, 43+3.46 dBA |
| Han 2015 | Fixed site SLM measurement | AWA6218C type, Hangzhou Aihua Instrument Co, Ltd, China. | Unknown | LAeq | 4 functional areas chosen (residential, commercial, transportation, construction). Stratified random sample of participants within each functional area. Frequency: once 10-min interval per hour, in working days (6:00-22:00, 22:00-6:00). Followed protocols in accordance with technical specifications. | Transportation hub area daytime: measured + SD (78.3+1.2 dB) |
| Liu 2017 | Fixed site SLM measurement | AZ® 8928 Digital SLM | 40, 8 public buildings, 6 commercial, 20 residential & 6 infrastructure projects; Each have 8 ref points | LAeq | Construction noise for residents surrounding the selected construction sites (within 100 m). Measurements conducted inside buildings. The 12 h construction noise (LAeq) was obtained by measuring the 10 min equivalent continuous. A weighted sound pressure level every 2 h for 12 h (from 08:00 to 20:00). | LAeq 60-80 dB |
| Wu 2019 | Fixed site SLM measurement | DT-8852 SLM | 42 | LAeq | Short-time noise measurements were conducted at the same time as questionnaires’ survey. For each building, noise levels were obtained 10-min interval in 7:00-10:00 and 17:30-21:00). SLM was mounted at 1.5 m above the ground and more than 2 m away from the wall, and with the window opened in the room. The acquisition positions were all facing the main road/bridge. | Ranges from 62-72 dB |
| Nazneen 2020 | Fixed site SLM measurement | Extech SLM | 57 | LAeq | Noise survey was conducted in urban zones of Peshawar in accordance with the ISO 1996-1:2016. Link and main roads were visited and areas with the similar land use structure were merged. Noise measurements were carried out on days without rain or strong wind and recorded each minute for 12 h (07:00-19:00). | 46.3-86.3 dB |
| Zamorano-Gonzalez 2021 | Fixed site SLM measurement | Type I SLM | 15 sites with highest traffic flow (count using video camera) were selected | LAeq | SLM on tripods at a height of 4 m off the ground and located at a distance of 3 m from any facade or wall. Noise measurement is carried out in three shifts of the day (6:30-7:30, 12:30-13:30, 17:30-18:30) from Monday to Friday. Participants: people living within 250 m of the streets/avenues in which the noise was measured. | 70.5-75.6 dB |
| Agarwal 2009 | Fixed site SLM measurement | Not mentioned | 10 intersections covering all the four major directions of the city with heavy traffic flow | LAeq | Measurement takes place in general from 8 a.m. to 8 p.m. | Leq range 72.6 - 83.9 dBA |
| Goswami, 2009 | Fixed site SLM measurement | Not mentioned | 6 | Lday | In each sited, noise measured in four different spots; 16 different time each periods during daytime (not specific), on 1st week of March, 2008. Each site selected to have typical road width, roadside housing pattern, traffic flowing pattern in different directions. | Range: 88 -122.9 dBA |
| Lepore, 2010 | Fixed site SLM measurement | Not mentioned | 2 (schools) | LAeq | Participants were drawn from two public schools defined as "noisy" and "quiet" based on their location (distance to commercial district and main streets). Noise levels at both schools were measured mid-morning between the outside entrance to the buildings and the school gate. | Noisy school: peak of 82 dBA; Quiet school: peak of 65 dBA over a five min interval |
| Goswami 2011 | Fixed site SLM measurement | Model SL-4001 | 5 | Lday | Noise levels were measured at each site at 16 different times in a day following standard procedure using calibrated sound pressure level (dB) meter during day-time, at the five different sites along the investigated road of 11 km. | Range 70.4-121.2 dB |
| Belojevic 2012 | Fixed site SLM measurement | Type 2250 “Brüel & Kjær” | 115, in the middle of 115 municipal streets. Measurements In front of 8 schools. | LAeq | Leq was measured in 5 intervals (8:00-10:00, 14:00-16:00,18:00-20:00, 22:00-0:00, 0:00-2:00) in Sep-Oct 2008. Interval:15 min; the speed of sampling:10 per second, with 9000 samples. For home: matching their address with noise levels in the given street. School: measured in 3 DT intervals (9:00-11:00, 12:00-14:00,15:00-17:00). Traffic density (i.e., the numbers of light and heavy vehicles per hour) was registered in all streets & in front of all schools. | 70.9 ± 5.7 dBA |
| Seabi 2015 | Fixed site SLM measurement | SVAN 955 Type 1 SLM and Rion NC74 acoustic calibrator | 5 (schools) | Leq | Measured the external noise surrounding five schools. Measurement was taken between 08:00-10:30 hours while tests were being taken. | 2011 exposed school: 55.2 dB, quiet school: 50.5-57.9 dB |
| Agarwal 2011 | Fixed site SLM measurement | SLM SC-30 (version 1.0-2.1) | Unknown | Ldn | SLM was mounted on a stand at a height of 1.2 m above the ground level with 7.5 m distance from the centre line of the road. The noise data were taken at the pre-selected study locations at which the traffic data and noise social survey were carried out. The traffic volume and noise monitoring were noted for 24 hours. | Leq range 73-86 |
| Onchang 2018 | Fixed site SLM measurement | RION Models NL-06 and NL-22 | 20 (10 for on-campus and 10 for off-campus) | LAeq | Noise levels at both on-campus and off-campus locations. Measurements were made every 15 min over a 24-h period situated 2 m from building facades. Time of measurement (07:00–19:00,19:00–22:00, 22:00–07:00). | Leq 24h, outside:64.7+4.8; Leq 24h insideL 60.3+ 1.8 |
| Paunovic 2013 | Fixed site SLM measurement | Noise level analyzer type 2250 "Brüel and Kjær" | 8 schools, 115 streets | Lden | measured in front of all 8 schools & all 115 streets during September/October 2008. School: 3 shift (9: 00 - 11: 00;12:00-14:00;15:00-17:00). Home: in the middle of streets closed to home, 3 shifts (10: 00 - 12: 00;14:00-16:00;18:00-20:00). 15min interval, speed of sampling: 10/sec, with 9000 samples. Leq was calculated for each school; a composite DT and NT Leq levels were calculated for each street. Also collected information on traffic density and accoustic insulation in buildings. | Lden: mean+SD 62.2+4.2 |
| Siddiqui 2015 | Fixed site SLM measurement | Standard SLM ST-85A | 5 | dB | Subjects were selected who were workers of different trades at these places. Measurements conducted at five busy intersection sites, and then participants were people who worked near those sites. | Ranges from 100 dB to 110 dB |
| Ravindra 2016 | Fixed site SLM measurement | “Type 2” Integrating SLM | 27 sites, outdoor, indoor, roadside and residential areas | LAeq | SLM was installed above 1.2 m using a tripod. Each location was monitored for 24 h including both day and night observations. Measurement duration: 6:00-10:00, 22:00-06:00. Sound level, diurnal pattern and weekend vs. weekday trend were also investigated | Ranges from 45-120 dB |
| Masoudzadeh 2017 | Fixed site SLM measurement | Not mentioned | Unknown | Unknown | Conducted based on the studies by Oveisi et al., Alizade (full text not found) et al. on noise pollution in Sari in 2007-2008. Participants not exposed to noise were from Keshavarz Blvd. where noise level was 18.84 dB. Participants exposed to noise were from Helal Ahmar Sq. where noise pollution level was 75.85 dB. | Non-exposed: 18.84 dB; exposed: 75.85 dB |
| Ana 2009 | Fixed site SLM measurement | Factory calibrated TECPEL Model 330 series SLM | 4 schools | LAeq | Measurements were conducted twice 9:00-10:00 & 13:00-14:00 at two points in the classroom (where students were seated & playground) for a period of five school days. Traffic density: manual count of the number of vehicles, around the school with the highest noise levels was also determined during the study period. Floor space and the sizes of potentially open doors and windows were determined. | 68.3–84.7 dBA |
| Banerjee 2009 | Fixed site SLM measurement | Not mentioned | 25 | LDN | 869 individuals were interviewed in person at 25 pre-selected locations, where traffic noise and classified vehicular count measurements were also conducted. Noise levels near interview sites were used for impact evaluation. | 80.62±15.88 dB(A) |
| Daruis 2014 | Fixed site SLM measurement | Calibrated SoundTrack LxT SLM. | 6, based on the distance from the source of the noise (the road) | LAeq, Lmax, L10 and L90 | Residential areas along 3km stretch of Paka Rd. Instrument installed on a tripod at 1.2 to 1.5 m in height and 1 m of distance from the wall. For road: measurement conducted on weekday nights from Sun. to Thu.;and weekend nights from Fri. to Sat., time duration (19:00-21:00, 21:00-23:00, 23:00-01:00, 01:00-03:00,03:00-06:00) for 40 minutes at each locations. For home: instrument was in the living room (19:00-21:00, 21:00-23:00), in the bedroom (23:00-01:00, 01:00-03:00,03:00-06:00). | 57.3 ± 4.723 dBA |
| Farooqi 2021 | Fixed site SLM measurement | Not mentioned | SLM model TES-1351B class 2103 | LAeq | SLMs had a frequency range of 20–8000 Hz and an accuracy of ±1.0 dB. Measurements taken during September 2019; Short-term road traffic noise collected continuously for 15 min at each location (LAeq15). Placing SLM at a tripod at the level of 1.7 m from the level of the pavement and a distance of 3 m from the noise reflecting surface. | Residential area Jhang: 70 ± 15.2; Chiniot: 88 ± 9.4 |
| Guoqing 2012 | Fixed site SLM measurement | Not mentioned | 39 set around the Hangzhou Xiaoshan International Airport | LWECPN, LDN | The measurement was performed for 24 h, recording effective perceived noise level (LEPN) of each single flight incident. LWECPN: calculated according to the average LEPN (according to number of flights during 6:00-19:00, 19:00-22:00, 22:00-6:00) of all flight incidents in a week, the number of flights in 3 different periods in a week. The points with the same values of 60, 65, 70, 75, 80 dB were found and plotted; All the points with the same value were connected, the graph of equal noise level contour was formed. | Ranges from 60 to 80 dB |
| Hamid 2019 | Fixed site SLM measurement | Digital SLM/OS-11 | 8 Locations | Leq | Daily monitoring was carried out for 3 consecutive days of each month for a duration of 8 months from July 2015 to February 2016, thrice a day (08:00-10:00, 13:00-15:00, 18:00-20:00). | 67.2 dB to 120.1 dB |
| Paunovic 2014 | Fixed site SLM measurement | Noise level analyzer type 2250 ‘Brüel & Kjær’ | 118, in the middle of 118 streets | Leq | Performed on weekdays (8:00-10:00, 14:00-16:00, 18:00-20:00, 22:00-0:00, 0:00-2:00) during Sep–Oct 2008. SLM was placed on pavement by the road, 3 m away from the nearest façade, at a height of 1.5 m. Measurement sites were located in the middle of the street between the crossroads away from other possible sources of noise. Interval: 15 min; speed of sampling: 10 per second, with 9000 samples. All respondents lived within 100 m of any measurement site in one street. | 67.35 ± 3.53 dB |
| Ristovska 2009 | Fixed site SLM measurement | Brüel & Kjaer 2260 Investigator type sound analyser | Unknown | Lday & Lnight & Lden | Community noise was measured for one week in the spring and one week in the autumn of 2006. Measurements were taken at 15-minute intervals for the day time (07:00-23:00, Lday) and two measurements were taken for the night time (23:00-07:00, Lnight) | Lday: 62±6.45 Lnight: 56±6.52 |
| Seabi 2012 | Fixed site SLM measurement | SVAN 955 Type 1 SLM and Rion NC74 acoustic calibrator. | 5 schools located in high or low aircraft noise exposure areas | Leq | Aircraft noise exposures. Noise measurements were taken during the testing period (8:00-10:00). | Experimental group mean 69 dB, max: 95 dB; Control group mean: 40 dB, max 54 dB |
| Seabi 2013 | Fixed site SLM measurement | SVAN 955 Type 1 SLM and Rion NC74 acoustic calibrator | 5 schools located in high or low aircraft noise exposure areas | Leq | Aircraft noise measurements were taken during the testing period (8:00 - 10:00) outside the classrooms in order measure aircraft noise levels. Participants were grouped as high noise (HN) group and low noise (LN) groups. No measurements were conducted at the children’s homes. | Wave 1: 54.4 to 55.3 Leq and 73.2–74.3 Lamax. Noise during Waves 2 and 3 when aircraft were gone: 55.2 Leq and maximum noise levels of 60.8 to 71.2 Lamax. |
| Gjestland 2015 | Fixed site SLM measurement | Calibrated class 1 sound level meter (RION NL-21/22) fitted with a rain protection windscreen (RION WS-16) | 64, for road traffic/aircraft | Leq | The ACN level was logged every second for 7 consecutive days. Flight data: obtained from the official airport website. RTN was measured curb-side outside the residences for a 24-h period. Measurements were made for each section of the road (between two intersections), and this noise level was used for all the residences on that block. Positioned on a tripod 1.5m above ground/roof. | ACN: 48.0-70.6 dB; RTN:60.9-83.1 dB |
| Nandanwar 2009 | Exposure assigned by a proxy variable | N/A | N/A | N/A | Exposure assigned by a proxy variable (living in close proximity to major road) | N/A |
| Ali 2018 | Exposure assigned by a proxy variable | N/A | N/A | N/A | Exposure assigned by a proxy variable (motorcycle riding and duration of riding) | N/A |
| Badihian 2020 | Self-reported | N/A | N/A | N/A | Based on self-reported questionnaire data on noise sensitivity | N/A |
| Dias 2021 | Self-reported | N/A | N/A | N/A | Based on self-reported questionnaire data on self-perception of noise, investigated by the question: “In your neighbourhood, does the noise bother you?” (yes, no, I don't know, no answer). | N/A |
| Dzhambov 2014 | Self-reported | N/A | N/A | N/A | Interview asking about the last 3 months when answering the questions about noise sensitivity and noise characteristics (e.g., type of noise exposure: continuous/intermittent). | N/A |
| Firdaus 2010 | Self-reported | N/A | N/A | N/A | Information included: sources of noise, causes of noise, intensity of noise | N/A |
| Ma 2018 | Self-reported | N/A | N/A | N/A | Measured by questionnaire: types of noise-pollution exposures and 5-point Likert scale measurement of noise perception | N/A |
| Legend: PE: Personal monitoring; LUR: Land use regression; IQR: Inter quartile range; SPL (sound pressure level); WT (wind turbine); DT (Day time); RMSE: Root mean squared error | | | | | | |

# Supplementary 8 Individual risk of bias assessments

## Table 8.1 Individual risk of bias assessments for the annoyance studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Paunovic 2014 | Rating | PH | PH | PL | PL | L | L | L | N/A | PL | PH |
|  | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups | There is insufficient information about the exposure assessment methods (temporal generalizability of measurements; if measurement conducted at same site over multiple days) to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit a judgment of low risk of bias, but there is indirect evidence which suggests that methods were robust | The study appropriately accounted for most but not all of the Tier I confounders or used appropriate statistical techniques; & some of the other Tier II confounders using appropriate statistical techniques. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for potential mediators such as duration of residence, bedroom window orientation, floor level, high noise sensitivity, etc. |  |
| Song 2016 | Rating | L | PL | PL | PH | L | L | L | L | PL | PH |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment & enrolment procedures, participation were sufficiently detailed to support the assertion that risk of selection effects was minimal. | There is insufficient information about the exposure assessment methods to permit a judgment of low risk of bias, but there is indirect evidence that suggests that methods were robust (used noise model predictions to assign exposures for health analysis, and assumption made that used measurements to validate models) | Insufficient information to permit a judgment of low risk of bias, but there is indirect evidence which suggests that methods were robust | The study evaluated some but not all of the Tier I confounders, & some but not all of the other Tier II confounders, therefore is expected to introduce substantial bias | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g., adjusting for noise sensitivity, visibility of wind turbine etc.) |  |
| Nguyen 2020 | Rating | PL | PH | PL | PH | L | L | L | L | PL | PH |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (moderate-low RMSE (error) between model estimated & measured noise levels; did not provide information on mean errors), as described by the criteria for a judgment of high risk of bias. | Assessed outcome questionnaire, constructed following the ICBEN method. | The study evaluated some but not all of the Tier I confounders, & some but not all of the other Tier II confounders, therefore is expected to introduce substantial bias | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for potential mediators such as noise sensitivity, length of residence, floor area, sound insulation |  |
| Trieu 2021 | Rating | PL | PH | PL | PH | L | L | L | L | PL | PH |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (Unclear if model estimates in 2017/2018 are comparable to measured estimates in 2014/2015 survey as did not provide an estimate of mean bias to know whether over or under-estimating (Though did find a low RMSE)), as described by the criteria for a judgment of high risk of bias. | Assessed outcome questionnaire, constructed following the ICBEN method. | The study evaluated some but not all of the Tier I confounders, & some but not all of the other Tier II confounders, therefore is expected to introduce substantial bias | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for such as noise sensitivity, green, etc. |  |
| Agarwal 2009 | Rating | PH | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups (e.g.no random sampling, in 10 selected sites only) | Uncertain how exposure information was obtained (duration, date/month etc.) | Uncertain how outcome information was obtained | The study did not account for or evaluate multiple Tier I confounders, & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Goswami, 2009 | Rating | H | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Significant details were lacking) | Uncertain how exposure information was obtained (uncertain how exposures assigned to participants) | Uncertain how outcome information was obtained | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Nanwar 2009 | Rating | H | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Significant details were lacking. | Not an accurate/objective measure of exposure, exposure assigned by a proxy variable (living in close proximity to major road) | Uncertain how outcome information was obtained | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Ristovska 2009 | Rating | PH | H | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups | Uncertain how exposure information was obtained (unclear how surveyed participants were assigned exposures, number of measurements sites, sampling scheme, duration of measurements, etc.) | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that methods were robust | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Banerjee 2009 | Rating | PH | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups | Uncertain how exposure information was obtained (equipment, temporal/spatial pattern, duration, repeat or not, sampling strategy of sites) | Low confidence in the accuracy, this did not provide details on questionnaire instrument; Uncertain how outcome information was obtained | The study did not account for or evaluate multiple Tier I confounders, & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Phan, 2010 | Rating | PL | H | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | Uncertain how exposure information was obtained (insufficient information on whether repeat 24-hour measurements were collected at sites, dates of measurement, height of SLM off the ground, distance to facades, etc). | Assessed outcome questionnaire, constructed following the ICBEN method. | The study did not account for or evaluate multiple Tier I confounders, & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Firdaus 2010 | Rating | H | H | H | H | L | L | L | L | L | H |
| Rationale | Significant details were lacking. | Not an accurate/objective measure of exposure (ascertained from self-report, and insufficient details on exposure assessment/questionnaires, etc.). | Low confidence in the accuracy, this did not provide details on questionnaire instrument; Uncertain how outcome information was obtained | The study did not account for or evaluate multiple Tier I confounders , & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Agarwal 2011 | Rating | PL | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | Uncertain how exposure information was obtained (did not provide information on spatial & temporal resolution; did not provide repeated measures) | Uncertain how outcome information was obtained | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Guoqing 2012 | Rating | PH | H | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups | There is low confidence in the accuracy of the exposure assessment methods (measurements conducted around airport, but surveyed participants could live as far as 6km away from the airport) | Insufficient information to permit a judgment of low risk of bias, but there is indirect evidence which suggests that methods were robust | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Seabi 2013 | Rating | PL | PH | PL | H | L | L | L | L | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias. Insufficient information on monitoring protocols, duration (i.e., over multiple days), and methods. | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that methods were robust | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study |  |
| Banerjee 2013 | Rating | PH | H | H | PH | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were inconsistent across groups, as described by the criteria for a judgment of high ROB. | Uncertain how exposure information was obtained (unclear about equipment, monitoring protocols, duration of monitoring at each site, whether measurements were collected on same day or different days, time of the year, etc.). | (assessed outcome use not-validated structured questionnaire) | Age, residential period (years), gender. The study evaluated some but not all of the Tier I confounders, & some but not all of the other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Daruis 2014 | Rating | PH | H | H | H | L | PH | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups | Uncertain how exposure information was obtained (unclear how far way the dwellings were from locations of measurement). | Low confidence in the accuracy, this did not provide details on questionnaire instrument; Uncertain how outcome information was obtained | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Gjestland 2015 | Rating | PL | H | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | Uncertain how exposure information was obtained (lack of information on spatial representativeness of measurements in relations to surveyed participants and spatial sampling of sites). | Assessed outcome questionnaire, constructed following the ICBEN method. | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study |  |
| Liu,2017 | Rating | PL | PL | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of low risk of bias, but there is indirect evidence that suggests that methods were robust, as described by the criteria for a judgment of low risk of bias. | Assessed outcome questionnaire, constructed following the ICBEN method. | The study did not account for or evaluate multiple Tier I confounders , & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Sieber,2018 | Rating | L | H | PL | H | L | L | L | L | L | H |
| Rationale | The descriptions of the population selection were sufficiently detailed, & adequate data were supplied on the distribution of relevant study sample & population characteristics to support the assertion that risk of selection effects was minimal. | There is low confidence in the accuracy of the exposure assessment methods (LUR model had very low predictive accuracy (r2)). | Assessed outcome questionnaire, constructed following the ICBEN method. | The study did not account for or evaluate multiple Tier I confounders , & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Onchang 2018 | Rating | PH | PH | PL | H | L | L | L | L | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups (only healthy University students were recruited; 27.0% male & 73.0% female) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (measurements collected over 1, 24hr period at each site do not represent long-term trends in area.), as described by the criteria for a judgment of high risk of bias. | Assessed outcome questionnaire, constructed following the ICBEN method. | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Wu 2019 | Rating | PL | H | H | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is low confidence in the accuracy of the exposure assessment methods (unclear how exposures were assigned to participants). | (assessed outcome use not-validated structured questionnaire) | The study did not account for or evaluate multiple Tier I confounders , & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Paiva 2019 | Rating | PH | PL | H | PL | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were inconsistent across groups (e.g. sample was selected based on their proximity to the evaluated streets. | There is insufficient information about the exposure assessment methods (insufficient information on noise mapping inputs and temporal resolution of measurements) to permit a judgment of low risk of bias, but there is indirect evidence that suggests that methods were robust (validated model with greatest different between the measured & simulated values as 3.5 dB), as described by the criteria for a judgment of low risk of bias. | Low confidence in the accuracy, this did not provide details on questionnaire instrument; Uncertain how outcome information was obtained | The adjusted model has considered age, sex, & income. did not consider tier-2 variables. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Nazneen 2020 | Rating | PH | H | H | PL | L | L | L | L | L | H |
| Rationale | Researcher select participants with specific profession, it may lead to selection bias e.g. Policemen may suffer from higher level of noise & higher mental distress. | Uncertain how exposure information was obtained (i.e., sites location, duration of measurements) | Low confidence in the accuracy. Uncertain how outcome information was obtained | The study appropriately accounted for most but not all of the Tier I confounders or used appropriate statistical techniques; & some of the other Tier II confounders using appropriate statistical techniques. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Zamorano-Gonzalez 2021 | Rating | PH | H | H | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit a judgment of high ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were inconsistent across groups, as described by the criteria for a judgment of high ROB. | There is low confidence in the accuracy of the exposure assessment methods (exposure assigned to each participant if they lived within 250m of the streets/avenues where SLM monitored, potentially introducing risk of exposure misclassification). | Uncertain how outcome information was obtained | The study did not account for or evaluate multiple Tier I confounders , & did not account for or evaluate multiple other Tier II confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Dias 2021 | Rating | PL | H | H | H | L | L | L | N/A | PL | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrolment procedures, & participation were consistent across groups (respond rate 97%; Table 1 compare self-reported health conditions for male & female population) | There is low confidence in the accuracy of the exposure assessment methods (high risk of exposure misclassification by subjective proxy measure). | Low confidence in the accuracy, this did not provide details on questionnaire instrument; Uncertain how outcome information was obtained | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript that are of interest in this review have been reported in the pre-specified way. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for potential mediators such as mental health, sleep quality etc. |  |
| Nguyen 2011 | Rating | PL | H | PL | H | L | L | L | n/a | L | H |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | Uncertain how exposure information was obtained (temporal/spatial pattern, duration, repeat or not, etc.) | Assessed outcome questionnaire, constructed following the ICBEN method | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All of the study’s pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |

## Table 8.2 Individual risk of bias assessments for the cardiovascular outcome studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Huang 2013 | Rating | PL | L | L | L | L | L | L | L | PL | PL |
| Rationale | Insufficient information to permit a judgment of low ROB, but there is indirect evidence which suggests that in/exclusion criteria, recruitment & enrollment procedures, & participation & follow-up rates were consistent across groups as described by the criteria for a judgment of low ROB | There is high confidence in the accuracy of the exposure assessment methods (using personal monitor continuously for 2 hours). | Outcomes were assessed and defined consistently across all study participants, objectively measuring heart rate variability (HRV) | This is a randomised, crossover study, important confounders such as age, gender, day of the week has been considered. SES and smoking are not necessarily needed | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for BMI.) |  |
| Paunovic 2013 | Rating | PH | PL | L | PL | L | L | L | L | PL | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence suggests participation were inconsistent across groups, response rate less than 60% | There is insufficient information about the exposure assessment methods to permit a judgment of low risk of bias (unclear whether measurements can represent long-term noise exposures and did not control for classroom location), but there is indirect evidence that suggests that methods were relatively robust (took both school and home noise exposures into account (i.e., time-activity patterns)), as described by the criteria for a judgment of low risk of bias. | Outcomes were assessed and defined consistently across all study participants, using objective measures on blood pressure levels | The adjusted mode did not account for parental SES status (tier-1 variables), but has considered other variables. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for lifestyle behaviours such as drinking soft drinks, watching television.) |  |
| Banerjee 2014a | Rating | PH | PH | PH | PL | L | L | L | L | PL | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (e.g. sample was selected from those with close proximity to traffic only) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (did not provide information on spatial/temporal resolution of emissions input data and estimated model outputs or whether models were validated with fixed-site measurement), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model did not considered SES (Tier 1) variables and did not consider some tier-2 variables such as smoking. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for lifestyle behaviours such as stress, BMI.) |  |
| Banerjee 2014b | Rating | PH | PH | PH | PL | L | L | L | L | PL | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (e.g. sample was selected from those with close proximity to traffic only) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (did not provide any information on the field measurement data used to validate models, or what the validation results of that were), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model did not considered SES (Tier 1) variables but has considered some tier-2 variables such as smoking | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for lifestyle behaviours such as stress, BMI.) |  |
| Dzhambov 2016b | Rating | PH | PH | PH | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups, as described by the criteria for high ROB (19% response rate, not random sampling) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The fully adjusted model is subjected to over-adjustment issue. |  |
| Gilani 2021 | Rating | PL | PL | PH | PL | L | L | L | L | PL | PH |
| Rationale | insufficient information to permit low ROB, indirect evidence suggests that participation were consistent across groups (e.g. Table 1 comparisons of characteristics between those included and the targeted population) | There is insufficient information about the exposure assessment methods (did not provide information on the temporal/spatial aspects related to the input traffic data) to permit a judgment of low risk of bias, but there is indirect evidence that suggests that methods were robust (high correlation of model predicted and measured data), as described by the criteria for a judgment of low risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model did not considered SES (Tier 1) variables and some tier-2 variables such as smoking | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for potential mediators such as stress levels, sleep quality etc. |  |
| Roca-Barcelo,2021 | Rating | L | PH | L | PL | L | L | L | L | L | PH |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation were sufficiently detailed to support the assertion that risk of selection effects was minimal (This is a study based on all census tracts that were partially or entirely affected by aircraft noise). | There is insufficient information about the exposure assessment methods (e.g., input data and their temporal resolutions) to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias. (Used two different models to assign exposures at different noise bands, and unclear how comparable models are). | Outcomes were assessed and defined consistently across all study participants, using reliable measures such as medical records. | The adjusted model considered age, sex and SES (Tier-1 variables) and some tier-2 variables such as area-level smoking (lung cancer as a proxy), ethnicity etc. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Lepore, 2010 | Rating | PH | H | L | PH | L | L | L | L | PH | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (e.g.no random sampling, in two selected schools only) | There is low confidence in the accuracy of the exposure assessment methods (exposures classified in a binary way: low vs high exposed school. Could be a risk of exposure misclassification (spatially) between children within schools if they are large with varying levels of exposure between classrooms as noise only measured at one point). | Outcomes were assessed and defined consistently across all study participants, objectively measuring blood pressure levels | The study evaluated some but not all of Tier I confounders, and some but not all of Tier II confounders, | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model did not adjust multiple potential confounders, as is expected to introduce substantial bias |  |
| Hamid 2019 | Rating | H | H | PH | H | L | PH | L | L | H | H |
| Rationale | There were indications from descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation, or data on the distribution of relevant study sample and population characteristics that risk of selection effects were substantial (only 50 participants from 8 areas) | Uncertain how exposure information was obtained (insufficient information on monitoring protocols and how exposures were assigned to surveyed participants). | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | This study potentially has substantial bias due to previous criteria. |  |
| Farooqi 2021 | Rating | H | H | PH | H | L | PH | L | L | H | H |
| Rationale | There were indications from descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation, or data on the distribution of relevant study sample and population characteristics that risk of selection effects were substantial (significant details were lacking) | Uncertain how exposure information was obtained (unclear how surveyed participants were assigned to exposures). | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | This study potentially has substantial bias due to previous criteria. |  |

## Table 8.3 Individual risk of bias assessments for the mental health studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Ma 2020 | Rating | L | L | PL | L | L | L | L | L | PL | PL |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation and follow-up rates were sufficiently detailed, and adequate data were supplied on the distribution of relevant study sample and population characteristics | There is high confidence in the accuracy of the exposure assessment methods (personal monitoring). | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for physical activity and air pollution |  |
| Tao 2020 | Rating | L | L | PL | L | L | L | L | L | PL | PL |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation and follow-up rates were sufficiently detailed, and adequate data were supplied on the distribution of relevant study sample and population characteristics | There is high confidence in the accuracy of the exposure assessment methods (using personal monitor continuously for 48 hours). | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for overall physical health status and the other on overall life satisfaction |  |
| Tao 2021 | Rating | L | L | PL | L | L | L | L | L | PL | PL |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation and follow-up rates were sufficiently detailed, and adequate data were supplied on the distribution of relevant study sample and population characteristics | There is high confidence in the accuracy of the exposure assessment methods (using personal monitor continuously for 48 hours). | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for physical activity and severity of the overstress symptom for the past week and availability of chronic physical diseases |  |
| Dzhambov 2017 | Rating | PH | PH | PL | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (only students were recruited) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (lacking information on spatial and temporal resolution of input data, and other model parameters), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for potential mediators such as noise sensitivity) |  |
| Dzhambov 2018a | Rating | PH | PH | PL | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (only students were recruited) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (lacking information on spatial and temporal resolution of input data, and other model parameters), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for time spent at home/day, duration of residence, and population density etc.) |  |
| Dzhambov 2014 | Rating | PH | H | PL | H | L | L | L | N/A | PH | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups | There is low confidence in the accuracy of the exposure assessment methods (high risk of exposure misclassification by subjective rating). | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Dzhambov 2018b | Rating | PH | PH | PL | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (only students were recruited) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (did not collect temporally representative measurements of long-term day-time exposures for LUR models; Small number of measurement sites for study area (~ 40), assigned day-time exposures based on residential address but university students likely spent most of the day at university), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for duration of residence, time spent at home/day, social cohesion, and physical activity etc.) |  |
| Dzhambov 2019 | Rating | PH | PH | PL | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (only students were recruited) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (did not collect temporally representative measurements of long-term day-time exposures for LUR models; Small number of measurement sites for study area (~ 40), assigned day-time exposures based on residential address but university students likely spent most of the day at university), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The final reporting model may have a potential over-adjustment issue (e.g. adjusting for population density, and university faculty) |  |
| Masoudzadeh 2017 | Rating | PH | H | PL | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (e.g. sample was selected from those with close proximity to traffic only) | There is low confidence in the accuracy of the exposure assessment methods, with very little information provided. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
| Ma 2018 | Rating | L | H | PL | L | L | L | L | L | PL | H |
| Rationale | The descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation and follow-up rates were sufficiently detailed, and adequate data were supplied on the distribution of relevant study sample and population characteristics | There is low confidence in the accuracy of the exposure assessment methods (not an accurate/objective measure of exposure and risk of recall bias). | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | Tier I important confounders such as age, gender, education has been considered. smoking are not necessarily needed | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The fully adjusted model is subjected to over-adjustment issue. |  |

## Table 8.4 Individual risk of bias assessments for the quality of sleep studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Stosic 2009 | Rating | PH | H | PH | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence suggests participation were inconsistent across groups, response rate less than 60% | Uncertain how exposure information was obtained insufficient information on duration, repeat measurements on different days, protocols, etc). | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study | The study appears to be free of other sources of bias. |  |
| Goswami 2011 | Rating | H | H | PH | H | L | PH | PH | N/A | L | H |
| Rationale | There were indications from descriptions of the source population, in/exclusion criteria, recruitment and enrolment procedures, participation, or data on the distribution of relevant study sample and population characteristics that risk of selection effects were substantial (significant details were lacking) | Uncertain how exposure information was obtained (unclear how exposures were assigned to participants and how sites were sampled). | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | Insufficient information was given to assess this | There is evidence that conflicts of interest are not capable of introducing ROB in the study | The study appears to be free of other sources of bias. |  |
| Han 2015 | Rating | PL | H | PH | H | L | L | L | L | L | H |
| Rationale | insufficient information to permit low ROB, indirect evidence suggests that participation were consistent across groups (e.g. Table 1 comparisons of characteristics between those included and the targeted population) | Uncertain how exposure information was obtained (insufficient information on number of sites, how participants exposures were assigned to these sites, as well as number of days monitored at each site). | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Ravindra 2016 | Rating | PH | PH | PH | H | L | L | L | N/A | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (study conducted in a hospital ) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study | The study appears to be free of other sources of bias. |  |

## Table 8.5 Individual risk of bias assessments for the cognitive outcomes studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Belojevic 2012 | Rating | PH | PH | PL | PH | L | L | L | L | L | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (Final sample:311, out of 1150 parents, children lived outside central Belgrade were excluded) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (unclear if they conducted repeated measurements at same sites over multiple days. Potential risk of exposure misclassification due to short-term measurement duration, and by assigning household exposures based on exposure level of street.), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The study evaluated some but not all of the important confounders (gender and SES), AND some but not all of Tier II confounders, | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Seabi 2015 | Rating | PL | PH | PL | PH | L | L | L | N/A | L | PH |
| Rationale | Insufficient information to permit low ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation and follow-up rates were consistent across groups as described by the criteria for low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (Unclear if measurements conducted on multiple days, or if classrooms were taken into account (e.g., if schools were large)), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The study evaluated some but not all of Tier I confounders, AND some but not all of Tier II confounders, | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |
|  | Rating | PL | PH | PL | H | L | L | L | N/A | L | H |
| Seabi 2012 | Rationale | Insufficient information to permit low ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation and follow-up rates were consistent across groups as described by the criteria for low ROB | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (insufficient information on monitoring protocols, duration, etc.), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit low ROB, indirect evidence which suggests that methods were robust, as described by the criteria for low ROB. | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | There is evidence that conflicts of interest are not capable of introducing ROB in the study. | The study appears to be free of other sources of bias. |  |

## Table 8.6 Individual risk of bias assessments for the others nonpriority outcomes studies.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Domain | Source population representation | Exposure assessment | Outcome assessment | confounding | blinding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias | Overall ROB |
| Dzhambov 2015 | Rating | PH | PH | PH | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups, as described by the criteria for high ROB (19% response rate, not random sampling) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (lacking information on spatial and temporal resolution of input data, and other model parameters), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The fully adjusted model is subjected to over-adjustment issue. |  |
| Dzhambov 2016a | Rating | PH | PH | PH | L | L | L | L | L | PH | PH |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups, as described by the criteria for high ROB (19% response rate, not random sampling) | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (lacking information on spatial and temporal resolution of input data, and other model parameters), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The adjusted model has considered age, sex, SES and other potentially confounding factors. | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The fully adjusted model is subjected to over-adjustment issue. |  |
| Ana 2009 | Rating | PH | PH | PH | H | L | PH | L | L | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation and follow-up rates were inconsistent across groups, as described by the criteria for high ROB. Did not provided respond rates, study population characteristic, etc. | There is insufficient information about the exposure assessment methods to permit a judgment of high risk of bias, but there is indirect evidence that suggests that methods were not robust (health analysis was conducted by comparing children in schools (not based on actual exposures) (crude exposure assignment)), as described by the criteria for a judgment of high risk of bias. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Siddiqui 2015 | Rating | H | H | PH | H | L | PH | L | L | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (hospital based study, only patients were recruited) | Uncertain how exposure information was obtained. | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only; insuffcient information on outcome assessment) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | Insufficient information was given to assess this | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |
| Ali 2018 | Rating | PH | H | PH | H | L | L | L | L | L | H |
| Rationale | Insufficient information to permit high ROB, indirect evidence which suggests that in/exclusion criteria, recruitment and enrolment procedures, and participation were inconsistent across groups (only healthy University students were recruited) | There is low confidence in the accuracy of the exposure assessment methods (Motorcycle riding and duration of riding is proxy for exposure) | Insufficient information to permit high ROB, indirect evidence which suggests that methods were not robust (e.g. using self-reported data only) | The study did not adjust for most important confounders | No blinding (as it is not an element in this type of study design) | All included participants have complete data on the studied outcomes | All pre-specified outcomes in the published manuscript. | The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study | The study appears to be free of other sources of bias. |  |

# Supplementary 9 Data extraction table

## Table 9.1 Full description of studies investigating noise effects on annoyance, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Region** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Paunovic 2014** | Cross-sectional | Belgrade, Serbia | 5861 | Final sample 5861, 3263 (55.7%) were women; mean age 42.91 ± 17.88 yrs | Fixed site SLM measurement | Annoyance | Questionnaire, constructed following the ICBEN method | Gender, age, duration of residence, bedroom window orientation, floor level, high noise sensitivity, Leq at daytime, presence of public transport at daytime, type of public transport at daytime. | Logistic regression | Significant predictors of high noise annoyance were the presence of public transport at daytime (Y/N) (OR = 1.47, 95% CI: 1.28, 1.70), and at night (Y/N) (OR = 1.39, 95% CI:1.20, 1.61). An increase of daytime noise level by 1 dBA significantly increased the probability of noise annoyance by 3% in both models | PH |
| **Song 2016** | Cross-sectional | Yueyang, China | 227 | 251questionnaires (227 valid). Mean age:57 ± 13.2, 54.6% were men | Fixed site SLM measurement & propagation-based model | Annoyance, sleep disturbance and self-reported illness (mental and CVDs） | Questionnaire adapted from Pedersen’s studies | N/A | Logistic regression | When LAeq increased per dBA, the odds for them being annoyed (highly annoyed) by wind turbine noise would increase 1.208 (95% CI: 1.060–1.378) times. | PH |
| **Nguyen 2020** | Cross-sectional | Ho Chi Minh City, Vietnam | 890(2014) 1109(2015) 1286(2015) 623(2017) 132(2018) | 68.5(2014) 85.3(2015) 98.8(2015) 95.8(2017) 83.3(2018) | Fixed site SLM measurement & propagation-based model | Annoyance | Questionnaire, constructed following the ICBEN method | Sex, age, noise sensitivity, length of residence, floor area, sound insulation | Logistic regression | For 1 dB change in Lden, noise annoyance increased 29.7% (95% CI: 1.248 1.349). The change effect due to the operational change is observed to decrease in the follow-up study with regard to annoyance and remains the same with regard to insomnia. Noise sensitivity is also a significant factor that affected the prevalence of annoyance and insomnia among the residents living around NBIA | PH |
| **Trieu 2021** | Cross-sectional | Ho Chi Minh City, Vietnam | 881 (survey2008) and 502 (survey2019) | Respond rate: 88% (survey2008) and 60% (survey2019) | Fixed site SLM measurement and noise model | Annoyance | Questionnaire, constructed following the ICBEN method | Sex, age, green and convenience evaluation of the residential areas, and noise sensitivity | Logistic regression | For 1 dB change in Lden, OR for noise annoyance was 1.25 (95% CI: 0.80 1.31). A gradual increase in the annoyance due to aircraft noise was not found. Results of multiple regression analysis indicated that differences in the reaction of the residents to noise in the two studies were significantly attributed to non-acoustic factors. | PH |
| **Agarwal 2009** | Cross-sectional | Jaipur City, India | 450 | Age and sex not mentioned | Fixed site SLM measurement | Annoyance | Questionnaire, MDS (Mean Dissatisfaction Score) | N/A | Correlation | 60% people were disturbing by noise and were aware adverse health effects of noise pollution. | H |
| Nguyen 2011 | cross-sectional | Hanoi/Ho Chi Minh City, Vietnam | Ho Chi Minh City: 1562 Hanoi:1397 | Age: 89% in Ho Chi Minh City and 92% in Hanoi range from 20 to 60 | Fixed site SLM measurement | Annoyance | Questionnaire, constructed following the ICBEN method | City (Hanoi/Ho Chi Minh City) | Multivariate logistic regression | For per IQR increase in Lden, annoyance increase by 14.8% (95% CI: 0.128, 0.168) | H |
| **Goswami, 2009** | Cross-sectional | Balasore, India | 212 | Age and sex not mentioned | Fixed site SLM measurement | Annoyance | Structured questionnaire (not specific) | N/A | Descriptive analysis | 63% respondents were not satisfied about the noise level in their places in Balasore. | H |
| **Nandanwar 2009** | Cross-sectional | Nagpur City, India | 378 | Aged 15-55 yrs, sex not mentioned | Proximity to major road as proxy | Annoyance, headaches, Nervousness, hearing damage | Structured questionnaire (not specific) | N/A | Descriptive analysis | More than half of the total sample population around five major intersections in Nagpur city expressed annoyance with traffic noise during daily activities. Of these, 29 % were “extremely” and 24 % “very much” annoyed, followed by 22 % to “some extent”, 19 % “ little” and 6% “none” annoyed | H |
| **Ristovska 2009** | Cross-sectional | Skopje, Macedonia | 510 | Final sample 510, 47.65 % men, mean age 37.34 yrs | Fixed site SLM measurement | Annoyance | Questionnaire according to the ISO/TS-15666 (8) | N/A | Correlation | Significant association between exposure to Lday in the range 61 db(A) to 65 db(A) and annoyance in the exposed population (chi-square =86.14; p<0.001 | H |
| **Banerjee 2009** | Cross-sectional | Asansol, India | 869 | Age and sex not mentioned | Fixed site SLM measurement | Percent highly annoyed | Questionnaire (Noise and Attitudinal Survey) | N/A | Correlation | Across different population densities the heavy vehicles (truck/bus) was observed to be most annoying source with higher values in the 3,000–20,000 population density level, constituting 66.75% of respondents. | H |
| **Phan, 2010** | Cross-sectional | Hanoi/Ho Chi Minh City, Vietnam | Hanoi:1503;Ho Chi Minh City:1471 | Aged 18-60+ yrs, Hanoi (male 48.7%) Ho Chi Minh City (male: 49.6%) | Fixed site SLM measurement | Annoyance; attitude/activities disturbance | Questionnaire, constructed following the ICBEN method | Frequency of window-opening, window orientation, city | Logistic regression | Road traffic noise characterize: high noise exposure and frequent, impulsive horn sounds; noise annoyance among Vietnamese not as severe as expected | H |
| **Firdaus 2010** | Cross-sectional | Delhi, India | 1693 | Age and sex not mentioned | Subjective rating of noise level | Annoyance, disturbance in sleep/communication | Questionnaire (not specific) | N/A | Correlation | In high-density zone, 28.16% respondents reported annoyance; 27.15% reported interference with communication and 2.72% are concerned with depression/mood swings/indigestion/hypertension, caused by the noise levels. | H |
| **Agarwal 2011** | Cross-sectional | Jaipur, India | 550 | Males: 35- 56%, females 44-64% at all sites. Age: 15- 65 yrs | Fixed site SLM measurement | Annoyance; Irritation; Headache; Hypertension; Loss of sleep | Structured questionnaire (not specific) | N/A | Linear regression | Noise annoyance correlated with different noise parameters. | H |
| **Guoqing 2012** | Cross-sectional | Hangzhou, China | 764 | 764 valid questionnaires, age and sex not mentioned | Fixed site SLM measurement | Annoyance | Questionnaire (social and Socio-Acoustics surveys) | N/A | Descriptive analysis | If LWECPN was 64.3 db (LDN was 51.4 db), then 15% respondents were highly annoyed. If LWECPN was 68.1 db (LDN was 55.0 db), then 25% respondents were highly annoyed. The annoyance threshold of aircraft noise (LWECPN) was 73.7 db, while the annoyance threshold of a single flight incident instantaneous noise level (lamax) was 72.9 db | H |
| **Seabi 2013** | Cohort | Durban International Airport, South Africa | 732 | 732 children mean age:11.1 at 2009. 649 (mean: 12.3) and 174 (mean: 13.3) reassessed in 2010 & 2011 | Fixed site SLM measurement | Annoyance, general health | Biographical Questionnaire; noise Annoyance: four-point Likert scale; GHQ | N/A | Multivariate analyses of variance, manovas | Aircraft noise exposure did not have adverse effects on the children’s self-reported health outcomes. Taken together, these findings suggest that chronic exposure to aircraft noise may have a lasting impact on children’s annoyance, but not on their subjective health rating. | H |
| **Banerjee 2013** | Cross-sectional | Asansol, India | 221 | Aged 19-59 yrs, quiet area (male:56.1%), noisy area (male: 42.11%) | Fixed site SLM measurement | Annoyance | Structured questionnaire (not specific) | Age, residential period (years), gender | Logistic regression | Self-reported annoyance was marked at levels above 65-70 db (A). A 67.5 db (AA) is suggested as a threshold level. The association was statistically significant for female subjects with the adjusted odds ratio being 2.35 (95% CI: 0.99-5.58). | H |
| **Daruis 2014** | Cross-sectional | Terengganu, Malaysia | 114 | 53 males, adults | Fixed site SLM measurement | Annoyance | Structured questionnaire (not specific) | N/A | Descriptive analysis | 61% felt that their residential area is noisy at night. Self-reported respondents' disturbance and depression of the level of traffic noise were also measured. Result only percentage. Most respondents reported that noise was extremely high in their residences at night and they felt disturbed and depressed about it every day. | H |
| **Gjestland 2015** | Cross-sectional | Five major cities, Vietnam | 7199 | 50%-98% | Fixed site SLM measurement | Annoyance | Questionnaire, constructed following the ICBEN method | N/A | Descriptive analysis | It was shown that the Vietnamese population was more tolerant to road traffic noise (5–10 db higher road traffic noise for the same annoyance) than the European and North American population, but similarly annoyed by aircraft noise | H |
| **Liu,2017** | Cross-sectional | Zhejiang,China | 1027 | Adult, mean age of 41.8, 52.97% were male | Fixed site SLM measurement | Annoyance | Questionnaire, constructed following the ICBEN method | N/A | Quadratic polynomial regression | The percentage of highly annoyed people affected increasing from 15%-20% to 30%-40% over the range. There also different levels of annoyance depending on the time of day, and the location and activities of those affected | H |
| **Sieber,2018** | Cross-sectional | Western Cape province, South Africa | 364 | Mean age 40.3± 9.7, 562 out of 600 eligible participant, final analysis sample is 364, 7.7% were male | LUR model | Noise sensitivity and noise annoyance | Questionnaire, constructed following the ICBEN method | N/A | Logistic regression | Compared to Switzerland, higher percentages of highly noise sensitive individuals (women:35.1% vs 26.9%; men: 25% vs 20.5%) and people highly annoyed to road traffic noise (women: 20.5% vs 12.4%; men: 17.9% vs 11.1%) were observed in South Africa. While in South Africa women were more annoyed to neighbourhood noise than in Switzerland (21.1% vs 9.4%), not the case among men. Multivariable logistic regression models showed that in both countries men tended to be less sensitive and less annoyed by noise. | H |
| **Onchang 2018** | Cross-sectional | Silpakorn, Thailand | 786 students | Aged 21.16 ± 1.44 off-campus (male:28.7%),on-campus (male: 24.7%) | Fixed site SLM measurement | Annoyance; GPA | Questionnaire, constructed following the ICBEN method | N/A | Logistic regression | GPA in the off-campus group compared to those for on campus residents with OR values ranging from 1.049 to 1.164. The most important noise-impacted factors affecting off-campus students’ cumulative GPA were reading and mental tasks (OR = 2.801). Rest disturbance had a positive influence on cumulative GPA for on-campus students. | H |
| **Wu 2019** | Cross-sectional | Guangzhou, China | 528 | 528 valid questionnaires collected, aged 10-80 yrs, 53% male | Fixed site SLM measurement | Annoyance; physical comfort | Structured questionnaire (not specific) | N/A | Linear regression | Around 70% of participants consider the traffic noise has negative effect, and about 60% of participants consider the noise has moderate or much higher impact on physical comfort. Around 65% of participants consider the noise had moderately or much higher impact on their psychological comfort | H |
| **Paiva 2019** | Cross-sectional | Brazil,São Paulo | 225 | Mean age 49.0 ±1.26) yrs old, 58.2% female | Propagation-based model | Annoyance, perception | Structured questionnaire (not specific) | Sex, age and income | Logistic regression | Noise-related annoyance was reported by 48.4% of the respondents. Associations were observed between living in areas exposed to traffic noise and feeling annoyed with this noise (p < 0.001). | H |
| **Nazneen 2020** | Cross-sectional | Peshawar, Pakistan | 2500 | Age and sex not mentioned | Fixed site SLM measurement | Headache, exhaustion, & psychological Symptoms | Questionnaires, overall opinion about noise developed by Tripathy (1999) | Age, gender, job | Structure equation modelling | Different models were prepared and a modified model obtained the acceptable model fit, i.e., chi-square 0.093, χ2/df 1.286, comparative fit index 0.986, goodness of fit index 0.966, normed fit index 0.943, Tucker-Lewis index 0.977, and root mean square error of approximation 0.034. The modified model gives not only the information about direct but also indirect effects of noise on the exposed population | H |
| **Zamorano-Gonzalez 2021** | Cross-sectional | Matamoros, Mexico | 2350 | 2,350 people were participated, aged 18-75 yrs, 41.4% were men | Fixed site SLM measurement | Noise perception | Structured questionnaire (not specific) | N/A | Pearson correlation | The overall perception of traffic noise annoyance identified that 1,131 participants (48.1%) responded “Yes” as they considered the noise annoying. Participants who responded “No” as well as those who responded “Do not know” resulted in a total of 1,219 people (51.9%). | H |
| **Dias 2021** | Cross-sectional | Brazil,Belo Horizonte | 3934 | 59.2% women (mean age 44.8 ±16.7), & 40.8% men (mean age 43.2 ±17.0) | Subjective rating of noise level | Perception | Structured questionnaire (not specific) | 0 to 4 yrs of schooling, single marital status, good traffic, no loud music, discussions, or late-night parties, no gunfire, good sleep quality, very good and good self-rated health, and no report of migraine, depression, mental disorder, or digestive disorder | Generalized Linear Latent and Mixed Models | The prevalence of noise annoyance was 47% for women and 39.8% for men. For both genders, noise annoyance was independently associated with bad traffic and the presence of loud music, discussions, and late-night parties | H |
| \*ACN: Aircraft noise; DT: Daytime; GHQ: General Health Questionnaire; GPA: grade point average; ICBEN: International Commission on Biological Effects of Noise; RTN: road traffic noise; LWECPN: weighted effective continuous perceived noise levels; N/A: not applicable; NBIA: Hanoi Noi Bai International Airport; NT: night-time; QoS: Quality of sleep; SLM: sound level meter; LUR: Land use regression. | | | | | | | | | | | | |

## Table 9.2 Full description of studies investigated noise effects on CVDs, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Country** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Huang 2013** | Crossover | Beijing, China | 40 | Aged 19-32 yrs, 17 males | Personal SLM monitoring | HRV | Measurement: Ambulatory electrocardiogram monitoring, (model MGY-H7). | Gender, age, BMI, hour of day, day of week, location, real-time temperature (Temp) and relative humidity (RH) | Mixed-effects regression models | Decrease of pNN50 and HF and increase of LFHFR were associated with per 1 dBA increment in noise (15-min moving average), with changes of −3.10% (95% CI, −4.56% to −1.62%), −1.71% (95% CI, −3.03% to −0.36%), and 2.49% (95% CI, 1.14–3.85%), respectively. | PL |
| **Paunovic 2013** | Cross-sectional | Belgrade, Serbia | 1113 | Children aged 7-11 yrs, 521 boys | Fixed site SLM measurement | BP | Measurement | Noise levels at schools at daytime, noise levels at home at nighttime, age, gender, BMI, heart rate, mother's education, family history of hypertension, eating snacks, watching television, orientation of child's room toward the street, windows glazing at home, and windows glazing at school. | Mixed linear model | Children attending schools with public transport running nearby had by 1.3 mmhg higher systolic pressure than did children from schools without public transport. | PH |
| **Banerjee 2014a** | Cross-sectional | Asansol, India | 909 | Response rate: 90.9 %, 533 females (Mean age: quiet area: 41.6) & 376 males (mean age quiet area: 43.2) | Propagation-based model | Coronary heart disease | Structured questionnaire (not specific) | Age, residence period, BMI, self-reported stress, bedroom window orientation, neighbourhood satisfaction, noise sensitivity, and presence of other chronic diseases. | Logistic regression | The adjusted OR for self-reported CHD was 1.72 (95 % CI: 1.36, 2.19) per 5 dB(A) increase of Lden (range 55–80 dB(A)). A gender-related risk difference was observed among male (OR 1.47 (1.07–2.02)) and female (OR 1.83 (1.27–2.65)) respondents. | PH |
| **Banerjee 2014b** | Cross-sectional | Asansol, India | 909 | Same as Banerjee 2014a | Propagation-based model | Hypertension | Structured questionnaire (not specific) | Age, BMI, other chronic disease, family history of cardiovascular disease for both genders, was adjusted for self-reported mental stress/tension, bedroom window orientation, and smoking habits for males. | Logistic regression | The adjusted OR for self-reported hypertension was 1.99 (95 % CI: 1.66, 2.39) per 5 db(A) increase of Lden (range 55.1–77.9). A gender-related risk difference was observed among the male (OR 1.81 (1.42–2.31)) and female (OR 2.18 (1.66–2.88)) respondents. | PH |
| **Dzhambov 2016b** | Cross-sectional | Plovdiv, Bulgaria | 513 | Response rate: 88.30%, mean age no IHD 34.71; IHD: 59.45, male no IHD: 33.67%; IHD: 68.57% | Propagation-based model | Ischemic heart disease  (IHD) | Self-reported doctor-diagnosis with IHD Questionnaire (Not specific) | Age, gender, ethnicity, SES, educational attainment, pack-years of smoking, lifetime occupational noise exposure, distance to major road, duration of residency, BMI, T2DM, arterial hypertension, noise sensitivity, sleep disturbance and bedroom location | Log-link Poisson regression | Lden ≥ 65 dB was associated with higher risk (RR=1.84, 95% CI: 0.61, 5.57) of IHD in long-term residents (≥ 20 years) | PH |
| **Gilani 2021** | Cross-sectional | Baramulla, India | 909 | Adults, Male:376, Females:533 | Propagation-based model | Noise sensitivity, self-reported stress levels, qos, coronary artery disease | Questionnaire, Weinstein’s Noise Sensitivity Scale | Stratified by gender; adjusted for age, body mass index, stress levels, residence period, chronic disease, sensitivity, physical activity, bedroom window orientation, sleep quality; | Logistic regression | Residents living in noisy areas were found to have a 2.25 times higher risk per 5 dB(A) increment in the noise levels (95% CI = 1.38 to 3.67). Males were at a higher risk of CAD (OR = 2.61; 95% CI = 1.84 to 3.72) as compared to females (OR = 2.07; 95% CI = 1.37–3.13). | PH |
| **Roca-Barcelo,2021** | Ecological | São Paulo‘s Congonhas airport, Brazil | 1,482,276 | Aged over 20 yrs, inhabitants exposed to aircraft noise > 50 db | Propagation-based model | CVD, stroke and CHD mortality | Calculated by census tract, according to the residential address at time of death | Age, sex, calendar year and area-level covariates including socioeconomic development, ethnicity, smoking and road traffic related noise and air pollution | Poisson regression model | After accounting for all covariates, areas exposed to the highest levels of noise (> 65 db) showed a relative risk (RR) for CVD and CHD of 1.06 (95% CI: 0.94; 1.20) and 1.11 (95% CI: 0.96; 1.27), respectively, compared to those exposed to reference noise levels (≤50 dB). The RR for stroke ranged between 1.05 (95% CI: 0.95, 1.16) and 0.91 (95% CI: 0.78;1.11) for all the noise levels assessed. | PH |
| **Lepore, 2010** | Cross-sectional | Pune, India | 189 | Children in 3rd & 6th grade, noisy (51.6% males) & quiet school (52.1% males) | Fixed site SLM measurement | SBP, DBP | Measurement | Sex, age, BMI, testing problems, and resting DBP | ANCOVA test | Relative to quiet-school children, noisy-school children had significantly lower increases in blood pressure when exposed to either an acute noise or non-noise stressor. | H |
| **Hamid 2019** | Cross-sectional | Lahore, Pakistan | 50 | Age and sex not mentioned | Fixed site SLM measurement | Respiratory diseases, skin problems, cardiovascular problems, hearing impairment | Questionnaire (not specific) | N/A | Descriptive analysis | Existence of various diseases among respondents due to deteriorating air quality and elevated noise levels | H |
| **Badihian 2020** | Cross-sectional | Iran, all provinces except Qom province | 14274 | A sample of 14,400 students, aged 7-18 years. Final sample 14,274 students were completed, mean age 12.28, 51% boys | Noise annoyance | Psychological distress (questionnaires) BP values were measured | Questionnaires, from the Global School-based Student Health Survey (GSHS); measurement | Age, gender, weight status, parental and family history of hypertension, sleep duration, added salt to table food, sedentary behaviours, passive smoking, cigarette and hookah smoking, waist circumference, history of low birthweight, and breastfeeding | Linear regression | DBP and MAP had positive correlations with noise annoyance (0.028, 95 % CI: 0.005 - 0.05 & 0.025, 95 % CI: 0.002, 0.04, respectively). Participants with higher psychological distress were 15 % more likely to experience abnormally high BP compared to those with normal psychological status or mild distresses (OR: 1.15, 95 % CI: 1.003, 1.34). | H |
| **Farooqi 2021** | Cross-sectional | Chiniot and Jhang, Punjab, Pakistan | 200 | Age and sex not mentioned | Fixed site SLM measurement | Hypertension, headache, sleep disturbances, and dizziness, tinnitus | Questionnaire (not specific) | Age, sex | Correlation | Respondents in Jhang and Chiniot suffered from many noise-related health problems such as annoyance (53 and 51%), depression (45 and 47%), dizziness (61 and 65%), headache (67 and 64%), hypertension (71 and 56%), hearing loss (53 and 56%), physiological stress (65 and 65%), sleeplessness (81 and 84%), and tinnitus (70 and 62%) due to noise, respectively. | H |
| \*BP: blood pressure; CAD: coronary artery disease; CHD: Coronary heart disease; CVD: cardiovascular disease; DBP: diastolic blood pressure; GHQ: General Health Questionnaire; GSHS: Global School-based Student Health Survey; HF: high frequency; HRV: Heart Rate Variability; IHD: Ischemic heart disease; LFHFR: low-to-high frequency power ratio; MAP: mean arterial blood pressure; OR: odds ratio; pNN50: percentage of differences between adjacent normal RR intervals that are larger than 50 ms; QOS: quality of sleep; SBP: systolic blood pressure; SLM: sound level meter; WNSS: Weinstein Noise Sensitivity Scale | | | | | | | | | | | |

## Table 9.3 Full description of studies investigated noise effects on mental health outcomes, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Region** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Ma 2020** | Cross-sectional | Beijing, China | 97 | Respond rate 82.91%, aged 40.58, Male: 45.36% | Personal SLM monitoring | Mental health | Questionnaires (Not specific) | Gender, age, household income, air pollution exposure and physical activity | Logistic regression | Individual-level noise exposure based on their space-time behaviours over a 24-h period (Leq,24h) was significantly and negatively associated with residents’ self-reported mental health, in both weekday (−0.93, 95% CI: −1.85, −0.02) and weekend models (−1.89, 95% CI: −3.39, −0.38). | PL |
| **Tao 2020** | Cross-sectional | Beijing, China | 101 | Respond rate 86.32% Male:47.5age 30–49: 62.4% | Personal SLM monitoring | Annoyance, psychological stress | Geographic Ecological Momentary Assessment | Gender, age, hukou status, marital status, employment status, household monthly income, and child availability in the family, activity locations, survey date, and health status | Structural equation model | With each additional increase of one standard deviation of measured noise level, the momentary stress level would rise by 7.2% through the mediation of momentary NA (95% CI: 0.058, 0.086) | PL |
| **Tao 2021** | Cross-sectional | Beijing, China | 101 | 47.5% male, 62.4% aged 30–49 yrs | Personal SLM monitoring | Stress | Geographic Ecological Momentary Assessment | Gender, age, education status, household registration status, monthly household income, lifestyles, and short- and long-term health statuses | Multilevel ordinal regression models | LAeq was positively associated with momentary stress levels but not statistically significant after controlling for several sets of covariates (OR = 1.15; 95% CI = (0.95,1.56) | PL |
| **Dzhambov 2017** | Cross-sectional | Plovdiv, Bulgaria | 399 | Aged 15-25 yrs, Mean age 17.89 yrs, 67.9% male | Propagation-based model | Annoyance, mental health | Questionnaire, constructed following the ICBEN method | Sex, age, ethnicity, socioeconomic status, and noise sensitivity | Ordinary least squares regression & structural equation modelling | There was no direct path linking Lden to GHQ-12 (β standardized = 0.004; 95% CI: − 0.08, 0.09; p = 0.928); The pathway “Lden → noise annoyance → GHQ-12” accounted for 68.5% of the total indirect path and was statistically significantly larger than the other paths (p < 0.05). | PH |
| **Dzhambov 2018a** | Cross-sectional | Plovdiv, Bulgaria | 399 | Aged 17.89, ranging 15–25 yrs Male: 67.9% | Propagation-based model | Mental health | General Health Questionnaire (GHQ-12) | Sex, age, ethnicity, noise sensitivity, socioeconomic status, educational institution, month of data collection, air pollution, presence of water bodies in the respective buffer, time spent at home/day, duration of residence, and population density within the 1000-m buffer | Least squares linear regression | Living in a neighbourhood deprived of trees (<5.84%) enhanced the negative effect of noise, whereas in neighbourhoods with higher tree cover density noise had no effect | PH |
| **Dzhambov 2018b** | Cross-sectional | Plovdiv, Bulgaria | 720 | Aged 21±3.0, ranging 18–35 yrs Male: 243 (33.8%) | LUR model | Mental health | General Health Questionnaire (GHQ) | Age, sex, ethnicity, individual-level economic status, duration of residence, time spent at home/day, population, month of data collection, potential mediators: environmental annoyance, sleep disturbance, restorative quality, social cohesion, and physical activity | Structural equation modelling | Observed association between higher LAeq and GHQ. LAeq was associated with ↑ annoyance, and through it with ↓ restorative quality, and then in turn with↓ physical activity, and thus with ↑ GHQ | PH |
| **Dzhambov 2019** | Cross-sectional | Plovdiv, Bulgaria | 437 | 620 invited, 581 agreed, excluded 52 and 92 did not live in Plovdiv, final sample 437 (70%) | LUR model | Mental health, depression and anxiety symptom | PHQ-9 & the GAD-7 scale, NA: a 5-point verbal scale; NS: Noise Sensitivity Scale Short Form (NSS-SF) | Gender, age, nationality, income adequacy, population density, and university faculty | Linear and logistic regression | LAeq was indirectly associated with higher PHQ-9/GAD-7 scores through higher NA, but only in the low NS group. The relationship between LAeq and NA was stronger in students reporting depression/anxiety. While high NS was associated with high NA even at low noise levels, LAeq contributed to NA only in students low on NS. | PH |
| **Dzhambov 2014** | Cross-sectional | Plovdiv, Bulgaria | 182 | Mean age 36.93 yrs, male 54.95% | Subjective rating | Mental health displaced aggression | Displaced Aggression Questionnaire (DAQ) | N/A | Linear regression | Hearing noises above the perceived normal threshold, higher noise sensitivity and continuous noises were associated with higher levels of DA (p<0.05). Low frequency and high intensity noises were also associated with higher DA scores. | H |
| **Masoudzadeh 2017** | Cross-sectional | Sari, Iran | 200 | Age and sex not mentioned | Fixed site SLM measurement | People’s mental health status | Questionaire, conducted by GHQ-28 scale and used for screening the mental disorders | N/A | T-test | Complaints about physical difficulties in traders’ group working in non-polluted environment are lower than traders’ group working in an environment exposed to noise pollution (P=0.036). Average depression including residents and traders in both non-polluted areas and areas exposed to pollution had no difference (P = 0.299) | H |
| **Ma 2018** | Cross-sectional | Beijing, China | 1125 | Respond rate:87.90% age range: 18-65 | Subjective rating | Mental health | Invalidated questionnaires | Age, gender, income, education, employment, marital status, residential status or migrants, housing tenure, housing satisfaction, traffic congestion and community attachment, distances to the nearest subway station, main road, restaurant and park | Bayesian multilevel logistic model | Perceived higher noise-pollution exposure is significantly associated with worse mental health, while physical environment variables seem to contribute little to variations in self-reported mental disorders, except for proximity to the main road. | H |
| **\***GAD-7: Generalised Anxiety Disorder Assessment; GHQ: General Health Questionnaire; ICBEN: International Commission on Biological Effects of Noise; LUR: land use regression; NA: noise annoyance; N/A: not applicable; NS: noise sensitivity; PHQ-9: Patient Health Questionnaire-9; SLM: sound level meter; LUR: Land use regression. | | | | | | | | | | | |

## Table 9.4 Full description of studies investigated noise effects on quality of sleep, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Region** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Stosic 2009** | Cross-sectional | Niš, Serbia | 911 | Adult, male: 42.6 % were men | Fixed site SLM measurement | QoS | Structured questionnaire (5 point Likert scales) | N/A | Correlation | Average noise sensitivity score of noisy area residents was significantly higher than of quiet area residents [(89.3±17.7) vs. (79.7±20.3), respectively; Mann-Whitney U test, Z=5.46, p<0.001]. | H |
| **Goswami 2011** | Cross-sectional | Balasore, India | 136 | Age and sex not mentioned | Fixed site SLM measurement | QoS, cardiovascular problems, vomiting, hypertension, restlessness etc. | Structured questionnaire (detail unknown) | N/A | Descriptive analysis | 53% respondents have identified that headache was the main health problem in them due to noise and 21% respondents have reported to have visited Otorhinolaryngologists for health advice. | H |
| **Han 2015** | Cross-sectional | Luzhou, China | 381 | Aged 7- 93 yrs, 159 male and 222 female | Fixed site SLM measurement | QoS, BP, subjective feeling of nervous system damage, attention | Pittsburgh Sleep Quality Index (PSQI) | N/A | ANOVA, Wilcoxon signed rank test and Student test | Sleep quality, prevalence of hypertension, and attention in transportation hub areas were significantly different from those in the commercial, construction, residential areas (P <0.05); only 24.46% of people knew the health hazards associated with noise; 64.57% of residents have adapted to the current noise environment | H |
| **Ravindra 2016** | Cross-sectional | North India | 100 | Over 18 yrs, sex not mentioned | Fixed site SLM measurement | Headache, QoS, hypertension | Questionnaire, constructed following the ICBEN method | N/A | Descriptive analysis | Most all the respondents (97%) regarded traffic as the major source of noise. About three-fourths (74%) reported irritation with loud noise whereas 40% of respondents reported headache due to noise. Less than one-third of respondents (29%) reported loss of sleep due to noise and 8% reported hypertension, which could be related to the disturbance caused due to noise. | H |
| \*BP: blood pressure; ICBEN: International Commission on Biological Effects of Noise; QoS: quality of sleep; SLM: sound level meter. | | | | | | | | | | | |

## Table 9.5 Full description of studies investigated noise effects on cognitive outcomes, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Region** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Belojevic 2012** | Cross-sectional | Belgrade, Serbia | 311 | Aged 7-11 yrs, 146 boys | Fixed site SLM measurement | EF | Questionnaire, five item scale adapted from the Attention Deficit Disorder Questionnaire | Gender, SES | Linear regression | No effects of ambient noise levels on EF; however, a significant interaction with sex was observed, which indicated adverse noise impacts on EF in boys | PH |
| **Seabi 2015** | Cohort | Durban International Airport, South Africa | 732 | 732 learners (mean age=11.1) yrs in 2009; 650 learners (mean age=12.3 yrs, 50.1% males) reassessed in 2010; 178 learners (mean age=13.1 yrs, 52% males) reassessed in 2011 | Fixed site SLM measurement | Reading comprehension | Questionnaire, Reading comprehension was measured with the Suffolk Reading Scale Level 2.22 | Gender, language and deprivation | Univariate analysis and repeated multiple analyses of covariance | The noise level taken in 2010 and 2011 was lower for both formerly exposed schools (Leq: 55.2; Lamax: 60.8–71.2) and quiet schools (Leq 50.5–57.9, Lamax 60.6–70.5). Did not observe significant associations of noise effects with reading comprehension | PH |
| **Seabi 2012** | Cross-sectional | Durban International Airport, South Africa | 437 | For 834 participants, mean age 11.9 yrs, ranging 9 -14 yrs, 322 (39%) males and 331 (40%) females. Out of 834, 141 excluded, final sample 693, | Fixed site SLM measurement | Reading comprehension | Questionnaires (Biographical questionnaire; Suffolk Reading Scale Level 2; Figure Analogies subtest of the Quantitative battery for Cognitive Abilities Test) | N/A | Linear regression | A lower reading comprehension for those 437 (52%) learners who were exposed to chronic aircraft noise (Leq: 63.5–69.9 dba, Lamax: 89.8–96.5 dba), compared to 337 (48%) learners from relatively quiet areas (Leq: 54.4–55.3; Lamax: 73.2–74.3) | H |
| \*EF: Executive Function; SLM: sound level meter; | | | | | | | | | | | |

## Table 9.6. Full description of studies investigated noise effects on other non-priority outcomes, ordered by risk of bias and year.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Design** | **Region** | **Population** | **Population characters (age & sex) (n/%)** | **Exposure assessment (type)** | **Health outcomes** | **Source of outcome data** | **Confounder** | **Statistical analysis** | **Key findings** | **Over all ROB** |
| **Dzhambov 2015** | Cross-sectional | Plovdiv, Bulgaria | 513 | Mean age: 36.45, female (63.94%) | Propagation-based model | Self-reported body mass index (BMI) | Questionnaire short form (NoiSeQSF) | Age, gender, ethnicity, SES, distance to major road, noise sensitivity, noise annoyance, sport, diet, smoking and sleep disturbance | Linear regression | Total effects were 0.48 kg/m2 (95% CI: 0.15-0.81) increase in BMI per 5 dB Lden in the whole sample and 0.63 kg/m2 (95% CI: 0.01-1.24) among long-term residents (20 years). The relative risks of obesity were 1.03 (95% CI: 1.01-1.05) among all participants and 1.05 (95% CI: 1.01-1.09) among long-term residents. | PH |
| **Dzhambov 2016a** | Cross-sectional | Plovdiv, Bulgaria | 513 | Aged 18-83 yrs, mean age: 36.45, 36.1% were males. | Propagation-based model | T2DM | Questionnaire, T2DM status was determined by self-reported doctor diagnosis | Age, gender, ethnicity, BMI, family history of T2DM, noise sensitivity, air pollution annoyance, noise annoyance, bedroom location, sleep disturbance, pack-years of smoking | Multivariate logistic and log-linear poisson regressions | T2DM was positively associated with exposures to l 71-80 dB (OR = 4.49, 95% CI: 1.38, 14.68), pm2.5 25.0-66.8 ug/m3 (OR = 1.32, 95% CI: 0.28, 6.24), benzo alpha pyrene 6.0-14.02 ng/m (OR = 1.76, 95% CI: 0.52, 5.98) and high road traffic (OR = 1.40, 95% CI: 0.48, 4.07). | PH |
| **Ana 2009** | Cross-sectional | Ibadan, Nigeria | 400 | Teenagers grades 10–12 (above 14 yrs of age), sex not mentioned | Fixed site SLM measurement | Tiredness, lack of concentration, irritability, hearing loss, deafness | Questionnaire (open and closed question) | N/A | Descriptive analysis | Over 60% of respondents reported that vehicular traffic was major source of noise, and over 70% complained being disturbed by noise. Three schools reported tiredness, and one school lack of concentration, as the most prevalent noise-related health problems | H |
| **Siddiqui 2015** | Cross-sectional | Karachi, Pakistan | 125 | 18 -47 yrs, males 84% (105) & females 16% (20) | Fixed site SLM measurement | Tinnitus, symptoms of NIHL, and community response to noise, annoyance | N/A | N/A | Descriptive analysis | Subjects exposed to noise for more than 12 hours per day were 36.8%. varying degree of hearing loss was evaluated in subjects where 17.6% were normal, 33.6% had mild hearing loss, 45.6% had moderate and 3.2% had moderately severe hearing loss. Traffic noise was found to bother 55.2% of subjects. | H |
| **Ali 2018** | Cross-sectional | Selangor, Malaysia | 301 motorcyclists | Aged 19- 25 yrs, 63.1% male | Motorcycle riding, and duration of riding as proxy for exposure | HRQOL | WNNS;WHOQOL-BREF | N/A | ANOVA | No significant difference in ns between males and females. Motorcycle driving experience for <4 years displayed a higher tendency toward ns. A significantly (p = 0.004) decreasing trend among low, moderate, and high ns with their respective HRQOL while a high ns showed significantly (p = 0.015) lower scores on the social domain of the QOL | H |
| **\***BMI: body mass index; GHQ: General Health Questionnaire ; N/A: not applicable; NIHL: Noise-induced hearing loss; NoiSeQSF: Noise Sensitivity Questionnaire Short Form; NS: noise sensitivity; QoL: Quality of life; SLM: sound level meter; T2DM: type 2 diabetes mellitus; WHOQOL-BREF: World Health Organization Quality of Life-BREF; WNNS: Weinstein noise sensitivity scale | | | | | | | | | | | |

# Supplementary 10 Assessment of downgrading factors for quality of evidence.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Health outcome** |  | **Risk of bias** | **Indirectness** | **Inconsistency** | **Imprecision** | **Publication bias** |
| **Annoyance** | **Rating** | -2 | 0 | 0 | -1 | 0 |
|  | **Rationale** | 21/25 studies were rated ‘high’ and 4/25 were ‘probably high’, mainly due to lack of confounding control and Exposure assessment. | Most studies examine adults or older adults. | Most studies show higher percentage of highly annoyance with noise exposure | most studies conducted simple correlation analysis rather than association type analysis; some CIs are very wide | No direct evidence of publication bias, other than those typically of influence |
| **Cardiovascular outcomes** | **Rating** | -2 | 0 | -1 | -1 | 0 |
|  | **Rationale** | 4/11 were rated ‘high’ and 7/11 studies were rated ‘probably high’, mainly due to population selection/exposure assessment | Most of these studies include the general population, though some focussed on children (7-18 years) | Heterogeneous in results – different cardiovascular outcome investigated | Some CIs are very wide | No direct evidence of publication bias, other than those typically of influence |
| **Mental health** | **Rating** | -1 | 0 | 0 | -1 | 0 |
|  | **Rationale** | 3/10 were rated ‘high’ and 4/10 were rated ‘probably high’ mainly due to mainly due to population selection/exposure assessment | Most of these studies include adults in the general population, though some focussed on college student (15-25 years) | Studies show variation in increased risk/null effect with noise | Some CIs are very wide | No direct evidence of publication bias, other than those typically of influence |
| **Quality of sleep** | **Rating** | -2 | 0 | 0 | -1 | 0 |
|  | **Rationale** | 4/4 were rated ‘high’, mainly due to lack of confounding control | Most of these studies include adults in the general population, one study not mentioned age/gender | Most studies show higher percentage of sleep disturbance with noise exposure | most studies conducted simple descriptive/correlation analysis rather than association type analysis | No direct evidence of publication bias, other than those typically of influence |
| **Cognitive outcomes** | **Rating** | -2 | -1 | -1 | -1 | 0 |
|  | **Rationale** | 1/3 were rated ‘high’ and 2/3 were rated ‘probably high’, mainly due to lack of confounding control | Most of these studies focussed on children | Studies show variation in increased risk/null effect with noise | Some CIs are very wide | No direct evidence of publication bias, other than those typically of influence |
| **Others** | **Rating** | -2 | 0 | 0 | -1 | 0 |
|  | **Rationale** | 3/5 were rated ‘high’ and 2/5 were rated ‘probably high’, mainly due to population selection/exposure assessment/lack of confounding control | Most of these studies include adults in the general population, though some focussed on teenagers/college student (15-25 years) | Studies show variation in increased risk/null effect with noise | most studies conducted simple descriptive/correlation analysis rather than association type analysis; Some CIs are very wide | No direct evidence of publication bias, other than those typically of influence |

# Supplementary 11 Assessment of upgrading factors for quality of evidence.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Health outcome** |  | **Large magnitude of effect** | **Dose-response** | **Residual Confounding Increases Confidence** |
| Annoyance | Rating | 0 | 0 | 0 |
|  | Rationale | Relatively modest magnitudes of effect | Little evidence of a dose-response effect | Not likely that residual confounding would underestimate results |
| Cardiovascular outcomes | Rating | 0 | 0 | 0 |
|  | Rationale | Some larger magnitudes of effect, but inconsistent | Inconsistent | Not likely that residual confounding would underestimate results |
| Mental health | Rating | 0 | 0 | 0 |
|  | Rationale | Relatively modest magnitudes of effect | Little evidence of a dose-response effect | Not likely that residual confounding would underestimate results |
| Quality of sleep | Rating | 0 | 0 | 0 |
|  | Rationale | Relatively modest magnitudes of effect | Inconsistent | Not likely that residual confounding would underestimate results |
| Cognitive outcomes | Rating | 0 | 0 | 0 |
|  | Rationale | Relatively modest magnitudes of effect | Inconsistent | Not likely that residual confounding would underestimate results |
| Others | Rating | 0 | 0 | 0 |
|  | Rationale | Relatively modest magnitudes of effect | Inconsistent | Not likely that residual confounding would underestimate results |

# Supplementary 12 Comparisons of exposure range extracted from systematic reviews supporting the 2018 WHO Environmental Noise Guidelines for the European Region (70-72, 74, 75) and the present review

|  |  |  |
| --- | --- | --- |
| **Table S13.** Comparisons of exposure range extracted from systematic reviews supporting the WHO Environmental Noise Guidelines for the European Region and the present review | | |
| Health outcomes | Exposure range | |
|  | WHO reviews | This review |
| Annoyance and perception | ACN: 11–74 dB LAeq, 24h;  12–78 dB Lden RTN: 10–82 dB LAeq,24h;  16–83 dB Lden WTN: 29-56 Lden/Ldn | ACN: 48.0–80.0 dB Leq  (some studies reported 44–81 dB Lden)  RTN: 46.3–86.3 dB LAeq  WTN: 44.8–50.4 LAeq |
| CVDs | RTN: 20–80 Lden | RTN: 67.2 dB–120.1dB Leq  (some studies reported 55–80 dB Lden) |
| Mental health | Not mentioned | RTN: 36– 97 dB LAeq  (some studies reported 55–80 dB Lden) |
| Quality of sleep | ERR valid for an Lnight range of 40–65 dB | RTN: 45–120 dBA LAeq: 24h |
| Cognitive outcomes | not mentioned | ACN: 50.5–95 dB LAeq |
| ACN: Aircraft noise; RTN: Road traffic noise; WTN: Wind turbine noise; ERR: Exposure response relationship; CVDs: Cardiovascular diseases. | | |
| Note that there was a heterogeneity of noise metrics (e.g., LAeq, Lden etc) reported across studies in our review. We summarised ranges corresponding to different noise sources, based on reported noise metric, though due to this heterogeneity, we are limited in the comparison we are able to make with the studies in the WHO reviews. Similarly, noise levels captured with different exposure assessment methods and time-scales may limit comparability. | | |

1. a Note: Limitations to GRADE’s risk of bias assessments as stated by GRADE: “First, empirical evidence supporting the criteria is limited. Attempts to show systematic difference between studies that meet and do not meet specific criteria have shown inconsistent results. Second, the relative weight one should put on the criteria remains uncertain. The GRADE approach is less comprehensive than many systems, emphasizing simplicity and parsimony over completeness. GRADE’s approach does not provide a quantitative rating of risk of bias. Although such a rating has advantages, we share with the Cochrane Collaboration methodologists a reluctance to provide a risk of bias score that, by its nature, must make questionable assumptions about the relative extent of bias associated with individual items and fails to consider the context of the individual items.” [↑](#footnote-ref-2)
2. GRADE includes a fourth type of indirectness that occurs when there are no direct (i.e., head-to-head) comparisons between two or more interventions of interest. This criterion is not relevant to our study. [↑](#footnote-ref-3)