PEER-REVIEWED · OPEN ACCESS
Sriwijaya Journal of Forensic and Medicolegal
Vol. 3, No. 2 (2025), pp. 104–114 · e-ISSN 2987-1530 · Article ID: 326
journal homepage: phlox.or.id/index.php/sjfm

Original Research Article

Age and Sex Associations With Body Mass Index in an Audited Postmortem CT Metadata Cohort

Rachmat Hidayat[1*], Muhammad Yoshandi[2]

[1]*Department of Medical Biology, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia*
[2]*Department of Health Sciences, Tembilahan Community Health Center, Tembilahan, Indonesia*

ARTICLE INFO

Article history
Received: September 18, 2025
Revised: October 20, 2025
Accepted: November 19, 2025
Available online: November 25, 2025
Published: November 30, 2025

Keywords
Anthropometry
Body mass index
Forensic radiology
Open data
Postmortem computed tomography

Corresponding author
Rachmat Hidayat
E-mail: dr.rachmat.hidayat@gmail.com
ORCID: 0000-0001-5770-1201

DOI
https://doi.org/10.59345/sjfm.v3i2.326

ABSTRACT

Background. Open postmortem computed tomography (PMCT) repositories support forensic research, but record duplication and incomplete metadata can distort anthropometric inference.

Objective. To examine age- and sex-related variation in body mass index (BMI) after curator-informed audit of VSDFullBody.

Methods. We performed a retrospective secondary analysis of 105 metadata rows. A calibration phantom was excluded; seven concordant duplicate pairs were collapsed; both records from four duplicate pairs with conflicting metadata were excluded; curator data reconciled six sex labels and recovered 13 BMI values. The primary cohort comprised 89 decedents, including 83 complete cases. BMI was regressed on age per 10 years and audited biological sex using HC3 standard errors; bootstrap, duplicate-choice, robust-regression, leave-one-out, head-coverage, and missing-outcome sensitivity analyses assessed stability.

Results. Mean BMI was 24.2 kg/m2. Each 10-year age increment was associated with 0.49 kg/m2 higher BMI (95% confidence interval [CI], 0.05-0.93; p=0.029). Adjusted BMI was 2.58 kg/m2 higher in male than female decedents (95% CI, 0.18-4.99; p=0.036); the bootstrap 95% CI was 0.34-4.88. Sensitivity estimates retained the same direction.

Conclusion. In this audited PMCT metadata cohort, age and male sex were associated with modestly higher BMI. Curator companion files materially increased analyzable completeness, supporting explicit metadata audit before reuse of open forensic imaging collections.

How to cite. Hidayat R, Yoshandi M. Age and sex associations with body mass index in an audited postmortem CT metadata cohort. Sriwijaya J Forensic Medicoleg. 2025;3(2):104–114. doi:10.59345/sjfm.v3i2.326

1. Introduction

Postmortem computed tomography (PMCT) has developed from an experimental adjunct into an established component of forensic imaging practice. It preserves a three-dimensional digital record before dissection, permits repeated review of anatomical findings, and can reveal skeletal injury, gas distribution, foreign material, and spatial relationships that may be difficult to retain after conventional autopsy. Recent prospective and observational comparisons show that PMCT and autopsy have different diagnostic strengths rather than being interchangeable in every circumstance, supporting a complementary and question-specific role in medicolegal investigation.[1–4] Optimization of forensic CT reconstruction further demonstrates that acquisition and postprocessing choices influence image quality and diagnostic potential.[5]

The breadth of contemporary PMCT research illustrates its practical range. Recent studies have evaluated lung findings with autopsy correlation, bleeding-source localization, dental identification, automated detection of fatal cerebral hemorrhage, interpretation by forensic pathologists, suicidal hanging, fatal blunt injury, and thermal skeletal injury.[6–13] Together, these investigations show that a durable PMCT record can support both case evaluation and reproducible method development, while the appropriate claim remains dependent on the anatomy, pathology, and reference standard studied.

Open repositories amplify this scientific value by making digital specimens available beyond the institution that generated them. The Virtual Skeleton Database (VSD) was designed as an open-access environment for biomedical research and collaboration, and its VSDFullBody collection provides anonymized whole-body cadaver CT archives with associated demographic and anthropometric metadata.[14,15] These resources are particularly relevant to forensic science, where local case volumes, governance requirements, scanner access, and the practical impossibility of repeating an examination can restrict replication.

The benefit of open access, however, depends on the analytical unit being correctly identified. A downloadable archive is not necessarily equivalent to a unique decedent. A repository may contain repeated scans, alternative reconstructions, incomplete uploads, calibration objects, or corrections maintained outside the primary metadata export. Treating each archive as an independent person can alter denominators, narrow uncertainty artificially, and give individuals with multiple files disproportionate influence. Conversely, excluding every incomplete record without consulting curator documentation can discard recoverable age, sex, height, or weight information. These issues are not merely clerical: they define which observations enter a model and therefore shape the estimates that readers interpret as evidence.

For secondary forensic imaging research, provenance includes the relationship between an image archive, its descriptive record, curator notes, and any later correction file. An analysis that links these layers can distinguish true missingness from information stored in a companion source, identify documented duplicate relations, and retain an auditable account of each transformation. This approach is especially important for an open collection whose users were not present during acquisition and cannot resolve discrepancies by returning to the original examination.

Age and sex are central descriptors in forensic imaging and identification research. Quantitative CT, three-dimensional skeletal reconstructions, and deep-learning approaches have recently been evaluated for age estimation across adult and subadult samples and for forensic sex estimation.[16–21] These studies also demonstrate that imaging-derived models are population-, anatomy-, and method-dependent, reinforcing the need to describe the age and sex structure of any reused forensic CT cohort before broader application.

Height, weight, and body mass index (BMI) provide a complementary anthropometric description. Recent postmortem research found meaningful error in visual and circumference-based weight estimation, whereas automated CT-based whole-body measurement showed close agreement with measured forensic weight in a dedicated validation cohort.[22,23] BMI expresses weight relative to squared height and offers a familiar summary, but it is not a direct measurement of body fat, muscle mass, visceral adiposity, or nutritional status.

Contemporary CT and population studies demonstrate why this boundary matters. Age and BMI have different relationships with visceral, subcutaneous, hepatic, marrow, and muscle adiposity; age- and sex-specific CT reference values vary across compartments; automated tools require validation across scanners and institutions; and anthropometric distributions vary with sex, health status, and population background.[24–29] Outcome-oriented CT studies further show that tissue measures may add information beyond BMI, while direct validation in a diverse clinical cohort confirms that BMI and CT-derived composition are related but not equivalent.[30–32] A metadata-only analysis therefore cannot be presented as voxel-derived body composition. Its valid contribution is to characterize recorded anthropometry, evaluate cohort coherence and completeness, and generate reproducible estimates that guide later image-based studies.

Postmortem samples introduce additional selection processes. Decedents referred for forensic examination are not a probability sample of the living population; referral patterns depend on jurisdiction, manner and cause of death, circumstances of discovery, and institutional practice. Postmortem interval, decomposition, medical intervention, trauma, fluid shifts, and incomplete documentation can affect the availability or meaning of anthropometric fields. Open dissemination adds a second selection layer because only some examinations are shared and only some records contain reusable metadata. Consequently, associations in an open PMCT collection describe the audited cohort and should not be converted into population prevalence, diagnostic thresholds, or causal statements.

The VSDFullBody repository is accompanied by a curator workbook that records empty archives, duplicate CTs, conflicting metadata, and recognized sex-label discrepancies. That companion source makes it possible to construct an explicit decedent-level audit rather than relying only on the user-facing metadata extraction. It also permits a substantive methodological question: whether the age- and sex-related BMI estimates remain stable when documented duplicate conflicts, restored anthropometry, influence, model form, and residual missing outcomes are examined transparently. Addressing that question can improve the immediate analysis while illustrating a reproducible strategy for reuse of other open forensic imaging collections.

The novelty of this study is the integration of curator-documented archive relationships and metadata corrections into a decedent-level anthropometric analysis of VSDFullBody, coupled with prespecified robustness checks that separate estimate magnitude, uncertainty, duplicate choice, influence, and missing-outcome assumptions. The primary aim was to estimate the independent associations of age at death and audited biological sex with BMI in the reconciled VSDFullBody cohort. The secondary aims were to quantify how curator reconciliation changed cohort completeness, characterize BMI availability by head-coverage category and sex, and evaluate the stability of the associations using heteroscedasticity-consistent inference, bootstrap resampling, alternative duplicate selections, robust regression, leave-one-out analysis, head-coverage adjustment, and pattern-mixture sensitivity analysis.

2. Methods

2.1 Study design and reporting framework

This study was a retrospective cross-sectional secondary analysis of an existing open PMCT metadata collection. The design was selected because the exposure variables and anthropometric outcome were recorded at the level of the archived examination, with no longitudinal follow-up or intervention. The analysis focused on associations within the available cohort and did not attempt to estimate causal effects or population prevalence. Planning, cohort construction, statistical analysis, and reporting were organized according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.[33]

2.2 Data sources and scope

The primary user-supplied file was researchpedia-dataset.csv, containing 105 rows derived from archives associated with the VSDFullBody record on Zenodo. Each row included an archive identifier, age information, a sex label, descriptive scan information, and anthropometric fields where available. The curator-linked VSD_Comments.xlsx workbook served as the authoritative companion for documented archive status, duplicate relations, corrected sex labels, and selected height and weight values.[15] The repository and companion workbook were treated as linked components of one data source rather than as competing datasets. No raw DICOM images, reconstructed CT volumes, segmentations, or voxel intensities were downloaded or analyzed. All conclusions therefore concern repository metadata and derived BMI rather than radiological diagnoses or CT tissue composition.

The unit of analysis was a unique decedent represented by a usable archive after reconciliation. Archive rows were retained as the initial audit units because duplicate relationships and curator notes were expressed through archive identifiers. The audit then mapped those rows to the final decedent-level cohort. This distinction prevented the raw number of downloadable records from being interpreted automatically as the number of individuals. The analytic dataset preserved the original identifier, curator-linked values, derived variables, reason for exclusion or collapse, duplicate group, and final inclusion status so that every decision could be traced from the raw metadata to the model input.

2.3 Eligibility and cohort audit

All 105 supplied rows entered the audit. Phantom001.zip was excluded because it represented a calibration phantom rather than a human decedent. Curator comments were parsed for explicit duplicate relations. Eleven duplicate pairs were identified. Seven pairs were classified as concordant when curator sex, age, height, and weight agreed within prespecified tolerances of no more than 1.1 years for age, 1.0 kg for weight, and 0.02 m for height. These tolerances allowed for rounding or minor recording differences while requiring the records to describe the same anthropometric profile. Each concordant pair was collapsed by retaining the lexicographically first archive identifier as the reproducible representative record.

Four documented duplicate pairs contained larger anthropometric discrepancies or sex-discordant information. For the primary cohort, both records from each conflicting pair were excluded because the available evidence did not justify privileging one record as the accurate representation of the decedent. This conservative rule avoided arbitrary selection and prevented unresolved metadata from entering the primary model. It also reduced information, so the effect of the decision was examined in two sensitivity cohorts: one retained the first member of every conflicting pair and the other retained the second. The concordant-pair and conflicting-pair rules were defined separately because a repeated scan with coherent metadata raises a dependence problem, whereas a repeated scan with incompatible metadata also raises a measurement-identification problem.

2.4 Metadata reconciliation and variable derivation

Curator values were applied through a documented hierarchy. When the curator workbook provided a sex value, that value replaced the embedded label for analysis; otherwise, the embedded value was retained. Six retained records changed sex classification through this reconciliation. The variable was termed audited biological sex because the source documentation provided female and male categories. The analysis did not infer gender identity and did not independently determine sex from CT morphology. Female was used as the regression reference group, making the reported sex coefficient the adjusted mean BMI difference for male relative to female decedents within the analytic cohort.

Height and weight were reconciled from curator fields when the primary extraction was missing or when the companion workbook documented the applicable value. BMI was recalculated for every eligible record using BMI = weight (kg) / [height (m)]^2. Recalculation applied one formula consistently and avoided mixing stored BMI values with newly derived values of unknown precision. Thirteen BMI values unavailable in the initial extraction became calculable after curator height and weight were incorporated, and one additional height value was reconciled. Implausible arithmetic results were checked against their component variables, while no value was removed merely because it was high or low. The retained outcome thus reflected documented anthropometry rather than an investigator-imposed BMI range.

Age at death was calculated from source age in days divided by 365.25. For regression, age was expressed per 10-year increment and centered at the complete-case mean. Scaling by decade produced a coefficient with a clinically and forensically interpretable unit, while centering made the model intercept correspond to the predicted BMI for a female decedent of average analytic age. Centering did not change the estimated age slope, sex difference, fitted values, or model fit. Head coverage was derived from the archive description and categorized as head included, no head, or unspecified. It was not considered a biological exposure; it was used to describe repository structure, evaluate outcome completeness, and assess whether a scan-description feature influenced the primary estimates.

2.5 Descriptive and bivariate analyses

Cohort derivation was summarized by counts at each audit stage and by the number of metadata corrections contributing to the final dataset. Continuous variables were described using mean and standard deviation together with median and interquartile range, allowing readers to assess central tendency without relying on a single distributional summary. Categorical variables were summarized as counts and percentages. Descriptive statistics were calculated for the complete-case cohort and stratified by audited sex where relevant. The number and proportion of records with BMI were also reported within sex and head-coverage categories.

The unadjusted female-male BMI comparison used Welch's two-sample t test, which does not require equal group variances. The difference was expressed as male minus female so that its direction matched the regression coefficient. Linear age-BMI association was described with Pearson's correlation and a 95% confidence interval. Spearman's rank correlation was added because it is less dependent on a linear relationship and the influence of exact extreme values. These bivariate analyses were descriptive companions to the multivariable model, not alternative primary tests. Differences between their p-values were interpreted through effect estimates, uncertainty, covariate adjustment, and distributional features rather than as contradictory binary outcomes.

2.6 Primary multivariable model

The primary analysis used ordinary least-squares regression with continuous BMI as the outcome and centered age per 10 years plus audited sex as predictors.

The model was intentionally parsimonious. The available dataset did not include sufficiently complete information on cause of death, ancestry, disease burden, postmortem interval, decomposition, treatment, or socioeconomic factors to support broader confounder adjustment. Adding variables simply because they were technically available would not resolve unmeasured selection and could reduce precision in this modest sample. The coefficients were therefore presented as mutually adjusted within-cohort associations rather than causal effects.

Heteroscedasticity-consistent HC3 standard errors were used for the primary 95% confidence intervals and p-values. HC3 applies greater leverage adjustment than simpler sandwich estimators and is appropriate when sample size is moderate and influential observations may be present. Conventional ordinary least-squares coefficients were retained because the estimand was the adjusted conditional mean difference, while robust standard errors reduced dependence of inference on constant residual variance. All tests were two-sided. The threshold $p<0.05$ was used descriptively and was not treated as a guarantee of replication, clinical importance, or absence of residual bias.

2.7 Model diagnostics

Model form and stability were assessed rather than assumed. A quadratic age term was compared with the linear specification using Akaike's information criterion to evaluate whether the additional curvature improved fit sufficiently to justify greater complexity. Residual normality was examined with a Shapiro-Wilk test and graphical review. The Breusch-Pagan procedure screened for heteroscedasticity. The condition number and variance-inflation measures assessed collinearity. Cook distance identified observations with a combination of leverage and residual magnitude, using 4/n as a conventional screening threshold rather than an automatic exclusion rule. No observation was deleted solely because it exceeded that threshold.

Leave-one-out analysis refitted the primary model after omitting each complete case in turn and recorded the range of the sex coefficient. This procedure directly addressed whether the adjusted sex estimate depended on one person. A stable coefficient range would not prove that the model was unbiased, but it would show that the observed association was not created by a single retained record. Diagnostic findings were interpreted jointly with the bootstrap and robust-regression analyses because each method addresses a different source of uncertainty.

2.8 Sensitivity and supplementary analyses

Nonparametric bootstrap analysis drew 5,000 samples with replacement from the complete-case cohort, refitted the same age-plus-sex model, and obtained a percentile 95% confidence interval for the sex coefficient. Resampling was performed at the reconciled decedent level, not the raw archive-row level. The bootstrap therefore represented sampling variability among audited individuals and did not reintroduce collapsed duplicates. Its interval was compared with the HC3 interval to evaluate whether the substantive conclusion depended on the large-sample approximation used by the sandwich estimator.

Three specification analyses evaluated cohort and model choices. First, head-coverage category was added to the primary model to assess whether a repository-description feature associated with BMI completeness also altered the age- and sex-related estimates. Second, the primary model was repeated in the two alternative duplicate-choice cohorts that retained either the first or the second member of each conflicting pair. Third, Huber robust regression reduced the influence of observations with large residuals. Agreement in direction and similar magnitude across these analyses was considered evidence of numerical stability, whereas statistical significance was evaluated separately because standard errors and estimands vary across methods.

Six audited decedents lacked BMI after reconciliation. Fisher exact tests evaluated whether completeness differed by audited sex or head-coverage category; exact tests were selected because some category counts were small. The missing outcome itself could not be recovered or verified from the available files, and its mechanism could not be identified from observed data. A deterministic pattern-mixture sensitivity analysis was therefore used to show how departures from the complete-case assumption would affect the sex estimate. Missing male BMI values were assigned their model-based predictions plus a common delta ranging from -30 to +4 kg/m^2, and the age-plus-sex model was refitted at each value. The analysis recorded the delta at which the 95% confidence interval crossed zero and the larger shift required to reverse the point estimate. These deltas are transparent stress tests, not estimates of the unobserved values.

2.9 Software, reproducibility, and ethics

Analyses were conducted in R version 4.5.3 using a fixed computational workflow. The submission package includes the analysis script, raw-input checksums, an audited record-level dataset, duplicate and exclusion logs, machine-readable tables, model summaries, and the figure source. Separating raw, reconciled, complete-case, and sensitivity cohorts allowed every reported denominator and estimate to be regenerated. Numerical values in the manuscript were transferred from those outputs rather than recalculated manually during document formatting.

The source collection is anonymized and publicly available under a Creative Commons Attribution-NonCommercial-ShareAlike license. The repository reports that the shared collection followed approval by the Cantonal Ethics Committee Bern.[15] Ethical approval for this secondary analysis was granted by the CMHC Ethics Committee Indonesia (No. 008/CMHC/Ethics/08/2025). No new contact with living participants or decedents' relatives occurred, and no new consent was obtained for this metadata-only secondary analysis. Original consent documentation was not available in the accessed materials and was therefore not inferred.

3. Results

3.1 Cohort construction and metadata reconciliation

The audit began with 105 archive-level rows. One row, Phantom001.zip, was removed because it represented a calibration object, leaving 104 human archive records for duplicate assessment. The curator companion identified 11 duplicate pairs. Collapsing the seven concordant pairs removed seven repeated representations while retaining one reproducible archive identifier for each decedent. Excluding both members of the four conflicting pairs removed a further eight records. The primary audited cohort consequently contained 89 unique decedents. The complete archive-to-decedent flow from 105 raw rows to 89 audited decedents and 83 complete cases is detailed in Table 3.

Metadata reconciliation altered both availability and classification. Curator values corrected six sex labels among retained records. Curator height and weight information made 13 BMI values calculable that had been unavailable in the embedded extraction, and one additional height value was reconciled. These changes increased the amount of usable anthropometric information without adding new individuals. After all available corrections were applied, six of the 89 audited decedents still lacked BMI. Thus, 83 decedents (93.3% of the audited cohort) had complete age, sex, and BMI and entered the primary regression; 28 were female and 55 were male.

3.2 Participant characteristics

Among the 83 complete cases, mean age was 53.2 years (SD 20.5). Mean BMI was 24.2 kg/m^2 (SD 4.7), indicating appreciable variation around the central value. Female decedents had a mean BMI of 22.7 kg/m^2 (SD 5.6), whereas male decedents had a mean BMI of 24.9 kg/m^2 (SD 4.1). The larger female standard deviation indicates that group means alone do not describe the full overlap between distributions; complete overall and sex-stratified anthropometric summaries are detailed in Table 1.

The unadjusted difference in mean BMI was 2.22 kg/m^2 higher for male than female decedents. Its Welch 95% confidence interval ranged from -0.20 to 4.64 kg/m^2, with p=0.072. The interval therefore included small differences in either direction while remaining compatible with a positive difference of practical size. This crude comparison did not account for age distribution. It was used to describe the sample rather than to override the adjusted analysis specified as primary.

3.3 Age and BMI in bivariate analyses

Pearson's correlation between age and BMI was 0.169 (95% CI, -0.049 to 0.371; p=0.127), representing a small positive linear association with substantial uncertainty. Spearman's rank correlation was 0.224 (p=0.042), also indicating that older decedents tended to occupy somewhat higher positions in the BMI distribution. The different p-values did not imply opposing effects: both coefficients were positive and modest. The distinction reflects the methods' different sensitivity to linear spacing, ranking, and the observed distribution.

3.4 Primary age- and sex-adjusted model

In the primary ordinary least-squares model, each 10-year increment in age was associated with 0.49 kg/m^2 higher BMI after adjustment for audited sex (HC3 95% CI, 0.05 to 0.93; p=0.029). Expressed over a 20-year difference, the fitted linear contrast would be approximately 0.98 kg/m^2, provided that the model is used within the observed age range. The coefficient describes a cross-sectional association among decedents and does not represent within-person aging or a causal change in body mass.

At the same modeled age, male decedents had a BMI 2.58 kg/m^2 higher on average than female decedents (HC3 95% CI, 0.18 to 4.99; p=0.036). The observed data and sex-specific fitted lines are shown in Figure 1A, while the complete primary and sensitivity-model coefficients are detailed in Table 2. The interval indicates that the data were compatible with a small positive adjusted difference as well as one approaching 5 kg/m^2. Accordingly, the point estimate is more informative than a significance label alone, while the interval conveys the remaining imprecision.

3.5 Model diagnostics and influence

Adding a quadratic age term did not improve Akaike's information criterion, supporting retention of the linear age specification for this dataset. The Breusch-Pagan test did not identify important heteroscedasticity (p=0.151), and the Shapiro-Wilk test did not indicate a strong departure from residual normality (p=0.425). The model condition number was 5.39, with no evidence that collinearity between age and sex destabilized the coefficients. These diagnostics supported the parsimonious model while HC3 inference remained the prespecified primary approach.

Six observations exceeded the conventional Cook-distance screening threshold of 4/n. They were retained because a threshold exceedance identifies observations for evaluation rather than proving error or invalidity. In leave-one-out models, the male-versus-female coefficient ranged from 2.15 to 2.93 kg/m^2. Every refit therefore retained a positive difference of similar order, showing that the primary point estimate was not generated by any single complete case.

3.6 Robustness across analytic specifications

The 5,000-resample bootstrap 95% confidence interval for the sex coefficient was 0.34 to 4.88 kg/m^2, close to the HC3 interval. Adding head-coverage category produced an adjusted sex estimate of 2.80 kg/m^2. When the first member of each conflicting duplicate pair was retained, the estimate was 2.48 kg/m^2; retaining the second produced an estimate of 2.88 kg/m^2. Huber robust regression yielded a positive estimate of 2.08 kg/m^2. These specifications are compared directly in Figure 1B and enumerated in Table 2. Their agreement in direction and broadly similar magnitude indicates that the main point estimate was not highly sensitive to the documented duplicate choice, head-coverage adjustment, or residual down-weighting.

3.7 BMI completeness and missing-outcome sensitivity

Six audited decedents lacked BMI. Completeness did not differ clearly by audited sex in Fisher's exact test (p=0.171), but it varied by head-coverage category (Fisher's exact p=0.017). The underlying denominators, available BMI counts, missing counts, and percentages for every sex and head-coverage category are detailed in Table 4. Head coverage was derived from archive descriptions and may represent differences in series selection or upload structure rather than a biological determinant of missing anthropometry. Its association with completeness supported including a head-coverage-adjusted specification and cautioned against assuming that the remaining missing outcomes were completely random.

In the deterministic pattern-mixture analysis, the adjusted sex association lost conventional statistical significance when missing BMI values were assigned values 2.5 kg/m^2 below their model-based predictions. The point estimate remained positive at that threshold. Reversal of the estimated direction required a much larger shift, near -22.5 kg/m^2. The full trajectory of the adjusted estimate and its 95% confidence interval across deltas from -30 to +4 kg/m^2 is shown in Figure 2. The two thresholds address different questions: a relatively modest departure could move the confidence interval across zero, whereas a far more extreme common shift was necessary to change which sex had the higher adjusted mean. These results demonstrate greater stability of direction than of statistical significance under adverse assumptions about the missing male BMI values.

Table 1. Characteristics of audited BMI-complete decedents.

Characteristic	Overall	Female	Male
Complete-case participants, n	83	28	55
Age, years, mean (SD)	53.2 (20.5)	58.1 (20.3)	50.7 (20.4)
Height, m, mean (SD)	1.7 (0.1)	1.7 (0.1)	1.8 (0.1)
Weight, kg, mean (SD)	73.0 (17.4)	62.7 (16.6)	78.2 (15.4)
BMI, kg/m2, mean (SD)	24.2 (4.7)	22.7 (5.6)	24.9 (4.1)
BMI, kg/m2, median (IQR)	24.4 (21.4-27.1)	23.5 (18.8-27.2)	24.6 (22.7-27.1)

Values are n, mean (SD), or median (IQR). BMI, body mass index; IQR, interquartile range; SD, standard deviation.

Table 2. Age and sex associations with BMI in primary and sensitivity models.

Model	Term	Estimate	95% CI	p	n
Primary audited cohort	Age, per 10 years	0.49	0.05 to 0.93	0.029	83
Primary audited cohort	Male vs female	2.58	0.18 to 4.99	0.036	83
Adjusted for head-coverage category	Male vs female	2.80	0.29 to 5.31	0.029	83
All records before duplicate handling	Male vs female	2.69	0.51 to 4.88	0.016	97
Deduplicated; first conflicting record retained	Male vs female	2.48	0.15 to 4.81	0.037	87
Deduplicated; second conflicting record retained	Male vs female	2.88	0.54 to 5.21	0.016	87

BMI, body mass index; CI, confidence interval. Primary and ordinary least-squares sensitivity intervals use HC3 standard errors; Huber robust regression is displayed in Figure 1.

Table 3. Cohort audit and metadata reconciliation flow.

Audit Item	Value
Downloaded archives represented in metadata	105
Calibration phantom excluded	1
Documented duplicate CT pairs	11
Concordant duplicate pairs collapsed	7
Conflicting duplicate pairs excluded	4
Unique decedent records in audited cohort	89
BMI values recovered from curator companion	13
BMI-complete decedents	83
BMI-missing decedents	6
Audited sex corrections applied	6

Conflicting duplicate pairs were excluded from the primary cohort; alternative one-record selections were retained for sensitivity analysis.

Table 4. BMI availability by audited sex and head-coverage category.

Stratum	Category	Total n	BMI available, n (%)	BMI missing, n (%)	Fisher exact p
Sex	Female	28	28 (100.0)	0 (0.0)	0.171
Sex	Male	61	55 (90.2)	6 (9.8)	
Head coverage	No head	32	27 (84.4)	5 (15.6)	0.017
Head coverage	Unspecified	33	33 (100.0)	0 (0.0)	
Head coverage	Head included	24	23 (95.8)	1 (4.2)	

Percentages use the audited cohort denominator within each category. Fisher exact p-values compare all categories within each stratum; the p-value is shown once per stratum. BMI, body mass index.

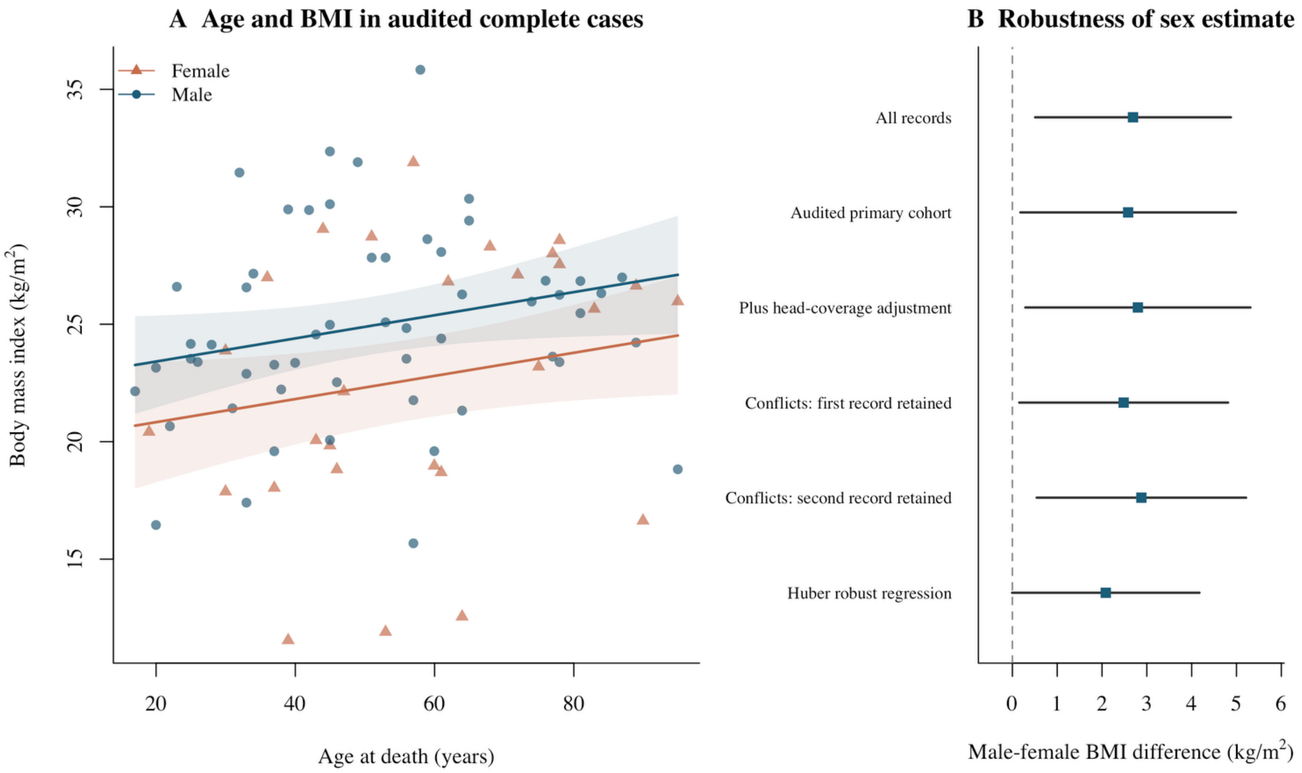

Figure 1. Age-BMI association and robustness of the adjusted sex estimate. Panel A shows audited complete cases and model-based sex-specific fitted lines with 95% confidence bands. Panel B compares the male-female BMI coefficient across duplicate-handling, head-coverage, and robust-regression specifications. BMI, body mass index.

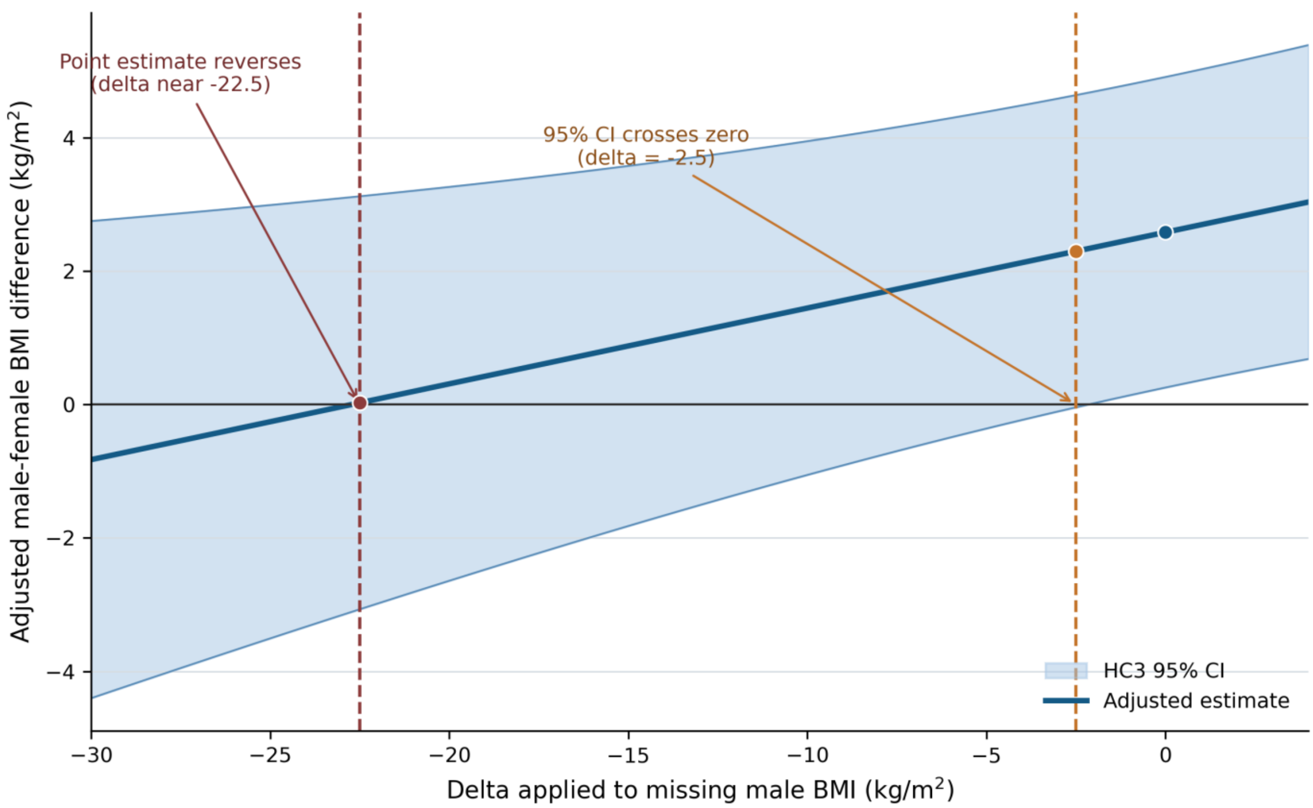

Figure 2. Pattern-mixture sensitivity of the adjusted male-female BMI estimate. The solid line shows the adjusted estimate after assigning each missing male BMI its model prediction plus the displayed delta; the shaded region shows the HC3 95% confidence interval. The interval crosses zero at delta=-2.5 kg/m², whereas point-estimate reversal occurs near delta=-22.5 kg/m². BMI, body mass index; CI, confidence interval.

.

4. Discussion

4.1 Principal findings

This curator-integrated analysis of an open PMCT metadata collection produced three principal findings, summarized across Tables 1-4 and Figures 1-2. First, the raw archive count did not represent the final number of unique analyzable decedents: removing a calibration phantom, reconciling 11 duplicate pairs, and applying curator corrections changed the cohort from 105 rows to 89 audited individuals, 83 of whom had complete age, sex, and BMI. Second, older age and male sex were independently associated with modestly higher BMI. The fitted age difference was 0.49 kg/m^2 per decade, and the adjusted male-female difference was 2.58 kg/m^2. Third, the sex estimate remained positive and of similar magnitude across bootstrap, head-coverage, duplicate-choice, Huber, and leave-one-out analyses, although missing-outcome assumptions affected whether its confidence interval excluded zero.

These findings contribute at two levels. Substantively, they describe the age- and sex-related anthropometric structure of the reconciled VSDFullBody cohort. Methodologically, they show how curator documentation changes denominators, subgroup classification, outcome availability, and the uncertainty attached to an open-data estimate. The second contribution is central because secondary imaging analyses often begin with downloadable files whose internal relationships are not self-evident. A transparent audit converts archive-level availability into a defensible person-level analytic cohort.

4.2 Interpretation of the sex association

The adjusted difference of 2.58 kg/m^2 indicates that male decedents had a higher mean BMI than female decedents at the same modeled age in this cohort. The confidence interval was wide enough to include both a small difference and a substantially larger one. This uncertainty is expected from an analytic sample of 83 with an unequal sex distribution and appreciable within-group variation. The estimate should therefore be interpreted as a cohort-specific quantitative contrast rather than a precise reference value for forensic populations.

The observed direction is consistent with recent large-scale anthropometric evidence showing higher average BMI and waist circumference among men in some populations, while women may have higher total body-fat percentage and subcutaneous adipose tissue.[28,29,32] CT-based normative work likewise demonstrates that sex differences depend on the compartment measured: BMI, visceral fat, subcutaneous fat, muscle area, and tissue attenuation are related but not interchangeable constructs.[24–26,30,31] The present metadata analysis cannot determine which tissue compartments produced the BMI contrast. A higher BMI among male decedents could reflect differences in lean mass, skeletal frame, adipose distribution, or combinations of these features.

The sex pattern differs from the illustrative population described in the original VSD report, where female BMI was slightly higher than male BMI in a selected femur-analysis sample.[14] This is not necessarily a contradiction. The earlier example and the present analysis involve different anatomical subsets, eligibility decisions, metadata processing, and possibly different balances of age and referral characteristics. The comparison instead underscores that a result from one derivative repository cohort should not be generalized to every archive or to the broader population without a reproducible cohort flow.

Sex-associated BMI differences can also reflect selection into a medicolegal sample. Manner and cause of death, hospitalization, chronic disease, frailty, occupational exposure, and circumstances of discovery may differ between female and male decedents. None of those variables was available for adjustment. The model therefore estimates an association conditional on age, not a biological effect attributable solely to sex. Consistent use of association language preserves this distinction and aligns the interpretation with what the data can support.

4.3 Interpretation of the age association

The age coefficient was statistically distinguishable from zero in the adjusted model but modest in magnitude. A fitted increase of 0.49 kg/m^2 per decade corresponds to approximately 2.45 kg/m^2 across 50 years, assuming linearity within the observed range. That difference may be relevant when characterizing the composition of a research repository, but it should not be described as rapid age-related change or direct evidence of an individual gaining weight over time. The cross-sectional coefficient combines differences among people born in different periods, survival, referral, and other cohort processes.

Population and imaging evidence demonstrates why a simple biological-aging interpretation would be incomplete. Large European cohorts report age-dependent changes in BMI and other anthropometric measures that differ by sex and health status.[28] CT and MRI studies further suggest that age can be associated with redistribution among visceral, subcutaneous, marrow, hepatic, and muscle fat compartments even after considering BMI.[24,25,31] The small positive age-BMI slope in VSDFullBody may therefore coexist with more complex unmeasured changes in body composition.

The Pearson and Spearman analyses reinforce the need to focus on magnitude rather than dichotomized significance. Both correlations were positive and small, but only the rank coefficient had $p<0.05$. Adjustment for sex in the primary model also produced a more precise age estimate than the unadjusted Pearson result. These patterns can arise when sex composition, non-normal spacing, or influential values obscure a modest monotonic relationship in a crude linear correlation. They do not justify selecting the most favorable test; rather, the prespecified regression provides the main estimate and the correlations describe complementary aspects of the data.

4.4 Meaning of the metadata audit

The curator workbook materially increased analyzability. Thirteen BMI values that were unavailable in the embedded extraction became calculable after documented height and weight were joined, and six retained sex labels were reconciled. Without those steps, the complete-case sample would have been smaller and subgroup composition would have differed. The audit therefore influenced both precision and the scientific meaning of the model. The complete effect of these decisions on cohort flow is detailed in Table 3. It also demonstrated that a repository's correction layer can be as important as the primary data table for valid reuse.

Duplicate handling had a smaller effect on the final sex coefficient than metadata recovery, but it remained essential. Counting repeated archives as independent observations would violate the assumption that rows contribute independent information. Selecting one conflicting record without a rule would make the result depend on an undocumented preference. Excluding both members of conflicting pairs was conservative and transparent; repeating the analysis with each alternative member then quantified the cost of that decision. The resulting estimates of 2.48 and 2.88 kg/m^2 bracketed the primary estimate closely, supporting its numerical stability.

This workflow operationalizes reusability and provenance in practical terms. Reuse does not end when a file can be downloaded. It requires a clear unit of observation, machine-readable identifiers, documented relationships among files, a hierarchy for corrections, and an audit trail connecting raw records to reported estimates. Repository curators can facilitate this process by publishing versioned correction tables and stable duplicate identifiers. Secondary analysts can reciprocate by reporting how those tables were joined and by sharing the resulting cohort logic without redistributing restricted content.

4.5 Statistical stability, uncertainty, and missingness

Several robustness analyses addressed distinct questions. HC3 inference reduced dependence on constant residual variance; the bootstrap evaluated resampling variability; Huber regression reduced the effect of large residuals; leave-one-out analysis tested dependence on individual records; head-coverage adjustment evaluated a repository-structure covariate; and alternative duplicate cohorts tested unresolved record choice. Agreement across these methods is more informative than repeating one model with minor cosmetic changes. The results consistently placed the male-female difference near 2-3 kg/m^2, as detailed in Table 2 and Figure 1.

Stability of the point estimate did not eliminate uncertainty. The primary HC3 interval began at 0.18 kg/m^2, and the pattern-mixture analysis showed that assigning missing male outcomes 2.5 kg/m^2 below prediction moved the interval across zero. Thus, evidence for the exact lower bound and conventional significance was sensitive to plausible adverse missingness. By contrast, reversing the point estimate required a common shift near -22.5 kg/m^2, far larger than the shift needed to affect significance. Reporting both thresholds distinguishes two scientifically different conclusions: the direction was comparatively robust, while the strength of evidence against a null difference was less robust.

The association between head coverage and BMI completeness suggests structured data availability, but it does not identify a missingness mechanism. The category-specific counts are detailed in Table 4, and the corresponding pattern-mixture stress test is shown in Figure 2. The description may proxy upload practice, scan protocol, or archive composition. It need not cause anthropometric missingness, and it does not reveal the BMI values that were absent. Pattern-mixture analysis is appropriate in this setting because it exposes how conclusions change over stated unobserved-value assumptions rather than treating model predictions as known truth. Multiple imputation would not remove the need for sensitivity analysis when the determinants of missing outcomes are incompletely observed.

4.6 Forensic and imaging implications

Forensic imaging repositories are frequently used to develop identification, trauma, and anthropometric methods. PMCT offers reproducible geometry and permanent digital review, while autopsy and ancillary tests retain advantages for some soft-tissue, microscopic, toxicological, and pathological findings. Comparative studies show that the two approaches identify overlapping but nonidentical findings and are often complementary.[1–4] The current study does not compare diagnostic modalities, but it benefits from the same principle: a digital examination becomes more valuable when its context and limitations are retained alongside the images.

The reconciled BMI distribution can inform planning for future VSDFullBody analyses. Investigators developing segmentation or anthropometric algorithms can see that the available complete-case cohort spans variation in age and BMI but is not balanced by sex and is not a population reference sample. They can prespecify whether a method requires voxel data, known measured weight, full anatomical coverage, or additional covariates. The six remaining missing BMI values and the description-based coverage categories should be considered when selecting training and validation subsets.

A future voxel-level study could move beyond the present claim ceiling by segmenting visceral and subcutaneous adipose tissue, skeletal muscle, and bone and then evaluating how these compartments relate to BMI, age, and sex. Normative CT research demonstrates that automated or semiautomated measurements can reveal sex- and age-specific patterns not visible in BMI alone.[24–27] Such work would require validation of segmentation quality, consistent anatomical landmarks, scanner and reconstruction information, and careful consideration of postmortem changes. It should not be inferred from metadata-only results.

Standardized postmortem anthropometry would also strengthen forensic reuse. Measured body weight and height, timing relative to imaging, postmortem interval, decomposition status, major fluid administration, amputation or tissue loss, and cause of death would help distinguish measurement variation from biological and procedural variation. Because body-weight estimation can be inaccurate in forensic practice, recording whether a value was measured, estimated, or transferred from documentation is important.[22,23] A shared data dictionary could preserve that provenance without exposing identifying information.

4.7 Strengths and limitations

The principal strength of this study is its alignment between question, available evidence, and analytic claim. The audit used curator-documented relationships rather than inferring duplicates solely from similar values. Corrections were applied through a reproducible hierarchy, and all denominators were traced from archive rows to complete cases. The primary model was parsimonious and accompanied by robust standard errors, bootstrap resampling, duplicate-choice analysis, head-coverage adjustment, influence assessment, robust regression, and a transparent missing-outcome stress test. Tables and machine-readable outputs allow numerical reconciliation.

The study also has important limitations. VSDFullBody is a convenience collection of forensic decedents rather than a probability sample, so its BMI distribution and coefficients cannot estimate population prevalence or serve as universal reference standards. The source sex variable was binary and did not represent gender identity or the full diversity of sex characteristics. The sample size limited precision and restricted the complexity of defensible models. Cause and manner of death, ancestry, clinical history, nutritional status, postmortem interval, decomposition, and acquisition details were unavailable or incomplete, leaving substantial potential for selection and residual confounding.

Anthropometric values were obtained from repository metadata rather than independently remeasured by the present investigator. The original procedures and measurement error were not documented uniformly, and four duplicate pairs remained irreconcilable. Excluding both records from those pairs reduced information, although the alternative selections produced similar estimates. Six audited records still lacked BMI, and their unobserved values could affect interval-based conclusions under adverse assumptions. The deterministic sensitivity analysis clarifies that vulnerability but cannot identify the true missing outcomes.

Finally, no CT voxel data were analyzed. BMI cannot describe tissue distribution and should not be treated as a radiological body-composition endpoint.[24,32] Postmortem changes, positioning, scan range, and reconstruction parameters might influence future voxel-based measurements even though they do not change the arithmetic definition of metadata BMI.[5,6] The present findings should therefore guide audit and cohort characterization, not substitute for direct image analysis or for validated living-population anthropometry.

5. Conclusion

After curator-informed reconciliation of a calibration object, documented duplicate relations, sex labels, and anthropometric fields, the VSDFullBody collection yielded 89 unique decedents and 83 complete cases. Older age and male sex were associated with modestly higher BMI, with adjusted estimates of 0.49 kg/m^2 per decade and 2.58 kg/m^2 for male relative to female decedents. The sex estimate remained positive across bootstrap, influence, robust-regression, head-coverage, and duplicate-choice analyses, although missing-outcome assumptions affected interval-based significance. These findings are cohort-specific anthropometric associations, not measures of CT tissue composition or population prevalence. Joining curator corrections before cohort definition and reporting an explicit archive-to-decedent audit can strengthen the reproducibility and interpretability of open forensic imaging research.

Declarations

Use of generative artificial intelligence. The authors declare that no generative artificial intelligence (AI) tools were used in the preparation, drafting, analysis, or editing of this manuscript.

Author contributions. Both authors (RH and MY) contributed to the study design, data curation and audit, analysis, and drafting, and both reviewed and approved the final manuscript.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest. The authors declare no competing interests.

Ethical approval. Ethical approval was granted by the CMHC Ethics Committee Indonesia (No. 008/CMHC/Ethics/08/2025).

Consent for publication. No new consent was obtained for this anonymized secondary open-data analysis; the accessed repository page does not state the original consent process.

Data availability. VSDFullBody and its curator companion workbook are available from Zenodo (doi:10.5281/zenodo.8270365) under the stated CC BY-NC-SA license. The analysis code, audited cohort, and machine-readable outputs accompany the submission package.

References

1. Lathrop SL, Wiest PW, Andrews SW, et al. Can computed tomography replace or supplement autopsy? J Forensic Sci. 2023;68(2):524-535. doi:10.1111/1556-4029.15217.
2. Mentink MG, Latten BGH, Bakers FCH, et al. Efficacy of postmortem CT and tissue sampling in establishing the cause of death in clinical practice: a prospective observational study. J Clin Pathol. 2024;77(4):259-265. doi:10.1136/jcp-2021-207946.
3. Speelman AC, Engel-Hills PC, Martin LJ, et al. Postmortem computed tomography plus forensic autopsy for determining the cause of death in child fatalities. Pediatr Radiol. 2022;52(13):2620-2629. doi:10.1007/s00247-022-05406-7.
4. Likkachai K, Wongwaisayawan S, Siriwes K, et al. Abbreviated injury scale-guided assessment of traumatic deaths: postmortem CT versus autopsy. Forensic Sci Int Synerg. 2025;10:100588. doi:10.1016/j.fsisyn.2025.100588.
5. Hansen PL, Leth PM, Nielsen PA, et al. Optimization of postprocessing parameters for abdominal forensic CT scans. Forensic Sci Int Synerg. 2024;8:100478. doi:10.1016/j.fsisyn.2024.100478.
6. O'Donnell C, Iles L, Woodford N, et al. Post-mortem CT lung findings at a medicolegal institute in SARS-CoV-2 RT-PCR positive cases with autopsy correlation. Forensic Sci Med Pathol. 2021;17(4):611-620. doi:10.1007/s12024-021-00389-7.
7. Chatzaraki V, Thali MJ, Ampanozi G, et al. Diagnostic accuracy of postmortem computed tomography for bleeding source determination in cases with hemoperitoneum. Int J Legal Med. 2021;135(2):593-603. doi:10.1007/s00414-020-02472-0.
8. Mizuno S, Ono S, Makino Y, et al. Validity of dental findings for identification by postmortem computed tomography. Forensic Sci Int. 2022;341:111507. doi:10.1016/j.forsciint.2022.111507.
9. Zirn A, Scheurer E, Lenz C, et al. Automated detection of fatal cerebral haemorrhage in postmortem CT data. Int J Legal Med. 2024;138(4):1391-1399. doi:10.1007/s00414-024-03183-6.
10. Makino Y, Unuma K, Nolte KB, et al. Accuracy of forensic pathologists in incorporating post-mortem CT (PMCT) in forensic death investigation. J Forensic Sci. 2022;67(6):2351-2359. doi:10.1111/1556-4029.15131.
11. Lyness JR, Collins AJ, Rutty JE, et al. Comparison of findings identified at traditional invasive autopsy and postmortem computed tomography in suicidal hangings. Int J Legal Med. 2022;136(6):1865-1881. doi:10.1007/s00414-022-02874-2.
12. Levin JH, Pecoraro A, Ochs V, et al. Characterization of fatal blunt injuries using postmortem computed tomography. J Trauma Acute Care Surg. 2023;95(2):186-190. doi:10.1097/TA.0000000000004012.
13. Hammarlebiod S, Farrugia A, Bierry G, et al. Thermal bone injuries: postmortem computed tomography findings in 25 cases. Int J Legal Med. 2022;136(1):219-227. doi:10.1007/s00414-021-02708-7.
14. Kistler M, Bonaretti S, Pfahrer M, et al. The Virtual Skeleton Database: an open access repository for biomedical research and collaboration. J Med Internet Res. 2013;15(11):e245. doi:10.2196/jmir.2930.
15. Kistler M. VSDFullBody: The Virtual Skeleton Database Full Body CT Collection. Version 1.0.0. Zenodo; 2013. doi:10.5281/zenodo.8270365.
16. Kondou H, Bandou R, Ichioka H, et al. Estimating the age at death for forensic cases using quantitative computed tomography. Forensic Sci Int. 2022;337:111367. doi:10.1016/j.forsciint.2022.111367.
17. Xiong J, Ma Y, Cao Y, et al. Age estimation by modified Suchey-Brooks method using three-dimensional reconstructed CT images of Chinese Han population. Leg Med (Tokyo). 2023;65:102304. doi:10.1016/j.legalmed.2023.102304.
18. Zangpo D, Uehara K, Kondo K, et al. Estimating age at death by Hausdorff distance analyses of the fourth lumbar vertebral bodies using 3D postmortem CT images. Forensic Sci Med Pathol. 2024;20(2):472-479. doi:10.1007/s12024-023-00620-7.
19. Lopatin O, Barszcz M, Wozniak KJ, et al. Skeletal and dental age estimation via postmortem computed tomography in

Polish subadults group. Int J Legal Med. 2023;137(4):1147-1159. doi:10.1007/s00414-023-03005-1.

20. Pham CV, Lee SJ, Kim SY, et al. Age estimation based on 3D post-mortem computed tomography images of mandible and femur using convolutional neural networks. PLoS One. 2021;16(5):e0251388. doi:10.1371/journal.pone.0251388.

21. Lye R, Min H, Dowling J, et al. Deep learning versus human assessors: forensic sex estimation from three-dimensional computed tomography scans. Sci Rep. 2024;14(1):30136. doi:10.1038/s41598-024-81718-y.

22. Merriam T, Enders M, Bolliger SA, et al. Anthropometric weight estimation is less accurate than visual weight estimation in forensic postmortem cases in a Swiss population. J Forensic Leg Med. 2024;108:102787. doi:10.1016/j.jflm.2024.102787.

23. Schenkl S, Muggenthaler H, Koch R, et al. Automated body weight measurement using computed tomography. Sci Rep. 2025;15(1):36110. doi:10.1038/s41598-025-24060-1.

24. Wang L, Liu Y, Li K, et al. Age and BMI have different effects on subcutaneous, visceral, liver, bone marrow, and muscle adiposity, as measured by CT and MRI. Obesity (Silver Spring). 2024;32(7):1339-1348. doi:10.1002/oby.24040.

25. Marquardt JP, Tonnesen PE, Mercaldo ND, et al. Subcutaneous and visceral adipose tissue reference values from the Framingham Heart Study thoracic and abdominal CT. Invest Radiol. 2025;60(2):95-104. doi:10.1097/RLI.0000000000001104.

26. Magudia K, Bridge CP, Bay CP, et al. Population-scale CT-based body composition analysis of a large outpatient population using deep learning to derive age-, sex-, and race-specific reference curves. Radiology. 2021;298(2):319-329. doi:10.1148/radiol.2020201640.

27. Pooler BD, Garrett JW, Southard AM, et al. Technical adequacy of fully automated artificial intelligence body composition tools: assessment in a heterogeneous sample of external CT examinations. AJR Am J Roentgenol. 2023;221(1):124-134. doi:10.2214/AJR.22.28745.

28. Stein MJ, Fischer B, Bohmann P, et al. Differences in anthropometric measures based on sex, age, and health status: findings from the German National Cohort (NAKO). Dtsch Arztebl Int. 2024;121(7):207-213. doi:10.3238/arztebl.m2024.0016.

29. Brath MSG, Sahakyan M, Mark EB, et al. Ethnic differences in CT derived abdominal body composition measures: a comparative retrospect pilot study between European and Inuit study population. Int J Circumpolar Health. 2024;83(1):2312663. doi:10.1080/22423982.2024.2312663.

30. Lee MH, Zea R, Garrett JW, et al. Abdominal CT body composition thresholds using automated AI tools for predicting 10-year adverse outcomes. Radiology. 2023;306(2):e220574. doi:10.1148/radiol.220574.

31. Rule AD, Grossardt BR, Weston AD, et al. Older tissue age derived from abdominal computed tomography biomarkers of muscle, fat, and bone is associated with chronic conditions and higher mortality. Mayo Clin Proc. 2024;99(6):878-890. doi:10.1016/j.mayocp.2023.09.021.

32. Sheean P, O'Connor P, Joyce C, et al. Validating the use of body mass index with computed tomography in a racially and ethnically diverse cohort of patients admitted with SARS-CoV-2. Nutr Clin Pract. 2024;39(5):1259-1269. doi:10.1002/ncp.11164.

33. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. doi:10.1371/journal.pmed.004029