# # Section 0: Overview

This Supplementary Information (SI) file presents additional details on our methodology and step-by-step implementation in R with annotated code scripts and source data reference.

Overall, this file is structured into two main sections, with the first section providing additional details and justifications for some of the core methodology we used (e.g., background on non-detects and missing data imputation, multilevel meta-analysis, and how the utilized method justifies our study) and the second section dealing with R implementation and preliminary results that are not provided in the main text. As such, this file is organized into 5 main outlines, each presenting SI Method (background and literature search), dataset preparation and exploration, missing data handling (non-detects, standard deviation, sample size), multilevel meta-analysis, and probabilistic risk assessment.

# # Section 1: SI Method

## Literature Search

Studies included in this study was a part of much comprehensive literature database compiled in support of national-scale, multi-pathway, multi-pesticide exposure and cumulative risk assessment in Ethiopia (see unregistered systematic review protocol here: xx).

***Search strategy***

Briefly, in order to identify studies reporting pesticide levels in environmental various matrices (e.g., air, water, soil) and foods in Ethiopia, 13 international and local databases, including gray literature sources, were searched on xx/xx (see Table S1; last updated on xx/xx). Our search strategy used combinations of terms related to pesticides (e.g., agrochemical, insecticide, fungicide, herbicide) and occurrence (e.g., pollution, exposure, monitoring, residue, contamination, air, soil, water) with geographic limitation to Ethiopia. Before undertaking the initial search, we evaluated and validated our strategy against a pre-defined benchmark set of 35 studies collected from multiple sources using Scopus database (Lagisz *et al.*, 2025). A list of benchmark studies used for search strategy validation was provided in Table S2. Our preliminary search retrieved all of the benchmark studies (100%), indicating the sensitivity and comprehensiveness of our terms. Then the validated search terms were tailored to each database based on its indexing system, controlled vocabulary, and search functionalities.

We identified a total of 1,539 studies across databases and repositories, including additional 50 studies identified through reference scanning of relevant studies and Google search. These records were then imported into EndNote, merged and exported as RefMan (RIS) file. Initially, 250 duplicates were removed using the *synthesisr* R package (Westgate and Grames, 2020) and then the remaining 1248 records were imported to Rayyan (<https://rayyan.ai/>), where additional 20 duplicates were removed.

Table S1: Lists of database and repositories, search strategies used, and retrieved studies.

|  |  |  |
| --- | --- | --- |
| **Database** | **Search strategy** | **Hits** |
| Web of Science | ((TS=(pesticide OR agrochemical OR insecticide OR fungicide OR herbicide)) AND TS=(pollution OR exposure OR monitoring OR residue OR contamination OR air OR soil OR water OR food)) AND TS=(Ethiopia) | 426 |
| Scopus | ( TITLE-ABS-KEY ( pesticide OR agrochemical OR insecticide OR fungicide OR herbicide OR organochlorine OR ocp OR ddt ) AND TITLE-ABS-KEY ( pollution OR exposure OR monitoring OR concentration OR level OR residue OR contamination OR air OR soil OR water OR food ) AND TITLE-ABS-KEY ( ethiopia ) AND NOT TITLE-ABS-KEY ( mosquito\*) ) | 478 |
| PubMed | (((pesticide[MeSH Terms] OR agrochemical[MeSH Terms] OR insecticide[Title/Abstract] OR fungicide[Title/Abstract] OR herbicide[Title/Abstract] OR organochlorine[Title/Abstract] OR ocp[Title/Abstract] OR ddt[Title/Abstract]) AND (pollution[Title/Abstract] OR exposure[Title/Abstract] OR monitoring[Title/Abstract] OR concentration[Title/Abstract] OR level[Title/Abstract] OR residue[Title/Abstract] OR contamination[Title/Abstract] OR air[Title/Abstract] OR soil[Title/Abstract] OR water[Title/Abstract] OR food[Title/Abstract])) AND (Ethiopia[Title/Abstract])) NOT (mosquito[MeSH Terms]) AND (2000:2025[pdat]) | 231 |
| OpenAlex | (pesticide OR agrochemical OR insecticide OR fungicide OR herbicide) (pollution OR exposure OR monitoring OR concentration OR level OR residue OR contamination OR air OR soil OR water OR food) Ethiopia | 990 |
| Google Scholar | allintitle: (pesticide OR agrochemical OR insecticide OR fungicide OR herbicide) (pollution OR exposure OR monitoring OR residue OR contamination OR air OR soil OR water OR food) Ethiopia | 63 |
| Semantic scholar | (pesticide) (air, water, soil, food pollution OR contamination) "Ethiopia" | 90 |
| OAIster | (pesticide OR agrochemical OR insecticide OR fungicide OR herbicide) AND (pollution OR exposure OR monitoring OR concentration OR level OR residue OR contamination OR air OR soil OR water OR food) AND Ethiopia | 116 |
| Local repositories | Addis Ababa University, Haramaya University, Jimma University, Bahir Dar University, Hawassa University, The University of Gondor (Pesticide) |  |

***Screening, Inclusion and Data Extraction***

Two-staged screening was conducted. First, the title/abstract of each record was assessed for relevance using Rayyan. At this stage, a total of 134 records were retained after exclusion of 1114 non-relevant records including commentaries and editorials (), reviews (), different country (), and unrelated focus (). We then retrieved the full-texts of all relevant studies and further assessed if original residue data in any Ethiopian matrix is presented, non-duplicate (the measurement across multiple studies), published in English or any local languages the authors understand (the detailed eligibility criteria is presented in attached protocol). At this stage, 74 studies were removed including . Finally, the remaining 60 studies were tagged according to their respective matrix type (e.g., air, water, food) and archived (at: xx).

For the purpose of present study, we included 40 studies specifically focusing on pesticides in food and provided quantitative residue data (e.g., mean, median, standard deviation). From these studies, we extracted study characteristics (first author, title, DOI, publication year), sample characteristics (food source type, analytical instrument, sample size, sample year), and pesticide data (name, summary statistics) (See SI Table S3). All analyzed pesticides, regardless of detection status, were extracted from included studies along with analytical instrument used and respective limits of detection and quantification (LOD/LOQ) (Table S4). Raw pesticide residue concentrations were prioritized (i.e., unique sample/location measurements), and if unavailable, summary statistics (e.g., mean and standard deviation) were extracted. Data from figures were extracted using PlotDigitizer.

Study identification, screening and inclusion as well as data extraction were primarily performed by one author (EMA) and independently verified by the remaining authors, with disagreements resolved through discussion. The overall process was summarized using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram as shown in SI Figure 1.

# Background and Justifications

## Handling non-detects (left-censorship)

The presence of non-detects, i.e., measurements below the analytical methods detection limit (also known as censored data) are common in environmental monitoring dataset, and this is also true for our residue dataset (41%). Another challenge is the difficulty to discern NDs from true zeros, i.e., whether NDs truly represent absence of pesticide residues or inability to measure the already presenting concentration above zero. The probability of zero expansion or zero-inflation in our dataset is high given the high proportion of NDs, however, properly handling the issue is difficult since we don't have explicit detection rates or information to rigorously estimate the detection rate. When a laboratory test did not detect a chemical, the chemical concentration in the sample could have any value between zero and the LOD.

In handling left-censored data, regulatory bodies recommend a traditional simple substitution methods for NDs (e.g., replacing with 1/2 DL or zero). However, studies have showed that the method is known to introduce significant bias in datasets with greater than 30% missing values. In this regard, most regulatory bodies suggest replacing NDs with zero, the detection limit (DL), or half the DL instead simply removing them (EFSA, 2010; US EPA, 2000a). However, this has been a center of debate for a long time (Helsel, 2006), and it has now been shown that such simple substitution introduce a significant, leading to inaccurate risk estimations especially when the proportion of ND is high. Results showed that the number of samples had a relatively limited impact on the accuracy and precision of estimates, but the degree of censoring had a large effect. When analysing a complex set of data, it was also shown that it is essential to identify possible sources of heterogeneity in a dataset, such as country of sample collection/origin, food group, laboratory, etc. Statistical analyses should either be conducted separately from these factors, or, to explicitly account for this heterogeneity, fixed/random effect ML models could be used (EFSA, 2010).

Given the substantial proportion of left-censored data (NDs), several advanced imputation methods were considered and compared: Zero-Inflated Lognormal (ZILN), Maximum Likelihood Estimation (MLE) with random draws, Regression on Order Statistics (ROS), and Kaplan–Meier (KM) imputation (Canales *et al.*, 2018; Gómez-Carracedo *et al.*, 2014; Lee *et al.*, 2024). In brief, ZILN method account for the possibility of true zero concentrations (complete absence) alongside concentrations below the DL. For NDs, a heuristic approach was used where 50% were randomly assigned as "true zeros" and the remaining were imputed through truncated lognormal random draws, based on the estimated parameters from the censored lognormal fit. We introduced ZILN, a method which assumes a mix of true zeros and lognormal values (Sang *et al.*, 2024), depending on the suspected likelihood of zero-expansion in our data (SI Figure 3), but it is important to note that unless there is reason to assume that a food does not contain a residue (which there is none in our case), it should be assumed that NDs may contain the actual pesticide (WHO/FAO, 2009). MLE method involved fitting a censored lognormal distribution to the data and then imputing NDs by drawing random values from the fitted distribution, truncated at the respective detection limits, reintroducing variability into the imputed non-detects. ROS method provides a semi-parametric fit under a lognormal assumption and is robust even with moderate censoring, estimating concentrations for NDs based on the observed values and their ranks. KM method is a non-parametric method that does not assume any underlying distribution and it uses the empirical cumulative distribution function derived from both detected and non-detected values to impute missing concentrations, where NDs were imputed by drawing from the estimated survival function below the respective DLs. A comparative analysis revealed that the KM imputation most closely preserved the original distribution of the detected concentrations, thus selected for subsequent analyses, while others underestimated central values by pulling imputed NDs towards lower concentrations.

## Handling Missing Data

It has been widely noted that meta-analyses in many disciplines commonly encounter missing and incompletely reported data in original publications, especially for variance measures (Parker *et al.*, 2016)(Gurevitch *et al.*, 2018). The most common approach to dealing with missing data is to delete cases containing missing observations. However, this approach reduces statistical power and increases estimation bias (Nakagawa and Freckleton, 2008). Various previous studies have suggested that multiple imputations can yield grand mean estimates that are less biased than those obtained from complete-case analyses (Azur *et al.*, 2011; Ian *et al.*, 2011; Kambach *et al.*, 2020). We derived missing residue means from geometric means (n = 8) and medians (n = 18), and SDs using modified Hozo method (n = 10) (SI Section 3).

For the remaining SD and sample size, we used multiple imputation using multivariate imputation by chained equations (MICE) (Buuren and Groothuis-Oudshoorn, 2011). MICE is popular due to its flexibility, and it has been shown to outperform classical data imputation methods (Pridham *et al.*, 2022). Our imputation model was run based on the assumption missingness is at random and highly depend on observed mean using predictive mean matching over 50 imputations and 20 iterations (SI Section 3.5). Convergence diagnostics confirmed the plausibility and robustness of imputed values (SI Figure 6). The final complete dataset was selected based on minimal absolute percentage error from observed values.

Typically Rubin’s rules are applied to combine the results from the multiple imputed datasets after analysis, however, given our large dataset, complexity of implemented MMA models, and computation limitations, we selected only one complete dataset with minimal absolute percentage error from observed values (SI Section 3.5).

## Multilevel Meta-analysis

Multilevel meta-analysis offers a powerful statistical technique for synthesizing such complex datasets, along with a better understanding of sources of variability, leading to more reliable and informative conclusions for environmental management and policy (Nakagawa *et al.*, 2023). Detailed background and practical applications of MMA are provided elsewhere (Assink and Wibbelink, 2016; Harrer *et al.*, 2021; Nakagawa *et al.*, 2023; Van den Noortgate *et al.*, 2013).

## Probabilistic Risk assessment

Traditional deterministic risk assessments often use single “point estimates” (e.g., average or worst-case values) and this could lead to an overly conservative, and potentially unrealistic, overall risk estimate. In contrast, PRA combined with Monte Carlo simulations is recommended in regulatory settings because it provides a scientifically robust, comprehensive and realistic risk estimates, generating range of outcomes that reflects true population variability and uncertainty. More details on PRA can be here (Flinders *et al.*, 2025; Khalid, 2023; US EPA, 2000b, 2014). Traditional deterministic risk assessments often rely on single-point estimates for exposure parameters, which inherently fail to capture the full extent of variability and uncertainty present within a population. In contrast, probabilistic models, such as those employing Monte Carlo simulations coupled with Latin Hypercube Sampling (LHS), provide a more realistic and comprehensive estimation of chemical intake and associated health risks. This approach explicitly accounts for the natural variations in food consumption patterns, pesticide residue concentrations, and individual body weights across a diverse population.

In dietary intake assessments, the concentration data used will depend on the nature of the specific intake assessment. The concentration of an ingredient or chemical constituent in food can be obtained from

**Pesticide Registration and Control Regulation, *Legal basis for assessing consumer exposure through food*.** *Schedule II – Article 1.1.4***.** *(The Ministry ... shall evaluate ...) the exposure of consumers and animals through their diet following the intended uses and under locally relevant conditions of use, and:*

*a. The pesticide shall not be registered if its intended use will lead to residue levels at harvest, slaughter or after storage or processing, as appropriate, which exceed the nationally established maximum residue limit (MRL) or a provisional MRL. b. In the absence of a nationally established MRL or provisional MRL, Codex Alimentarius MRLs shall apply, if established for the commodity and pesticide under review. c. Taking into account all registered uses of the pesticide, the intended use shall not be authorized if the estimated total dietary exposure exceeds the Acceptable Daily Intake (ADI) or the Acute Reference Dose (ARfD).*

As Ethiopia is a member of CODEX Alimentarius, the CODEX MRLs (CXLs) are used as a basis for risk assessment. Where CODEX MRLs do not cover the use of a plant protection product in Ethiopia, no national MRL will be set as appropriate national Ethiopian legislation is currently not in place. For the chronic consumer risk assessment the WHO-GEMS IESTI model version revision 14 is used using the GEMS food consumer cluster diets from August 2006. The model has been slightly adapted to Ethiopian circumstances by marking irrelevant commodities in red and by adding teff as a commodity. The model ***IEDI\_calculation\_Ethiopia.xltm*** enables the calculation of the International Estimated Daily Intake (IEDI) based on estimated Supervised Trial Median Residue (STMR, STMR-P) or the Maximum Residue Level (MRL) for relevant commodities. It summarizes the total intake in mg/person/day and calculates the total intake as percentage of the Acceptable Daily Intake (ADI) for 13 world food clusters. Food cluster A is considered appropriate for Ethiopia. A scientifi c evaluation system for the registration of pesticides in Ethiopia

# Novelty and Impact

The novelty of our methodological framework is justified by: (i) comparing four methods and selecting the best-performing approach to minimize bias in left-censored data (i.e., Zero-Inflated Lognormal, MLE, Regression on Order Statistics, KM); (ii) incorporating RVE into MMA models to quantify heterogeneity while also adjusting for small-sample biases; (iii) utilizing 20,932 unique dietary records from 7,527 households to derive nationally representative food consumption rates; and (iv) being the first study of its kind in Africa, and among the few in the Global South, to provide comprehensive and representative dietary pesticide risk estimates.

First, our analytical methodological innovations are novel. We present one of the first known implementations in a low-income country setting of a full pipeline combining three-level multilevel meta-analysis with population-wide probabilistic risk assessment via Monte Carlo–based Latin Hypercube Sampling. This allows us to accurately capture hierarchical variance (sampling, within-study, and between-study) and generate distributions of dietary exposure and health risk—not mere point estimates. Furthermore, we rigorously compared and adopted a Kaplan–Meier nonparametric estimator for handling heavily left-censored residue data; this is rare in environmental meta-analyses and outperforms standard approaches (e.g., ZILN, MLE, ROS) in preserving distributional integrity. Simultaneously, we used MICE with predictive mean matching to impute missing sample sizes and variances, minimizing bias and maximizing data retention.

Second, the study addresses a long-standing geographic data void—Ethiopia and much of sub-Saharan Africa lack national dietary pesticide risk assessments. We synthesized over 2,200 unique residue measurements from 40 local studies covering 18 food subgroups, filling a critical information gap. Our analysis revealed persistent exposure to legacy organochlorine pesticides (e.g., DDT, endosulfan), especially in high-risk foods like khat and honey—highlighting ongoing public health exposures despite global bans.

Third, the impact pathways are substantial. By identifying priority pesticide–food combinations and quantifying non-cancer (HQ) and cancer risk burdens, we provide actionable evidence to inform Ethiopian regulators in policymaking—such as revising MRLs, targeting monitoring, or phasing out high-risk chemicals. Moreover, our methodology offers a replicable template for LMICs and other data-scarce regions, showcasing how advanced statistical frameworks (e.g., KM, MICE, MMA, PRA) can enable robust exposure analysis even with fragmented datasets. This has implications for global food safety frameworks and the evolution of risk analysis in environmental health.

In summary, our integration of cutting-edge missing data handling, hierarchical meta-analysis, and probabilistic risk modeling within a poorly-studied LMIC context contributes uniquely to exposure science. It advances methodological rigor, fills continental data gaps, and delivers a policy-ready assessment with wider applicability to other global-South settings—thereby amplifying both the novelty and impact of dietary pesticide exposure research.

References

Assink, M. and Wibbelink, C.J.M. (2016), “Fitting three-level meta-analytic models in R: A step-by-step tutorial”, *The Quantitative Methods for Psychology*, Vol. 12 No. 3, pp. 154–174, doi: 10.20982/tqmp.12.3.p154.

Azur, M.J., Stuart, E.A., Frangakis, C. and Leaf, P.J. (2011), “Multiple imputation by chained equations: What is it and how does it work?”, *International Journal of Methods in Psychiatric Research*, Vol. 20 No. 1, pp. 40–49, doi: 10.1002/mpr.329.

Buuren, S. van and Groothuis-Oudshoorn, K. (2011), “mice : Multivariate Imputation by Chained Equations in R”, *Journal of Statistical Software*, Vol. 45 No. 3, pp. 1–67, doi: 10.18637/jss.v045.i03.

Canales, R.A., Wilson, A.M., Pearce-Walker, J.I., Verhougstraete, M.P. and Reynolds, K.A. (2018), “Methods for handling left-censored data in quantitative microbial risk assessment”, *Applied and Environmental Microbiology*, Vol. 84 No. 20, doi: 10.1128/AEM.01203-18.

EFSA. (2010), “Management of left‐censored data in dietary exposure assessment of chemical substances”, *EFSA Journal*, Vol. 8 No. 3, doi: 10.2903/j.efsa.2010.1557.

Flinders, C., Barnhart, B., Morrison, E.B., Anderson, P.D. and Landis, W.G. (2025), “Probabilistic approaches for risk assessment and regulatory criteria development: current applications, gaps, and opportunities”, *Integrated Environmental Assessment and Management*, doi: 10.1093/inteam/vjaf016.

Gómez-Carracedo, M.P., Andrade, J.M., López-Mahía, P., Muniategui, S. and Prada, D. (2014), “A practical comparison of single and multiple imputation methods to handle complex missing data in air quality datasets”, *Chemometrics and Intelligent Laboratory Systems*, Vol. 134, pp. 23–33, doi: 10.1016/j.chemolab.2014.02.007.

Gurevitch, J., Koricheva, J., Nakagawa, S. and Stewart, G. (2018), “Meta-analysis and the science of research synthesis”, *Nature*, Vol. 555 No. 7695, pp. 175–182, doi: 10.1038/nature25753.

Harrer, M., Cuijpers, P., Furukawa, T.A. and Ebert, D.D. (2021), *Doing Meta-Analysis With R: A Hands-On Guide*, 1st ed., Chapman & Hall/CRC Press, Boca Raton, FL and London.

Helsel, D.R. (2006), “Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it”, *Chemosphere*, Vol. 65 No. 11, pp. 2434–2439, doi: 10.1016/j.chemosphere.2006.04.051.

Ian, R.W., Patrick, R. and W., A.M. (2011), “Multiple imputation using chained equations: Issues and guidance for practice”, *Statistics in Medicine*, Vol. 30 No. 4, pp. 377–399.

Kambach, S., Bruelheide, H., Gerstner, K., Gurevitch, J., Beckmann, M. and Seppelt, R. (2020), “Consequences of multiple imputation of missing standard deviations and sample sizes in meta-analysis”, *Ecology and Evolution*, Vol. 10 No. 20, pp. 11699–11712, doi: 10.1002/ece3.6806.

Khalid, M. (2023), “Monte Carlo analysis for probabilistic risk assessment”, *Encyclopedia of Toxicology, Fourth Edition: Volume 1-9*, Vol. 6, pp. V6-517-V6-522, doi: 10.1016/B978-0-12-824315-2.00109-3.

Lee, M., Saha, A., Sundaram, R., Albert, P.S. and Zhao, S. (2024), “Accommodating detection limits of multiple exposures in environmental mixture analyses: an overview of statistical approaches”, *Environmental Health: A Global Access Science Source*, Vol. 23 No. 1, doi: 10.1186/s12940-024-01088-w.

Nakagawa, S. and Freckleton, R.P. (2008), “Missing inaction: the dangers of ignoring missing data”, *Trends in Ecology and Evolution*, Vol. 23 No. 11, pp. 592–596, doi: 10.1016/j.tree.2008.06.014.

Nakagawa, S., Yang, Y., Macartney, E.L., Spake, R. and Lagisz, M. (2023), “Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences”, *Environmental Evidence*, Vol. 12 No. 1, p. 8, doi: 10.1186/s13750-023-00301-6.

Van den Noortgate, W., López-López, J.A., Marín-Martínez, F. and Sánchez-Meca, J. (2013), “Three-level meta-analysis of dependent effect sizes”, *Behavior Research Methods*, Vol. 45 No. 2, pp. 576–594, doi: 10.3758/s13428-012-0261-6.

Parker, T.H., Nakagawa, S. and Gurevitch, J. (2016), “Promoting transparency in evolutionary biology and ecology”, *Ecology Letters*, Vol. 19 No. 7, pp. 726–728, doi: 10.1111/ele.12610.

Pridham, G., Rockwood, K. and Rutenberg, A. (2022), “Strategies for handling missing data that improve Frailty Index estimation and predictive power: lessons from the NHANES dataset”, *GeroScience*, Vol. 44 No. 2, pp. 897–923, doi: 10.1007/s11357-021-00489-w.

Sang, C., Niu, Y., Gao, Q., Zhang, J., An, W., Shao, B. and Yang, M. (2024), “Characterizing the cumulative health risks of 19 kinds of pesticides in Chinese food from the cancer and non-cancer perspective”, *Journal of Environmental Management*, Vol. 351, p. 119813, doi: 10.1016/j.jenvman.2023.119813.

US EPA. (2000a), “Assigning values to non-detected/non-quantified pesticide residues in human health food exposure assessments”, No. 6047, pp. 1–25.

US EPA. (2000b), “Risk characterization Handbook”, No. U.S. Environmental Protection Agency, p. 57.

US EPA. (2014), “Risk Assessment Forum White Paper: Probabilistic Risk Assessment Methods and Case Studies. EPA/100/R-14/004”, No. July, p. 98.

WHO/FAO. (2009), “Principles and Methods for the Risk Assessment of Chemicals in Food”, *International Journal of Environmental Studies*, pp. 1–7.