SOCILEXIA
Journal of Social Sciences and Humanities

SOCILEXIA: Journal of Social Sciences and Humanities
Vol. 1, No. 4, 2026
E-ISSN: 3164-9610
punggawajournal.com/socilexia

The World's Most Dangerous Country May Be a Statistical Illusion: Reassessing the Evidence Behind Homicide Rankings

Zul Khaidir Kadir[1]
[1] Universitas Muslim Indonesia, Indonesia
Korespondensi: zulkhaidir.kadir@umi.ac.id

Informasi Artikel
Riwayat artikel:
Received Sept 15th, 2026
Revised Oct 24th, 2026
Accepted Nov 1th, 2026

Keyword:
Homicide measurement; International rankings; Recording rules; Statistical comparability.

ABSTRAK

Published homicide rankings are often treated as stable comparisons of lethal violence across countries, although their underlying observations may come from different recording systems. This critical documentary review examines what existing evidence establishes about national rank instability and which stronger claims remain unsupported. It draws on official UNODC and WHO documentation and published quantitative studies, with a descriptive re-tabulation of selected results from a 2026 comparison of 41 European countries. That study associated input-stage police recording with counts approximately 13 to 15 percent higher than later-stage recording systems. Model-adjusted homicide rankings shifted by 4.2 places on average. Published examples include movement of ten places for Portugal and eleven for Germany. These are secondary findings, not new estimates or direct WHO–UNODC rank comparisons. The review distinguishes rate disagreement, model-adjusted rank movement, and error relative to an unknown true homicide incidence. It finds that the evidence warrants caution about definitive league tables but cannot establish a corrected global order or identify the world's most dangerous country. A modest documentary method is sufficient to clarify these limits. An original global rank audit would require a separately verified country-year matrix and matched source provenance.

© 2026 The Authors. Published by Punggawa Legacy Center. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

PENDAHULUAN

An international homicide ranking appears to answer a straightforward question about where lethal violence is most common. The number used for that answer, however, has already passed through a recording system before it reaches a comparative table. In one country, an investigation may determine the police classification. Elsewhere, the official count may begin with the first recorded suspicion. A third jurisdiction may provide medically certified deaths rather than police cases. These procedures can produce different figures without establishing that any institution has deliberately misreported a death. Homicide remains an unusually valuable indicator in comparative criminology, yet the value of the indicator should not be confused with the stability of every national position. LaFree (2021) explained the prominence of homicide in international criminology. Bennett and Lynch (1990) identified that problem in the wider comparison of crime indicators.

Countries separated by a small difference in homicide rates can exchange places after a modest change in the numerator. A shift from fifth to ninth place can attract public attention even when the recorded difference amounts to a fraction of one death per 100,000 people. Rank is therefore more sensitive to the spacing of observations than a league table usually makes apparent. Bennett (2004) regarded measurement comparability as a basic condition of comparative criminal justice research. His concern becomes unusually concrete when the outcome is expressed as the name of a country rather than as a distribution of recorded rates. The review by Koeppel et al. (2015) reaches related concerns.

Two official homicide indicators do not describe identical objects. UNODC's intentional homicide framework concerns unlawful killings committed with intent, while WHO mortality data classify medically certified deaths by cause. Police records may be organized around incidents or offenses. Mortality data concern deceased persons. Recording can also occur at different stages of investigation. The International Classification of Crime for Statistical Purposes addresses some definitional differences, but national implementation remains uneven (Bisogno et al., 2015). Harmonized labels cannot by themselves eliminate differences in counting practice, as Harrendorf (2012, 2018) has argued. Killias and Rau (2000) considered European comparability. A source comparison must therefore begin with the institutions and units producing each rate.

Andersson and Kazemian (2018) compared UN and WHO homicide data and found differences in magnitude alongside broadly similar outcomes across the countries they studied. Rogers and Pridemore (2023) subsequently demonstrated that changing homicide measures could alter the size, significance, and even direction of cross-national statistical associations. Their findings concern substantive inferences, rather than the public interpretation of a country ranking. Neither result permits the assumption that one official source is a universal standard. The more useful question is what a particular source comparison establishes and how much of the difference can safely be attributed to administrative measurement. The disagreement deserves closer attention because apparently similar conclusions can conceal differences in the specific countries and years on which they rest.

Buil-Gil, Pina-Sánchez, and Aebi (2026) provide direct evidence that national recording rules are associated with reported homicide levels. Their study of 41 European countries found that systems registering homicide when police first receive a report recorded approximately 13 to 15 percent more homicides than systems using later stages, after accounting for the measures specified in their models. The authors also adjusted homicide rates for recording differences. Their adjusted ranking moved an average of 4.2 positions, with several countries undergoing much larger changes. The present article cannot claim that international homicide rank changes have never been estimated. That contribution already exists, and any further analysis must be distinguished from it. Its adjusted ranking already answers a question that the initial draft had presented as untouched by previous empirical work.

A global comparison raises an additional problem that a European adjustment study cannot settle. WHO's Mortality Database is based on submitted civil registration records, while UNODC series may draw on criminal justice records, mortality records, or a combination of sources. Two internationally published numbers are not independent measurements merely because the hosting organizations differ. In a comparison of 81 countries, Kadir (2026e) found that close correlation between UNODC and WHO rates did not guarantee agreement in their reported levels. WHO Global Health Estimates are also distinct from the raw Mortality Database, as Kanis et al. (2017) and Mathers (2020) explain. The distinction becomes decisive if a researcher describes disagreement between series as previously concealed murder or interprets agreement as proof that both are correct.

Published work already establishes that homicide indicators are sensitive to source construction. What remains difficult is translating this evidence into a defensible interpretation of country ranks without implying that the true incidence of homicide has been recovered. The relevant distinction is between observed differences in rates, observed or adjusted changes in ordinal position, and error relative to a true but unobserved quantity. The first two can be described from suitable records or published results. The last requires stronger evidence. Calling every shift a correction would make the statistical audit depend on an untested premise about which source is right.

Existing comparative evidence sets the limits of credible homicide-ranking claims. It focuses on the correspondence between police and mortality statistics, the observed associations with recording practices, and the national rank changes reported after statistical adjustment. The evidence comes from published studies and official methodological documents rather than new field observations. The purpose is to assess the strength of the ranking claim, not to produce a replacement global league table. A credible conclusion must preserve the difference between published European results and any statement about all countries.

RESEARCH METHODS

A critical documentary review was undertaken using published quantitative studies of cross-national homicide measurement and the methodological documentation of UNODC and WHO. Studies were selected for their direct treatment of alternative homicide sources, recording rules, administrative

definitions, or national ranking results. Particular attention was given to the 2026 European comparison by Buil-Gil and colleagues, alongside the earlier UN–WHO comparison by Andersson and Kazemian (2018) and the synthesis by Rogers and Pridemore (2023). Official database documents were used to check the meaning of the indicators. No survey, interview, case-level extraction, or original country-year merge was performed. The review retained the authors' own labels for statistical adjustments so that modelled positions were not confused with raw source comparisons.

Published figures were transcribed with their stated populations, denominators, and comparison types. Descriptive re-tabulation was limited to the recorded classification counts and the named country rank changes reported by Buil-Gil et al. (2026). The direction and absolute size of each documented change were obtained by simple subtraction of the published positions. The source-specific mean correspondence ratios and correlations between correspondence ratios were reported without treating them as new estimates. Tables and graphs identify the original study. No Spearman coefficient for UNODC and WHO rankings was computed, because a verified matched-year global matrix was not available.

Interpretation was restricted to what the selected evidence could establish. Differences between mortality and criminal justice rates were treated as source discrepancies rather than validated errors. Published adjusted positions were described as model-dependent results, not as confirmed true ranks. Comparisons involving WHO Global Health Estimates were kept distinct from the raw WHO Mortality Database. The review examined how definitions, recording stage, and source overlap could affect the use of rankings. This method supports critical interpretation and descriptive secondary analysis. It cannot estimate a new global ranking, identify the world's most dangerous country, or attribute source disagreement to a specific cause for any country.

RESULTS AND DISCUSSION

Homicide Indicators Describe Different Administrative Events

UNODC defines intentional homicide through unlawful killing with an intentional element, yet its international records may originate from domestic criminal justice or public health systems. WHO Mortality Database observations are compiled from cause-of-death information submitted through civil registration. The two indicators therefore have distinct administrative histories even when both concern fatal violence. A difference between them may arise before any comparison is made. Classification, incident counting, date of registration, and completeness are not properties of the underlying violent event. They are properties of the statistical procedure through which that event is represented. Harrendorf (2018) explains why such distinctions matter for cross-national criminal justice indicators. Enzmann and Podana (2010) describe comparable source differences.

Medical certification answers a question about the cause of death. Criminal justice recording often answers a question about an alleged or established unlawful offense. Cases can move between categories as investigations develop. The moment at which a count is taken therefore matters. A death initially coded as having undetermined intent may later be treated as homicide, while a police investigation may cease to classify an incident as a suspected intentional killing. A yearly rate may also be affected by the administrative timing of these decisions. Official publication gives the resulting number authority, but it does not make these pathways equivalent. Gruszczyńska and Heiskanen (2012) discuss this problem.

The source difference becomes harder to interpret where the two international series share a domestic input. The UNODC methodological annex recognizes the use of public health records for homicide statistics. A WHO mortality record and a UNODC entry may then trace back to the same registration system. Similar values would not amount to two independent confirmations. In other countries, genuinely separate systems may produce divergent rates without either being demonstrably incorrect. The United Nations and WHO documentation should therefore accompany any statistical comparison of the two series. A comparison based only on agency names misses the critical fact of where each observation originated. A separate methodological analysis by Kadir (2026j) examines source provenance and the possible effects of different recording practices. Estienne and Morabito (2016) report related variation.

The comparison needs to distinguish intentional homicide from the wider category of violent death. Death certificates can include external causes assigned without a criminal judgment. A police database can retain an apparent intentional killing before a court has ruled on criminal responsibility.

Neither administrative record should be mistaken for a completed judicial finding. The distinction matters because judicial outcomes may take years to settle, whereas national homicide indicators are compiled annually. Kadir (2026d) likewise warns against treating fatal outcomes as sufficient evidence of a distinct causal process. Treating one record as legally final and the other as medically tentative distorts what each source actually observes.

The way an institution revises a record also deserves attention. Annual police counts may be updated after an investigation, while mortality files can change when death certificates are amended. These changes need not occur in the same reporting year. Without archived data releases, a researcher may unknowingly compare different stages of administrative revision. A discrepancy that appears to reflect national legal practice might instead reflect the date on which a dataset was downloaded. Official series therefore need both an observation year and, where relevant, a publication version.

Table 1. Statistical objects used in the documentary assessment

Dimension	Police or Criminal Justice Records	Mortality Registration
Primary statistical object	Unlawful intentional killing or victim record, depending on series	Medically certified death and cause-of-death code
Administrative decision	Initial suspicion or later case classification	Death registration and certification
Source status	UNODC may use police or public health records	WHO Mortality uses raw submitted civil registration
Immediate caution	Victim and incident units must be distinguished	Raw and modelled WHO series must be distinguished

What the European recording study has already established

Buil-Gil et al. (2026) examined 41 European countries over the period 1998 to 2022. They classified national police recording stages into 22 input systems, 12 process systems, and seven output systems. These are published counts, not observations independently reconstructed here. The shares are approximately 54, 29, and 17 percent. Figure 1 displays that distribution. The differences matter because an early recording stage may capture suspected cases that a later procedural stage would exclude. This is a plausible measurement mechanism, although the counts alone do not establish the size of its influence on any national homicide rate. The recording stage was established from country metadata, rather than inferred from the size of each national homicide rate. That distinction matters for interpreting the regression results.

Their regression analysis associated input-stage recording with homicide counts approximately 13 to 15 percent higher than those observed under later recording systems, conditional on the models' comparison variables. The estimate is an association across the studied observations. It should not be applied as a universal correction to countries with input-stage recording. The result nevertheless disputes the assumption that homicide is almost fully insulated from counting rules because the offense is severe. A difference of this magnitude may be relevant to closely spaced positions in a homicide ranking. Whether the ranks change depends on their original separation. Bartram et al. (2024) caution against unwarranted extrapolation.

Legal classification varied alongside the point of recording. Of 38 countries with the required metadata, seven included every homicide category specified for inclusion. Eleven of 37 countries with exclusion information removed all categories specified for exclusion. Seventeen of the 41 countries treated euthanasia as homicide in their national recording arrangements. Figure 2 presents those published counts without merging their different denominators. The figures cannot be interpreted as percentages of the same set of countries, and they cannot reveal whether one system records more actual lethal events. They establish an obstacle to direct equivalence between legal classifications. Missing classification metadata could also be concentrated in particular institutions. The available published counts do not establish whether that selection occurred, so the proportions must remain descriptive.

Several classification decisions have consequences for rates that standardization per 100,000 inhabitants cannot remove. Buil-Gil and colleagues reported a positive association between inclusion of terrorism-related deaths and recorded homicides in their models. Decisions about assisted or instigated suicide also mattered under some specifications. An international label can therefore conceal national differences at the boundary of the category. The issue is administrative comparability, not an assertion that states deliberately manipulate homicide. Boivin and Cordeau (2011) provide a wider explanation of how discretion enters police statistics, while Bisogno et al. (2015) describe the harmonization that international classification seeks to achieve. Comparisons from other offenses are relevant. Von Hofer (2000) examined the consequences of statistical construction, while Chon and Clifford (2021) studied the influence of legal classification on recorded rape rates. These analogies do not estimate homicide error.

The frequency of a recording rule should not be interpreted as its statistical effect. Twenty-two countries using input-stage records does not imply that 22 countries understate or overstate homicide. The direction of any observed difference depends on what cases the recording process admits and later removes. Regression can estimate an association after the variables included in the model have been considered. It cannot reveal unobserved case histories from national totals. The 13 to 15 percent estimate is consequently stronger evidence about patterned administrative differences than about the true number of deaths.

The European comparison also makes the regional limitation concrete. Its countries share international reporting institutions to a greater extent than many states elsewhere, yet their definitions remain heterogeneous. The authors' findings do not establish that a larger difference must exist outside Europe. That possibility is plausible but untested by their data. A global article should treat the European result as a reason to investigate other regions, rather than as a basis for assigning a larger unmeasured distortion to them. The latter claim would turn an empirical regional limitation into an unsupported global conclusion.

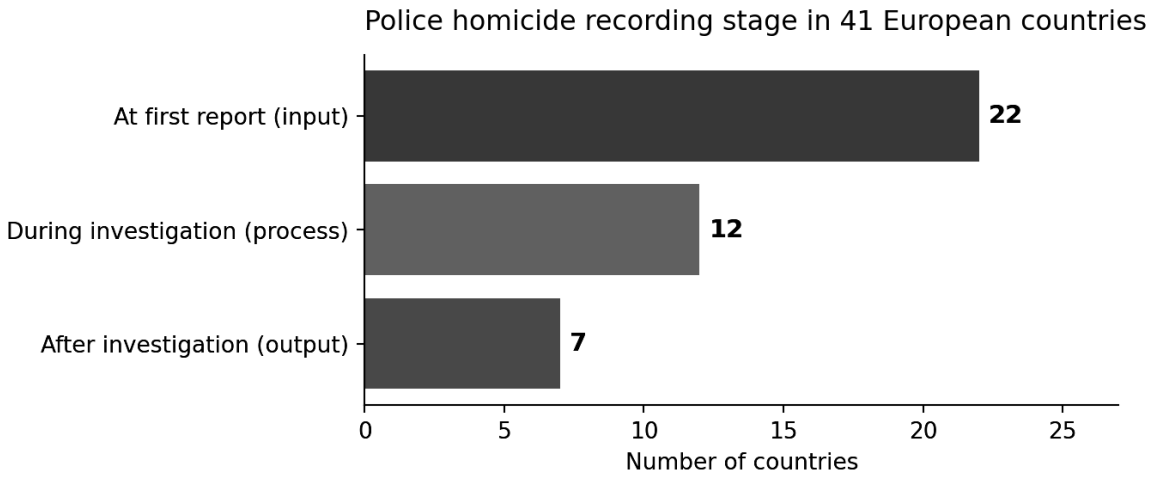


Figure 1. National recording-stage counts reproduced from Buil-Gil et al. (2026), based on 41 European countries.

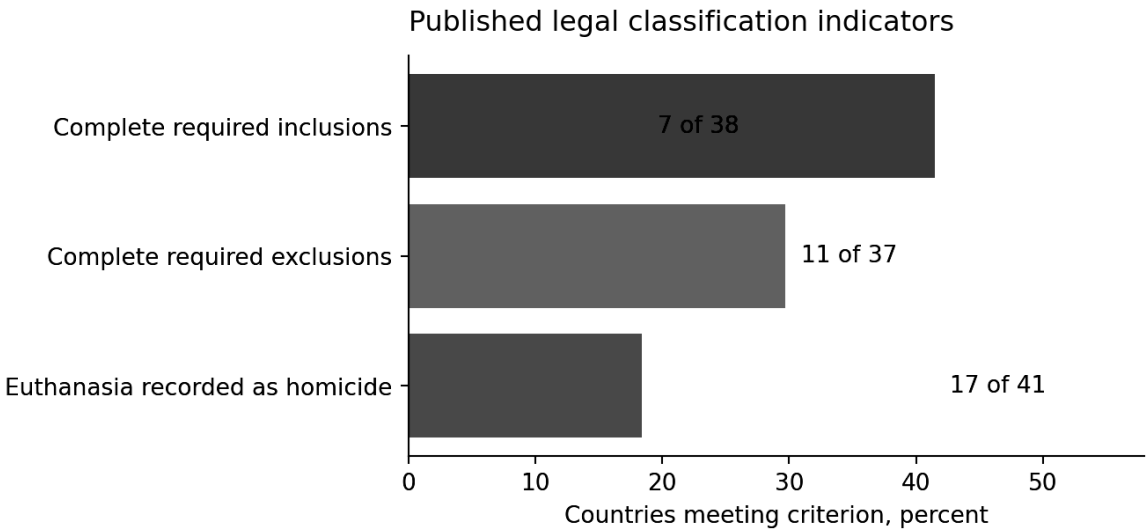


Figure 2. Re-tabulation of distinct legal-classification counts from Buil-Gil et al. (2026). Denominators are not identical.

Table 2. Published quantitative results retained with their original interpretation

Published Measure	Value	Meaning and Boundary
European sample	41 countries	Countries analyzed by Buil-Gil et al. (2026)
Input, process, output recording	22, 12, 7 countries	Descriptive classification counts
Association with input recording	Approximately 13 to 15 percent higher	Conditional association, not a universal correction
Mean adjusted rank movement	4.2 positions	Model-adjusted European ranks, not global raw-source change
Mean absolute adjusted-rate change	0.31 per 100,000	Reported mean magnitude of change after adjustment
Eurostat compared with raw WHO Mortality	1.31 mean ratio	Level correspondence, not a rank statistic

The Ranking Shifts Are Real Published Findings, Not New Estimates

The most consequential finding for the present question has already been produced. After adjusting European police homicide rates for source correspondence and selected recording rules, Buil-Gil et al. (2026) reported a mean absolute shift of 4.2 positions in country rankings. Their adjustment also changed average homicide rates by 0.31 deaths per 100,000 inhabitants in absolute terms. These outcomes were generated through the authors' statistical models. They are not direct rank changes obtained by replacing a UNODC column with a WHO Mortality column. Any claim that the present review independently measured them would be false.

Portugal moved from twenty-seventh position in the original European ranking to seventeenth after the full adjustment, a rise of ten places. Iceland changed from thirty-sixth to twenty-sixth and North Macedonia from eighteenth to ninth. Slovenia moved from thirtieth to twenty-fourth. Albania rose from sixth to second. These changes are large enough to alter the narrative attached to some countries within the sample. They do not demonstrate that any country's violence was inaccurately counted by precisely the amount implied by its adjustment. The model seeks to address differences in recording, but the corrected position remains conditional on that modelling strategy. The ten-place

movement can be calculated from the published positions, but the underlying correction cannot be replicated from two rank numbers alone.

The changes did not run in one direction. Slovakia fell from twelfth to twenty-third position, Germany from twenty-second to thirty-third, and France from seventeenth to twenty-fifth. Figure 3 displays the reported starting and adjusted positions for the named countries. Table 3 gives the signed movements. A favorable movement in a ranked list should not be described as a decline in the number of victims. It may occur because the relative position of other countries changes. Likewise, an apparently unfavorable movement need not reflect an increase in lethal violence. The comparison concerns revised estimates of place within a particular group of European countries.

The published 4.2-place mean shift is informative but cannot stand alone as a measure of the global problem. Sample composition shapes every rank, and different adjustment models can generate different positions from the same administrative observations. A model that standardizes recording stage is doing more than comparing two raw sources. Its answer is conditional on estimated coefficients and on the model used to make the adjustment. The existence of sizable movement establishes sensitivity under that procedure. It does not reveal what would happen in an unobserved global comparison, or how the results would change under a different set of inclusion rules.

The position of an individual country should be interpreted alongside the number of countries in the ranking. An eleven-place change among roughly forty countries is a different statement from eleven places among more than 150. The European adjustment excluded some jurisdictions where full data were unavailable. Thus, even the modelled comparison operates within a specified sample. Transplanting the exact movement into a global league table would alter its denominator and meaning. A dramatic number is persuasive only when its reference population is visible.

An adjusted country rank may change because an estimated correction raises its own rate, because the adjustment lowers the rates of its neighbors, or because both processes occur together. Published rank changes do not by themselves separate these routes. Country positions have relational rather than individual properties. A claim that a state became more dangerous after adjustment would therefore be misleading. No new victims were observed during the calculation. The change is in the modelled ordering of an existing set of observations.

Table 3. Selected published changes in European country positions after adjustment

Country	Original rank	Adjusted rank	Places moved
Portugal	27	17	+10
Iceland	36	26	+10
North Macedonia	18	9	+9
Slovenia	30	24	+6
Albania	6	2	+4
Slovakia	12	23	-11
Germany	22	33	-11
France	17	25	-8

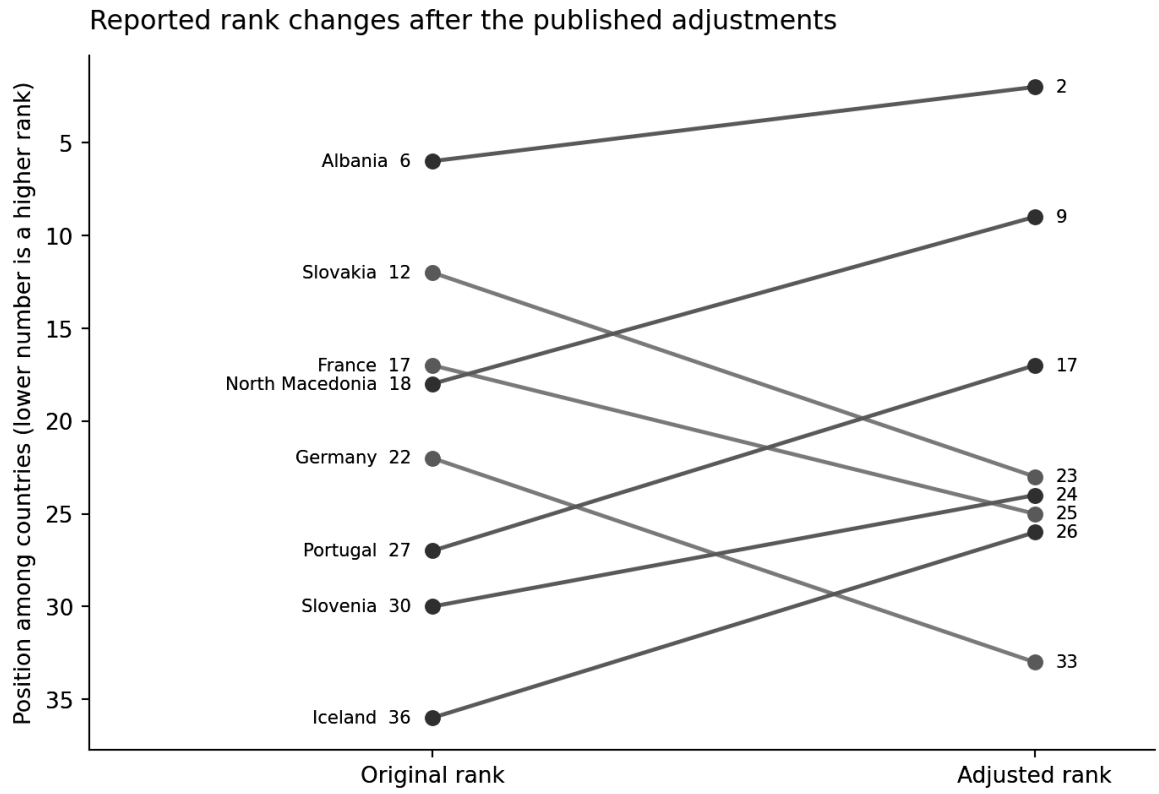


Figure 3. Original and model-adjusted ranks reported in Buil-Gil et al. (2026), with movements plotted here for the named countries.

Correspondence Ratios Are Not Country Rank Correlations

The source correspondence results from the same study illustrate another form of sensitivity. Eurostat counts averaged 0.90 of WHO Global Health Estimates and 0.98 of IHME estimates, while averaging 1.31 times the raw WHO Mortality records. These are published mean ratios that concern source levels. Figure 4 reproduces the three values with their source labels. The apparent direction of discrepancy changes with the comparison series. Compared with a modelled mortality estimate, police counts tend to appear lower. Compared with raw registered mortality in the countries studied, they tend to appear higher.

The authors also reported correlations of 0.53, 0.58, and 0.71 between different correspondence ratios. The first value compares ratios based on WHO Global Health Estimates and WHO Mortality. The second concerns WHO Mortality and IHME. The third concerns IHME and WHO Global Health Estimates. These are not Spearman coefficients comparing national homicide rankings. Substituting them for rank agreement statistics would be a serious reporting error. Their limited interpretation is still valuable, since a country's relative discrepancy can look different depending on which source supplies the alternative measure.

Andersson and Kazemian (2018) reported differences in cross-national homicide levels while finding some similarities in their substantive comparative results. Rogers and Pridemore (2023) found more consequential variation across homicide measures in their own analytical designs. The two findings need not be reconciled into a single verdict about reliability. They differ in sample, period, dependent variable construction, and modelling choices. Their joint implication is narrower. A source comparison may matter more for some uses of homicide statistics than for others. Rankings, temporal trends, and explanations based on structural covariates should be examined separately.

A high correlation between two national homicide rates would not show that every rank is stable. Countries can be close in numerical value but separated by several ordinal places after a small revision. At the same time, two measures can differ considerably in level while preserving nearly the same ordering. A scatterplot and a rank-change plot therefore ask different questions. The former concerns agreement in the reported measurement. The latter concerns movement within a sample. Neither establishes the underlying number of unlawful killings with certainty. The country-specific

displacement still has to be read beside the numerical rates. A move of several places among closely spaced countries may correspond to a small difference in mortality burden, while a smaller rank move can accompany a large change in estimated rates. Brunton-Smith et al. (2024) assess related measurement uncertainty.

Figure 5 shows the absolute movements extracted from the eight published country examples. It adds no new homicide estimate. The plot is useful because it places the reported ten- and eleven-place changes beside smaller movements in a form that can be compared directly. It must not be interpreted as a population distribution of all 41 countries. Countries were chosen because the source article reported their starting and adjusted ranks. A descriptive graphic becomes misleading the moment a selected subset is presented as if it were a random or complete sample.

A source-correspondence ratio of 1.31 means that one published rate series exceeds another on average under the study's method of aggregation. It does not mean that precisely 31 percent of murders are missing from mortality registration. The ratio may combine differences in cause classification, reporting completeness, data revision, and the populations covered. The absence of a validated reference series makes the source discrepancy informative but not self-explanatory. Calling it an undercount requires evidence that identifies which deaths ought to have appeared in a particular administrative record.

The finding that correspondence ratios correlate only moderately also cautions against replacing one series with whichever estimate produces the most dramatic result. Different global mortality estimates can share inputs and still use different adjustment procedures. A high degree of overlap in sources limits the value of apparent agreement as an independent validation. Mathers (2020) provides a detailed account of why global cause-of-death estimation uses statistical adjustment. The presence of such adjustment is neither a defect nor proof of accuracy. Its implications depend on the research question.

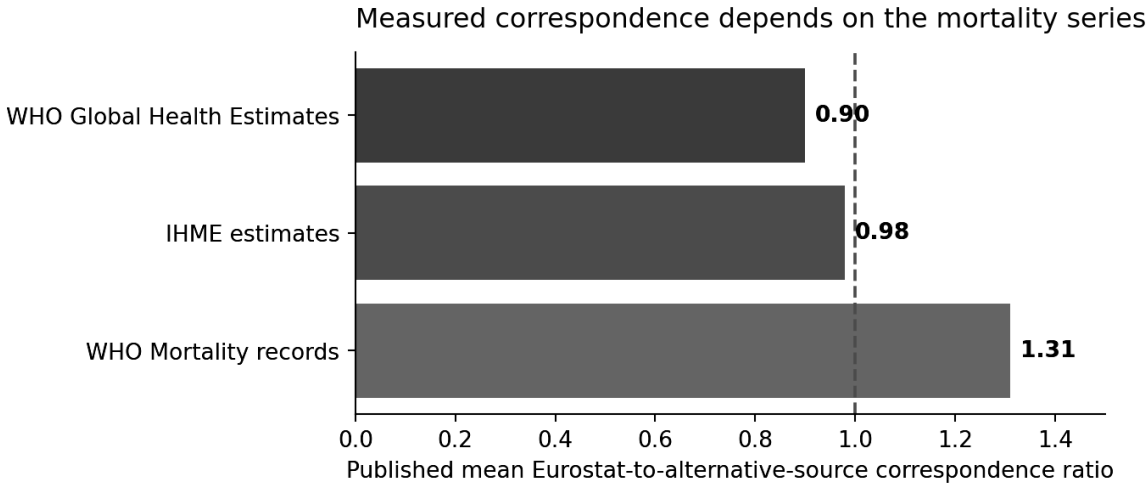


Figure 4. Mean source-correspondence ratios published in Buil-Gil et al. (2026). These are not country rank correlations.

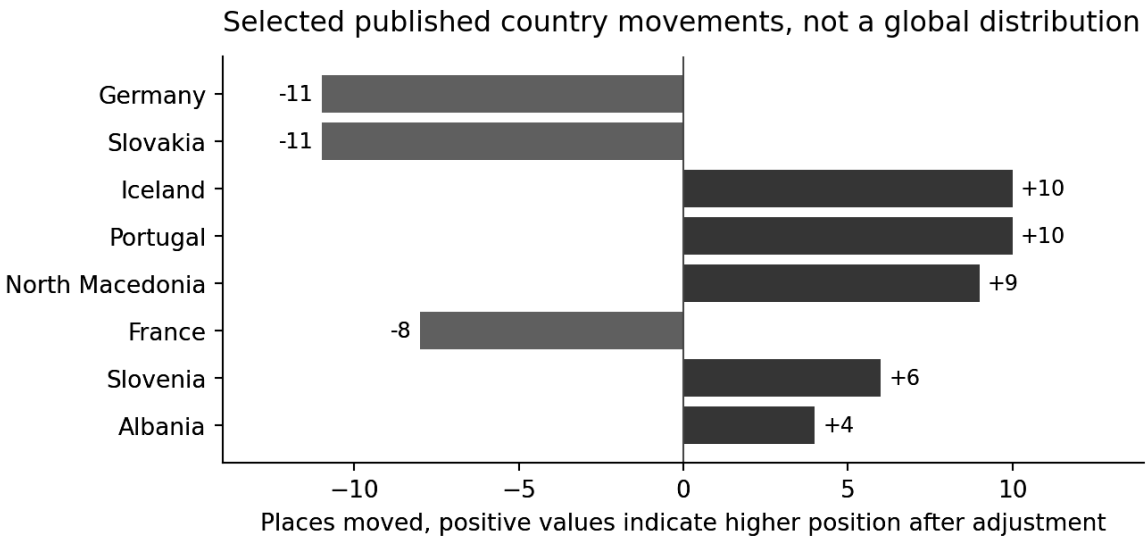


Figure 5. Descriptive arithmetic re-tabulation of selected country rank movements published by Buil-Gil et al. (2026). Positive values indicate movement toward rank 1.

Global Inference Stops Where Comparable Observations Stop

WHO's Mortality Database contains medically certified, country-submitted deaths. Its public documentation states that the recorded deaths are not automatically corrected for incomplete civil registration. This is different from the WHO Global Health Estimates, which can incorporate adjustments and information from multiple sources. Kanis et al. (2017) warned that using estimated homicide indicators without considering their construction can complicate cross-national inference. Treating a raw mortality series and an estimated global series as interchangeable would create an apparent cross-source comparison that does not have a common measurement basis.

Countries differ in the years for which complete mortality records are available. A chart that silently selects the last available year for one country and a newer year for another cannot isolate source disagreement. Temporal change in homicide may account for part of the difference. The problem is especially serious in countries undergoing sharp short-term changes in violence. The primary year must be fixed before a rank comparison is attempted, and the underlying record must be checked rather than inferred from a dashboard label. WHO documentation also identifies differences in population inputs, which require attention when rates are compared. Marshall and Block (2004) identified this constraint.

The eligible country set forms part of the result. Excluding countries without complete records may produce a sample dominated by better-resourced statistical systems. Such a finding cannot be generalized to all states, especially where gaps in death registration are most serious. The European Homicide Monitor discussion by Kivivuori, Liem, and Markwalder (2024) illustrates why detailed common coding is difficult even among countries cooperating on homicide research. Limited coverage can improve comparability within the available group while narrowing the reach of the conclusion. Better documentation can improve within-sample comparability while narrowing external scope. In a six-country comparison, Kadir (2026b) also identified registration completeness as a limit on interpreting national differences. The appropriate heading would describe the observed group rather than promise coverage of nations that could not be included. Bijleveld (2023) addresses the continuing shortage of genuinely comparable European crime data.

A global rank may also conceal changes in the comparison group. A country can move upward because another nation was omitted, even if its homicide rate stayed exactly the same. Rankings calculated using different national sets are therefore not directly comparable. The relevant requirement is to construct both lists from an identical country-year sample. When stricter quality criteria are applied, both lists have to be recalculated within the restricted sample, and the new number of countries must be reported. This is a simple rule, but it prevents a change in eligibility from being mistaken for a change in homicide measurement.

The phrase 'world's most dangerous' is not equivalent to 'highest recorded intentional homicide rate'. Homicide captures one form of fatal interpersonal violence. It does not measure all hazards, violent

events, exposure to crime, or individual likelihood of victimization. Countries also differ in their age structures and in the distribution of risk among residents. A national rate cannot describe all of these differences. Comparing homicide with recorded serious assault, Kadir (2026c) also found that the lethal share of recorded violence varied widely between countries. The headline can function as a question about the misuse of rankings, but the analysis should not reproduce the same misuse by treating homicide alone as a complete measure of national danger.

The choice of a reference year is consequential even in a descriptive comparison. Some countries experience prolonged stability in intentional homicide, while others undergo sharp changes after political crises or shifts in criminal markets. Comparing figures from different calendar years can introduce real changes in violence into a study that claims to test recording differences. A common year is therefore a minimum requirement rather than a sophisticated statistical remedy. Where one source has no observation for that year, the correct response may simply be to exclude the country and disclose the loss of coverage.

Country comparability also depends on territorial coverage. Some international sources publish national figures while others report selected jurisdictions or treat a territory separately. A comparison that merges these units can generate a rate difference without any dispute over the definition of homicide. The problem extends to countries that have undergone political or administrative boundary changes. Harmonizing country names in a spreadsheet does not resolve it. The underlying jurisdiction and the population to which the death count refers must be checked before any rank is assigned. Kadir (2026f) found comparable limitations when mapping homicide across Indonesian administrative and police reporting units.

The legal and statistical boundaries of the homicide numerator

Different legal definitions can affect international rankings even where the statistical rate is expressed in a common unit. Police records classify offenses according to domestic legal and investigative processes. Mortality registration generally uses international cause-of-death codes. These two routes may disagree about deaths associated with police intervention, terrorism, or disputed intention. The ICCS creates a common framework for intentional homicide, but its categories do not automatically resolve how national agencies assign each case. International classification may reduce variation while leaving administrative practice heterogeneous. Kadir (2026i) makes a related distinction between the social circumstances of homicide and the legal categories applied to it in Indonesia.

Recorded events can change status over time. An unexplained death may later be recognized as a homicide once forensic findings emerge. A suspected homicide may be removed from police counts if the investigation establishes a different cause. The size of such changes depends on procedures that are rarely visible in published international rates. Researchers cannot infer them from the net difference between WHO and UNODC alone. Kadir (2026a) also explains how investigative clearance filters the cases from which detailed offender information can be drawn. Where incident-level linkage is absent, the discrepancy identifies the need for explanation without telling us which cases account for it.

The numerator also requires careful attention to the unit counted. A single incident can involve several victims, while police systems may report events under rules that differ by offense or person. Mortality statistics necessarily concern deaths. A comparison should verify that the criminal justice series counts victims rather than incidents before it is converted to a victimization rate. Different rules can otherwise create disagreement that has little to do with whether the deaths were recognized. Standardizing by population only addresses the denominator and cannot correct a mismatch between victims and incidents.

Denominator consistency should not be dismissed as a technical detail. Population estimates may come from national agencies or international population series. A rate calculated from the same death count can move slightly when the denominator changes. The researcher should either calculate comparable crude rates from counts and one population source or document why the published rates are sufficiently aligned. Age-standardized mortality rates should not be placed beside crude intentional homicide rates as if both used the same underlying denominator. Even a carefully documented numerator cannot rescue a comparison that mixes these measures. Population comparability is part of the validity problem, rather than an optional improvement to the graphic. Kadir (2026h) demonstrates the stakes of matching population denominators and distinguishing reported zeros from missing prison-homicide data.

International homicide rates are frequently discussed as if the population were homogeneous. In reality, fatal violence is distributed unevenly by age and sex, and those distributions differ among countries. A country with a younger population can have a different crude rate even if age-specific risks are similar to those elsewhere. This does not make crude homicide rates invalid. They answer a question about the burden within the actual population. Age-standardized measures answer a different question about comparable risk structures. Mixing the two within one ranking would add a demographic adjustment to a source comparison without making it explicit.

Public-health cause categories include information that criminal justice datasets may not collect in the same way. The coding of undetermined intent is especially relevant because uncertainty about the circumstances of death may persist after certification. Andersson and Kazemian (2018) considered unclassified deaths in their assessment of homicide statistics. Their treatment illustrates why these deaths should be discussed without automatically reclassifying them as murder. The uncertainty category can signal an information problem, but the number of uncertain deaths does not directly measure hidden homicide.

Why Ordinary Descriptive Methods are Enough for a Limited Claim

A simple cross-source audit can be rigorous when the underlying records are available. Rank changes can be calculated by sorting two complete and comparable rate columns for the same countries and year. Spearman correlation can then describe the broad agreement in ordering. The signed difference in rank identifies direction, while the absolute difference describes movement. Absolute and relative rate differences answer separate questions about scale. None of these calculations requires fieldwork, complex regression, or Bayesian estimation. They also do not estimate the true number of unrecorded deaths.

Country-year matching matters more than software complexity. The same country names, territories, years, and rate definitions must enter both columns. Records that rely on a shared original input require identification, because two matching rates may not represent independent evidence. The full available precision should be retained before ranks are calculated. Genuine ties need a consistent treatment, and relative differences are undefined when the chosen denominator is zero. These decisions are straightforward enough to explain in a short Method section. Their omission would weaken even an otherwise sophisticated analysis. Fairbrother (2014) discusses the different requirements of longitudinal comparative models, a level of inference not attempted here.

No such merged global file has been independently assembled for this review. It would be misleading to place a new global Spearman coefficient beside the published European results, or to name the most misranked country without the relevant observations. The present paper therefore remains a critical secondary analysis rather than an original global quantitative test. This restriction is substantive. It separates what can be demonstrated from the material examined from conclusions that depend on data not yet processed. Hirschauer et al. (2020) also caution that conventional inferential language should not be detached from the sampling design.

A new original contribution could still be produced from a matched global dataset, but it would have to go beyond showing that ranks sometimes change. Buil-Gil et al. (2026) already documented that result in Europe after statistical adjustment. A more distinct comparison would examine raw, same-year UNODC and WHO positions while tracking source provenance and the sensitivity of ranks to country eligibility. Even then, it should be presented as a descriptive source audit rather than a correction of the true national homicide order.

The case for a simple method depends on discipline in what it claims. A Spearman coefficient can summarize whether two observed country orders are broadly similar, but it cannot tell a researcher which source is true. An absolute rank shift can identify unusually mobile positions, yet a large shift may occur where rates are tightly clustered. A percentage discrepancy can become unstable when the baseline homicide rate is very small. Descriptive measures should therefore be reported together with the observed rates and the countries included. Simplicity removes unnecessary modelling assumptions without removing the need for sound data construction.

A stronger statistical method would not repair missing provenance. Bayesian ranks, for example, require assumptions about the distribution of the underlying observations and the information carried by each source. The original draft invoked that family of methods before reliable global inputs were available. Retaining the label would have made the method more ambitious than the analysis could

support. A documentary review can instead make a limited claim about what already published comparisons establish. The boundary is explicit, which is preferable to creating an impression of precision through an unexecuted model.

Implications for Theory, Policy, and Comparative Publication

Homicide figures are often outcomes in studies of structural and institutional factors. If the construction of that outcome varies with the institutions included among those predictors, estimated relationships may depend partly on measurement. Rogers and Pridemore (2023) demonstrate that changes in homicide source can alter cross-national model results. Dawson (2018) provides one macro-comparative example. Kadir (2026g) studied recorded homicide as an outcome associated with subnational spending priorities, illustrating the same need to define the observed measure. The issue is not that international statistical models are inherently invalid. Researchers need to justify the source used and assess whether a plausible alternative changes the conclusion. Measurement differences can bias the estimated contribution of institutions when the same institutions also influence how police classify deaths. Nivette (2011) identified weaknesses in cross-national homicide predictors that make this measurement issue especially consequential.

Governments may use a country's rank as evidence of performance even when the underlying difference from neighboring countries is small. A position in a table is politically legible, but it is often a poor substitute for changes in victim numbers and rates over time. Public reporting should state the source, year, and set of included countries. It should also disclose material changes when another defensible source is used. A national position can be informative without being definitive. Presenting an adjusted ranking as the only correct one replaces one unexamined certainty with another. Tonry (2015) questioned the conditions that make cross-national criminal justice comparisons useful.

The published European evidence is strong enough to reject a simplistic claim that homicide rankings are immune to administrative recording differences. It is not strong enough to identify a false global leader. The boundaries of the evidence should therefore determine the boundaries of the manuscript's contribution. A descriptive documentary analysis can organize what is known and expose unwarranted leaps from rate disagreement to allegations of hidden homicide. It cannot compete with a new matched-country analysis on the question of global rank magnitude. The article gains credibility by preserving that difference rather than concealing it behind an ambitious title. van Dijk et al. (2022) provide a separate demonstration of how international crime patterns are constrained by the source used.

A reviewer may ask why the European study needs another article describing its results. Displaying the same values in another graph is insufficient. Its contribution is a critical assessment of the kinds of ranking claims that the published evidence permits, together with explicit separation of raw-source disagreement from model-adjusted displacement. This is a narrower scholarly object than the original manuscript promised. Its value depends on a precise and nonrepetitive argument, not on the number of sources cited or the number of figures included. The editorial response is to narrow the research claim until it fits the observed evidence, rather than disguise the absence of original country-level results with added literature.

A policy brief may compare countries with rates differing only slightly and still emphasize their ordinal positions. Such emphasis can convey a degree of precision unsupported by administrative data. Reported homicide levels, the size of the differences between countries, and changes in the result when the source changes should remain visible. An international ranking might then be treated as a provisional summary rather than as a definitive judgment. This reform does not require sophisticated computing. It requires readers to receive more information than a list of positions alone.

The question for a future original study is no longer whether recording rules can alter homicide rankings. The 2026 European study already answers that question. A distinct global contribution would document source-specific rank shifts among the same eligible countries, identify where changes arise from differences in rates rather than sample selection, and report the stability of comparisons at different parts of the homicide distribution. These tasks are technically modest but demand an auditable global file. Until such a file exists, a secondary review must acknowledge that its strongest numerical result belongs to an earlier investigation.

CONCLUSION

Published comparisons establish that homicide recording systems differ and that these differences can affect reported national rates. The 2026 European study also documents substantial rank movement after adjustment for selected recording practices. An average shift of 4.2 places is meaningful evidence against treating every national position as fixed. It does not prove that the adjusted rank is the true rank, and the named national movements cannot be extrapolated to the entire world. The stronger conclusion is that statistical rules influence the interpretation of international homicide rankings. A country-level comparison remains conditional on the source, the statistical object, and the set of countries admitted to the ranking.

The available documentary material cannot identify the world's most dangerous country or establish how often UNODC and WHO produce different positions in a single global year. Those findings require an independently checked matched dataset. The present review offers a narrower contribution by separating published adjusted rankings from raw cross-source comparisons and by identifying where measurement uncertainty is too great for definitive claims. Such restraint does not weaken the argument. It prevents an article criticizing the misleading certainty of homicide league tables from creating another misleading figure of its own.

REFERENCES

Andersson, C., & Kazemian, L. (2018). Reliability and validity of cross-national homicide data — A comparison of UN and WHO data. *International Journal of Comparative and Applied Criminal Justice*, *42*(4), 287–302. https://doi.org/10.1080/01924036.2017.1370676

Bartram, D., Alaimo, L. S., Avery, E., Bardo, A., Di Bella, E., Binder, M., Botha, F., Fachelli, S., Gatto, A., Lu, J., Okulicz-Kozaryn, A., Sanchez, A., & Tani, M. (2024). Towards the next fifty years of social indicators research — Some guidance for authors. *Social Indicators Research*, *174*, 1–17. https://doi.org/10.1007/s11205-024-03401-z

Bennett, R. R. (2004). Comparative criminology and criminal justice research — The state of our knowledge. *Justice Quarterly*, *21*(1), 1–21. https://doi.org/10.1080/07418820400095721

Bennett, R. R., & Lynch, J. P. (1990). Does a difference make a difference? Comparing cross-national crime indicators. *Criminology*, *28*(1), 153–181. https://doi.org/10.1111/j.1745-9125.1990.tb01322.x

Bijleveld, C. (2023). European criminology needs European data — The case for a pan-European crime and victimization survey. *European Journal of Criminology*, *20*(3), 785–791. https://doi.org/10.1177/14773708231174658

Bisogno, E., Dawson-Faber, J., & Jandl, M. (2015). The international classification of crime for statistical purposes — A new instrument to improve comparative criminological research. *European Journal of Criminology*, *12*(5), 535–550. https://doi.org/10.1177/1477370815600609

Boivin, R., & Cordeau, G. (2011). Measuring the impact of police discretion on official crime statistics — A research note. *Police Quarterly*, *14*(2), 186–203. https://doi.org/10.1177/1098611111404179

Brunton-Smith, I., Cernat, A., Pina-Sánchez, J., & Buil-Gil, D. (2024). Estimating the reliability of crime data in geographic areas. *The British Journal of Criminology*, *64*(6), 1347–1361. https://doi.org/10.1093/bjc/azae018

Buil-Gil, D., Pina-Sánchez, J., & Aebi, M. F. (2026). The effect of counting rules on cross-national comparisons of homicide. *Social Indicators Research*, *181*, Article 44. https://doi.org/10.1007/s11205-025-03802-8

Chon, D. S., & Clifford, J. E. (2021). The impacts of international rape laws upon official rape rates. *International Journal of Offender Therapy and Comparative Criminology*, *65*, 244–261. https://doi.org/10.1177/0306624X20952399

Dawson, A. (2018). Police legitimacy and homicide — A macro-comparative analysis. *Social Forces*, *97*(2), 841–866. https://doi.org/10.1093/sf/soy043

Enzmann, D., & Podana, Z. (2010). Official crime statistics and survey data — Comparing trends of youth violence between 2000 and 2006 in cities of the Czech Republic, Germany, Poland, Russia, and Slovenia. *European Journal on Criminal Policy and Research*, *16*, 191–205. https://doi.org/10.1007/s10610-010-9121-z

Estienne, E., & Morabito, M. (2016). Understanding differences in crime reporting practices — A comparative approach. *International Journal of Comparative and Applied Criminal Justice*, *40*(2), 123–143. https://doi.org/10.1080/01924036.2015.1086397

Fairbrother, M. (2014). Two multilevel modeling techniques for analyzing comparative longitudinal survey datasets. *Political Science Research and Methods*, *2*(1), 119–140. https://doi.org/10.1017/psrm.2013.24

Gruszczyńska, B. Z., & Heiskanen, M. (2012). Trends in police-recorded offences. *European Journal on Criminal Policy and Research*, *18*, 83–102. https://doi.org/10.1007/s10610-011-9160-0

Harrendorf, S. (2012). Offence definitions in the European sourcebook of crime and criminal justice statistics and their influence on data quality and comparability. *European Journal on Criminal Policy and Research*, *18*(1), 23–53. https://doi.org/10.1007/s10610-011-9161-z

Harrendorf, S. (2018). Prospects, problems, and pitfalls in comparative analyses of criminal justice data. *Crime and Justice*, *47*(1), 159–207. https://doi.org/10.1086/696042

Hirschauer, N., Grüner, D., Mußhoff, O., Becker, C., & Jantsch, A. (2020). Can p-values be meaningfully interpreted without random sampling? *Statistics Surveys*, *14*, 71–91. https://doi.org/10.1214/20-SS129

Kadir, Z. K. (2026a). The cleared-case canon – Investigative selection as a structural limit on general homicide theory building. *LEXARA – Journal of Law and Legal Studies*, *1*(2), 1–10. https://doi.org/10.67707/lexara.v1i2.197

Kadir, Z. K. (2026b). Does Islam explain homicide? A paired cross-national comparison of religion, state capacity, and lethal violence. *SIMURGA – Journal of Islamic Studies and Muslim Societies*, *1*(1), 1–9. punggawajournal.com/simurga/article/view/89

Kadir, Z. K. (2026c). The geography of lethal violence escalation – Cross-national differences in recorded homicide and serious assault across structural contexts. *Punggawa Law Review*, *1*(4), 10–21. https://doi.org/10.67707/plr.v1i4.158

Kadir, Z. K. (2026d). The homicide exceptionalism problem – When fatal outcome is mistaken for etiological distinctiveness. *MULTIVA – Journal of Multidisciplinary Research*, *1*(1), 1–9. punggawajournal.com/multiva/article/view/182

Kadir, Z. K. (2026e). The homicide visibility gap – Cross-national discrepancies between UNODC-recorded intentional homicide and mortality-based violence. *Punggawa Global Research – Jurnal Multidisiplin*, *1*(4), 52–63. https://doi.org/10.67707/pgr.v1i4.157

Kadir, Z. K. (2026f). Mapping recorded homicide in Indonesia – An ecological criminological analysis using BPS data. *Punggawa Law Review*, *1*(2), 9–18. punggawajournal.com/plr/article/view/21

Kadir, Z. K. (2026g). Recorded homicide as lethal violence and subnational spending priorities in Indonesia. *RATIONA – Journal of Public Affairs, Government, and Society*, *1*(1), 1–11. punggawajournal.com/rationa/article/view/187

Kadir, Z. K. (2026h). Safer outside than inside? The prison homicide paradox and the problem of reported zero. *SOCILEXIA – Journal of Social Sciences and Humanities*, *1*(3). https://doi.org/10.67707/socilexia.v1i3.303

Kadir, Z. K. (2026i). The social anatomy of homicide – Relationships, situations, and legal categories in Indonesia. *SOCILEXIA – Journal of Social Sciences and Humanities*, *1*(1), 1–10. punggawajournal.com/socilexia/article/view/115

Kadir, Z. K. (2026j). The state does not count the same dead – The politics of homicide measurement across criminal justice and public health systems. *Punggawa Global Research – Jurnal Multidisiplin*, *1*(5), 24–34. https://doi.org/10.67707/pgr.v1i5.245

Kanis, S., Messner, S. F., Eisner, M. P., & Heitmeyer, W. (2017). A cautionary note about the use of estimated homicide data for cross-national research. *Homicide Studies*, *21*(4), 312–324. https://doi.org/10.1177/1088767917715670

Killias, M., & Rau, W. (2000). The European sourcebook of crime and criminal justice statistics — A new tool in assessing crime and policy issues in comparative and empirical perspective. *European Journal on Criminal Policy and Research*, *8*, 3–12. https://doi.org/10.1023/A — 1008767709296

Kivivuori, J., Liem, M., & Markwalder, N. (2024). European homicide monitor — Research, new developments, and future. *Journal of Contemporary Criminal Justice*, *40*(3), 431–445. https://doi.org/10.1177/10439862241253386

Koeppel, M. D. H., Rhineberger-Dunn, G. M., & Mack, K. Y. (2015). Cross-national homicide — A review of the current literature. *International Journal of Comparative and Applied Criminal Justice*, *39*(1), 47–85. https://doi.org/10.1080/01924036.2013.836676

LaFree, G. (2021). Progress and obstacles in the internationalization of criminology. *International Criminology*, *1*, 58–69. https://doi.org/10.1007/s43576-021-00005-2

Marshall, I. H., & Block, C. R. (2004). Maximizing the availability of cross-national data on homicide. *Homicide Studies*, *8*(3), 267–310. https://doi.org/10.1177/1088767904265455

Mathers, C. D. (2020). History of global burden of disease assessment at the World Health Organization. *Archives of Public Health*, *78*, Article 77. https://doi.org/10.1186/s13690-020-00458-3

Nivette, A. E. (2011). Cross-national predictors of crime — A meta-analysis. *Homicide Studies*, *15*, 103–111. https://doi.org/10.1177/1088767911406397

Rogers, M. L., & Pridemore, W. A. (2023). A review and analysis of the impact of homicide measurement on cross-national research. *Annual Review of Criminology*, *6*, 447–470. https://doi.org/10.1146/annurev-criminol-030521-102909

Tonry, M. (2015). Is cross-national and comparative research on the criminal justice system useful? *European Journal of Criminology*, *12*(4), 505–516. https://doi.org/10.1177/1477370815581699

van Dijk, J., Nieuwbeerta, P., & Joudo Larsen, J. (2022). Global crime patterns — An analysis of survey data from 166 countries around the world, 2006–2019. *Journal of Quantitative Criminology*, *38*, 793–827. https://doi.org/10.1007/s10940-021-09501-0

von Hofer, H. (2000). Crime statistics as constructs — The case of Swedish rape statistics. *European Journal on Criminal Policy and Research*, *8*(1), 77–89. https://doi.org/10.1023/A — 1008713631586

United Nations Office on Drugs and Crime. (2025). UNODC intentional homicide data [Data set]. dataunodc.un.org/

World Health Organization. (2026). WHO Mortality Database [Data set]. platform.who.int/mortality

World Health Organization. (2026). Global Health Estimates 2023 — Deaths by cause, age and sex, by country and region, 2000–2023 [Statistical release]. www.who.int/data/gho/data/themes/mortality-and-global-health-estimates